CONFIDENTIAL UNTIL PUBLISHED
Evidence Review Group's Report

Zynteglo for treating transfusion-dependent beta-thalassaemia

Produced by	CRD and CHE Technology Assessment Group, University of York, Heslington, York, YO10 5DD
Authors	Lindsay Claxton, Research Fellow, CRD
	Mark Corbett, Research Fellow, CRD
	Matthew Walton, Research Fellow, CRD
	Peter Murphy, Research Fellow, CRD
	Sumayya Anwer, Research Fellow, CRD
	Melissa Harden, Information Specialist, CRD
	Sofia Dias, Professor of Health Technology Assessment, CRD
Correspondence to	Sofia Dias, Professor of Health Technology Assessment, CRD, University of York, York YO10 5DD
Date completed	20/01/2020

Source of funding

This report was commissioned by the NIHR Systematic Reviews Programme as project number 18/54/11.

Declared competing interests of the authors

None

Acknowledgements

We thank Professor Sally Kinsey at Leeds Teaching Hospitals NHS Trust for her expert advice throughout this project.

Rider on responsibility for report

The views expressed in this report are those of the authors and not necessarily those of the NIHR HTA Programme. Any errors are the responsibility of the authors.

This report should be referenced as follows:

Claxton L, Corbett M, Walton M, Murphy P, Anwer S, Harden M, Dias S. Zynteglo for treating transfusion-dependent beta-thalassaemia: A Single Technology Appraisal. CRD and CHE, University of York, Technology Assessment Group, 2020.

Contributions of authors

Lindsay Claxton, Peter Murphy and Matthew Walton wrote the cost effectiveness sections and conducted the ERG economic analyses. Lindsay Claxton provided advice, commented on drafts of the report and took overall responsibility for the cost effectiveness sections. Mark Corbett and Sumayya Anwer wrote the clinical effectiveness sections of the report. Mark Corbett took overall responsibility for the clinical effectiveness sections. Melissa Harden wrote the sections on the search strategies. Sofia Dias provided advice, commented on drafts of the report and took overall responsibility for the report.

Note on the text

All commercial-in-confidence (CIC) data have been highlighted in blue and underlined, all academic-in-confidence (AIC) data are highlighted in yellow and underlined, all depersonalised data (DPD) are highlighted in pink and underlined.

Copyright statement

Table of Contents

Table of Tables

Table of Figures

List of abbreviations

CS	Company submission
AE	Adverse event
BBTS	British Blood Transfusion Society
BSH	British Society for Haematology
CMR	Cardiovascular magnetic resonance
CSR	Clinical study report
DICE	Discretely integrated condition event
DSA	Deterministic sensitivity analysis
EMA	European Medicines Agency
eMIT	Electronic market information tool
ERG	Evidence review group
G-CSF	Granulocyte-colony stimulating factor
GvHD	Graft-versus-host disease
Hb	Haemoglobin
HCV	Hepatitis C virus
HES	Hospital episode statistics
HLA	Human leukocyte antigen
HLA	Human leukocyte antigen
HRQoL	Health related quality of Life
HSC	Haematopoietic stem cells
HSCT	Haematopoietic stem cell transplant
HST	Highly specialised technology
ICER	Incremental cost-effectiveness ratio
ITT	Intention to treat
LIC	Liver iron content
LVV	Lentiviral vector
MRI	Magnetic resonance imaging
NICE	National Institute for Health and Care Excellence
PAS	Patient access scheme

PASLU	Patient access scheme liaison unit
PB VCN	Peripheral blood vector copy number
PFC	Point for clarification
PSA	Probabilistic sensitivity analysis
PSS	Personal social services
QALY	Quality adjusted life-year
RCPath	The Royal College of Pathologists
SAE	Serious adverse event
SLR	Systematic literature review
SmPC	Summary of product characteristics
SMR	Standardised mortality ratio
SoC	Standard of care
TDT	Transfusion dependent thalassaemia
TI	Transfusion independent
TP	Transplant Population
UKTS	United Kingdom Thalassaemia Society
VCN	Vector copy number
VOD	Veno-occlusive disease
WTP	Willingness-to-pay

1 Summary

1.1 Critique of the decision problem in the company's submission

Both the NICE scope and the Zynteglo marketing authorisation refer to a ≥ 12 years β-thalassaemia population for which hematopoietic stem cell transplantation (HSCT) is appropriate, but a matched related donor is not available. The ERG considers that the reference to HSCT being 'appropriate' should primarily be interpreted as relating to fitness to receive conditioning chemotherapy prior to an autologous cell therapy (such as Zynteglo) – rather than allogeneic HSCT, because for the overwhelming majority of NHS patients ≥ 12 years allogeneic HSCT will not be considered appropriate. Allogeneic HSCT is not recommended for adults in the UK, and it is very rarely considered in patients aged ≥12 years, as the risks outweigh the benefits. In light of this, the ERG has some concerns about the interpretation of the restriction of "a matched related donor is not available" in the ≥12 years β-thalassaemia population, as it could leave some patients without a viable transplant option if they are ≥12 years old but have a matched related donor.

For the Zynteglo clinical trials, transfusion-dependent β-thalassaemia (TDT) was defined as requiring 8 or more transfusions per year or ≥100ml /kg/year of packed red blood cells. Patients excluded from the trials were those with evidence of severe iron overload, or with hepatitis B or C, or other clinically significant active infections. The ERG's clinical adviser thought these were appropriate criteria.

The EMA SmPC noted that only a few patients homozygous for IVS-I-110 or IVS-I-5 mutations were included in the Zynteglo studies. As part of the conditional license, the EMA require the company to submit interim and final data from such patients with severe non-β^0/β^0 genotypes. These patients may be under-represented in the Zynteglo trials, considering their prevalence in the NHS population.

The intervention in the company submission was the same as that specified in the final scope. Patients may only be treated with Zynteglo once. None of the Zynteglo trials had a control arm, so the company identified data on comparator therapies through systematic reviews and their own studies. The systematic reviews were often too limited in their search dates to identify suitable studies so additional 'targeted reviews' were undertaken, but detailed methods were not reported.

The outcomes specified in the company submission (CS) matched the NICE scope, with the exception of 'symptoms of anaemia'. However, symptoms of anaemia are a consequence of low haemoglobin levels and both total haemoglobin levels and haemoglobin A (HbA^{T87Q}) levels were reported as outcomes in the CS.

1.2 Summary of clinical effectiveness evidence submitted by the company

The clinical effectiveness data on Zynteglo presented in the CS were derived from two phase 1/2 studies in patients with any genotype (studies HGB-204 and HGB-205), and one ongoing phase 3

study in patients specifically with non-β^0/β^0 genotypes (HGB-207). Of the patients recruited, [redacted] were Asian, [redacted] were white and [redacted] were classed as 'other' race.

For the primary outcome – transfusion independence (TI) – a response rate of 83% (20/24 patients) was seen in the 'transfusion evaluable' population. No events for loss of TI have been recorded so far with patients showing generally stable levels of vector copy number (the average number of vector copies per cell) in peripheral blood and HbA^{T87Q} (the haemoglobin derived from the modified stem cells). Of the four patients who did not achieve TI in the transplant population, two had 'substantial' transfusion reductions (≥60% reduction in frequency) and two were transfusion dependent.

For adults with baseline HRQoL measurements, the CS noted a general trend of [redacted]. The company concluded that the EQ-5D data collected in the Zynteglo trials may not accurately reflect the HRQoL of patients treated with Zynteglo.

The company used summary data from both their own studies and identified published studies to derive outcomes for patients receiving comparator treatments (transfusions and chelation) in the economic model. One identified study led the company to conclude that the conditioning chemotherapy associated with Zynteglo therapy would increase infertility by 24% in men and 57% in women when compared to transfusion-dependent patients.

1.3 Summary of the ERG's critique of clinical effectiveness evidence submitted

The company stated that 31 articles were included "for qualitative synthesis" in their systematic review but no such synthesis of these articles was presented. The CS synthesis focussed on the trials HGB-204, HGB-205, and HGB-207 but no reference was made to existing conference abstracts on study HGB-212, and no unpublished data on this trial were presented.

The ethnicity distribution of the Zynteglo trial population does not appear to well represent the UK TDT population. In the UK, around 10-15% of patients with thalassaemia are white, while [redacted] of the trial patients were white. This may mean that the proportion of the trial population with a severe non-β^0/β^0 IVS-I-5 genotype is not representative of the UK setting, since IVS-I-5 mutations are quite common in thalassaemia patients of Indian or Pakistani descent. Furthermore, patients with an IVS-I-110 severe non-β^0/β^0 genotype were excluded from HGB-207. While most patients appear to respond well to Zynteglo, the small subgroup of patients who are homozygous for these mutations, or heterozygous for IVS-I-110 or IVS-I-5 together with a β^0 mutation, appeared less likely to achieve transfusion independence than other non-β^0/β^0 genotype patients. The ERG is concerned that there is very little evidence (both now and expected in the future) for a key subgroup of patients seen in the

NHS – those with IVS-I-5 mutations – and believes that the possible impact of heterogeneity across trials on transfusion outcomes should have been considered by the company.

The Zynteglo manufacturing processes have evolved during the trial programme with the aim of [redacted]. It is possible that [redacted] may increase the probability of achieving TI in patients with severe non-β^0/β^0 genotypes (compared to the previous processes used in the trial programme) but only results from the study HGB-212 and further data from HGB-207 can resolve this uncertainty. The primary outcome of HGB-212 is transfusion reduction (in HGB-207 it was transfusion independence) suggesting lower expectations of a TI response in patients with more severe genotypes.

The trials results remain immature, and the number of patients treated is small, so uncertainty exists regarding the persistence of the Zynteglo treatment effect, and the possibility of adverse events in the medium-to-long term. Zynteglo appears to have an acceptable short-term safety profile.

The company's systematic reviews to identify comparator group data were restricted to studies published from 2007 onwards and review articles appear to have been excluded. These criteria proved to be too restrictive as the company had to undertake additional "targeted reviews" to identify sufficient evidence. The justification for selecting specific studies from these reviews was often not presented. The ERG considers that the applicability of the comparator data is not optimal as some studies do not reflect the improvements in TDT patient treatment and management achieved over the last 10-20 years. This was exacerbated by the limited or absent critiques of the likely limitations of many studies, and their possible implications. The company commissioned a 'chart review' of the medical records of UK patients with TDT. However, the Chart Review also had limitations in how well it reflected the Zynteglo trial population [redacted]

1.4 Summary of cost effectiveness submitted evidence by the company

The company's submission of economic evidence included a systematic review to identify previous economic analyses. One overall search was used to identify studies on the cost-effectiveness of treatments, health-related quality of life (HRQoL), and cost and resource utilisation in patients with TDT. The review identified five cost-effectiveness models and one cost-of-illness model. The remaining nine publications were resource use/cost studies. Two of the identified cost-effectiveness analyses were used to inform clinical parameters in the company's model.

The company presented a *de novo* economic analysis of Zynteglo compared with the standard of care (SoC), consisting of blood transfusions and iron chelation therapy, in transfusion-dependent beta thalassaemia. Treatment effectiveness was assessed through the achievement of transfusion-independence. Total costs and QALYs were assessed over a lifetime time horizon, and discounted at a rate of 1.5% for each arm.

The model used a discrete event simulation structure, implemented through the discretely integrated condition event (DICE) simulation framework. The model structure was driven by the transfusion status of patients, which determined their tissue-specific iron levels. A patient's iron level then determines the risk of developing complications attributable to iron overload, also influences mortality risk, quality of life, and chelation requirements.

Zynteglo, as implemented in the model, consists of three stages, each comprising distinct processes and treatment costs. These include patient stem cell mobilisation and apheresis, myeloablative conditioning, and administration of the transduced cells. The comparator considered in the company's model was 'current care', which consists of regular transfusions and iron chelation therapy. In the economic model, patients were allocated to either oral (deferasirox, deferiprone), subcutaneous (desferrioxamine) or a combination of two different chelation therapies.

The clinical effectiveness of Zynteglo in the model is assessed by the proportion of patients achieving transfusion independence. This was estimated from pooled results in all non-β^0/β^0 genotype subjects from studies HGB -204, HGB-205, and HGB-207. The engraftment procedure was assumed to be successful in all patients. In the long-term, it was assumed that there would be no loss of graft, i.e. transfusion independence status was permanent, and that transfusion-reduced patients would not experience an increase in the need for transfusions or return to transfusion dependence over time.

Modelled baseline levels of iron overload are based on the mean of the Chart Review population, with each simulated patient randomly assigned to an overload risk category (low, medium, high) for the cardiac, liver, and endocrine systems. To predict the complications of iron overload, the model uses literature-based rates and risk equations to estimate the rate of developing complications based on distribution of iron levels in the heart, liver, and serum (ferritin).

In the absence of direct long-term survival data for the transfusion dependent population after treatment with Zynteglo, the company used a range of external sources to predict long-term survival. The company applied a standardised mortality ratio based on transfusion-dependence status to general population mortality rates. Mortality following the development of cardiac complications was modelled separately, to account for the specific impact of this aspect of the condition.

The health-related quality of life (HRQoL) estimates used in the company's base-case analysis for patients achieving transfusion independence were based on a vignette study conducted by the company. Owing to a perceived lack of appropriate HRQoL data for transfusion dependent patients, the company commissioned the Chart Review of the medical records of UK TDT patients. Patients' EQ-5D-3L and EQ-5D-Y questionnaire data were used to generate utility scores. The mean of these scores was applied to transfusion dependent patients in the model. The model also incorporates disutility increments associated with chelation, infertility, cardiac complications, liver complications, and endocrine complications.

Resource use and costs include: drug acquisition costs; Zynteglo pre-treatment and administration costs; post-Zynteglo infusion monitoring costs; treatment and monitoring costs for blood transfusions and iron chelation therapy; costs associated with managing iron overload-related complications; and the costs of managing adverse events associated with iron chelation therapy.

The company found Zynteglo to be more costly (cost difference of [redacted]) and more effective (13.14 QALYs gain) compared to SoC. The deterministic base case incremental cost-effectiveness ratio (ICER) for Zynteglo was [redacted] per QALY gained, the mean probabilistic ICER was [redacted] per QALY gained. The majority of the additional QALYs were generated as a result of additional life years; however, improvements in quality of life also had a considerable impact on the incremental QALYs. The company reported that the most influential parameters in the one-way sensitivity analysis included the iron chelation acquisition cost and the distribution of oral iron chelation therapies across patients.

1.5 Summary of the ERG's critique of cost effectiveness evidence submitted

The ERG highlights a significant number of issues that contribute to uncertainty in the cost-effectiveness results presented by the company.

By reflecting iron overload complications in only the liver, cardiac and endocrine systems, the company's model may have omitted important aspects of the condition, such as splenectomy and the development of osteoporosis. These were included in other cost-effectiveness analyses for this condition, and clinical advice suggested that improvements in patient care mean treating the consequences of TDT in older patients (such as osteoporosis) is increasingly important.

The ERG did not consider the model to adequately correlate patient characteristics and key outcomes. Historic iron levels were not considered in rate of complication development, or time to iron normalisation, which was assumed to be four years in TI patients regardless of prior iron levels. Furthermore, the risk of developing iron overload-related complications depended only upon the current iron load category, rather than accounting for the severity or duration of historic iron overload,

or the patient's age. This meant that very few patients in the Zynteglo arm developed cardiac complications, as they were assumed to be at no risk after four years.

The ERG identified several issues regarding the composition of the modelled population. These include the age, weight, iron load, comorbidities, and genotype of patients. The age of patients included in the Chart Review (the source of comparator HRQoL data and resource use) did not match the NICE Scope or the Zynteglo trial population. A large proportion of Chart Review patients were aged over 35, with some over 60 years of age, had a number of co-morbidities that would have precluded treatment with Zynteglo, and had other complications whose disutilities would be double counted by the model, thereby introducing bias in favour of Zynteglo. Modelled patient weight was assumed to be [redacted] throughout the lifetime of both paediatric and adult patients, which the ERG considered an oversimplification and an overestimation. Finally, the ERG considered the modelled population to underrepresent severe non-β^0/β^0 genotypes covered by the marketing authorisation, and may be more highly prevalent in the UK.

The intervention as implemented in the economic model matches the product licence. However, the distribution of chelating agents may not represent current clinical practice in this population. The relatively recent development of the evidence around the safety and efficacy of using a combination of agents, means there may not yet be a consensus on best clinical practice, adding further uncertainty.

The ERG has a number of concerns regarding the company's justification for the use of the non-reference case discount rate of 1.5% in the economic evaluation. The company argue that Zynteglo restores people who would otherwise die or have a very severely impaired life to full or near full health. The ERG highlighted the age of the literature cited in support of this assumption, and identified recent sources stating that patients optimally managed with currently available therapies could have a near-normal life expectancy. A number of evidence sources, including the company's own Chart Review and HGB trials, supported the notion that the impact of TDT and current management on HRQoL was not as severe as argued by the company. Furthermore, the ERG did not consider there to be sufficient evidence to conclude with certainty that Zynteglo restores individuals' health to full or near full in terms of length and quality. It is also therefore uncertain whether or not Zynteglo will commit the NHS to significant irrecoverable costs.

The ERG highlighted a number of uncertainties in the modelled treatment effectiveness, including uncertainty around the generalisability of the trials to the UK, given the potential underrepresentation of IVS-I-110 or IVS-I-5 genotypes. This may impact the overall rate of achieving transfusion independence. The ERG also highlight that there is insufficient evidence to support the assumption of permanent engraftment and an indefinite treatment effect in all patients. Furthermore, the ERG considers that cited evidence did not support the assumption of iron normalisation in all transfusion

independent patients, and that the four year time frame may be too optimistic. The source of assumptions surrounding the modelling of complications from iron overload and the mortality rate of transfusion dependent patients also may have introduced uncertainty.

A number of issues regarding the modelled HRQoL result in uncertainty in the company's results. Firstly, an inappropriate value set was selected as the basis of the general population utility estimates. The company selected a subset which excluded all individuals with a history of a health condition meaning that the baseline utility of a patient aged 75 was higher than someone aged 30 in the general population. Secondly, the utility of transfusion dependent patients was based on the full Chart Review population, despite its demographics differing substantively from patients included in the Zynteglo trials. As a result, the ERG requested the company re-analyse HRQoL data in the Chart Review, limiting the population to only those aged from 12-35 years, and excluding patients with comorbidities already separately accounted for in the model. This resulted in a higher utility for TDT patients, comparable to other literature-derived estimates.

The economic model failed to account for patients withdrawing from treatment during the pre-transplant stage. The company's submission showed one patient from the ITT population in HGB-204 discontinued Zynteglo due to inadequate stem cell mobilisation. The potential costs of this are not captured in the model, which prospectively selects only those who successfully received Zynteglo infusion. Furthermore, uncertainty remains around the cost of chelation therapy, given the uncertainty around the weight of modelled patients.

1.6 ERG commentary on the robustness of evidence submitted by the company

1.6.1 Strengths

The effectiveness of Zynteglo was compared against SoC (transfusions and iron chelation), and the outcomes assessed were appropriate.

In the company's economic model, the impact of a range of uncertainties were explored using sensitivity and scenario analysis. The company also addressed numerous additional uncertainties in response to ERG requests and clarifications. Where there was a lack of appropriate data for use in the submission, the company also endeavoured to generate appropriate evidence, notably commissioning of the Chart Review of the medical records of patients with TDT.

1.7 Weaknesses and areas of uncertainty

The main weaknesses and areas of uncertainty identified by the ERG include:

The representativeness of the trial population

The Zynteglo trial results have uncertain applicability to the population likely to receive Zynteglo in the NHS as the trial population might under-represent certain genotypes which are prevalent in the UK.

Heterogeneity of effect

The level/extent of heterogeneity of effect (i.e. in achieving TI) remains an area of uncertainty. Heterogeneity based on genotype and the evolving manufacturing process is not addressed in the evidence.

Immaturity of the data

The trials results are still immature, and the number of patients treated is small, so uncertainty exists regarding the longevity of the Zynteglo treatment effect, and the possibility of adverse events in the medium-to-long term.

Model structure and parameters

There is a lack of interaction between patient characteristics and key outcomes in the economic model. The model uses mean cohort values for all patients and fails to adequately account for interaction between age, weight, event risk, and events themselves. Uncertainty also exists regarding the validity/appropriateness of many of the comparator data used in the model.

The use of the Chart Review

The Chart Review provided the source of many of the model inputs, including HRQoL data for transfusion dependent patients, the proportion of patients on chelation therapy and the average weight of patients. However, the Chart Review included patients outside of the NICE scope as well as patients with comorbidities, which would preclude them from treatment with Zynteglo.

The non-reference case discount rate

The company's use of a non-reference case discount rate of 1.5% in the economic evaluation has a considerable impact on the incremental costs and QALYs of Zynteglo compared to SoC. The 1.5% discount rate leads to a substantial underestimate of the company's ICER. Considerable uncertainty remains regarding the company's justification for using this discount rate.

The modelled HRQoL

Due to the use of utility values for transfusion independent patients being derived from a vignette study; the selection of an inappropriate general population utility value set; and the use of utility values for transfusion dependent patients with questionable internal and external validity, considerable uncertainties remain regarding the HRQoL of Zynteglo and SoC patients.

1.8 Summary of exploratory and sensitivity analyses undertaken by the ERG

The key uncertainties addressed by the ERG scenario analyses relate to:

- The baseline characteristics of the modelled population
- The discount rate
- The use of the Chart Review as the modelled comparator
- The modelled utility decrements

The company presented additional analyses as part of their points for clarification response which include age category specific body weights; alternative proportions of chelation therapy, and alternative utility values based on the age- and comorbidity-adjusted Chart Review reanalysis.

The results of these scenario analyses, including the ERG's alternative base-case are summarised in Table 1. Due to time constraints and the nature of the model structure, deterministic ICERs are presented throughout unless otherwise stated.

The ERG alternative base-case analysis incorporated a number of alternative assumptions, a number of which were also explored by the company in scenario analyses. The changes made by the ERG include:

- Alternative discount rate of 3.5%,
- Age category specific body weight (paediatric and adult),
- 20% of the population have hypogonadism at baseline,
- Age-adjusted proportions of chelation type from the Chart Review,
- Age and comorbidity-adjusted utility values from the Chart Review,
- Age-adjustment of utilities based on values for the full general population,
- Alternative utility decrement for transfusion independent patients on subcutaneous chelation ,
- electronic market information tool (eMIT) drug acquisition costs.

Under the ERG's alternative set of assumptions, the ICER for Zynteglo versus SoC is [redacted] per QALY gained.

A scenario analysis is provided on the ERG's base case in which a discount rate of 1.5% is used for costs and QALYs. The resulting deterministic ICER is [redacted] per QALY gained. Further analyses undertaken by the ERG on their alternative base-case suggested that the mortality rate for transfusion-dependent patients was also an influential parameter in the analysis.

Table 1 Summary of ERG exploratory analyses

	Incremental costs	**Incremental QALYs**	**ICER (£/QALY)**
Base case	[redacted]	13.13	[redacted]
Scenario 1: Age category-specific body weight (paediatric and adult)	[redacted]	13.13	[redacted]
Scenario 2: 20% of population have hypogonadism at baseline	[redacted]	12.98	[redacted]
Scenario 3: Adjusting clinical effectiveness data for underrepresented genotypes	[redacted]	12.45	[redacted]
Scenario 4: Adjusted chelation therapy distribution	[redacted]	13.13	[redacted]
Scenario 5: Engraftment failure – 1%	[redacted]	13.03	[redacted]
Scenario 6: Engraftment failure – 5%	[redacted]	11.81	[redacted]
Scenario 7: 5% relapse every 10 years	[redacted]	11.26	[redacted]
Scenario 8: 10% relapse every 10 years	[redacted]	9.51	[redacted]
Scenario 9: Alternative SMR of 2 (transfusion dependent)	[redacted]	11.99	[redacted]
Scenario 10: Time to iron normalisation – 5 years	[redacted]	12.85	[redacted]
Scenario 11: Time to iron normalisation – 7 years	[redacted]	12.01	[redacted]
Scenario 12: Time to iron normalisation – 10 years	[redacted]	11.28	[redacted]
Scenario 13: Patients with normalised levels of iron face a residual risk of developing iron overload-related complications	[redacted]	9.13	[redacted]
Scenario 14: Alternative discount rate - 3.5%	[redacted]	7.29	[redacted]
Scenario 15: Age-related disutilities taken from full Ara and Brazier population	[redacted]	11.32	[redacted]
Scenario 16: Age- and comorbidity-adjusted Chart Review utility values	[redacted]	8.92	[redacted]
Scenario 17: Cumulative impact of 15 and 16	[redacted]	7.11	[redacted]
Scenario 18: Subcutaneous chelation therapy decrement for TI patients	[redacted]	13.07	[redacted]
Scenario 19: No infertility disutility	[redacted]	13.71	[redacted]
Scenario 20: eMIT drug acquisition costs	[redacted]	13.13	[redacted]
Scenario 1: Age category-specific body weight (paediatric and adult)	[redacted]	13.13	[redacted]
ERG Alternative base case analysis (deterministic)	[redacted]	**3.05**	[redacted]
ERG Alternative base case scenario analysis - 1.5% discount rate for costs and outcomes	[redacted]	6.71	[redacted]
ERG Alternative base case scenario analysis - SMR of 2 for transfusion dependent patients	[redacted]	2.48	[redacted]

2 Background

2.1 Description of the technology being appraised

Zynteglo is a gene therapy that provides functional β-globin to patients with TDT using the patient's own cells ex-vivo to correct the underlying cause of the disease. Zynteglo is an autologous $CD34^+$ cell enriched population that contains haematopoietic stem cells (HSCs) transduced with lentiviral vector (LVV) encoding the $\beta^{A\text{-}T87Q}$ globin gene. Zynteglo is administered as a single intravenous infusion, where copies of the functional $\beta^{A\text{-}T87Q}$ establish a population of undifferentiated, long-term HSCs in the bone marrow, integrating the β-globin gene into the patient's genome.

Zynteglo should be administered in a specialised treatment centre by physicians with experience in treating patients with TDT and HSC transplantation.[1] The company expressed a preference that treatment centres are co-located with haemoglobinopathy medical expertise as patients will need to be heavily chelated before treatment, to prevent complications. The processes involved in manufacturing and administering Zynteglo include mobilisation and apheresis to harvest stem cells from the patient, cryopreservation of stem cells and a back-up collection, shipping between treatment centres and the manufacturing facility, purification of stem cells and transduction of cells using a viral vector. The treatment process from mobilisation to the end of the hospital stay lasts 13-19 weeks and depends greatly on the number of cycles of mobilisation and apheresis the patient undergoes.

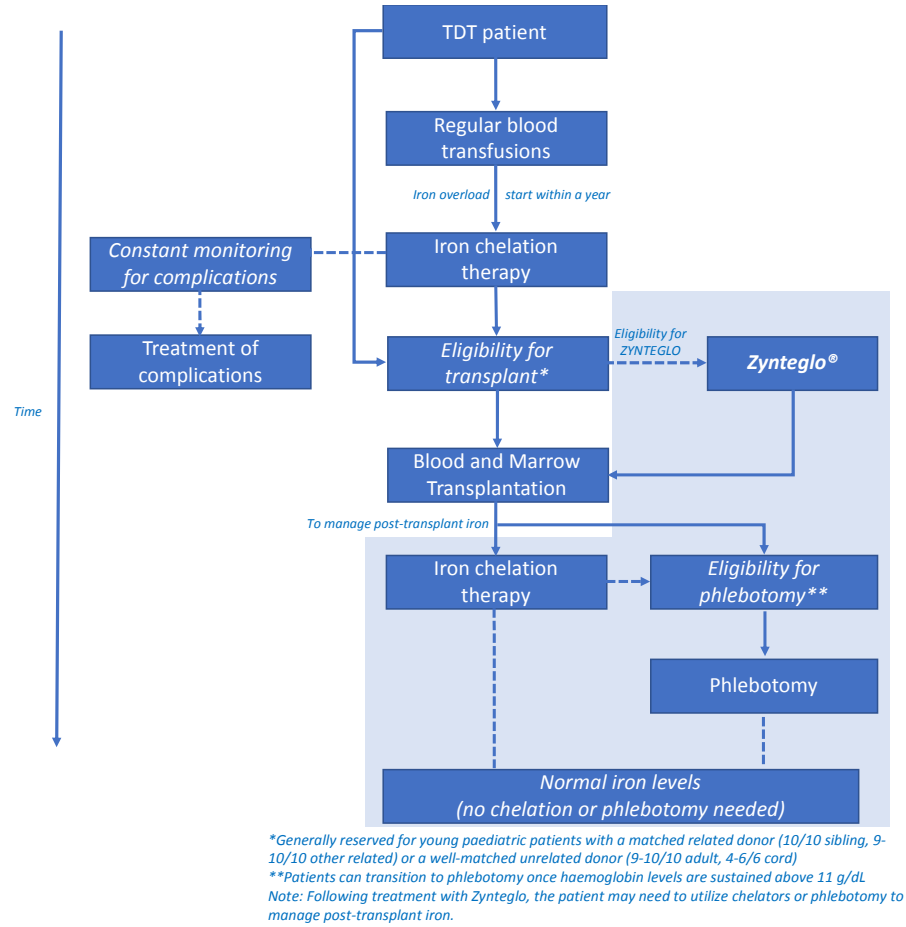
Prior to the infusion of Zynteglo, patients undergo full myeloablative conditioning using chemotherapy with busulfan to destroy the existing bone marrow stem cells. A description of the treatment process is detailed in Table 2 (p12-17) of the CS. Post infusion, iron loading in patients is managed through either phlebotomy – if the patient's unsupported haemoglobin levels reach a certain threshold – or chelation therapy if phlebotomy is not feasible. This is necessary because while transfusion-independent patients do not experience additional iron loading after infusion, previously stored iron levels need to be actively reduced.

In their submission on Zynteglo, NHS England emphasised the importance of establishing the infrastructure necessary to support treatment implementation. While the techniques used are not novel or exclusive to Zynteglo, there may be a need to expand existing services, such as apheresis, which is currently already used to gather stem cells, albeit not in TDT patients undergoing HSCT. Cell management strategies, tailored to the specific treatment would need to be established, as well as a capacity to store back-up copies of stem cells. The submissions from the Royal College of Pathologists (RCPath) and British Society for Haematology (BSH) stated that additional testing (which is not offered at present) would be required to determine patient eligibility for the treatment, though this is funded by the company.

2.2 The health condition and position of the technology in the treatment pathway

The company submission provided an overview of the standard care pathway for TDT, adapted from UKTS's guidelines (Figure 3, p 21 of CS). In Figure 14 (p 44 of the CS), reproduced here as Figure 1, the company also highlighted the position Zynteglo would occupy in the treatment pathway, where it is presented as an option for patients eligible for HSCT but without a matched, related donor.

Figure 1 Treatment pathway for TDT patients including Zynteglo (Fig 14, CS)



*Generally reserved for young paediatric patients with a matched related donor (10/10 sibling, 9-10/10 other related) or a well-matched unrelated donor (9-10/10 adult, 4-6/6 cord)
**Patients can transition to phlebotomy once haemoglobin levels are sustained above 11 g/dL
Note: Following treatment with Zynteglo, the patient may need to utilize chelators or phlebotomy to manage post-transplant iron.

The scope of the population as ascertained by the company is patients aged over 12 years with TDT with a non-β^0/β^0 genotype. The maximum age a patient can receive Zynteglo is unspecified, though it is unlikely that many older patients would receive Zynteglo since eligibility is linked to the physical condition of the patient and iron damage tends to accumulate with age. Currently the oldest patients treated with Zynteglo were [redacted] years old. Zynteglo is also a potential treatment for older patients who are ineligible for allogeneic-HSCT according to current guidelines, due to an increased risk of graft rejection and graft-versus-host disease (GvHD).

3 Critique of company's definition of decision problem

3.1 Population

The population specified in the NICE scope was "*People aged 12 years and over with transfusion-dependent beta-thalassaemia with a non-β^0/β^0 genotype, who are eligible for hematopoietic stem cell transplantation but do not have access to a matched related donor*". This is very similar to the wording of Zynteglo's licensed indication: "*...patients 12 years and older with TDT who do not have a β^0/β^0 genotype, for whom HSC transplantation is appropriate but a human leukocyte antigen (HLA)-matched related HSC donor is not available*." For the Zynteglo clinical trials, TDT was defined as requiring 8 or more transfusions per year or ≥100ml /kg/year of packed red blood cells, which the ERG's clinical adviser thought was an appropriate definition.

The reference to HSC transplantation in the marketing authorisation and the population described in the NICE scope can be read as mainly encompassing autologous HSC therapies – such as Zynteglo – rather than allogeneic HSCT, and relates to patient fitness to receive conditioning chemotherapy. Allogeneic HSCT is not recommended for adults in the UK, regardless of donor availability, and is rarely considered in patients ≥ 12 years, due to the risks outweighing the potential therapeutic benefit. The risks associated with allogeneic HSCT are greater than those associated with autologous HSCT therapies (such as Zynteglo) because the former carries the risk of severe immune reaction from GvHD and the risk of infection as a result of taking immunosuppressive agents. Autologous transplants do not carry these particular risks.

The summary of product characteristics (SmPC) for Zynteglo states:

"*Contraindications to the mobilisation agents and the myeloablative conditioning agent must be considered*", adding that:

"*HSC transplantation with myeloablative conditioning is not appropriate for patients with TDT who have evidence of severely elevated iron in the heart i.e., patients with cardiac T2* <10 msec by magnetic resonance imaging (MRI). MRI of the liver should be performed on all patients prior to myeloablative conditioning. It is recommended that patients with MRI results demonstrating liver iron content ≥15 mg/g undergo liver biopsy for further evaluation. If the liver biopsy demonstrates bridging fibrosis, cirrhosis, or active hepatitis, HSC transplantation with myeloablative conditioning is not appropriate*."

The Zynteglo trials excluded patients with cardiac T2* <10 msec by MRI and patients with evidence of liver disease defined by MRI evidence or specific elevated liver function tests (e.g. 3 times the upper limit of normal). Patients with any other evidence of severe iron overload, and patients with hepatitis B, hepatitis C or other clinically significant active infections were also ineligible for the

trials. The ERG's clinical adviser thought these were appropriate criteria for considering fitness to receive Zynteglo, since the main risks from the myeloablative conditioning would be liver failure and infections such as sepsis (which the heart must be strong enough to withstand).

The mutations which cause β-thalassaemia are regionally specific, based on four groups: Mediterranean, Asian Indian, Southeast Asian, and sub-Saharan African. Each country in a group displays a few common alleles but also a larger number of alleles are found at much lower gene frequencies.[2] It is important to consider this heterogeneity of mutations because certain mutations – such as IVS-I-110 and IVS-I-5 – are associated with dramatically reduced β-globin production, behaving phenotypically as β^0 genotypes (despite being classed as β^+genotypes). Patients with a β^0/β^0 genotype are not covered by the marketing authorisation, but 'severe' non-β^0/β^0 genotypes are covered (such as homozygous IVS-I-110 and IVS-I-5 genotypes, and IVS-I-110/β^0 or IVS-I-5 /β^0 genotypes). This despite the fact that in Zynteglo trial HGB-207, IVS-I-110 mutations were considered as being "equivalent to a β^0 mutation" and were grounds for exclusion from the study, if paired with another IVS-I-110 mutation, or a β^0 mutation.

Moreover, there is heterogeneity of response to Zynteglo in patients with β^0/β^0 genotypes (not covered by the MA) and patients with severe non-β^0/β^0 genotypes deemed equivalent to a β^0 mutation (which are covered by the MA), when compared with patients with other non-β^0/β^0 genotypes (see Section 4.2.2). Severe non-β^0 mutations are quite prevalent in UK patients, based on a study which found IVS-I-5 to be the most common mutation (22.5% and part of the 'Asian Indian' group of mutations) and IVS-I-110 the fourth most common mutation (5.5%, and part of the 'Mediterranean' group of mutations) in 1,712 unrelated β-thalassaemia carriers who required screening for antenatal diagnosis in the UK.[2] The ERG also asked the company to comment (in a point for clarification) on whether the proportion of patients with IVS-I-110 or IVS-I-5 mutations in the Zynteglo studies adequately reflects the proportion likely to be eligible in the NHS. The company stated that specific genotype data from a bluebird bio sponsored study of genotyping adult patients with β-thalassaemia from the Manchester centre for Genomic Medicine found 4 of 14 patients with non-β^0/β^0 genotypes with an underlying IVS-I-110 or IVS-I-5 mutation. The EMA's SmPC noted that only a few patients homozygous for IVS-I-110 or IVS-I-5 were included in the Zynteglo studies. As part of the conditional license the EMA require the company to "submit interim and final data from patients with a severe non-β^0/β^0 genotype such as IVS-I-110 included in Study HGB-212". Although there are prevalence estimates for individual mutations in the UK, the prevalence of severe non-β^0/β^0 genotypes appears unclear. The ERG also has concerns that severe non-β^0/β^0 patients who had IVS-I-110 mutations were excluded from one of the pivotal Zynteglo trials (HGB-207) which contributed to the submission efficacy data, even though these patients are covered by the marketing authorisation. It is acknowledged that an ongoing trial (HGB-212) is studying such patients but these data were deemed by the company to be

too immature to contribute to the CS. Of the Zynteglo trial cohort 'transplant population', [redacted] had severe non-β^0/β^0 genotypes.

Routine genotype testing is not part of usual NHS practice. The introduction of testing will be needed to identify patients who may be eligible for Zynteglo, which will have implications for infrastructure. The ERG notes that costs for genotype testing are incurred by the manufacturer (Table 58, CS).

3.2 Intervention

The intervention in the CS was as specified in the scope: Zynteglo gene therapy (previously known as LentiGlobin) which is administered as a single-dose and which should only be administered once. The therapy involves transplantation of autologous $CD34^+$ haematopoetic stem cells which have been transduced by a lentiviral vector encoding the $\beta^{A\text{-}T87Q}$-globin gene. The minimum recommended dose is 5.0 x 10^6 $CD34^+$ cells/kg.

Zynteglo therapy is comprised of multiple interacting components and processes and should therefore be considered a complex intervention. The CS reported that the complete treatment process lasts 13-19 weeks; Table 2 of the CS presented details of the various stages involved, which are:

1. Mobilisation and apheresis, in which stem cells are mobilised from the bone marrow using granulocyte-colony stimulating factor (G-CSF) and plerixafor, and are harvested via apheresis. This also includes collection of back-up cells for rescue treatment. This is followed by shipment of cells to the manufacturing facility for stem cell processing.
2. While the extracted cells are undergoing processing, the second stage, pre-treatment and myeloablative conditioning, occurs. Prior to treatment with Zynteglo, patients undergo hypertransfusion to maintain haemoglobin levels during the period in which transfusions are stopped, iron chelation is also discontinued prior to myeloablative conditioning. Patients then begin prophylaxis for veno-occlusive disease (VOD) with ursodeoxycholic acid, and for seizures using clonazepam. When the transduced cell product has been successfully manufactured and received by the administration site, patients undergo myeloablative conditioning using busulfan over the course of four days.
3. The third stage comprises the administration of the transduced cells, which is performed in a <30 minute intravenous infusion in a specialist treatment centre. This is followed by an inpatient stay of 21-42 days until engraftment of the infused cells has occurred.

Where necessary, patients may need to undergo one or more additional cycles of mobilisation and apheresis, separated by at least 14 days, in order to obtain enough cells for manufacture.

The EMA designated Zynteglo as an orphan medicine in 2013. Marketing authorisation for Zynteglo was granted by the EMA on 29th May 2019. The approval is conditional, meaning the company must

provide the EMA with results of ongoing studies to allow annual assessment of effectiveness and safety data (beginning on 29th May 2020). Zynteglo was evaluated through the EMA Adaptive Pathways programme so data will continue to be generated and re-evaluated by the EMA.

Gene therapies are different to pharmacological therapies in that they are often not fixed, but may change over time. It was evident from section B.2.1.2 of the CS that this is happening with Zynteglo (and is discussed further in this report in Section 4.2.1).

3.3 Comparators

The scope comparators were defined by NICE as "*established clinical management including blood transfusions and chelating agents*". None of the Zynteglo trials had a control arm, so data from conventional direct comparisons with Zynteglo were not available. However, the CS noted that, for the primary and key secondary outcomes of transfusion independence and transfusion reduction, a comparator cohort of patients receiving transfusions and chelation therapy would not be appropriate since patients with TDT receiving established clinical management do not spontaneously achieve transfusion independence or have significant reductions in their transfusion requirements. Consequently, the company utilised the before-and-after treatment dataset from the single-arm Zynteglo trials population. The ERG concurs with this approach for these transfusion outcomes.

For the remaining outcomes the company undertook systematic reviews to identify appropriate comparator datasets of TDT patients to evaluate:

- The clinical safety of iron chelation and transfusion therapies
- The impact of TDT and iron chelation and transfusion therapies on health-related quality of life
- How iron overload-related complication rates vary by iron levels

To obtain further comparator data the company also undertook an observational Chart Review study of UK TDT patients with the aim of describing: transfusion requirements, patient demographics and baseline clinical characteristics, routine management, clinical outcomes, quality of life and complications related to iron overload and iron chelation therapy.

Comparator data were therefore derived from a variety of sources, depending on the outcome in question. The relevance and usefulness of a given comparator study depends largely on how closely the cohort matches either the Zynteglo trial cohort, or an NHS cohort, in important factors which can affect outcomes – such as fitness for autologous HSCT transplantation and applicability of clinical management to current NHS practice, particularly with respect to chelation therapies and iron-overload monitoring. A critique of the identification of comparator data is in Section 4.3.

3.4 Outcomes

The outcomes specified in the CS matched the NICE scope outcomes with the exception of 'symptoms of anaemia'. Although the company stated this was covered in Table 1 of the CS the ERG could not find outcome data on symptoms of anaemia in the CS. Nevertheless, the ERG does not see this as an important issue since symptoms of anaemia are a consequence of low haemoglobin levels and both total haemoglobin levels and haemoglobin A (HbA^{T87Q}) levels were reported as outcomes in the CS. Only one small paragraph of the CS was used to describe growth and development outcomes. The ERG therefore requested more detailed results via a point for clarification (see Section 4.2.2.4). The primary efficacy outcomes of the single-arm Zynteglo trials varied, reflecting both the various stages of product development and the differing populations recruited (see Section 4.2.2).

4 Clinical Effectiveness

This section contains a critique of the methods of the systematic reviews of clinical effectiveness and safety data on Zynteglo and on comparators, followed by a description and critique of the included studies.

4.1 Critique of the methods of the Zynteglo review

The CS did not include a systematic review of Zynteglo studies as the company stated that this was not required since no Zynteglo studies had been conducted outside of bluebird bio. In a point for clarification (PfC) the ERG requested a systematic review of Zynteglo studies, which was subsequently provided. Separate reviews were also conducted to identify outcome data for patients receiving established clinical management (i.e. blood transfusions and chelation). These are discussed in Section 4.3).

4.1.1 Searches

Searches to identify studies of Zynteglo for the treatment of TDT were not provided in the original CS. In the points for clarification, the ERG requested that the company provide a systematic review to demonstrate that they had identified all studies of Zynteglo (Question A23, p20). The company provided a systematic literature search document containing details of the searches in Appendix G of their clarification response.

The company reported searches of the following databases on 13th November 2019: PubMed, Embase (Ovid) and the Cochrane Library (Ovid). Retrieval was limited to English language studies. In addition, the following conference websites were searched on 19th November 2019 to identify more recent conference abstracts not yet available via Embase: American Society of Hematology (ASH 2018, 2019), European Hematology Association (EHA 2019), British Blood Transfusion Society (BBTS), European Society for Blood and Marrow Transplantation (EBMT 2019) and the International Society of Blood Transfusion (ISBT 2018).

The database search strategies were clearly reported in Table 1 of Appendix G of the clarification response, with 67 studies found in total after deduplication. The terms used within the strategies for Zynteglo could potentially have been expanded to include title and abstract searches of bb305 and T87Q to ensure comprehensive retrieval of all relevant studies. Retrieval may also have been improved by searching for the specific trial codes (HGB-204, HGB-207, HGB-212) and the trial names (Northstar), particularly in EMBASE, to ensure that all relevant conference abstracts were retrieved. Indexing terms were not included in the search of PubMed and the Cochrane Library. This was most likely because both databases do not contain any specific indexing terms for Zynteglo yet. However, to ensure studies were not missed by the searches, it would have been appropriate to include some broader indexing terms for Zynteglo, such as genetic therapy, lentivirus, and beta-

globins and combine those with terms for thalassaemia. A similar approach could have been adopted for the search of Embase. Although the search of Embase included the indexing term lentiglobin bb305 this has only been available since 2017. Prior to this, broader indexing terms would have been used to index articles about Zynteglo, therefore it would have been beneficial to include some of the broader indexing terms in the strategy. The ERG ran a search in EMBASE (Ovid) including these additional terms detailed above, but did not identify and further relevant studies of Zynteglo.

The search for conference abstracts via conference websites was clearly reported in Table 2 (of Appendix G of the clarification response), detailing the specific URLs, search terms and results. However, the search of the BBTS website was missing from the table. Additional searches by the ERG identified 3 relevant conference abstracts.[3-5] These 3 conference abstracts do not appear to have been identified by the searches provided by the company.

A search of trial registers was not reported in the systematic literature search document, therefore it was not clear to the ERG that all ongoing or completed but not yet published trials of Zynteglo for beta thalassaemia had been identified. The ERG carried out searches of ClinicalTrials.gov, the WHO International Clinical Trials Registry Platform and the EU Trials Register to identify any trial register records of Zynteglo for beta-thalassaemia. The ERG searches were carried out on 18th October 2019 and retrieved 75 records: no previously unidentified trials were found.

4.1.2 Inclusion criteria

The Zynteglo review eligibility criteria were presented in Appendix G (Table 3) of the company's response to the ERG's points for clarification. These were suitably broad to allow identification of relevant studies. Appropriate screening methods were reported as being used for minimising the possibility of reviewer errors and biases affecting the final list of studies included.

4.1.3 Critique of data extraction

As mentioned in Section 4.1 the company did not initially undertake a systematic review of Zynteglo studies. The review document (Appendix G of CS) provided by the company as a response to an ERG point of clarification only went as far as the study selection phase, and did not include data extraction tables.

4.1.4 Quality assessment

The company presented the results of quality assessment for Zynteglo trials HGB-204, HGB-205 and HGB-207 as Tables 1-3 in Appendix D of the company submission. No quality assessment was provided for the ongoing study HGB-212, which is at an earlier stage than the ongoing HGB-207 study. The company did not indicate which tool had been used for quality assessment, but the ERG identified it as a modified version of the GATE framework[6] which is recommended by NICE for evaluating studies of public health interventions.[7] The methods of quality appraisal were not

described, including what the symbols represented in the "response" column of the CS Appendix D tables. To provide clarity when interpreting the company's quality assessment results the ERG looked into how the GATE framework is used. The appraisal checklist consisted of items categorised into five sections: section 1 attempts to assess external validity, whereas sections 2-4 aim to assess internal validity. Each item was assigned one of the five possible responses described in Table 2 and the reasoning behind the decision was explained in the corresponding "comments" column. In the final section, each study was awarded a grade for the overall quality with respect to the internal and external validity individually. The validity grades are described in Table 3.

Table 2 List of possible responses for the NICE quality appraisal checklist

Response	Description
++	Indicates that for that particular aspect of the study design, the study has been designed or conducted in such a way as to minimise the risk of bias.
+	Indicates that either the answer to the checklist question is not clear from the way the study is reported, or that the study may not have addressed all potential sources of bias for that particular aspect of study design.
-	Should be reserved for those aspects of the study design in which significant sources or bias may persist.
Not reported (NR)	Should be reserved for those aspects in which the study under review fails to report how they have (or might have) been considered.
Not applicable (NA)	Should be reserved for those study design aspects that are not applicable given the study design under review (for example, allocation concealment would not be applicable for case control studies).

Table 3 Description of overall study quality grading for validity

Response	Description
++	All or most of the checklist criteria have been fulfilled where they have not been fulfilled the conclusions are very unlikely to alter.
+	Some of the checklist criteria have been fulfilled, where they have not been fulfilled, or not adequately described, the conclusions are unlikely to alter
-	Few or no checklist criteria have been fulfilled and the conclusions are likely or very likely to alter.

Overall, the results of the quality assessment were not very informative or particularly well-reported. Details about the quality assessment process, such as how many researchers were involved and whether their appraisals were conducted independently, were not provided. Additionally, many items on the checklist were assigned a response without providing sufficient explanation to justify the decision. This lack of transparency in reporting the quality appraisal process means the possibility of errors or bias affecting the assessments cannot be ruled out. Some obvious inconsistencies were identified in the appraisal results. Firstly, the heading for 'Section 2' in all three appraisals was

assigned a quality assessment response. Secondly, identical explanations in the comments section were assigned different responses across studies. For instance, when asked in item 1.2 whether the eligible population was representative of the source population, the comment for all three studies was "Yes, the criteria set represent TDT patients". For studies HGB-204 and HGB-207, this item was given a response of '++', whereas for HGB-205 was given '+'. Similar inconsistencies also occurred in items 1.3, 3.4, 4.2, and 4.4.

There are no tools specifically to assess the quality of single-arm studies. The checklist used by the company is not strictly appropriate either, as many of the items were deemed 'not applicable' for all three studies. Although section 1 of the checklist assesses the external validity of the study population, the information provided was quite basic and did not describe how representative the study population would be to the eligible NHS population (e.g. in terms of genotypic and ethnicity distributions) or which effect modifiers/confounders may potentially be important.

4.1.5 Evidence synthesis of Zynteglo studies

The company stated that 31 articles were included "for qualitative synthesis" in their Zynteglo systematic review (Appendix G of the clarification response). However, no such synthesis of these articles was presented. This is likely a consequence of the company not initially undertaking a systematic review of Zynteglo studies. The CS synthesis instead focussed on the trials HGB-204, HGB-205 and HGB-207. No references were made to conference abstracts on study HGB-212.

The CS did not present a description nor a rationale regarding how data from the three Zynteglo trials were synthesised. Data from the three studies were pooled directly, without adjustment. Given that the included studies all had single-arms and small sample sizes the ERG can understand why this approach was adopted. However, the ERG believes that the issue of population and intervention heterogeneity across trials should have been considered and its possible impact discussed (see Section 4.2 of this report).

The CS reported that the individual trials were analysed according to:

- Intention-to-Treat (ITT) Population datasets i.e. all subjects who initiated any study procedures, beginning with mobilisation, and
- Transplant Population (TP) datasets i.e. all subjects who received Zynteglo treatment

The CS added that the ITT population was the primary population for the analysis of safety parameters and the TP was the primary population for the analysis of efficacy and pharmacodynamic parameters. The ERG requested ITT data on the flow of participants from screening to receipt of Zynteglo, including any reasons for withdrawal. The ERG noted that the safety dataset appears to use a different definition of ITT as it includes patients who signed informed consent forms but were yet to

begin the cell mobilisation stage. The ERG considers this to be the true ITT dataset i.e. all patients who signed informed consent to be included in the study.

4.2 Critique of trials of the technology of interest, their analysis and interpretation

4.2.1 Overview of the Zynteglo trials

The submission referred to two phase 1/2 Zynteglo trials in patients with any genotype (HGB-204, aka 'Northstar' and HGB-205), one ongoing phase 3 trial in patients specifically with non-β^0/β^0 genotypes (HGB-207, aka 'Northstar-2'), and one ongoing phase 3 trial in patients with "both non-β^0/β^0 and β^0/β^0" genotypes (HGB-212, aka 'Northstar-3'). All the studies were single-arm, open-label trials and are broadly summarised in Table 4 (adapted from Table 8 of the CS). Patients from these trials were followed up long-term in a separate study (LTF-303).

Study HGB-212 is the most recently initiated study in the Zynteglo trials programme but very limited study details were presented in the CS. The ERG requested further methods details on study HGB-212 via a point of clarification request. These were provided (see Table 4), with the company adding that the "data was not yet mature to provide meaningful results" for use in the submission. This seems an inconsistent approach as the company did report some immature results for ongoing study HGB-207 (p89, CS). The ERG also notes that two conference abstracts exist which report interim results for HGB-212.[3, 4] The most recent abstract indicated that, so far four patients had been recruited who had severe non-β^0/β^0 genotypes: two β^0/IVS-I-110, and two homozygous IVS-I-110.[4] The ERG believes that early efficacy data for these patients may help to resolve some of the uncertainty about the efficacy of Zynteglo in patients with severe non-β^0/β^0 genotypes.

The company stated (in Table 1 of their PfC response) that the HGB-212 study population was of TDT patients of "any genotype". When the ERG examined the relevant clinicaltrials.gov record and conference abstracts (one identified in the company's systematic review[4] and one identified by the ERG[3]) it appears that this study was, more specifically, of patients with "either a β^0 or IVS-I-110 mutation at both alleles".[3, 4] However, on the clinicaltrials.gov record, the exclusion criterion which relates to genotype reads: "Presence of a mutation characterized as other than β^0 (e.g., β^+, β^E, β^C)". It appears therefore that in study HGB-212 an IVS-I-110 mutation was deemed "equivalent to a β^0 mutation", as was the case in study HGB-207. However, given that "*certain β^+ genotypes such as the IVS-I-110 and IVS-I-5 mutations are associated with dramatically reduced β-globin production behaving phenoptypically as a β^0 genotype despite being grouped non-β^0/β^0 genotypes*" (CS, p19) it is unclear why IVS-I-5 mutations were not also specified as being eligible for study HGB-212.

The CS reported basic baseline data in terms of the race of patients included in the Zynteglo trials: [redacted] Asian, [redacted] white and [redacted] other (Table 18, p75 of CS). In the UK the largest group of patients with thalassaemia are those of Pakistani ethnicity and around 10-15% of patients are white.[8] Patients

in the Zynteglo trials were recruited in: Australia, France and the US in HGB-204; France in HGB-205 and Germany, France, Greece, Italy, Thailand, the UK and the US in HGB-207. The table of baseline characteristics presented in the CS was limited in allowing comparison of characteristics across studies because it also included data from six patients aged <12 years (from study HGB-207) who are not covered by the marketing authorisation. Also, no data were presented on the different genotypes across studies. In a point for clarification the ERG requested genotype data. The company stated that the genotypes included in the submission for the ITT population were: 6 β^{+}/β^{+} subjects (of which [redacted] are IVS-I-110), 15 βE/β0 subjects and 10 β^{0}/β^{+} subjects. The proportions of β^{E}/β^{0} and β^{0}/β^{+} patients [redacted]. The clinical implications of this are uncertain, although the CS baseline characteristics table suggested that [redacted] (though this was not a statistically significant difference).

Another source of heterogeneity across studies is the Zynteglo manufacturing process, which has changed over time. [redacted] due to "inherent subject-to-subject variability in transduction efficiency of HSCs". Ultimately, the changes in manufacturing processes were driven by a desire to [redacted].[9] This again highlights the importance of considering severity of genotype as an important subgroup. This has the potential to introduce heterogeneity of results across trials – with severe genotype response rates possibly varying depending on the manufacturing process used. Similarly, the ERG notes the dose difference between the early trials (≥3 x 10^{6} CD34+ cells/kg in HGB-204 and HGB-205) and later trials (≥5 x 10^{6} CD34+ cells/kg in HGB-207 and HGB-212) which may also be a source of heterogeneity of effect. On p189 of the CS it was reported that in the clinical studies, [redacted] of patients underwent one additional cycle of mobilisation and apheresis and [redacted] underwent a total of three cycles, in order to obtain more cells for additional manufacture.

Table 4 Overview of the Zynteglo trials (adapted from Table 8, CS)

Study, status, manufacturing process	**Population eligibility criteria, dose, and planned n (actual n for completed studies)**	**Primary endpoint(s)**
HGB-204 (Northstar) Completed February 2018 [redacted]	TDT (any genotype) Age ≥18 to ≤35 years Dose: ≥3 x 10^6 CD34+ cells/kg n=18 (10 non-β^0/β^0 genotype and 8 β^0/β^0 genotype)	**Safety:** Success and kinetics of engraftment, survival, incidence of AEs, monitoring for RCL, and insertional mutagenesis leading to clonal dominance or leukemia **Efficacy:** Proportion of patients with sustained production ≥2.0 g/dL HbA containing HbA^{T87Q} for the 6 months between months 18 and 24 post-transplant
HGB-205 Completed February 2019 [redacted]	TDT (any genotype) or severe SCD Age ≥5 to ≤35 years Dose: ≥3 x 10^6 CD34+ cells/kg n=7 (4 non-β^0/β^0 genotype TDT patients and 3 severe sickle cell disease)	**Safety:** success and kinetics of engraftment, survival, incidence of AEs, monitoring for RCL and insertional oncogenesis leading to clonal dominance or leukemia **Efficacy (TDT):** RBC transfusion requirements post-transplant; in-patient hospitalisation days (post-transplant discharge)
HGB-207 (Northstar-2) Ongoing [redacted]	TDT, non-β^0/β^0 genotype Age ≤50 years Dose: ≥5 x 10^6 CD34+ cells/kg n=15 (all belong to non-β^0/β^0 genotype)	**Efficacy:** Proportion of patients achieving transfusion independence (TI).
HGB-212 (Northstar-3) **Ongoing** [redacted]	TDT patients with "either a β0 or IVS-I-110 mutation at both alleles" Age ≤50 years Dose: ≥5 x 10^6 CD34+ cells/kg n=14 (5 non-β^0/β^0 and 9 β^0/β^0 genotype TDT patients)	**Efficacy: Proportion of patients achieving transfusion reduction, defined as demonstration of a ≥60% reduction in the annualised volume of packed red blood cells transfusion requirements (in mL/kg) in the post-treatment time period from 12 months post-drug product infusion through Month 24 (approximately a 12-month period), compared to the annualized mL/kg pRBC transfusion requirement during the 2 years prior to study enrollment.**
LTF-303 Ongoing	TDT: long-term follow up of patients in the above studies	Long-term safety and efficacy

4.2.2 Summary of the pooled dataset effectiveness results

The CS reported that the data presented related to a data-cut of June 2019, although this was contradicted in Table 21 of the CS, which stated that the data from HGB-207 were from an interim analysis dated 22nd February 2018. Results from p89 of the CS indicated that the June 2019 data were used for the primary endpoint (transfusion independence).

4.2.2.1 Transfusion and haemoglobin outcomes

The primary outcome was stated as the proportion of patients achieving transfusion independence. This was defined as a weighted average Hb ≥9 g/dL without any RBC transfusions for ≥12 months at any time during the study after Zynteglo transfusion. Twenty-four patients were classed as being "TI-evaluable" (i.e. patients who had either completed their parent study, or achieved TI, or won't achieve TI due to insufficient remaining follow up time). For the transplant population the rate of TI was 83% (20/24 patients). No results were presented for the ITT population. The ERG estimates this as [redacted]. For the 20 TI patients, no events for 'loss of TI' have been recorded so far.

[redacted]

[redacted]

Genotype heterogeneity in TI response

In light of the possibility of a heterogeneity of response based on genotype (see Section 4.2.1), the ERG requested transfusion independence and transfusion reduction results data for all patients with IVS-I-110 or IVS-I-5 mutations via a PfC. The ERG cross-referenced the company's response with data in the respective trial clinical study reports (CSRs) and in the CS. The aim was to collate the outcomes on all patients with severe non-β^0/β^0 genotypes. The company's response to the PfC identified different IVS-I-110 or IVS-I-5 patients from those previously identified by the company in their submission and from the patients identified by the ERG from the CSRs. The differences are documented in Table 5 and indicate the CSRs and the CS as being the best sources of data to investigate Zynteglo's efficacy in severe non-β^0/β^0 genotypes.

Table 5 Patients in Zynteglo studies with an IVS-I-5 or IVS-I-110 mutation

Study	Data from CSR or CS	Data from PFC response	Conclusion
HGB-204	[redacted]	[redacted]	Use CSR data
HGB-205	[redacted]	[redacted]	Use CSR data
HGB-207	[redacted]	[redacted]	Use CSR data – patients with severe non-β^0/β^0 genotypes which incorporated IVS-I-110 were excluded from HGB-207

Of the [redacted] patients identified in Table 5 as having severe non-β^0/β^0 genotypes – i.e. homozygous for IVS-I-110 or IVS-I-5, or heterozygous for IVS-I-110 or IVS-I-5 together with a β^0 mutation – [redacted]. An initial reading of these data suggests that the likelihood of achieving TI in patients with severe non-β^0/β^0 genotypes [redacted] to that for patients with β^0/β^0

genotypes i.e. [redacted]. For comparison, Study HGB-204 included 8 patients with β^0/β^0 genotypes (these patients are not included in the marketing authorisation). [redacted].[10]

However, the changes in the manufacturing processes, and doses, across the trials (discussed previously in Section 4.2.1) should also be considered with respect to efficacy. The CS points out (p89) the importance of the combination of transgene expression (i.e. Zynteglo-derived β-globin) and endogenous β-globin production in determining the probability of achieving TI, and that high levels of gene-derived haemoglobin production are needed to achieve TI (p170). [redacted], although study HGB-207 is still ongoing and is recruiting a more selective population than HGB-204. Nevertheless, this suggests that Zynteglo [redacted].

The TI response of the two IVS-I-5 patients with severe non-β^0/β^0 genotypes in study HGB-207 is encouraging, although patients with severe IVS-I-110 genotypes are excluded from this study. Therefore, the uncertainty surrounding the proportion of patients likely to achieve TI in the important subgroup of severe non-β^0/β^0 genotypes will only be clarified by results from the ongoing study HGB-212 (and further results from HGB-207). However, it is unclear whether IVS-I-5 patients are eligible for study HGB-212 so the number of patients recruited with severe non-β^0/β^0 genotypes may be small. The severe non-β^0/β^0 genotype subgroup appears likely to be important as taken together, the IVS-I-5 (in particular) and the IVS-I-110 alleles are quite common in the TDT UK population (see Section 3.1). With these issues in mind it is also worth noting that the primary outcome of study HGB-212 is transfusion reduction (in HGB-207 it was transfusion independence) suggesting lower expectations of a TI response in patients with more severe genotypes.

New data from studies HGB-207 and HGB-212 may also be complicated by the fact that the results from some patients may be affected by [redacted]

[redacted]. The CS stated (on p52) that the EMA conditional marketing authorisation decision included an approved commercial manufacturing process that "increased the acceptable range of transduction parameters compared to those used for the majority of subjects within studies HGB-207 and HGB-212".

[redacted]. The company has therefore proposed to the EMA that

[redacted].

Figures 22 and 23 of the CS indicate

[redacted]. Beyond this time point the patient numbers are very limited so there is some uncertainty about the longevity of transduced HSC engraftment in bone marrow and subsequent expression of HbA^{T87Q}.

4.2.2.2 Post-Zynteglo Iron levels

The company assumed an iron normalisation period of 4 years, following treatment with Zynteglo. However, results from section B.2.6.4 of the CS indicated this to be a simplistic assumption with wide variation in results across the small number of TI patients who had results at the 4-year timepoint. Data on iron levels are only available for a limited number of patients, as they have not yet been analysed for HGB-207. For example, some patients had an increase in liver iron content (Table 31, CS)[redacted] when compared with baseline levels. Table 33 in the CS shows that serum ferritin levels at 4 years were still clinically significantly high in some patients; the mean serum ferritin level was 1437ng/ml and the median 937 ng/ml for the 7 patients with available data. The company also stated that, out of the 11 non-β^0/β^0 patients that have achieved TI from Studies HGB-204 and HGB-205, all continued to have "normal" cardiac T2* values at their last follow-up, and maintain cardiac T2* values well above 20 msec (p105, CS). However, the company describe a cut-off of "normal" cardiac T2* of 40 msec, which only [redacted] patient appeared to achieve, and the lower limit of cardiac T2* was not significantly above 20 msecs, at [redacted] at 48 months. Post-transplant chelation guidelines were provided in the study protocol. However, resumption of iron chelation therapy was done at the investigator's discretion and in accordance with institutional protocols so there will have been variation in the timings and intensities of chelation.

4.2.2.3 Health-related quality of life

The Zynteglo health-related quality of life (HRQoL) data reported in the CS were limited because several patients from studies HGB-204 and HGB-205 did not have baseline measurements. The HRQoL tools used within and between trials varied widely, with the following being used: EQ-5D-3L, EQ-5D-Y, SF-36v2, PedsQL, and FACT-BMT. For adults with baseline HRQoL measurements, the CS noted a general trend of [redacted]. The company concluded that the EQ-5D data collected in the Zynteglo trials may not accurately reflect the HRQoL of patients treated with Zynteglo. Later in the submission (p182) the company discussed the problematic issues of ceiling effects and adaptation bias when evaluating HRQoL in patients with β-thalassemia. These also [redacted] from the results of the UK patient (and caregivers) preference study the company undertook, which suggested that TDT patient uptake of Zynteglo treatment would be somewhat limited. The study was an online survey of [redacted] TDT patients and [redacted] caregivers (total n=[redacted]), [redacted] of which had been living with beta-thalassaemia for [redacted] years or more. Of the [redacted] survey responders only [redacted] of patients agreed with the statement "Beta thalassaemia significantly impacts my quality of life" and only [redacted] indicated they would immediately accept a referral (to see a transplant specialist) and accept Zynteglo, were it offered.

4.2.2.4 Growth & Development

The company's results on growth and development outcomes in patients aged <18 years were restricted to one short descriptive paragraph (p113) so the ERG requested data on the growth and development endpoints detailed on p72 of the CS. [redacted] subjects underwent Tanner staging at screening: [redacted] males ranging from 8 to 15 years of age, and [redacted] females ranging from 5 to 17 years of age. [redacted] subjects were considered pre-pubertal at the time of Zynteglo Infusion: [redacted] males and [redacted] female. [redacted] subjects ([redacted]) were undergoing puberty, and one was post-pubertal at the time of drug product infusion. The other [redacted] subjects who underwent Tanner staging did not have data for the screening visit. For the [redacted] subjects with data after screening, [redacted] showed a [redacted] over time, with the exception of [redacted] who was already post-pubertal at the time of screening. Of note, a [redacted], was assessed as Stage II/I pubic hair/genitalia at screening, then assessed as Stage I/I pubic hair/genitalia at Month 6 but had [redacted] to Stage III/III pubic hair/genitalia by Month 18.

4.2.3 Adverse Events

Reporting clarity by company

At the FAC stage, the company acknowledged inconsistencies in the CS and PFC response regarding the population included in the safety assessment. Some confusion whether the AEs reported were exhaustive remains, due to differences in cut-off points used for ongoing studies and the lack of a CSR for HGB-212. A systematic review requested by the ERG via a PFC question referenced a

conference abstract reporting interim results of study HGB-212,[4] which gave an overview of AEs observed up to April 2019. One patient, genotype not reported, experienced congestive heart failure where the left ventricle ejection fraction fell. This was associated with the worsening of cardiac iron pre-engraftment. No further information was provided for this patient either to recommend inclusion or exclusion in AEs or to detail the patient's condition. The ERG considered it likely that this patient had a β^0/β^0 genotype, as the company excluded β^0/β^0 patients from the intention-to-treat (ITT) population. However, the ERG questions the rationale to include only non-β^0/β^0 patients for safety assessments, since in HGB-207 a severe non-β^0 allele (IVS-I-110, in particular) was considered equivalent to a β^0 mutation, which is a relationship that would apply in both directions.

Adverse Event data

Adverse event data were reported on pages 114-143 of the CS. Adverse events (AE) were assessed in the ITT population of 35 patients with non- β^0/β^0 genotypes aged ≥ 12 years from studies HGB-204 (n=11), HGB-205 (n=4) and HGB-207 (n=16) and HGB-212 (n=4). TDT patients with the β^0/β^0 genotype in HGB-204 were excluded by the company from the safety data as they were not part of the population of interest.

In Table 37 of the submission, the company reported the incidence of all AEs categorised by system organ class and pre-decided intervals in the treatment period. AEs have not been reported for all 35 patients in some of the later time intervals since not all patients in the ongoing HGB-207 and HGB-212 studies had progressed to these later stages.

[redacted] of all patients experienced at least one AE and [redacted] experienced at least one serious adverse event (SAE); [redacted] reported as being Zynteglo related.

[redacted] The fewest

[redacted] reported after 24 months post Zynteglo infusion. In addition to standard safety outcomes, patients were monitored for any sign of AEs caused by gene therapy using the BB305 vector.

Safety of the BB305 lentiviral vector

There were no safety issues related to the BB305 LVV in HGB-204, HGB-205 or HGB-207.[11, 12] The company did not mention HGB-212 in the relevant section of the CS. There was no evidence of clonal dominance (which indicates whether the insertion of a therapeutic gene could be a precursor to a haematological malignancy). No malignancies related to LVV were reported, including leukaemia or

lymphoma where the longest follow-up was 61.3 months after drug infusion. One patient who received Zynteglo experienced an occurrence of HIV-1 infection, but this was later confirmed to be wild-type HIV-1 and not due to lentiviral vector (LVV) recombination.

Engraftment and transplant-related complications

Engraftment of the gene-modified autologous cells was successful in all patients treated with Zynteglo across all the clinical trials up to the latest follow-up at 61.3 months.[11, 13, 14] There were no incidents of transplant-related mortality, graft rejection or GvHD.[14, 15]

Mobilisation and apheresis

Most AEs attributed to mobilisation and apheresis in TDT patients occurred [redacted]. AEs attributed to mobilisation and apheresis were summarised in Table 38 (pp 128-129) of the CS. Most were non-serious events: [redacted]. [redacted] serious adverse events (SAEs) were attributed to mobilisation and apheresis, an event each of: [redacted]

Conditioning

AEs attributed to busulfan conditioning that occur in at least 3 patients were summarised in Table 39 (pp 131-134) of the CS. Most events were not serious and expected from treatment with an alkylating agent such as busulfan, according to the prescribing information including [redacted].[16] Some observed AEs that were not part of the prescribing information were: [redacted] have been documented as potential side effects of busulfan. According to the latest data available, [redacted] patients reported at least one AE attributed to conditioning where [redacted]

Adverse events by ≥ Grade 3 severity

The incidence of all ≥ Grade 3 AEs were summarised in Table 40 (pp 137-139) of the CS. [redacted] experienced at least 1 AE ≥ Grade 3. [redacted]

[redacted]

There [redacted] drug-product related Grade 3 SAE in [redacted].

[redacted]

Serious adverse events (SAEs)

The incidence of all SAEs in TDT patients was presented in Table 41 (pp 141-142) of the CS. The overall survival was 100% and no transplant-related mortality was observed. Approximately [redacted] of the patients experienced an SAE,

[redacted] was deemed Zynteglo-related. [redacted] experienced SAEs prior to neutrophil engraftment which were attributed to study procedures, mobilisation, apheresis, or reasons unknown.

[redacted] which were all resolved

[redacted]. There were

[redacted]. In addition to the

[redacted] All cases were resolved with defibrotide.

[redacted]

4.2.4 Summary

Observations on how the Zynteglo trial programme has evolved – such as study variation in production processes, doses, eligibility criteria, primary outcomes, and results – suggests an ongoing drive to improve transfusion independence response rates. While most patients appear to respond well to Zynteglo, some patients, in particular those with a genotype deemed very similar or equivalent to a β^0/β^0 genotype (i.e. "severe" non-β^0/β^0 genotypes, who are covered by the marketing authorisation) and patients with a β^0/β^0 genotype (who are not covered by the marketing authorisation) tend not to

respond as well as other genotypes. [redacted] may increase the probability of achieving TI in patients with severe non-β^0/β^0 genotypes (compared to [redacted]) but only results from ongoing study HGB-212 and further data from the ongoing HGB-207 study can help to resolve this uncertainty. The company did not submit results for HGB-212, stating that the data were too immature, but the ERG considers that initial results on total Hb levels, HbA^{T87Q} levels and number and frequency of transfusions, for each of the non-β^0/β^0 patients would nevertheless be useful and could be submitted at the technical engagement stage of the appraisal.

The ethnicity distribution of the Zynteglo trial population is not a particularly good representation of the UK TDT population. The main implication of this is the possibility that the proportion of patients with severe non-β^0/β^0 genotypes in the trials is not representative of the UK setting. This was exacerbated in one trial by an eligibility criterion which excluded patients with a specific severe non-β^0/β^0 genotype. The longest follow up period for an individual patient is 5 years. Consequently, there is uncertainty about whether Zynteglo confers benefit to patients in the much longer term i.e. whether it truly is a curative therapy. It is possible that the transformed cells may not persist long-term. This may mean that some patients revert to needing regular transfusions long after they have achieved transfusion independence. Such future outcomes could also open the possibility of re-treatment with Zynteglo, although re-treatment is not currently part of the license. Zynteglo appears to have an acceptable safety profile in the short-term, though uncertainty exists about its long-term safety.

4.3 Critique of the systematic reviews for comparator data and Zynteglo proxy data

Only single-arm data were available on the effectiveness and safety of Zynteglo. Therefore, to compare the outcomes of patients receiving Zynteglo with those receiving transfusions and chelation therapy (in the economic model) the company undertook several systematic reviews to identify appropriate comparator datasets of TDT patients to evaluate:

- The clinical safety of iron chelation and transfusion therapies
- How iron overload-related complication rates vary by iron levels

The company provided the 120-page systematic review as a supplementary document (EVA-20726-04).[17]

To obtain further comparator data the company also undertook a "Chart Review" observational study of UK TDT patients with the aim of describing:

- Transfusion requirements,
- Patient demographics and baseline clinical characteristics,
- Routine management,
- Clinical outcomes,
- Complications related to iron overload and iron chelation therapy,
- Impact of TDT on quality of life.

This study was submitted by the company as a draft manuscript.[18]

4.3.1 Searches for comparator studies

The company conducted two searches for evidence on the following:

- What is the clinical safety of iron chelation and transfusion therapies in patients with TDT,
- How do iron overload-related complication rates vary by iron levels in TDT in Europe.

The searches for the above reviews were included in the 2018 report by Evidera in Section 3.1, p. 14-15.[19] These searches were updated in 2019 and included in the 2019 report by Evidera, with a description of the searches on p5 and full search strategies contained in Appendix D, p111-115.[17]

The following databases were searched in May 2017: MEDLINE (via PubMed.com and Embase.com) and Embase (via Embase.com). Retrieval was limited to English language studies with an abstract, published from 2007 to 31st December 2017. The searches of MEDLINE and Embase were updated in April 2019 along with a search of CENTRAL via the Cochrane Library (searched from 2007 to 1st April 2019).

Specific conferences taking place from 2015 onwards were searched via Embase.com to identify relevant conference abstracts or posters: American Society of Hematology (ASH), European Hematology Association (EHA), British Blood Transfusion Society (BBTS), European Society for Blood and Marrow Transplantation (EBMT) and the International Society of Blood Transfusion. In addition, the following online conference websites were searched: ASH (2018), EHA (2018) and the International Society of Blood Transfusion (2018). Further unpublished studies were identified through searches of ClinicalTrials.gov and the WHO International Clinical Trials Registry Platform.

The databases and other sources searched were appropriate to locate both published and unpublished studies of comparator treatments (iron chelation and transfusion therapies), and iron overload-related complications in patients with TDT. However, the Cochrane Database of Systematic Reviews (CDSR) is missing from Table D.1. Data Sources (p111. Appendix D, 2019 Evidera report) although it does appear to have been searched (Table D.6, p114, 2019 Evidera report). The results from CDSR could have identified any Cochrane Reviews relevant for SLRs 1 and 2.

The search strategies for all databases (MEDLINE, Embase, and the Cochrane Library) were structured appropriately using terms for thalassaemia combined with terms for either transfusion or iron chelation therapies. The term ‘thalassemia major’ has been included within the intervention terms in most of the database search strategies. This appears to be a mistake, though it would not have caused relevant studies to be missed. A lack of truncation was noted throughout all database strategies for some search terms – thalassemia and thalassaemia, anemia and anaemia. These terms could have been truncated as follows: thalassemi$, thalassaemi$, anemi$, anaemi$, to allow maximal retrieval of relevant records that use the same word stem in the title and abstract but have different endings eg: thalassaemic, thalassaemias, anaemic, anaemias etc.

Subject headings for iron chelation therapies and blood transfusion were missing from the MEDLINE, Embase and Cochrane Library search strategies. It is usual practice for systematic review searches to include both text word searches of the title and abstract fields as well as relevant subject headings to ensure all relevant studies are retrieved.

The search strategies for EMBASE were limited by publication type to records that have been assigned as articles or articles in press e.g. line #7, Table 2, p. 15 in the 2018 report by Evidera.[19] This may have omitted errata or corrections to published articles as well as other publication types such as book chapters, short surveys, reviews and conference papers.

Some minor reporting errors were found by the ERG relating to the searches. The PRISMA flow diagram (Figure 1, p. 9, 2019 report by Evidera), had typing errors in the first box – the 2017 searches

of PubMed retrieved 2398 results and the 2017 searches of Embase retrieved 2372 results.[17] The flow diagram was also missing the search results obtained from the Cochrane Library.

4.3.2 Selection of comparator studies

The eligibility criteria for the systematic reviews of comparator studies were presented in Table 1 of the separate systematic review document.[17] The screening methods used were appropriate for minimising the possibility of reviewer errors and biases affecting the final list of included studies. The eligibility criteria appeared broadly appropriate, although it appeared that systematic reviews were not eligible, which seems like an oversight. Moreover, the selection approach when several studies were identified was not reported, and little was presented in terms of a synthesis. The ERG considers that the validity of some of the comparator data is uncertain, as the justification for selecting specific studies from these review for use in the CS was often not presented.

4.3.3 Quality assessment of comparator studies

The quality of the single-arm studies in isolation is only one aspect of the critical appraisal of the clinical evidence submitted by the company. The appropriateness and relevance of the comparator studies which were used in the model should also be justified, but this was not included in the systematic literature review (SLR).[17, 19] The quality of some of the included studies was assessed using the Cochrane Risk of Bias Tool 2.0,[20] though no explanations were provided for the decisions made regarding bias judgements.[17] The review and the submission did not discuss how the characteristics of comparator study populations compared with the corresponding Zynteglo trials cohort, so it was not easy to evaluate the level of appropriateness of the comparator studies selected. This was particularly important for studies which reported outcomes likely to be affected by the way chelation therapy was managed. Similarly, external validity was largely overlooked for comparators studies i.e. generally, there was a lack of consideration and discussion about whether interventions and populations were adequately representative the TDT population in the UK.

4.3.4 Results

Mortality

Cardiac iron overload is the major cause of death in β-thalassaemia. The company's systematic review found cardiac-related mortality to be reported in 18 studies[17] though the CS added that no studies reported mortality based on the presence or absence of cardiac complications. The CS stated that two economic studies 'identified in the literature' reported cardiac and non-cardiac mortality. It is unclear how these studies were identified as they were not mentioned in the systematic review.[17] For transfusion-dependent patients without cardiac disease, a standardised mortality ratio (SMR) of 3.9 was used based on a paper published in 1996 which was referenced and used in a 2007 economic study.[21] The ERG considers the 1996 paper[22] to be an obsolete reference in terms of its relevance to current NHS practice because subcutaneous chelation was used, whereas the introduction of oral iron

chelators (deferiprone and deferasirox) in recent years has led to improvements in mortality rates in β-thalassaemia patients. Subcutaneous chelation has been found to be associated with poorer compliance than oral chelators[21] and improved compliance of chelation therapy has been found to be associated with the avoidance of complications associated with iron overload and subsequently survival and quality of life may approach a normal pattern. Some evidence exists showing that oral chelators have a protective effect on the heart compared with subcutaneous chelation.[23, 24]

In the model, patients who acquire cardiac disease were assumed to have an annual mortality rate of 13%. This was based on a study of 52 patients with β-thalassemia and heart failure who were treated in the mid-to-late 1990s.[25] This study is also not reflective of current UK practice and is therefore outdated. It is likely to overestimate cardiac mortality, based on recent evidence.[26] This is because of the impact of both the aforementioned introduction of oral chelation therapies and the introduction (in 1999 in the UK) of T2* cardiovascular magnetic resonance (CMR) for identifying myocardial siderosis. T2* CMR was applied rapidly in clinical management from 2000 as the benefits of direct visualisation of cardiac siderosis as a guide to the need for intensified iron chelation therapy, and a means of assessing response became clear.[27] The ERG's clinical adviser also stated that patients picked up as having an iron problem via T2* CMR would subsequently be managed by a cardiologist, which should improve their cardiac outcomes. The company's own review also reported that "adults aged 20 to 40 years old between 2000 and 2009 had a lower risk of cardiac mortality compared to similar adults in 1990 to 1999".[17] An assumed SMR of 1.25 was used to model for transfusion independent patients. This was based on the potential impact of myeloablative conditioning chemotherapy. In summary, the model inputs for mortality risk in patients who are transfusion-dependent is likely to have been overestimated and the risk in patients who achieve TI is uncertain since there is little robust evidence to support it.

Iron overload in TDT patients

In the company's systematic review 56 publications were identified which evaluated the burden of illness of iron overload in TDT.[17]

Cardiac complications of iron overload

The CS stated that 17 studies were identified which reported clinical measurements and outcomes relating to cardiac disease. A study by Pepe et al was used to inform annual rates of complications based on it having a large sample size (n=481), a good duration of follow-up (mean of 58 months), and the explicit provision of hazard ratios by myocardial T2*.[28] The ERG considers this to be a reasonable source of data to use, although notes that the study was conducted in a white population. There is therefore some uncertainty about the applicability of its results to a UK TDT population since in the UK only around 10-15% of TDT patients are white.[8].

Liver complications of iron overload

The CS stated that 18 studies of liver-related complications were identified, but none provided relationships between liver iron levels and complication rates. Therefore, a further 'targeted literature review' was conducted but no further details were given about this. The study selected was published in 2002 and was acknowledged in the CS as being outdated, though it did give separate results for patients with or without hepatitis C.[29] Given the previously discussed improvements in TDT patient care since 2002, the data from this study are likely to overestimate the rate of liver complications for patients with high iron levels. However, in addition to reporting by iron levels, the study reported liver complications by presence of hepatitis C infection. Many liver complications developed in patients in the past are a result of hepatitis C infections from historical blood transfusions carrying the virus. The use of these data was thought to mitigate the fact that this study is somewhat outdated, since it can be assumed that future rates of liver complications will be closer to those of patients without hepatitis C virus due to current blood donation screening practices. The study found that HCV negative patients with low or moderate iron levels remained fibrosis progression-free, but HCV negative patients with high iron had a median time to fibrosis progression of 100 days.

Other complications of iron overload

One study was used to predict diabetes and hypogonadism in adults using myocardial and serum iron levels.[30] The study was retrospective with case note and electronic data collated for the period 1999 to 2010. It reported a diabetes prevalence rate of 41% and a 67% rate for hypogonadism. These are notably higher than the rates reported in the company's Chart Review of UK patients (see Section 4.3.5).

Adverse events of chelation therapies

Although the company undertook a systematic review of the safety of iron chelation therapies it did not use the results to inform Table 51 of the CS which presented the probability of adverse events with iron chelation therapy. Instead, data were used from the prescribing information for deferasirox[31] and deferiprone[32] and from a Cochrane review for desferrioxamine.[33] The CS did not state why it did not use data from its own systematic review of adverse events and instead used alternative sources, including the Cochrane review (reviews appeared to be excluded from the company's systematic review, SLR 1).

4.3.5 Observational 'Chart Review' of UK TDT patients

The manufacturer also conducted and submitted a report of an observational study to understand the current real-world routine management of patients with TDT in the UK.[18] It aimed to describe: transfusion requirements, patient demographics and baseline clinical characteristics, routine management, clinical outcomes, impact of TDT on quality of life and complications related to iron overload and iron chelation therapy.

The study included 165 patients, a third of which were over 40 years old, whereas the oldest patient treated with Zynteglo in the trials was 34 years old. Patients had a mean of 13.5 transfusion episodes (20% of patients had 16 or more transfusions per year) and a median of 32.4 units of blood transfused per year. It is noteworthy that although several patients were excluded from the Zynteglo trials for having comorbidities – advanced liver disease or cardiac disease – such patients were not excluded from this chart review study. This was reflected in the proportion of patients who had high liver iron concentrations (29%). The prevalence of diabetes in the Chart Review was 13%, hypogonadism prevalence was 7%, and hypogonadotropic hypogonadism prevalence was 20%. These are much lower prevalences than those reported in the study identified in the company's systematic review (see Section 4.3.4) which were used to derive risks of developing these conditions in the model.

Another difference between the Chart Review and the Zynteglo trials is that [redacted] Pistoia et al (2019), in a study of 671 regularly transfused β-thalassaemia-major patients, found that the homozygous β^+/β^+ patients showed less myocardial iron overload and a concordant better global heart function when compared to the more severe groups (β^+/β^0 and β^0/β^0) and that the β^+/β^0 patients showed significantly higher global heart T2* values than the β^0/β^0 patients ($p < 0.05$).[34]

4.3.6 Zynteglo proxy data

The company also stated that they undertook reviews to identify proxy data for expected longer-term outcomes on Zynteglo. Studies of allogeneic post-transplant patients were used to identify proxy data for:

- Infertility and gonadal function, and
- Time to iron normalisation

The CS reported that three studies were identified using a "targeted review" on infertility and gonadal function (p171). Although it was unclear whether these were identified systematically, the CS stated that they had generally consistent findings on fertility rates. The company justified the use of one particular study based on it being the only study specifically in stem-cell transplanted thalassaemic survivors with infertility measured by gonadal dysfunction.[35] Based on this study, and on UK HES data, the company assumed that the conditioning chemotherapy associated with Zynteglo therapy would increase infertility by 24% in men and 57% in women (compared to SoC TDT patients).

Another 'targeted literature review' was used to identify two papers on iron store changes post allogeneic HSCT in thalassaemia patients. The company stated that they 'conservatively' used the Chaudhury et al 2017 study to support an assumption of a 4-year iron normalisation period following Zynteglo therapy.[36] However, this study (of 176 patients) reported that at ≥4 years post-transplant median ferritin levels were 870 ng/mL (range, 52 to 6847). Based on upper ranges of 'normal' levels being 200ng/mL in females and 300ng/mL in males the ERG does not consider this to be a conservative estimate.

4.3.7 Summary

The company's systematic reviews were restricted to studies published from 2007 onwards and review articles appeared to have been excluded. These criteria proved to be too restrictive since, in subject areas where suitable studies were not identified, the company had to undertake additional "targeted reviews" to identify studies. There was a lack of transparency about how these additional studies were identified and selected, meaning it was not possible for the ERG to make a judgement on whether the most appropriate evidence was used. This was exacerbated by the limited, or absent, critiques of the strengths and weaknesses of these studies. The ERG understands that it is often impractical to use systematic review methods to identify all sources of data for an economic evaluation, but transparency on the methods used to identify and select studies is nevertheless very important. It was also unclear why the company sometimes did not make use of their own systematic review results (e.g. for adverse events of chelation therapy). Consequently, the ERG has concerns that some model inputs for TDT patients receiving routine care – such as mortality risk and the rate of liver complications in patients with high iron levels – are likely to be overestimates.

A difference between the comparator studies and the population eligible for Zynteglo therapy is the genotype restriction (with Zynteglo). The comparator studies did not restrict by genotype and will have included some β^0/β^0 patients. It is possible that comparator study cohorts may have achieved better outcomes if patients with β^0/β^0 genotypes had been excluded.

4.4 Conclusions of the clinical effectiveness section

The Zynteglo trial results have somewhat limited applicability to the population likely to receive Zynteglo in the NHS as the trial population might under-represent certain genotypes which are prevalent in the UK. This is important because while most patients appear to respond well to Zynteglo, some patients tend not to respond as well as others. In particular, patients with a genotype deemed very similar or equivalent to a β^0/β^0 genotype (i.e. "severe" non-β^0/β^0 genotypes, which are covered by the marketing authorisation) and patients with a β^0/β^0 genotype (who are not covered by the marketing authorisation) appear less likely to achieve transfusion independence status than patients with other genotypes.

The Zynteglo manufacturing processes have evolved during the trial programme with the aim of maximising response rates. It is possible that [redacted] may increase the probability of achieving TI in patients with severe non-β^0/β^0 genotypes (compared to the previous processes used in the trial programme) but only results from the ongoing study HGB-212 and further data from the HGB-207 study (also ongoing) can resolve this uncertainty. The trials results are still quite immature, and the number of patients treated is small, so uncertainty exists regarding the longevity of Zynteglo and regarding the possibility of adverse events in the medium-to-long term.

A limitation of the company's systematic reviews to identify comparator group data was a lack of transparency about how studies were identified and selected, meaning it was not possible for the ERG to make a judgement on whether the most appropriate studies were used. The ERG also has concerns about the applicability of many of the studies which were selected as being appropriate (for providing model parameter data). Many studies are out of date and do not reflect the improvements in TDT patient treatment and monitoring achieved over the last 10-20 years. The company's own 'Chart Review' study of UK TDT patients also had limitations in how well it reflected the Zynteglo trial population. Specifically, the Chart Review study did not exclude patients with a β^0/β^0 genotype nor patients with important comorbidities (such as advanced liver disease or cardiac disease). Evidence identified from a study of thalassaemia patients who successfully underwent allogeneic stem cell transplants suggests that the conditioning chemotherapy associated with Zynteglo therapy would increase infertility when compared to TDT patients: by 24% in men and 57% in women.

5 Cost Effectiveness

This section focuses on the economic evidence submitted by the company and the additional information provided in response to the ERG's clarification questions. The submission was subject to a critical review on the basis of the company's report and by direct examination of the executable model. The critical appraisal was conducted with the aid of a checklist to assess the quality of the economic evaluation and a narrative review to highlight key assumptions and uncertainties (Appendix 1: Drummond Checklist).

5.1 ERG comment on company's review of cost-effectiveness evidence

The CS describes a systematic literature review that was conducted to identify economic studies (Section B.3.1, p. 148). One overall search was used to identify studies on the cost-effectiveness of treatments, HRQoL, and cost and resource utilisation in patients with TDT.

5.1.1 Searches

Searches were initially undertaken in 2017 and updated in 2019.[17, 19] The databases and other sources searched were appropriate to locate both published and unpublished studies of cost-effectiveness, HRQoL, and cost and resource use relating to transfusion-dependent beta-thalassaemia. The search strategies for all databases (MEDLINE, Embase, PsycINFO, and EconLit) were structured appropriately. Some minor reporting errors were found by the ERG relating to the searches. Full details of the search strategy used are provided in Appendix G of the CS. Further critique of the company's searches are provided in Appendix 2: Critique of the company's search strategies for cost-effectiveness evidence

5.1.2 Inclusion/exclusion criteria used for study selection

The company did not describe the eligibility criteria for study selection in their systematic review of cost-effectiveness studies. The search was described as a 'collective search strategy', which sought to identify cost-effectiveness studies, health related quality of life studies, and cost and resource use studies. The company stated that only studies evaluating blood transfusions and chelating agents were included, and that the 'geography' was limited to the US, France, Italy, Germany, Greece, and the UK. It is unclear why the company only considered these locations, and whether studies that drew efficacy data from outside these countries were also included.

The CS also discusses the results of a targeted search of cost-effectiveness studies evaluating allogeneic HSCT in any country; however, the methods are not described.

5.1.3 Studies included and excluded in the cost effectiveness review

According to the PRISMA diagrams presented in Appendix G of the company submission, the original review conducted by the company in May 2017 identified a total of 3,161 potentially relevant

studies, of which 2,920 were excluded at the primary screening stage. The remaining 241 studies underwent full-text assessment for eligibility. Seventy-two of these studies were excluded for a number of reasons reported in Figure 1 of CS Appendix G, and a further 154 studies were excluded as 'non-economic articles', producing a total of 15 studies. The company updated this review in April 2019, screening a total of 1,298 records for inclusion, none of which were found to be relevant.

The review identified five cost-effectiveness models and one cost-of-illness model. The remaining nine publications were resource use/cost studies. The company provided a brief description of the structures and assumptions of three cost-effectiveness studies conducted in the UK, one in the US, and one in Italy, and presents a summary of each in CS Table 43. The two cost-effectiveness studies identified through the 'targeted review' of included allogeneic HSCT for patients with thalassaemia major were undertaken in India[37] and Thailand[38], and were described in detail. Two of the identified cost-effectiveness analyses[21, 39] were used to inform clinical parameters in the company model.

5.1.4 Conclusions of the cost effectiveness review

The company makes no overall assessment of the appropriateness of inputs and assumptions adopted in the five cost-effectiveness studies identified in the initial review. They notably exclude some key assumptions and inputs considered in previous UK models from their own model, e.g. the development of osteoporosis. The company state that while the two studies from India and Thailand were unlikely to inform the parameters for the present evaluation as the intervention was allo-HSCT, the control arm comprising transfusion and chelation therapy could be useful for model validation. However, these studies were considered by the company to be particularly relevant to the Zynteglo model, and were used to guide their model structure and the selection of complications associated with iron overload.

The ERG considers the company's aggregation of the results of the systematic reviews of cost-effectiveness studies, costs and resource use, to be inappropriate and contrary to the principles of the PRISMA statement for transparency in reporting. Because the review methods were not reported, it is unclear whether all relevant literature was identified and included in the review.

However, as there are unlikely to be any studies which assess the cost-effectiveness of Zynteglo in a TDT population, the ERG consider the *de novo* cost-effectiveness analysis reported in the CS to be the most relevant source of evidence to address the present decision problem.

5.2 ERG's summary and critique of company's submitted economic evaluation

The company presented a *de novo* economic analysis of Zynteglo compared with standard care, consisting of blood transfusions and iron chelation therapy, in TDT. Effectiveness of treatment was

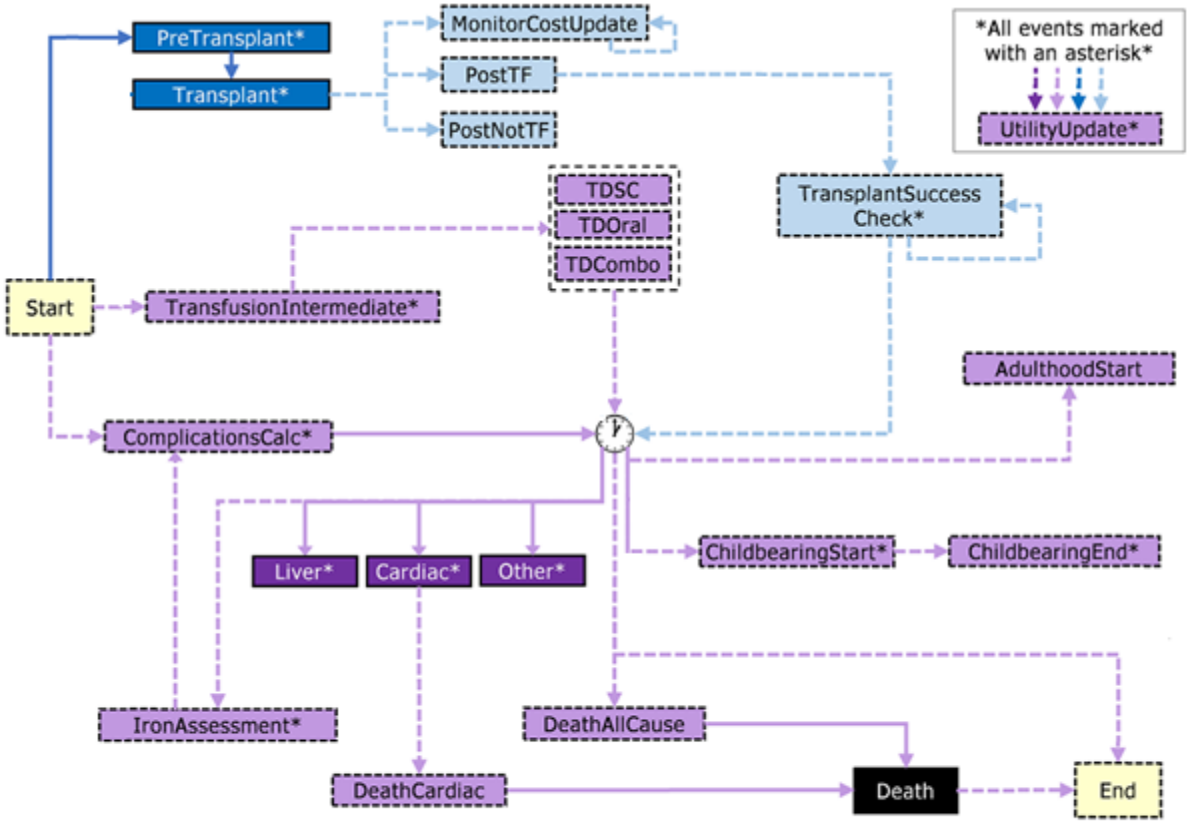
assessed through the achievement of transfusion-independence. Total costs and QALYs were assessed for each arm over a lifetime time horizon, and discounted at a rate of 1.5%.

The company submitted two economic models: one at the start of the appraisal, and an updated and corrected model following the clarification stage. The updated model included changes made by the company to some of the base case assumptions, as a result of clarification questions from the ERG. The ERG also identified modelling errors throughout the appraisal process, resulting in the company providing a number of iterations of the model containing various model corrections. All results presented in this section reflect the final, corrected version of the model provided appear less likely to achieve transfusion independence status than patients with other genotypes, and will not reflect the results presented in the original company submission or PFC.

5.2.1 Model structure

The company developed a *de novo* Excel-based model using a discrete event simulation structure, implemented through the discretely integrated condition event (DICE) simulation framework[40]. The basis of the model structure was driven by the transfusion status of patients, which determined their tissue-specific iron levels. A patient's iron level drives their risk of developing complications attributable to iron overload, and also influences mortality risk, quality of life and chelation requirements. Complications were assumed to be associated with the cardiac, liver and endocrine systems, as these were considered to be the organs most affected by iron overload. The presence of any of these complications was associated with an additional impact on quality of life, management costs and, in the case of cardiac disease, excess mortality risk.

Figure 2 Model schematic (Fig 2, clarification response)



Transfusion status

Patients enter the model as transfusion dependent, receiving a mean of 13.5 transfusions per year. In the Zynteglo arm, after infusion of transduced cells, patients are partitioned based on treatment success: they may become transfusion independent, have a clinically meaningful reduction in transfusion frequency (i.e. 'transfusion-reduced') or continue to receive the same number of transfusions as they did prior to receiving Zynteglo. Transfusion independence is defined as being free from transfusions for 12 months and having sustained haemoglobin levels above 9 g/dL. The rates at

which these outcomes occurred were based on Zynteglo trial data (HGB-204, HGB-205, and HGB-207) for non-β^0/β^0 patients, discussed in Section 4.2.2 and Section 5.2.6. In the model, TI patients were considered to be independent from transfusions (if achieved) beginning at 12 months post-transplant. In the standard of care arm, i.e. transfusions and iron chelation therapy, patients were assumed to remain transfusion dependent for the duration of the model time horizon.

While the model contains the functionality for patients to experience engraftment failure, or loss of graft and relapse after initially successful treatment with Zynteglo (i.e. losing their TI status), the company assumed that this would not occur in any patient.

Iron levels

Baseline iron levels, applied to patients in both treatment arms in the model, were estimated from a Chart Review of TDT patients in the UK, which was funded by the company (Section 4.3.5).[18] The baseline iron overload risk category for each simulated patient was randomly selected from the starting distribution (for details of this distribution see Section 5.2.6.3). Iron levels in the cardiac, liver and endocrine systems were based on the myocardial T2*, LIC and serum ferritin assessments respectively, and patients were classed as being either in a low, moderate or high risk category.

If transfusion independence was achieved after Zynteglo treatment, patients were assumed to achieve normalised iron levels after four years. Patients with substantially reduced transfusions (defined as at least a 60% reduction) were assumed to achieve reduced, but not normalised, levels of iron relative to their baseline levels. A distribution of lower iron levels was assumed by the company and applied in the model from one year onwards.

For patients receiving standard care, the baseline levels of iron were applied for the duration of the patients' lifetime; that is, iron levels in these patients did not increase or decrease over time, with the company justification that they represented the population mean iron levels for a TDT population and would be representative throughout their lifetime. These patients do not achieve a reduction or independence from transfusions at any time.

Risk of complications

The times to each complication event (cardiac, liver, and endocrine) were estimated at baseline, and were determined by the patient's assigned iron overload risk category. If patients achieved a reduction or independence from transfusions after treatment with Zynteglo, their risk of complications was reduced in line with their reduced iron overload category. Patients who were transfusion dependent maintained their risk throughout the model time horizon. After the iron normalisation period, TI patients were assumed to have no risk of additional iron overload or new iron-related complications: if the model estimated their time to developing a complication would occur after the date when iron normalisation would occur (4 years), then the risk was assumed to no longer apply and the

complication would not occur. If the time to developing a complication were to be reached before the end of the iron normalisation period, then the complication would occur. A similar method was applied for transfusion-reduced patients, although instead of patients facing no risk after four years, their time to the complication was re-estimated at one year, taking into account time already elapsed since the start of the model, based on the lower levels of iron.

Transfusion and chelation

In the standard care arm, patients received regular transfusions and chelation therapy to remove excess iron. Each patient was randomly allocated to receive either oral, subcutaneous, or a combination of oral and subcutaneous iron chelation therapy, using a distribution that was estimated from the Chart Review (Section 5.2.4).

Since excess iron can persist for several years after Zynteglo infusion and engraftment, iron chelation therapy and/or phlebotomy was continued following treatment. Transfusion independent patients continued to receive ongoing chelation and/or phlebotomy up until the end of the iron normalisation period, and experience the costs and the HRQoL impact associated with the mode of iron normalisation therapy received. Patients who became "transfusion reduced" following Zynteglo infusion were assumed to continue blood transfusions and iron chelation therapy, but with fewer transfusions and lower chelation therapy exposure, compared to the period before transplant.

Patient profiles

The discrete event simulation structure employed in the model is evaluated stochastically on a patient-level basis to produce estimates of the expected costs and benefits across the specified patient population. The model runs a number of 'profiles', which are hypothetical patients defined by age and gender, with each profile weighted to reflect the distribution of patients in the eligible treatment population (Section 3.1).

The model originally submitted by the company was based on [redacted] unique profiles reflecting the distribution of age and gender combinations between [redacted], and estimated results generated from 100 random samples of these profiles. Profiles in an updated company model were based on gender and three age bands of child, young adult, and adult, and the model estimated results generated from 600 samples. Patient age was used to determine the mortality rate (i.e. the time to death), quality of life, and some treatment related costs.

Time-to-event values are sampled for individual patients from probability distributions. Events that occurred stochastically were: time to development of complications due to iron overload in the cardiac, liver and endocrine systems, and time to death. Zynteglo treatment success and subsequent transfusion status was also assigned stochastically for patients in the Zynteglo arm. Baseline patient characteristics that were assigned randomly included organ-specific iron levels and type of chelation

agent. Generally, each stochastically-generated input was estimated independently of each other; for example, the risk of a patient having high cardiac iron was not linked to the likelihood that they would also have high liver iron, although the same random seed was used to estimate baseline and future iron levels for a given organ system, which enforced a degree of correlation between the two.

5.2.1.1 ERG comment

Omission of key elements of the condition

The company model, reflecting iron overload complications only in the liver, cardiac and endocrine organ systems, may have omitted some other important complications of beta thalassaemia, such as splenectomy and the development of osteoporosis. These have been present in other cost-effectiveness analyses for this condition, and the clinical advisor to the ERG suggested that, since TDT patients were living longer at present due to improvements in overall patient care, the management of osteoporosis is of increasing importance to older TDT patients. However, osteoporosis is associated with a low cost impact, and is considered to be a late developing complication with a more heavily discounted impact.

Diabetes and hypogonadism were modelled together within one category, both being complications associated with the endocrine system. However, in the model, it is not possible to develop both conditions concurrently. While the risk of developing each condition is estimated separately, the model considers only the one that occurs first. The unit cost of "endocrine complications" is a weighted average cost of treating each condition for each patient rather than a sum of each cost. The ERG requested that the company amend the model so that they are considered individually; however, the company did not do this as they considered that adding this complexity into the model would not result in significantly different results but would add to the level of uncertainty.

As it is likely that a successful transplant will lower the risk of complications, the omission of these aspects from the patient pathway is conservative, as it may underestimate the impact of these conditions upon costs and quality of life in transfusion dependent patients, i.e. those on the standard of care.

Timing of TI status after Zynteglo

In the model, patients who achieved transfusion independence were assumed to do so beginning at 12 months. As illustrated in Table 24 in the CS, TI-evaluable patients in the trials became transfusion independent from 12 to 24 months (time to reach TI ranged from [redacted] months), with a mean time to TI of [redacted]. While the timing of achievement of TI status has little impact in the company's base case model, where the cost of Zynteglo is subject to a simple discount applied at the time of treatment, it could be problematic should an outcomes-based payment scheme for

Zynteglo be introduced, when the timing of the assessment for the first outcome-based payment would be important.

Lack of correlation between patient characteristics and key outcomes

Although the model generated a range of patient characteristics to define the population, only age and gender were correlated with one another, and their interaction with other outcomes was limited. These limitations undermine the face validity of the model.

For example, in the company's base-case analysis, age was not linked to patient weight or iron level in each organ system, with the model taking the mean cohort value for all patients (Section 5.2.6). Patient weight in particular is strongly associated with age, although it is less clear how iron levels change over time.

There was also a lack of interaction between patient characteristics and event risk, and between the events themselves. Iron levels and the development of complications in each of the three organ systems were modelled independently of each other, and did not account for patient history to determine the rate of future events (e.g. time with iron levels to determine complications rate, or mortality). The time to iron normalisation was always modelled as being four years in TI patients, regardless of the patient's prior iron levels. The company justified this assumption as their trial data did not indicate a noticeable trend in length of time to move from higher iron levels to lower or normal iron levels post Zynteglo.

Overly complicated model structure

The company justified the use of an individual patient modelling approach as it has the possibility to allow the timing and order to vary between the various organ-specific iron overload events that could occur. They also considered that if the model took a cohort approach, an unfeasible number of health states would be required to model the range of iron levels in the three organ systems. A patient-level approach allows for interactions between variables to be captured, and patient history to be accounted for. The ERG considers that the model developed by the company adopted an overly complex structure but modelled outcomes in a simplistic manner. The model did not fully exploit the benefits of the patient-level approach, since outcomes in each organ system were modelled independently, and patient history was not always accounted for e.g. in the estimation of iron overload related complications. However, in order to benefit from the additional level of complexity that the patient-level model structure allows, it is necessary to be able to quantify the complex processes and relationships between patient characteristics and outcomes to inform the model. In many cases, contemporary data do not exist and it not possible to undertake these more realistic analyses.

Development of iron overload related complications

The risk of developing iron overload-related complications was based on a simplistic approach, which depended only on the iron load category, and did not consider the history or duration of iron overload or the patient's age.

It is possible that the model underestimates the risk of developing cardiac complications in the Zynteglo arm. Since the model did not take into account the patient's historic iron levels, very few patients in the Zynteglo arm developed cardiac complications (10% in the Zynteglo arm, compared to 41% for standard care). Patients who achieve transfusion independence experience normalised iron levels after four years in the model, at which point their complication risk is reassessed and fixed to zero. Therefore, transfusion independent patients are only at risk of such complications during the first four years of the model. However, some iron damage to the cardiac system is irreversible, and so patients who have developed some degree of damage before iron levels return to normal may continue to be at risk of complications, albeit to a lesser extent (discussed in more detail in Section 5.2.6.4).

Since the model did not take patients' age into account when estimating the time to development of complications, the actual age of onset varied considerably between patients, and lacked clinical plausibility. For example, cardiac complications typically present when a patient is in their 20s to 30s.[41] However, inspection of the sampled results for each patient for the company base case analysis showed that the age of onset of cardiac complications occurred equally at any age. This means that the modelled patient disease trajectory lacked face validity in this respect.

However, changing the model to account for patient history would require a fundamental change to the structure, and incorporation of data that does not presently exist. The company performed an SLR and search of real-world data sources to identify evidence of the relationship between iron levels at Time X and subsequent clinical events (i.e. complications) at Time Y, and it was not deemed possible to generate the required evidence in time for the appraisal.

5.2.2 The company's economic evaluation compared with the NICE reference case checklist

Table 6 summarises the ERG's assessment of whether the company's economic evaluation meets NICE's reference case and other methodological recommendations.

Table 6 Comparison of company's economic evaluation with NICE reference case

Attribute	Reference Case	Included in CS	Comment on whether *de novo* evaluation meets requirements of NICE reference case
Defining the decision problem	The scope developed by NICE	Yes	
Comparator(s)	As listed in the scope developed by NICE	Yes	
Perspective on costs	NHS and PSS	Partly	Only NHS costs have been considered.
Perspective on outcomes	All direct health effects, whether for patients or, when relevant, carers	Yes	The model considered QALY benefits to treated individuals.
Type of economic evaluation	Cost-utility analysis	Yes	
Time horizon	Sufficient to capture important differences in costs and outcomes between the technologies being compared.	Yes	The economic model uses a lifetime horizon, which caps survival at age 100.
Synthesis of evidence on health effects	Systematic review	Yes	
Source of data for measurement of health-related quality of life	Health effects should be expressed in QALYs. The EQ-5D is the preferred measure of health-related quality of life in adults.	Partly	Utilities were populated using a combination of sources, including the Chart Review study which used EQ-5D-3L. These values were used to estimate QALYs for untreated patients. Utilities derived from time trade-off interviews with the general public were used to estimate post-treatment utility scores. For some utility decrements, HUI-2 values were treated as equivalent to EQ-5D.
Source of data for measurement of health-related quality of life	Reported directly by patients and/or carers	Partly	Estimates of post-treatment HRQoL elicited from unaffected members of the public in TTO interviews. Utility of untreated patients based on EQ-5D. Disutilities associated with treatment and complications derived from TTO interviews with the US/UK public. Some estimated by healthcare professionals.

Source of preference data for valuation of changes in health-related quality of life	Representative sample of the UK population	Partly	Value set used was based only on healthy members of the UK general public.
Equity considerations	An additional QALY has the same weight regardless of the other characteristics of the individuals receiving the health benefit	Yes	No special weighting undertaken.
Evidence on resource use and costs	Costs should relate to NHS and PSS resources and should be valued using the prices relevant to the NHS and PSS.	Partly	Costs under control of personal and social services not considered. eMIT prices were not used.
Discounting	3.5% on costs and health benefits	No	Costs and benefits were discounted at 1.5% per annum.

5.2.3 Population

The patient population included in the cost-effectiveness analysis was patients 12 years and older with TDT who do not have a β^0/β^0 genotype. Section 3.1 provides further details on the population described in the licensed indication for Zynteglo.

The age and gender distribution of patients considered in the company's economic model were based on the [redacted] patients evaluable for transfusion independence from the clinical studies, HGB-204, HGB-205, and HGB-207. Clinical evidence on the effectiveness of Zynteglo is drawn from this analysis, see Section 3.1 for details. Age and gender were used to estimate mortality rates and baseline utility values, and some cost inputs.

The license for Zynteglo further describes that patients should be suitable for HSC transplantation. This specifically excludes patients with evidence of liver disease and patients with severely elevated cardiac iron (T2* <10 msec). This was also the exclusion criteria in the Zynteglo trials. The economic model, therefore, omitted patients with any iron overload-related complications at baseline (although these may develop at a later stage).

Other patient characteristics, such as weight and baseline iron levels, were based on the company's Chart Review (Section 4.3.5). The company considered the Chart Review to be a robust source reflecting current UK practice, and preferred its usage for these parameters over those of patients in their trials. Patient weight was specified in the model in order to calculate the costs of chelating agents which involve weight-based dosing. Mean patient weight was estimated [redacted], and this value was applied to all patients in the model. The baseline distribution of iron levels, presented in Table 7,

categorised patients as normal, low, moderate, or high iron based on serum ferritin, LIC, and myocardial T2* . Since patients with high cardiac iron loading would not be eligible for Zynteglo therapy, the distribution of Myocardial T2* was adjusted so no patients had high iron.

Table 7 Distribution of iron loading for transfusion-dependent patients (Table 48, CS)

Iron damage at baseline	Serum ferritin	LIC)	Myocardial T2*
Low iron	[redacted]	61%	88%
Moderate iron	[redacted]	23%	12%
High iron	[redacted]	16%	0%
Serum Ferritin: low iron, ≤1,000 ng/mL; moderate iron, 1,000-2,500 ng/mL; high iron, >2,500 ng/mL Liver Iron Concentration (LIC): low iron, <7 mg/g; moderate iron, 7-15 mg/g; high iron, ≥15 mg/g Myocardial T2*: low iron, >20 ms; moderate iron, 10-20 ms; high iron, <10 ms			

5.2.3.1 ERG comment

Several inconsistences between the populations in the Chart Review and the trial populations are described in Section 4.3.5. This impacts upon the estimation of baseline iron load and patient age. This also has implications for the estimation of HRQoL for standard care patients (Section 5.2.7) and the distribution of chelation therapy (Section 5.2.4), which were also based on the analyses of the Chart Review data.

Additionally, the ERG has a number of concerns around the modelled patient weight, comorbidities, and genotype distribution at baseline.

Age

The ERG considered that the most appropriate age range of patients was that of the patients in the Zynteglo trials, which were used to define the patient profiles in the model. However, a number of model inputs were derived from the company's Chart Review study, which included patients aged up to [redacted], whereas the trial populations from which Zynteglo data were derived was limited to patients aged 12-35 (see Section 5.2.4 for further detail). Table 8 illustrates the differences in the age distribution of the trial population whose efficacy data was used in the model, and the Chart Review. Most notably, [redacted] of the Chart Review population were over the age of 30 [redacted], whereas only 8.3% of the trial population were aged over 30 (all of whom were aged <35).

The characteristics, outcomes and HRQoL of patients in the older age ranges may not be reflective of patients under current clinical practice, due to changes in how patients have been managed over the last decade. They are also less likely to be eligible for Zynteglo treatment. However, the Chart Review analysis was not adjusted to account for this.

Table 8 Age distribution of patients in included trials vs Chart Review

Patient Age Distribution (Years)	Company base-case (clinical studies HGB-204, 205, 207) (n=[redacted])	Chart Review population (n = 165)
<12	[redacted]	[redacted]
12<18	[redacted]	[redacted]
18<30	[redacted]	[redacted]
30<40	[redacted]	[redacted]
40<50	[redacted]	[redacted]
50<60	[redacted]	[redacted]
≥60	[redacted]	[redacted]

Iron loading at baseline

The ERG considers the distribution of iron loading, derived from the Chart Review, to be broadly in line with the population based on estimates provided in the literature.[42] However, there were some differences between iron levels of the Chart Review population and the trial population; for example, more patients in the Zynteglo trials had low levels of cardiac iron (Table 32 of CS). As discussed in Section 4.3.5, this may due to the inclusion of a [redacted], who have been shown to be associated with higher levels of iron overload.[34] It may also be attributable to the patients in the Chart Review population who are older than those in the Zynteglo trials, who began management of iron overload before chelation and monitoring practices improved.

Baseline iron levels determine the rate at which patients develop complications, and so an overestimation of iron levels at baseline will, in turn, overestimate the rate at which complications develop. This will impact the standard of care arm more so than the Zynteglo arm, as the patients who achieve reduced levels of transfusions will develop lower rates of iron overload and, therefore, experience fewer complications (Section 5.2.6.4).

Patient weight

In the economic model, patient weight was used to calculate the dosage, and thus the cost of chelating agents. The company's base-case model fails to account for any interaction between age and weight, and assumes a constant weight of [redacted] throughout the patient's lifetime, which is likely to overestimate the weight of paediatric patients, placing a modelled 12 year old in the 98th percentile in the UK.[43, 44] Furthermore, the mean weight of patients included in the Zynteglo trials was [redacted]. An overestimation of patient weight in certain patient profiles will inflate the costs of standard of care and subsequently underestimate the ICER of Zynteglo. The company provided a scenario analysis which reduced the average weight of paediatric profiles in response to a request by the ERG. The scenario

adopted a weight of [redacted] for paediatric patients and [redacted] for adult patients, and the results can be seen in Section 5.2.10.3.

Iron overload-related complications

The modelled population excludes patients with the presence of iron overload-related complications at baseline. However, the company response to the ERG's clarification questions stated that "*individuals with TDT would not specifically be excluded from eligibility for Zynteglo due to either diabetes or hypogonadotrophic hypogonadism*." The Chart Review population included 20% of patients with hypogonadotrophic hypogonadism, [redacted] with diabetes and [redacted] with hypogonadism. The ERG believes patients with these complications should be included in the model to reflect baseline comorbidities in the population receiving Zynteglo. However, the number that would have endocrine disorders at baseline is difficult to determine. It is possible that the patients with these endocrine complications in the Chart Review also have other complications that preclude them from treatment with Zynteglo, but neither the CS nor the draft publication provided to the ERG provides this level of detail. These data are also not available from the Zynteglo trials, to provide a comparison with the Chart Review data.

The Chart Review population included patients with complications and high cardiac iron levels that would contraindicate treatment with Zynteglo. A number of important model parameters used in the model were based on the analysis of the Chart Review. This discrepancy in populations had a greater impact on the standard care arm, where the HRQoL throughout the modelled time period was based upon these analyses. This is discussed further in Section 5.2.7.

Representativeness of the distribution of genotypes

It is important to consider the heterogeneity of mutations covered by the marketing authorisation and whether the modelled population represents the distribution of genotypes that are present in the UK TDT population. As is described in Section 3.1, the trial population potentially underrepresents the severe non-β^0/β^0 mutations which are covered by the marketing authorisation and are prevalent in the UK. The underrepresentation of these difficult to treat genotypes with a potentially poorer prognosis may result in overestimated transfusion-related benefits for the Zynteglo-treated population, and underestimation of uncertainty in the cost-effectiveness results.

The Chart Review population does not exclude patients with β^0/β^0 mutations; however, clinical advice to the ERG suggested that other characteristics of TDT, such as HRQoL and transfusion and chelation requirements, are not influenced by genotype. Therefore, the analyses conducted by the company on their Chart Review data may not need to be adjusted for this characteristic.

5.2.4 Interventions and comparators

Zynteglo treatment

The intervention, as implemented in the model, consists of three stages, each of which comprises distinct processes and treatment costs. Further detail on the Zynteglo treatment process can be found in Section 3.2. In brief, these include mobilisation and apheresis of the patient's stem cells, hypertransfusion, pre-treatment and myeloablative conditioning. After these steps have been completed, the transduced cells (the Zynteglo product) are infused into the patient. This is followed by an inpatient stay of 21-42 days until engraftment of the infused cells has occurred.

The minimum recommended dose of Zynteglo is 5.0×10^6 CD34+ cells/kg (proportion of successfully transduced LVV+ cells varied between [redacted] in the HGB-207 study). In cases where multiple rounds of mobilisations were required, patients would receive more than one product lot of Zynteglo, and these were administrated in succession and considered one dose. Other drugs and dosages used in the executable model during steps 1 and 2 of the Zynteglo process are as follows:

- Filgrastim (Zarzio®) 1.0 million units/kg/day for 7 days;
- Plerixafor (Mozobil®), 0.24 mg/kg/day for 2 days;
- Busulfan, 0.8 mg/kg every 6 hours for 4 days;
- Ursodeoxycholic acid, 10mg/kg/day for 21 days;
- Clonazepam, 1mg/kg every 6 hours for 5.5 days.

Current standard of care

The comparator considered in the company's model was 'current care' for patients without an available HLA-matched related HSC donor. The current standard of care, therefore, consists of regular transfusions and iron chelation therapy, as detailed in Section 3.3. The company used the Chart Review described in Section 4.3.5 to define the management of standard of care (SoC) patients in their economic model. To determine the cost of chelation therapy in the model, patients were allocated to either oral (deferasirox, deferiprone), subcutaneous (desferrioxamine) or a combination of oral and subcutaneous chelation therapies, and incurred a weighted acquisition and monitoring cost for the iron chelating agents according to the proportion of patients on each therapy in the Chart Review (see Table 9), where [redacted] of patients received a combination of oral and subcutaneous chelation therapies.

The ERG requested that the company provide a reanalysis of the distribution of iron chelating agents excluding patients who would not have been eligible for enrolment in the Zynteglo trials. The results of this analysis are presented in Table 9. This appeared to show that [redacted]

[redacted]. These younger patients also represented [redacted] of those on two different oral chelation therapies in combination.

Table 9 Distribution of chelating agents in the company model (Table 44, CS)

Iron chelator	Mode of Administration	Distribution in full Chart Review population n (%)	Distribution in age 12-35 Chart Review reanalysis n (%)
Deferasirox	Oral	[redacted]	[redacted]
Deferiprone	Oral	[redacted]	[redacted]
Desferrioxamine	Subcutaneous	[redacted]	[redacted]
Deferiprone and Desferrioxamine	Oral and Subcutaneous	[redacted]	[redacted]
Deferiprone and Deferasirox	Oral	[redacted]	[redacted]
Deferasirox and Desferrioxamine	Oral and Subcutaneous	[redacted]	[redacted]

Standard of care patients were modelled to receive 13.5 transfusions per year, with those aged ≤18 years assumed to receive one pRBC unit per transfusion, while those aged >18 receive two units per infusion.

5.2.4.1 ERG comment

The ERG considers the intervention as implemented in the economic model to be in line with the licence. The comparator, i.e. blood transfusions and iron chelation therapy, is appropriate and in line with current practice in this population.

The ERG were concerned that the distribution of chelating agents may not be representative of those used in this population in practice. The relatively recent development of the evidence base around oral chelators and the safety and efficacy of combination therapy has resulted in uncertainty and inconsistency in clinical practice.

Clinical advice suggested that the iron load of older patients may no longer be adequately controlled by a single chelating agent, and thus these patients may be more likely to receive combination therapy. However, this did not appear to be reflected in the reanalysed Chart Review population, where a greater proportion of patients aged 12-35 were receiving combination therapy than the unrestricted population overall. It may be that the distributions in recent medical records are unstable and progressing towards combination therapy as clinical understanding and confidence with these drugs improves. This remains an area of uncertainty in the model.

It appears that in the model it is not possible for patients to receive two oral chelation therapies. These patients accounted for [redacted] in the company's analysis and [redacted] in the re-analysis of the Chart Review data. Therefore, the company model may underestimate the cost of chelation in a small proportion of patients.

5.2.5 Perspective, time horizon and discounting

The company's analysis adopted an NHS perspective only, and did not consider any costs incurred by Personal Social Services (PSS), which is not the perspective preferred in the NICE Methods guide.[7]

A lifetime horizon of 100 years was chosen as it was considered sufficient to capture all relevant differences in costs and benefits between the comparators. The ERG considers this an appropriate time horizon, as it is very unlikely that any patients would remain alive beyond this time period.

The economic model presented in the CS used a non-reference case discount rate of 1.5% for both cost and outcomes.

5.2.5.1 ERG comment

The ERG has a number of substantive concerns regarding the company's justification for the use of the non-reference case discount rate of 1.5% in the economic evaluation. The company's arguments are discussed in turn in the sections below, with specific focus on the implications for the economic analysis presented.

Zynteglo restores people who would otherwise die or have a very severely impaired life to full or near full health

The ERG has concerns with the company's position that the eligible population would otherwise die or have a very severely impaired life, and that Zynteglo restores these people to near or full health.

The ERG disputes the notion that without Zynteglo patients would otherwise die, given the uncertainty in the long term mortality of patients treated with transfusions and chelation. Estimates of life expectancy must be based upon current clinical management, and mortality figures cited by the company are based on up to 50-year old data. It is the ERG's understanding that evidence on projected life expectancy for patients treated optimally with current management strategies and therapies do not exist (see Section 5.2.6). Existing studies present limited follow-up data and enrol patients managed with different techniques and chelators increasing the uncertainty in the external validity of these results.[41] The lack of long term and generalisable survival data raises concerns regarding the statement that patients would "otherwise die". Furthermore, in the most recent edition of *Standards* published by the UK Thalassaemia Society,[45] it is stated that "the expectation is that well monitored and chelated patients will have a near normal life expectancy", which casts further doubt on the claim that conventionally treated patient would die.

The company's assertion that patients undergoing transfusion and chelation would otherwise have a severely impaired quality of life (HRQoL) is also not supported by existing evidence. As the company state in their CS, two studies in the literature derived utilities using EQ-5D and reported values of 0.86 and 0.87 for those with TDT on chelation therapy.[46, 47] These values suggest that TDT is associated with only a modest reduction in quality of life, and are supported by the analysis of 16-35 year olds with TDT in the Chart Review, whose HRQoL was reported as [redacted] (see Section 5.2.7). Data included in the UK Patient Preference Report provided by the company also seem not to support the assertion that patients on transfusions and chelation have a very severely impaired HRQoL.[48] Results showed 61% of TDT patients disagreed that beta thalassaemia significantly impacted upon their quality of life. Patients recruited to the Zynteglo trials also generally rated their HRQoL as similar to that of the general population prior to treatment.

Furthermore, in the most recent edition of *Standards for the Clinical Care of Children and Adults with Thalassaemia in the UK,*[45] it is stated that "discussion with the family must stress the usually excellent outcomes for children and adults managed conventionally with transfusion and chelation, now that monitoring for iron overload is more accurate and chelation choices are wider." The ERG would argue that it is, therefore, not likely that patients not treated with Zynteglo would otherwise have a "severely impaired" HRQoL, which is argued by the company to be comparable to patients with advanced stage cancer.

Finally, the ERG has concerns about the statement that Zynteglo restores the health of those treated to full or near full health. The ERG considers there to be insufficient evidence to conclude that Zynteglo restores individuals' health to full or near full, both in terms of length and quality. The trial data captured in HGB-204, HGB-205 and HGB-207 and provided in the CS does not provide sufficiently mature follow-up data to support this assumption. The company considered the HRQoL data captured in the trials was uncertain and potentially subject to bias and as a result, a vignette study was instead used. This evidence was based upon the UK general population's views on the descriptions of a pre- and post-Zynteglo health state, rather than from patients with beta-thalassaemia in the trial, as is recommended in the NICE reference case. As such, this evidence cannot be used to support the statement that Zynteglo restores people to full or near-full health.

Benefits are long term

As described in Section 5.2.6 and above, the ERG believes there is insufficient evidence to conclude with certainty that benefits of treatment will be truly life-long. Durable clinical efficacy has been demonstrated to 61.3 months in a small number of patients; however, with no data beyond this, the necessarily lifelong benefits of Zynteglo have not yet been demonstrated. This uncertainty applies to both persistence of transduced cells, and ongoing sufficient haemoglobin production of the graft, i.e. patients remaining transfusion-independent. On the subject of persistence of transduced CD34 cells vs

non-transduced cells, the company stated that peripheral blood vector copy number (PB VCN) serves as a surrogate marker for persistence of transduced HSCs and there had been no significant changes in PB VCN from Month 6 to last follow-up (Section 4.2.1). The ERG accepts that these results are promising but the evidence to suggest the life-long persistence of the graft in all patients is as yet unavailable.

Zynteglo will not commit the NHS to significant irrecoverable costs

The ERG contests the company's assertion that Zynteglo will not commit the NHS to significant irrecoverable costs. The ERG considers the [redacted] acquisition cost of Zynteglo to be significant, and one which could not be recovered by the NHS if the engraftment fails at any point, with patients reverting to transfusion and chelation therapy. As outlined in Section 5.2.4, there is limited long-term evidence of graft durability or transfusion independence, or reduction in iron loading. Should patients become transfusion-dependent at some point after Zynteglo, then not only would the full cost of Zynteglo have been incurred as well as any potential adverse events, such as infertility, leukaemia, and lymphoma, these costs would then be additional to the ongoing costs of transfusions and chelation. The ERG appreciates that such a scenario may be pessimistic; but the potential cost to the NHS associated with anything but an indefinite treatment effect in all patients could be large.

It is the ERG's understanding that the company have submitted a complex PAS to the Patient Access Scheme Liaison Unit (PASLU). The details of this PAS have not been finalised, however, the ERG understands the PAS consists of outcomes-based payments over time conditional on patients remaining transfusion-independent. The ERG has concerns over a PAS based on patient outcomes and length of time over which these outcomes will be judged. If payments are to be staggered over a time period [redacted], evidence from the clinical trials will not be able to address the uncertainty that there may be significant irrecoverable costs should patients become transfusion-dependent after the end of the payment period. In this case, the NHS will have paid the full acquisition cost and not received the lifetime benefits required to plausibly make Zynteglo a cost-effective treatment option. The timing and magnitude of the first payment must also be considered, when the complex PAS has been finalised.

Precedent

The company cite precedent as justification for the use of a 1.5% discount rate in the appraisal of gene therapies. The company submission states that as the lower discount rate of 1.5% was used in the appraisal of Strimvelis,[49] this should also apply to Zynteglo. Strimvelis, used for the treatment of adenosine deaminase deficiency–severe combined immunodeficiency, was evaluated under the highly specialised technology (HST) programme.[49] Within the NICE methods process guide for HSTs, the use of a non-reference case discount of 1.5% for both costs and health effects is permitted. In addition

to this, the HST guidance states that the use of the lower discount rate is for cases in which treatment has a very high likelihood of restoring people who would otherwise die or have a very severely impaired life to full or near full health, over a period of 30 years or more. In the appraisal of Strimvelis, overall survival data was available for up to 13 years. However, the NICE committee concluded that it was uncertain whether Strimvelis met the criteria to use a discount rate of 1.5% given uncertainties in whether the long-term benefits of treatment would be achieved.

As a result, the ERG believes there is little justification for the use of the non-reference case discount rate of 1.5%, and disputes all arguments put forward by the company in support of this assumption. The impact of using the reference case discount of 3.5% for costs and outcomes rather than 1.5% is explored in scenario analysis in Section 6.3.8.

5.2.6 Treatment effectiveness and extrapolation

5.2.6.1 Transfusion dependence

The clinical effectiveness of Zynteglo in the model is measured by the proportion of patients achieving transfusion independence. Transfusion independence is defined as being free from transfusions for 12 months and having sustained haemoglobin levels of above 9 g/dL. This was estimated from pooled results in all non-β^0/β^0 genotype subjects from studies HGB -204, HGB-205 and HGB-207 (see Section 4.2.2 for further details of the trials and the clinical data). A significant reduction in transfusions was stated as being a 60% reduction.

Of the 24 patients evaluable for TI status at the time of the submission, a total of 20 patients achieved TI (83.3%). Of the [redacted] patients who did not achieve TI, [redacted] ([redacted]) experienced significantly reduced transfusions. Of the remaining [redacted] ([redacted]) patients modelled as being transfusion-dependent, [redacted] remained transfusion-dependent following treatment with Zynteglo, and [redacted], and was modelled as being transfusion-dependent. In the original economic model, the company had assumed that the [redacted] patients who did not achieve TI were transfusion-reduced, but this was updated to the assumptions above following a clarification question from the ERG.

ERG comment

The ERG has some concerns regarding the applicability of these results to UK clinical practice, which were highlighted in Section 3.1. These pertained to the representation of the more severe genotypes in the trial population, and the refinement of the manufacturing process over time.

The trial population may not be generalizable to the UK, where patients with the IVS-I-110 or IVS-I-5 mutations, who are quite prevalent in the UK (Section 3.1), may be underrepresented. This will

potentially have an impact on the estimation of the rate of achieving transfusion independence. As these genotypes are known to be associated with reduced β-globin production, they may be less likely to achieve transfusion independence. The ERG has implemented an exploratory scenario in Section 6.3.2 whereby the proportion of patients who are transfusion independent has been re-estimated in accordance with the expected prevalence of IVS-I genotypes in the UK population.

These issues may be resolved upon further data collection, with the publication of the results from the HGB-212 trial enrolling the IVS-1-110 patients and using the optimised manufacturing process. At this point, it is difficult to predict the direction in which any future effectiveness estimates may go: as the procedure is optimised going forward, outcomes after Zynteglo treatment may be more optimistic than is currently estimated; however, the inclusion of the patients with more severe genotypes may imply less optimistic outcomes. The rate of transplant success was demonstrated to be a key driver of the model in the company's deterministic sensitivity analysis (Section 5.2.9.3), and so it is imperative that this uncertainty is considered during the decision making process.

5.2.6.2 Engraftment success and graft durability

The engraftment procedure was assumed to be successful in all patients, as there were no engraftment rejections (failures) in the trials. However, the need to collect back-up cells for rescue treatment, acknowledges that such a risk exists. It is possible that it occurs in such a small number of patients that it has not yet occurred during the limited data collection to date. It is also possible that moving from the trial setting to general practice, the risk will become more apparent.

In the long-term, it was assumed that there would be no loss of graft, that is, no patients would lose their transfusion independence status either by experiencing reduced haemoglobin levels or a return to transfusion, and that transfusion-reduced patients would not experience an increase in the need for transfusions or return to transfusion dependency over time.

ERG comment

There is currently insufficient available follow-up data for patients in the trials to determine whether permanent long-term engraftment occurs. The duration of transfusion independence was censored at the last Hb assessment and no events for loss of transfusion independence have yet been recorded (page 92 of CS). The clinical advisor to the ERG considered that the assumption of permanent engraftment among all patients was potentially over-optimistic.

[redacted] (page 89 of CS, discussed in Section 4.2.1). The company claim that the integrated transgene is stable in the HSCs, as no significant changes have been observed in PB VCN (peripheral blood vector copy number, used as a surrogate

marker for persistence of transfused cells) from month 6 to last follow up. The ERG considers that it is reasonable to assume the transgene is stable as long as TI status persists; follow-up data collected in LTF-303 will provide more evidence to support this assumption.

However, there is concern around the consequences of returning the patient's original cells (i.e. those that are not gene-corrected) in the Zynteglo product back to the patient, and that it may lead to decreased haemoglobin production and a return to a need for transfusions.

[redacted]

5.2.6.3 Organ-specific iron overload

Baseline levels of iron overload (discussed in Section 4.3.5), were based on the mean of the population studied in the Chart Review, with each simulated patient randomly assigned to an overload risk category (low, medium, high risk) for the cardiac, liver and endocrine systems. As mentioned in Section 5.2, these were assigned independently to each other and to the patient's age. Post-baseline, iron levels were dependent on the patient's transfusion status.

Transfusion-dependent patients
Iron levels in patients who are transfusion dependent were assumed to remain at their baseline levels throughout the time horizon of the model. These included all patients in the standard care arm, and the proportion of patients receiving Zynteglo who remained transfusion dependent.

Transfusion-independent patients
Transfusion-independent patients were assumed to achieve normalised iron levels in all organ systems by four years from initial treatment with Zynteglo.

Since the current evidence for long-term changes in iron levels following an auto-HSCT procedure such as Zynteglo is limited, the company searched for data on iron levels in patients that have received allogeneic-HSCT to inform the model. The company identified two papers that described the time course of iron store changes post-transplant specific to patients with TDT; however, the data provided by these studies was limited. An RCT comparing patients on deferasirox (n=12) to those receiving phlebotomy (n=14) at one year after HSCT, found reductions in LIC after one year; however these remained within the "moderate risk" category after one year [50]. One retrospective study with a longer follow-up period followed a small number of paediatric patients for four years, and found a 47% reduction in median ferritin levels after four years [36]. The company stated that the results

of this study supported an assumption of a 4-year iron normalisation period following Zynteglo therapy (p172 of CS).

Reduced transfusions

Patients who achieved meaningful transfusion reduction were assumed to also achieve reduced levels of iron, on the basis that ongoing chelation can "'catch up' on the existing iron stores and start reducing iron" (page 172 of CS).

In the model, reduced levels of iron were assumed to be achieved one year after initial Zynteglo treatment. Patients were assigned to either the low or medium overload risk categories, with relatively more patients having low iron levels compared to moderate iron levels (Table 10). It was assumed that no patient would achieve normalised iron levels, given the ongoing need for transfusions, and there would be no patients in the high iron categories. However, it was not clear how the proportions of patients in the low and moderate categories were derived.

Table 10 Distribution of iron loading for transfusion reduced patients (adapted from Table 49, CS)

Iron Normalisation at 4 Years Post-transplant	Distribution of Iron Loading			
	Serum Ferritin	LIC	Myocardial T2*	
			LIC ≥15 mg/g (High)	LIC <15 mg/g (Moderate/ Low)
Transfusion-reduced patients				
Normalised Iron	0%	0%	0%	0%
Low Iron	48%	74.5%	92.5%	92.5%
Moderate Iron	53%	25.5%	7.5%	7.5%
High Iron	0%	0%	0%	0%
Serum Ferritin: low iron, ≤1,000 ng/mL; moderate iron, 1,000-2,500 ng/mL; high iron, >2,500 ng/mL Liver Iron Concentration: low iron, <7 mg/g; moderate iron, 7-15 mg/g; high iron, ≥15 mg/g Myocardial T2*: low iron, >20 ms; moderate iron, 10-20 ms; high iron, <10 ms				

ERG comment

The ERG considers that the studies identified and described by the company did not support their assumption of iron normalisation in all patients,[36, 50] and that an assumed four year time to normalisation may be too optimistic. While both studies demonstrate that iron levels do reduce following an allo-HSCT, neither study demonstrates a time point by which all patients achieve normalised levels of iron. The authors of one study quoted that "only ≥4 years did ferritin levels begin to trend downward with a median 870 ng/mL".[36] The studies themselves are associated with a number of limitations, namely that they do not report on iron levels in all organ systems included in the model and they include only small patient numbers. In Chaudhury (2017), there was large variation in iron levels post-transplant, as observed by the wide confidence intervals reported for each data point.[36]

However, the studies are of limited generalizability to Zynteglo, since they included patients of a younger age and there is some evidence to suggest that age is significantly associated with post-transplant iron levels.[51] Therefore, older patients eligible for auto-HSCT may take even longer to achieve normalised iron levels.

However, the data collected to date for iron normalisation in patients who achieve transfusion independence following Zynteglo are too limited to determine the rate at which it occurs (Section 4.2.2.2). At the 48 month follow up, there remained a number of patients with moderate to high levels of serum ferritin and LIC (Table 11).

Table 11 Iron distribution (adapted from Table 22, clarification response)

Iron level	Serum Ferritin	Liver Iron	Cardiac T2*
Baseline iron levels (trial population, n=[redacted])			
Low	[redacted]	[redacted]	[redacted]
Moderate	[redacted]	[redacted]	[redacted]
High	[redacted]	[redacted]	[redacted]
Iron levels at 48 months (trial population, status of transfusion independence, n=[redacted])			
Low	[redacted]	[redacted]	[redacted]
Moderate	[redacted]	[redacted]	[redacted]
High	[redacted]	[redacted]	[redacted]
Serum Ferritin: low iron, ≤1,000 ng/mL; moderate iron, 1,000-2,500 ng/mL; high iron, >2,500 ng/mL Liver Iron Concentration: low iron, <7 mg/g; moderate iron, 7-15 mg/g; high iron, ≥15 mg/g Myocardial T2*: low iron, >20 ms; moderate iron, 10-20 ms; high iron, <10 ms			

The model also does not contain a function of iron decline over time for those who experience transfusion independence or reduction. Post-transplant iron levels are not linked to patient characteristics or pre-transplant iron levels. There is no modelled gradual decline in iron throughout the normalisation period, as iron levels are allocated at baseline and then re-assigned to post-transplant (normalised or otherwise) levels at the end of the normalisation period. This method overestimates exposure to iron overload during the duration of the normalisation period, which results in greater exposure to risk of iron-related complications, resulting in a conservative estimate of cost-effectiveness.

5.2.6.4 Complications from iron overload

To predict the complications of iron overload, the model uses literature-based rates and risk equations to estimate the rate of developing complications based on distribution of iron levels in the heart, liver, and serum (ferritin), provided in Table 12. The company undertook a systematic literature review to identify complication rates (Section 4.1.5). Cardiac disease outcomes were reported in 17 studies,

liver-related complications in 18 studies, and a single study was identified for diabetes and hypogonadism.

The study selected to provide the rate of developing cardiac complications reported that 2.2% of patients with low iron had heart failure, compared to 3.9% with moderate iron and 12.5% with high iron.[52] The mean time to heart failure onset was reported as 24.81 months (appearing at a mean age of 32.11 years), and this was used to estimate an annualised rate for each iron overload risk category. The study also reported the proportion of patients who developed arrhythmias and pulmonary hypertension, although only the rates for heart failure were used in the model.

Angelucci (2002) was the only study identified by the company that reported liver-related complications by iron level.[29] These data were used to estimate an annual rate of liver complications for high iron. Patients with medium and low risk from iron overload were assumed not to have a risk of developing liver complications.

The systematic review identified a single source which followed up 92 patients in the UK for signs of endocrinopathies.[30] The study reported a prevalence rate for diabetes of 41% and a rate of 67% for hypogonadism, the two most common endocrinopathies. Historic iron levels, including the patients' worst myocardial T2 and LIC levels and their mean serum ferritin levels, were analysed for associations with endocrinopathies. The rate of the development of these conditions was found to be significantly associated with cardiac iron (myocardial T2) as well as age. The company extracted odds ratios for variables included in their multivariate analyses, and estimated a risk equation to predict diabetes and hypogonadism based on age, myocardial and serum iron levels. As noted in Section 5.2.1.1, patients could develop diabetes or hypogonadism but not both concurrently.

Table 12 Predicting complications of iron overload (Table 50, CS)

	Annual Rate to Develop Complication		
Iron Loading	**Cardiac Complications**	**Liver Complications**	**Source:**
Low Iron	0.011	0.000	Pepe *et al.*, Angelucci *et al.*[28, 53]
Moderate Iron	0.019	0.000	
High Iron	0.065	0.083	
	Risk Equation for Other Complications		
Coefficient Names	**Diabetes Mellitus**	**Hypogonadism**	**Source:**
Intercept	-6.642	-2.921	Ang *et al.*[30]
Myocardial T2	2.960	1.361	
Age	0.095	0.095	
Ferritin	2.695	1.065	
Duration of Follow-up	8 years	8 years	

ERG comment

The ERG was concerned that the total impact of developing cardiac complications is not accounted for, and that there is an inconsistency with how cardiac complications are modelled. Only the rate of heart failure from the study was applied in the model; however, the costs of managing cardiac complications reflected a range of conditions including those with arrhythmia, which accounted for almost half of cardiac complications identified in the study.[21, 52] However, it is possible that the use of data from this study means that the model overestimates the proportion of cardiac complications that result in excess mortality. The study is considered to be outdated, with patients first scanned between 2008 and 2012. As discussed in Section 4.3.4, improved monitoring of cardiac symptoms may result in increased detection of abnormalities, but these patients are able to be managed early and optimally, resulting in fewer of the more severe complications and ultimately improved survival.

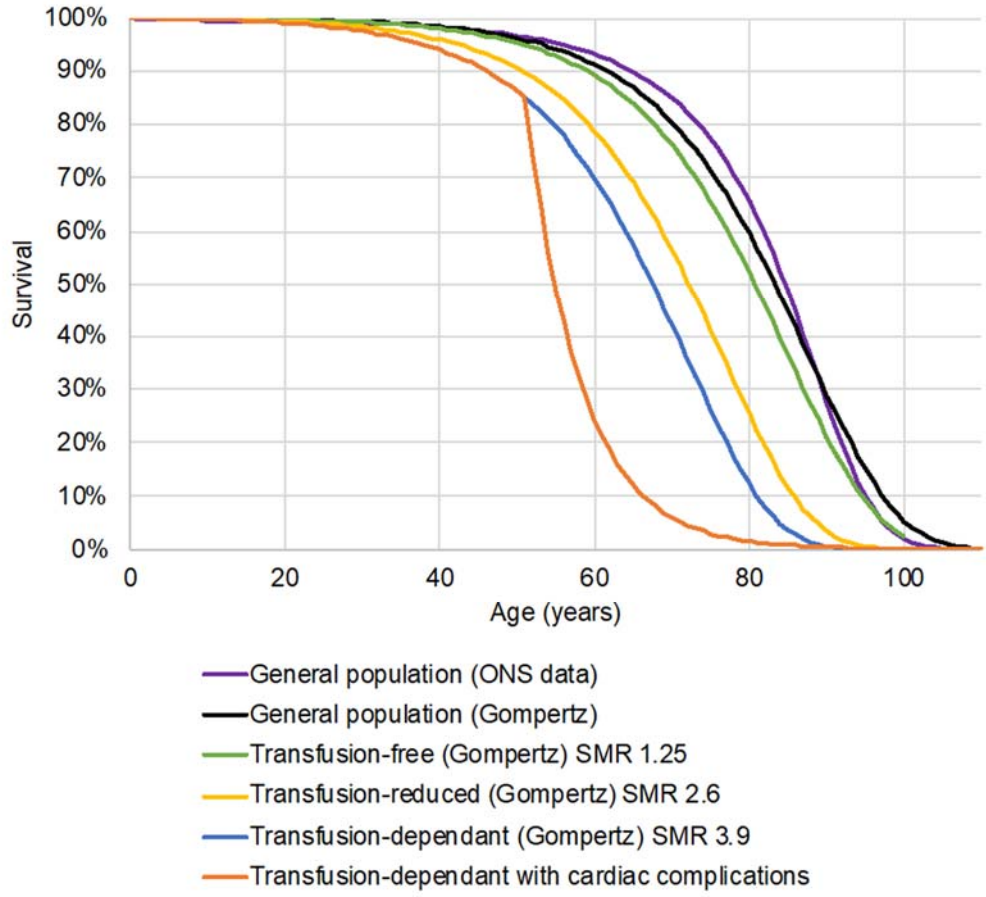
There was additional concern regarding the assumption that transfusion independent patients who achieve normalised iron levels are no longer at risk of developing iron overload-related complications. However, current evidence is limited and often contradictory. There may be a degree of pre-existing irreversible damage before treatment with Zynteglo, albeit sufficiently small to allow for eligibility for treatment, which could theoretically result in a long-term risk of developing complications. Looking to experience from the allo-HSCT population, occurrence of late hepatic, endocrine and cardiovascular complications, related to past and residual iron overload, have been acknowledged.[54] This is considered particularly relevant when transplant is performed in older children, adolescents or adults, and in patients who have received inadequate chelation therapy before HSCT. However, a French retrospective study of 99 patients found that very few patients developed a cardiac insufficiency, although no cardiac MRI was available at the onset of cardiac symptoms to allow investigation of a possible cardiac iron overload.[51, 55]

The long-term consequences of iron damage in patients who achieve transfusion independence remains an uncertainty in this analysis, and it is possible that the elevated risk of mortality modelled for these patients (Section 5.2.6.5) captures some of the impact of any complications that occur.

5.2.6.5 Mortality

In the absence of direct long-term survival data for a TDT population after treatment with Zynteglo, it was necessary to use a range of external sources to predict long-term survival. To model mortality in each treatment arm, the company applied an SMR based on transfusion-dependency status to the general population mortality. Mortality following the development of cardiac complications was modelled separately, to account for the specific impact of this aspect of the condition.

Figure 3 Illustration of SMRs applied to the general population of England and Wales (Fig 32, CS)



Age-related mortality

Age-related mortality was modelled using a Gompertz curve fit to life tables for England and Wales.[56] The purpose of fitting a curve rather than applying the survival data directly in the model is that it allows the application of a modifier, or a mortality hazard ratio, to adjust the survival for patients with TDT. Figure 3 allows for a visual inspection of fit between the fitted and observed survival estimates: the Gompertz provides an underestimate of survival between ages 60 and 85 and an overestimate thereafter, although both tend to zero at approximately the same time. No other survival models were explored.

Transfusion-independent mortality

Transfusion-independent patients in the model were assumed to have survival that is slightly worse than that of patients in the general population, with an SMR assumed to be 1.25. The company justified the use of an assumed rather than a literature-derived value due to insufficient evidence of the natural history following transplant in patients with thalassaemia, as no adult transplants have been endorsed or funded in the UK to date. The SMR was assumed to capture the potential mortality impact of myeloablative conditioning associated with the Zynteglo procedure.

Transfusion-dependent mortality

To estimate mortality in patients with cardiac complications and transfusion-dependent patients (without cardiac complications), two cost-effectiveness analyses identified by the company (described in Section 5.1 and Section 4.3.4) provided information for survival in TDT patients. An SMR of 3.9 was reported in a cost-effectiveness analysis of deferasirox and desferrioxamine in TDT, for patients who did not experience any cardiac complications.[21] This used data reported by a US study, which estimated mortality from 257 TDT patients on desferrioxamine followed between 1965 and 1994.[22] The SMR was estimated from the mortality of these patients relative to US 1998 mortality data.

In the absence of a specific SMR for patients with reduced transfusions, it was assumed that the mortality ratio was the mid-point of transition-dependent (3.9) and transfusion-independent patients (1.25), resulting in an assumed SMR of 2.6.

Cardiac mortality

The company states that cardiac dysfunction is considered to be the leading life-limiting complication in patients with TDT, due to cardiac iron accumulation. To capture the importance of cardiac mortality for patients with thalassaemia, mortality associated with the development of cardiac complications was modelled separately. All patients who acquire cardiac disease (regardless of transfusion status) were modelled as experiencing an annual mortality rate of 13%, based on a study which reported 48% survival after heart failure at 5 years.[25] The two cost-effectiveness analyses identified by the company both applied this rate in their analyses.[21, 39]

ERG comment

Mortality in transfusion-independent patients

The ERG considers that it is appropriate to model an elevated mortality for these patients compared to the general population, to capture the potential mortality impact of myeloablative conditioning associated with the Zynteglo procedure. While they will have had no iron-related complications to date (otherwise they would not be eligible for Zynteglo), they may have iron overload-related damage to a certain extent that might occur prior to transplant, which may be associated with an additional mortality risk. While there is no data available to determine the extent of this, the ERG considers that the SMR used by the company appears plausible.

Mortality in transfusion-dependent patients

The approach taken to model mortality was linked to transfusion status rather than the patients' iron levels. With greater clinician experience, alongside improved monitoring and iron chelation practices observed over the last decade, iron levels in TDT are more likely to be well controlled, and transfusion-dependent patients will have better mortality than is predicted by the older studies.

The study used to model transfusion-dependent mortality in this appraisal was outdated and of limited generalisability to the present decision problem, being based on a sample of subcutaneous chelation patients.[22] Further still, patients in the current decision problem would have lower iron levels and fewer iron overload-related complications than an unrestricted TDT population, with eligibility for Zynteglo requiring patients to be sufficiently fit to undergo the procedure.

There is some additional uncertainty regarding the estimation of the SMR value for transfusion-dependent patients. The SMR value for TDT, estimated by Delea from survival data reported by Gabutti, was "consistent with the range of parameters identified in Karnon et al (2012)" who also based their mortality rates on the Gabutti study.[21, 22, 39] Though the Karnon study did not provide a preferred point estimate, the range of RRs was between 1.1 and 3.91: the SMR applied by the company falls at the more pessimistic end of this range.[39] In the absence of a more contemporary source, the ERG has explored the impact of assuming lower SMRs in Section 6.3.6.

The same SMR was applied to patients in the standard care arm and the patients who remained transfusion-dependent after treatment with Zynteglo. Since patients are exposed to additional mortality risks due to the myeloablative conditioning procedure that occurs prior to Zynteglo, it is possible that transfusion-dependent patients after Zynteglo treatment may have an excess mortality that is higher than transfusion-dependent patients in the standard care arm.

The ERG noted that in the company's economic model, the majority of patients who developed cardiac complications died from said complications. The ERG considers that mortality related to cardiac disease is likely overestimated. As described in Section 4.3.4, improvements in monitoring of cardiac symptoms have led to improved patient outcomes after developing cardiac complications.

5.2.6.6 Adverse events of treatment

Associated with Zynteglo treatment

Patients face a risk of becoming infertile due to the use of myeloablative conditioning with transplantation. Data on infertility and sub-fertility following allogeneic HSCT is used as a proxy for possible infertility and sub-fertility following Zynteglo therapy. A study in Thailand[35] in stem-cell-transplanted, thalassaemic survivors, with infertility measured by gonadal dysfunction, reported 48% of men and 77% of women are infertile in a population receiving a mixture of standard and reduced intensity conditioning regimens. The company's UK HES analysis has indicated that 23.9% of male patients with TDT have testicular dysfunction and 19.5% of female patients have ovarian dysfunction.[57] The difference between these rates with TDT and the rates of infertility following myeloablative conditioning are applied in the model, to represent the incremental impact of infertility following myeloablative conditioning on TDT patients who receive Zynteglo. Thus, it is assumed that Zynteglo increases infertility by 24% in men and 57% in women.

Other adverse events specifically associated with Zynteglo (Section 4.2.3), are not explicitly modelled as it is assumed that the cost impact of the adverse events is captured by administration, hospitalisation and ongoing monitoring costs, and that the quality of life impact is reflected in the utility decrement associated with transplantation.

Associated with iron chelation

AEs associated with chelation therapy were included in the model if they occurred with an annual incidence of >5%. For deferiprone and deferasirox, AE data was taken from the respective prescribing information for each drug.[31, 32] For desferrioxamine, AE data were obtained from a published systematic literature review.[33] Rates are provided in Table 51 in the CS.

As with the adverse events of Zynteglo treatment, the model assumes that each adverse event disutility is captured in the health state utility. Costs of managing AEs are described in Section 5.2.8.

ERG comment

Treatment with Zynteglo is theoretically associated with the risk of insertional mutagenesis, potentially leading to development of malignancy. As per the SmPC for Zynteglo, patients are required to be monitored annually for signs of leukaemia or lymphoma.[58] The risk of developing leukaemia has not been included in the economic analysis, presumably due to a lack of sufficiently long-term data in sufficient numbers of patients to allow the estimation of malignancy risk to be estimated.

It was unclear what grade of adverse events were modelled for chelation therapy. Typically, only events of grade 3 and above are included in an economic analysis, because these have significant consequences on costs and HRQoL, and typically require hospitalisation. The wording of the description of the events in the publication suggests that the rates presented are for events of all grades, and the costs applied to these AEs are very low.

5.2.7 Health related quality of life

The company considered three main sources of evidence on HRQoL in the parameterisation of the economic model. These comprised the Zynteglo trials; a vignette study, in which time trade-off interviews were conducted with members of the general public to elicit utilities associated with TDT and its treatment; and a review of the medical records of UK TDT patients commissioned by the company (referred to as the Chart Review).[18, 59] Utility values estimated from the vignette study and the Chart Review were applied in the economic model, and were supplemented by values sourced from a number of published studies. For each health state and event, the company estimated the associated utility decrement, which was applied to general population utility norms in order to capture the impact of ageing.

Modelled patients incur a HRQoL decrement based on their transfusion status. Transfusion reduced patients incurred a disutility of 0.13, which was based on an assumption of a linear improvement in HRQoL between transfusion dependence and independence. In the company's model, transfusion reduced patients achieve a mean reduction of [redacted] in the number of transfusions per year. The company therefore assume the TD decrement of 0.27 is also reduced by [redacted] for TR patients.

A summary of the utility decrements applied in the model are provided in Table 13.

Table 13 Utility decrements used in the company base-case model (Table 56, CS)

Health state	Company base-case decrement	Assumption source
Zynteglo treatment	0.31	Vignette study
Up to one year post-Zynteglo	0.31	Vignette study
Transfusion-independent	0.02	Vignette study
Transfusion-reduced	0.13	Assumption
Transfusion-dependent	0.27	Chart Review [18]
Infertility	0.07	Busnelli *et al*., 2015; Scotland *et al*., 2011 [60, 61]
Cardiac Complications	0.11	Karnon *et al*., 2012 [39]
Liver Complications	0.067	Assumption
Endocrine Complications	0.074	Karnon *et al*., 2012 [39]

5.2.7.1 Trial-based utility values

HRQoL evidence was collected in the HGB-204, HGB-205, and HGB-207 studies (see Section 4.2.2). The company reported that for many patients in HGB-204 and HGB-205, baseline assessments were not performed. High baseline utility values were observed, and the general trend in change over time was of an initial reduction in HRQoL and a [redacted] by 12 months following receipt of Zynteglo. The company considered the baseline values to be "artificially high", which may be due to an "adjustment bias", making it difficult to detect a treatment benefit following Zynteglo. As such, the company elected not to apply the utility values estimated from EQ-5D data collected in their trials in the economic model.

5.2.7.2 Health state utility values

HRQoL associated with Zynteglo

The company conducted a vignette study to inform a number of important assumptions on quality of life of patients with TDT, including HRQoL in the time following treatment with gene therapy (Zynteglo) and of those achieving transfusion independence (following treatment with Zynteglo).[59] This study consisted of time trade-off interviews with members of the general public who were unaffected by TDT (n=207). Participants were asked to value eight TDT-related health state vignettes

which described disease and treatment scenarios in terms of a series of health events (symptoms, treatment, adverse events etc.).

The vignette study indicated a utility of 0.93 for transfusion independent patients, which the company concluded to be 0.02 lower than the general population value for this age group (40 to 45 years) using a set of population utility norms that were reported for individuals without comorbidities and no history of any health condition.[62]

The vignette study also estimated that patients would have a utility value of [redacted] in the year after treatment with Zynteglo. The public rated a vignette which describes the transplant and recovery process after gene therapy. The ERG consider this value to be appropriate and largely in line with the magnitude of disutility applied in previous appraisals for HSCT, given that Zynteglo would not be associated with GvHD.

HRQoL associated with TDT

The company commissioned a review of the medical records of 165 patients with TDT in the UK (the Chart Review), in response to the perceived lack of appropriate HRQoL data (see Section 4.3.5).[18] Utility scores were generated using the EQ-5D-3L questionnaire scores for [redacted] adult (≥16 years) patients, and the EQ-5D-Y for [redacted] patients aged 7-15 years. The mean of all patients aged ≥16 included in the Chart Review was estimated as 0.69, resulting in a utility decrement estimated by the company and applied to transfusion-dependent patients in the model of 0.27. Similarly to the transfusion independent utility value estimated from the vignette study, the decrement was estimated using a set of population utility norms that were reported for individuals without comorbidities (0.96). The company concluded that due to the similarity of this figure with that generated in the vignette study, this utility was appropriate and robust. Using the general population utility which is preferred by the ERG (0.93), the company's preferred disutility would be [redacted].[62]

The vignette study also provided an estimate for TDT, although the company considered that patient-elicited utility values were more appropriate for inclusion in the model. The public valued a health state describing life with TDT treated with transfusions and oral chelation agents with a utility score of 0.73; for those on transfusions and subcutaneous chelation agents this was 0.63. While the utilities associated with oral and subcutaneous chelation in TDT were not used directly in the company's base-case, they provided important support for their argument that patients with TDT have a 'very severely impaired life'.

ERG comment

Adjustment of HRQoL for age

Age-related decrements in quality of life were not appropriately captured in the company's model in

two fundamental ways. Firstly, an inappropriate value set was selected as the basis of the general population utility estimates. The company used age-stratified EQ-5D scores derived from a large population study by Ara and Brazier,[62] which has extensive precedent in previous NICE technology appraisals. However, they selected a subset of this population, which excludes all individuals with a history of a health condition, and with any ongoing health conditions. This is clearly inappropriate; by selecting out only individuals in perfect health in each age group, the company implicitly assume that patients with thalassaemia will never develop any other unrelated health condition during their lifetime. These patients will essentially remain in perfect health until death, excepting any previously incurred treatment related decrements associated with their TDT. This had an important effect upon the HRQoL of modelled patients, meaning that the baseline utility of a patient aged 75 would be higher than someone aged 30 in the general population. Compared to the correct general population value set, (i.e. inclusive of all individuals irrespective of health status), the modelled age-based utilities were 0.08 higher at age 40, 0.11 at age 60, and 0.19 at age 80.

Secondly, a modelling error meant that utilities were not regularly adjusted for age beyond the end of the first year of the model. This error was corrected by the company upon request by the ERG (Section 5.2.10).

HRQoL of transfusion dependent patients

The company's submission acknowledged that there is considerable existing evidence on the quality of life of patients with thalassaemia. One study elicited utilities from 196 TDT patients using the EQ-5D questionnaire.[46] These patients had a utility of 0.86 (95% CI 0.83 – 0.89), which, along with the other literature-derived estimates, the company described as 'unexpectedly high', and that they didn't appear to reflect the 'considerable burden' of the disease and its treatment. The relatively high HRQoL estimates derived from TDT patients were dismissed as being due to 'adaptation to life with a chronic condition', and therefore the company deem them not to truly represent the burden of the disease. However, the ERG disagrees that these values are too high, and that the HRQoL of TDT patients has been underestimated by the company. A patient preference report submitted by the company found that it is anticipated that only [redacted] of patients would be willing to accept treatment with Zynteglo, and only [redacted] said that beta thalassaemia significantly impacted upon their quality of life.[48] This may support the idea that current management is acceptable to most patients – rather than TDT resulting in a 'very severely impaired life', as argued by the company (Section 5.2.5).

In the clarification questions submitted to the company, the ERG questioned the use of the utility derived from the whole Chart Review population (0.69) for a number of reasons. Firstly, the demographics of the patients whose data were collected in the Chart Review appeared to differ substantively from those included in the Zynteglo trials. Table 8 in Section 5.2.3 illustrates the differences in the age distribution of the trial population whose efficacy data was used in the model,

and the Chart Review from which the modelled HRQoL values were derived. As utilities were adjusted over time to reflect the natural decline in HRQoL associated with age, the model would also have double counted the effects of aging due to the HRQoL having originally been derived from an older population.

Due to the long-term effects of transfusions and chelation therapy, and historic differences in disease management, it might be expected that the HRQoL of older patients to be lower than that of optimally managed younger patients. Therefore, the use of a mean HRQoL based on an older population for transfusion dependent patients in the model may be inappropriate. The ERG requested that the company re-analyse the HRQoL data in the Chart Review, limiting the population to only those aged from 12-35 years, and by matching to the age distribution in the three Zynteglo trials, i.e. the base-case model distribution.

The ERG also requested that, for this analysis, the company exclude patients with 'high' cardiac T2*, and patients whose existing co-morbidities were already accounted for separately in the model (to avoid double counting), or would otherwise be precluded from enrolling in the Zynteglo trials. Results were only available for TDT patients aged 16-35 who completed the EQ-5D-3L questionnaire, in whom the mean utility was [redacted] Thus, the most appropriate decrement applied in the model for TDT patients would be [redacted], i.e. the general population utility for people aged <30, 0.9383, minus the mean utility of patients aged 16-35 from the Chart Review, [redacted]. This value may be an underestimate, as at least one participant in this group rated their HRQoL as [redacted] which may be a reporting error and could have had an effect on the mean estimate.

Treatment-related utility decrements

The model assumes that only the frequency of transfusions has an effect on the disutility incurred, and a constant decrement is applied regardless of the type of chelation therapy received. This appears to be inconsistent with the results of the company's vignette study, in which the utility of TDT patients treated with oral chelation was rated 0.10 higher than for those on subcutaneous chelation (0.73 vs 0.63).

Zynteglo patients continue to receive iron chelation therapy during the 'iron normalisation' period, which, in the company's model, lasts for four years from Zynteglo infusion. However, utility decrements associated with the chelation therapies themselves were not applied in this period. It may therefore be plausible that transfusion independent patients treated with subcutaneous iron chelation therapy should incur a disutility for the duration of the iron normalisation period. A scenario applying a disutility to TI patients still on subcutaneous chelation therapy is presented in Section 6.3.9.3.

5.2.7.3 Complications and adverse event-related disutilities

While the model justifiably assumed that disutilities associated with individual adverse events would be captured in the utilities, the company separately applied decrements for broad groups of co-morbidities: infertility due to myeloablative conditioning, and for complications that develop due to iron overload including cardiac complications, liver complications, and endocrine complications (see Table 13 for a summary of utility decrements).

A disutility of 0.11 for cardiac complications was sourced from Karnon *et al*. (2012), a cost-utility analysis of deferasirox in TDT, which reported a decrement of 0.114 based on a 1993 study by Fryback *et al*.[39, 63] This study comprised time trade-off interviews with 1,356 people from a single US town between 1988 and 1990. The company also cite this cost-effectiveness study as the source of disutility estimates for endocrine complications, which again are originally from Fryback *et al.* In the company's model, the onset of a first endocrine complication (diabetes, hypogonadism etc.) is associated with a life-long utility decrement of 0.074. In the model presented by Karnon *et al*., diabetes was associated with a life-long utility decrement of 0.133, and all other endocrine disorders considered were each assumed to incur an additional decrement of 0.074, based on an assumption that each was worth half the disutility of diabetes.

A utility decrement of 0.07 associated with infertility was applied during 'childbearing years' – classed by the company as being between the ages of 15 and 45. This value was ultimately derived from a US study in which a committee used the HUI-2 scale to rate the quality of life of a female made infertile by gonorrhoea infection who wanted to have children.[64, 65]

ERG comment

Disutility associated with endocrine complications

Given more recent studies in which the EQ-5D was administered directly to patients with diabetes, a disutility of 0.133 for diabetes may be too large, particularly given advances in management of the condition over the 30 years since this value was elicited.[66, 67] The ERG therefore considers the company's preferred value of 0.074 to be reasonable for diabetes, but that other important endocrine disorders should have been associated with a separate disutility.

Disutility associated with infertility

While the value applied for infertility has precedent in previous cost-effectiveness models of fertility treatment, the ERG questions the appropriateness of its use in this disease area. The impact of infertility upon HRQoL is poorly understood and not typically well captured using EQ-5D. There is also a lack of reliable literature appropriately disentangling the effect of infertility itself from its most common causes and associated co-morbidities (e.g. endometriosis). As such, the ERG explores the

impact of removing the infertility-related decrement from the cost-effectiveness analysis in Section 6.3.9.4.

The ERG also questions the time frame within which the infertility decrement is applied; this appears to have been selected arbitrarily, and it is uncertain whether patients aged 15 would consider their lives to be adversely affected by an inability to have children. The studies cited in support of this value sought to capture the impact of infertility in individuals actively seeking fertility treatment, most of whom were around the age of 40. There may therefore be many patients aged >45 whose lives were adversely affected by an inability to have children. Equally, it cannot be assumed that all individuals of childbearing age necessarily incur the same disutility based on people seeking to address their infertility through assisted reproductive technology.

5.2.8 Resources and costs

The company's model adopted an NHS cost perspective. The costs included in the model comprised:

- Zynteglo acquisition, work-up and administration costs,
- Post-transplant monitoring costs,
- Blood transfusions,
- Iron chelation treatment and their associated administration costs,
- Monitoring costs for transfusion and chelation therapy,
- Management of chelation-related AEs,
- Management of iron overload-related complications.

Unit costs were sourced from a number of national sources, including NHS Reference Costs, the British National Formulary (BNF) and from the Personal Social Services Research Unit (PSSRU).[68-70] The key costs included in the model are summarised in Table 14, and a detailed description is provided in Section B.3.5 of the CS.

The required tests prior to Zynteglo treatment were based on the protocol for Study HGB-207. The company noted that a cost has been assigned to each test, despite not all being required. The costs associated with mobilisation and apheresis were sourced from a French budget impact analysis,[71] since the collection of relevant costs by bluebird bio were still ongoing at the time of the submission. A proportion of patients required multiple rounds of mobilisation, based on data from HGB-204, -205 and -207. Post-transplant monitoring was informed by the protocol for a bluebird bio registry, REG-501.

A micro-costing approach for transfusions was based on the costing statement for NICE guideline on blood transfusion.[72] Dosing of the iron chelation therapies was estimated from their respective

SmPCs. The cost of managing AEs associated with oral chelation was estimated at £16.33 per year, with £3.26 for subcutaneous chelation.

Table 14 Summary of the costs included in the economic model

Description of cost	Patients ≤18 years	Patients >18 years	Source
Zynteglo			
Zynteglo acquisition cost	[redacted]	[redacted]	bluebird bio, includes PAS
Pre-transplant cost	£27,057	£27,130	
Transplant-related costs	£34,539	£18,529	NHS Reference Costs[68]
Post-transplant monitoring costs – Years 1 & 2	£1,128	£1,128	NHS Reference Costs[68]
Post-transplant monitoring costs – Years 3 & 4	£927	£927	NHS Reference Costs[68]
Blood transfusions and chelation			
Annual treatment of blood transfusions and subcutaneous chelation therapy (transfusion-dependent)	£18,168	£23,287	NHS Reference Costs,[68] Agrawal *et al.,*[73] NICE guideline on blood transfusions,[72] PSSRU 2018,[70] BNF,[69] Bentley et al.,[74] Karnon *et al.*[75]
Annual treatment of blood transfusions and oral chelation therapy (transfusion-dependent)	£22,538	£24,772	
Annual treatment of blood transfusions and subcutaneous chelation therapy (transfusion-reduced)	£15,362	£16,290	
Annual treatment of blood transfusions and oral chelation therapy (transfusion-reduced)	£14,532	£15,460	
Health state costs			
Cardiac complications – Year 1	£6,871		Karnon *et al.*[39]
Cardiac complications – Year 2+	£3,534		Karnon *et al.*[39]
Liver complications	£1,754		NHS Reference Costs[68]
Endocrine complications (diabetes or hypogonadism)	£1,557		Karnon *et al.*[39]

5.2.8.1 ERG comment

The ERG has a number of concerns regarding the costs used in the economic model. These include the a number of minor costs associated with Zynteglo treatment that may have been omitted from the analysis, the use of BNF drug acquisition costs rather than eMIT costs, and the uncertainty in the cost of managing iron overload-related complications.

Underestimation of some of the costs of Zynteglo treatment

The ERG believes there may be costs associated with Zynteglo that have not been captured in the economic model. However, the ERG believes the impact of this on the company's base-case ICER to be limited, as they account for a small proportion of the total costs.

The model only estimates the cost of Zynteglo treatment for those patients who are successfully infused. There may be patients who initiate pre-transplant testing but are found not to be eligible for treatment with Zynteglo, or who experience failure of mobilisation, whose costs are not account for.

The revised CONSORT flow diagram presented in the company's response to clarification questions shows that one patient from the ITT population in HGB-204 discontinued Zynteglo due to inadequate stem cell mobilisation (Section 4.2.1). The costs that this patient incurred would include those associated with hypertransfusion, pre-treatment tests, and the costs of mobilisation and apheresis. In the company's base case, the pre-transplant costs in the Zynteglo arm are [redacted], and so a failure during the pre-transplant stage could result in irrecoverable costs for which there are no associated benefits. The CONSORT flow diagrams also show that, upon screening, patients in HGB-204 and 207 were found not to be eligible for Zynteglo and did not progress to gene therapy treatment (Section 4.2.1). The total cost of pre-treatment stage tests for such patients ≤18 and >18 are £2,629 and £2,539, respectively.

The CS includes thalassaemia genotyping in the listed pre-transplant tests, and states that the cost of genotyping is to be incurred by bluebird bio. As a result, the economic model assumes no cost of genomic testing. In the NHS, the recording of genotype appears to be already undertaken in some centres but practice is varied. (Section 3.1) In the company's Chart Review of medical records, [redacted] of patients had a recorded genotype, although no details were provided as to how genotype was determined. The clinical advisor to the ERG noted that it was not routinely undertaken at their centre.[76] The ERG also has concerns about the practicalities of the costs of genomic testing being incurred by the company, as testing is moving towards a new centralised system and it might be challenging for bluebird bio to incur the costs. As a result, the ERG notes that the company's model could result in underestimated associated costs to the NHS of Zynteglo treatment.

Additional resource use that may be associated with Zynteglo treatment and has been omitted from the model includes: the cost of repeating all childhood vaccinations 6 months post-transplant; and the long-term annual monitoring of transfusion independent patients for signs of leukaemia and lymphoma.[1, 59] The appropriate length of hypertransfusion prior to Zynteglo infusion is unclear, as the source study reported a duration of 3 months compared to one month applied in the company model, although the ERG noted that this study was based in France and may not be generalizable to UK practice.[71]

eMIT drug costs

The company sourced the unit costs of drugs in the model from the BNF.[69] The costs for a number of these, including the subcutaneous iron chelating agent desferrioxamine and two drugs used during the pre-transplant stage of Zynteglo treatment, can also be obtained from the electronic market information tool (eMIT).[77] This provides information on the average price paid by the NHS for pharmaceuticals, which can differ from the list prices listed in the BNF, and is seen as a more accurate and up to date indicator of costs. The difference between the BNF and the eMIT unit costs are presented in Table 15.

Despite a small difference in the BNF and eMIT drug acquisition costs of desferrioxamine, the ERG believes the overestimation of the costs could be important, given patients treated with chelation therapy remain on this treatment for life. The ERG considers eMIT to be a more appropriate source of drug acquisition costs, where available, and presents a scenario analysis in Section 6.3.10 using these values.

Table 15 Pre-transplant drug acquisition costs

Drug	Source in CS	BNF	eMIT
Busulfan	Assuming recommended daily dose 0.8mg/kg for 4 days (SmPC). Busulfan 60mg/10ml concentrate for solution for infusion vials (Accord Healthcare Ltd)	£1529.50	£386.14
Ursodeoxycholic acid	Assuming recommended dose of 10 mg/kg/day for 21 days (SmPC); unit cost of Ursofalk (Dr. Falk Pharma UK Ltd) 150mg tablets (60 tablets per package)	£19.02	£8.46
Desferrioxamine	Assuming 6 times per week, dose 40 mg/kg.	£46.63	£34.06

Cost of managing iron overload-related complications

There were some minor limitations associated with the estimation of the costs of managing iron overload-related complications. The company obtained the cost of cardiac complications from a cost-effectiveness study of chelation therapy and inflated to 2018 prices.[39] This study used a value reported by an older cost-effectiveness analysis,[21] where the costs of managing cardiac complications were estimated from US health insurance and pharmacy claims for patients with a diagnosis of cardiomyopathy, conduction disorders, cardiac dysrhythmias, or heart failure. It is unclear how the US health insurance costs were converted to UK-specific costs, and the ERG has concerns regarding how generalisable these costs are to current NHS practice. They are also inconsistent with how cardiac complications are modelled, where only the time to heart failure is captured.

As noted previously (Section 5.2.1), the endocrine complication was modelled as the first occurrence of either diabetes or hypogonadism and these conditions could not occur concurrently. In this respect, the total cost of endocrine complications will be underestimated in the model, as only one condition is accounted for. However, the total endocrine cost may be overestimated due to differences in prevalence rates between diabetes and hypogonadism. Given the range in the costs of managing endocrine complications (annual costs: diabetes (£4,497), hypogonadism (£487), hypoparathyroidism (£233) and hypothyroidism (£30)), the (unweighted) average endocrine cost will underestimate the cost of managing diabetes and overestimate the cost of hypogonadism in the model. Since hypogonadism is associated with a higher prevalence rate than diabetes in the model (41% and 67%, respectively), it is more likely to be the first (and therefore, only) occurring event.[30] As a result, if a

patient only develops hypogonadism, their total cost will be overestimated. Additionally, the use of hypo- and hypoparathyroidism costs in the average cost means that it is inconsistent with the rate of endocrine complications, where only the time to hypogonadism and diabetes are captured.

5.2.9 Cost effectiveness results

Zynteglo has a confidential patient access scheme (PAS), comprising a simple discounted price per patient of [redacted]. This is a discount of approximately [redacted] on the list price. The original company submission considered a complex PAS, but this had not been included in the cost-effectiveness results reported below as it is still under development and has not yet been approved at the time of the submission. The results of the original and uncorrected company base-case ICER cannot be reproduced in this report due to the inclusion of currently unconfirmed commercial agreements. No interventions that comprise the Zynteglo workup, administration and monitoring costs or those used in transfusions and iron chelation therapy have an associated PAS.

The cost effectiveness results outlined in this section are provided from a corrected and updated company analysis following the ERG's clarification questions and subsequent model corrections. The results presented below include the simple PAS discount for Zynteglo.

5.2.9.1 Base case results

Table 16 presents the base-case deterministic analysis of Zynteglo. These results are based on the mean costs and QALYs of 600 patient profiles generated by the model.

It shows that Zynteglo was associated with increased costs (cost difference of [redacted]) and was more effective (gain of 13.14 QALYs), compared with standard care (transfusions and iron chelation therapy). The company's base-case ICER was [redacted] per QALY.

In the Zynteglo arm, treatment costs of Zynteglo accounted for 88% of total costs, of which 84% were Zynteglo acquisition costs. Of the remaining costs, the item that accounted for the largest proportion of costs were the transfusion and iron chelation.

In the SoC arm, the costs of transfusion and iron chelation therapy accounted for 95% of the total costs. Oral chelation therapy plus transfusions accounted for the largest proportion of this, making up 65% of the total SoC costs. Of the remaining costs, iron overloading complications accounted for the largest proportion.

Table 16 Company model base-case results (Table 2, clarification response)

Technologies	Total costs	Total LYs	Total QALYs	Incremental costs	Incremental QALYs	ICER (£/QALY)
SoC	[redacted]	37.79	17.20			
Zynteglo	[redacted]	53.63	30.34	[redacted]	13.14	[redacted]

CS, company submission; ICER, incremental cost-effectiveness ratio; LYs, life-years (undiscounted); QALYs, quality-adjusted life-years (discounted); PAS, patient access scheme; SoC, standard of care (consisting of transfusions and iron chelation therapy)
These results were provided by the company on 20th December 2019

5.2.9.2 Probabilistic sensitivity analysis

The company performed a probabilistic sensitivity analysis (PSA) by running 1,000 iterations of the 600 profiles produced in the economic model. In each iteration, the model drew inputs from defined distributions for selected parameters. The mean probabilistic ICER of Zynteglo was [redacted] per QALY gained versus standard care. The cost-effectiveness plane showing the results of the PSA can be seen in Figure 4.

The company attributed the small differences between the results of the deterministic and the probabilistic analyses to the difference in the age distribution sampled in each analysis. The PSA allowed ages of up to 50 years, whereas only patients up to the age of 34 are included in the efficacy population.

The ERG has concerns regarding the 1,000 iterations used in the PSA and whether this accurately characterises decision uncertainty. Running 1,000 iterations required considerable computation time and due to time constraints and the nature of the company's model structure, the ERG was unable to re-run the model with an increased number of iterations.

Table 17 Company model probabilistic cost-effectiveness results (adapted from Table 3, clarification response)

Technologies	Total costs [95% CI]	Total LYs [95% CI]	Total QALYs [95% CI]	Incremental costs	Incremental QALYs	ICER (£/QALY)
SoC	[redacted]	39.03 [31.41 ; 46.63]	17.51 [14.10 ; 20.69]			
Zynteglo	[redacted]	53.82 [52.52 ; 55.01]	30.39 [29.63 ; 31.11]	[redacted]	12.89	[redacted]

CS, company submission; ICER, incremental cost-effectiveness ratio; LYs, life-years (undiscounted); QALYs, quality-adjusted life-years (discounted); PAS, patient access scheme; SoC, standard of care (consisting of transfusions and iron chelation therapy)
These results were provided by the company on 20th December 2019

Figure 4 Cost-effectiveness plane for Zynteglo (generated from the company model)

The probability that Zynteglo is the most cost-effective treatment option at WTP threshold of £30,000 is [redacted]. The cost-effectiveness acceptability curve for both comparators is provided in Figure 5.

Figure 5 Cost-effectiveness acceptability curve for Zynteglo and transfusions and chelation therapy (Fig 9, clarification response)

The results show that there is a difference of [redacted] per QALY between the deterministic and probabilistic ICERs. The average incremental QALYs gained with Zynteglo compared to transfusions and chelation therapy was [redacted], which was [redacted] QALYs less than in the deterministic analysis.

5.2.9.3 Deterministic sensitivity analysis

The company presented a series of deterministic sensitivity analyses (DSA) in the form of univariate sensitivity analyses to assess the impact of varying key model input parameters upon the ICER. The DSA inputs can be seen in the company's economic model.

A tornado diagram summarising the most influential parameters as reported by the company is presented in Figure 6. The results indicate that varying the rate of transplant success, the acquisition cost of the chelation therapy and the distribution of oral chelation therapies have the greatest impact on the ICER. The age distribution used in the model was also a driver of the model's results. In the base case, transplant success is 83.3%, however reducing the transplant success to 66% results in an increase of [redacted] in the ICER.

Due to an error in the implementation of the DSA in the company's model, the results provided in Figure 6 are based on an ERG corrected version of the model. For more detail see Section 5.2.10.

Figure 6 Univariate sensitivity analysis for Zynteglo vs standard care (generated from the company model, containing ERG corrections)

5.2.9.4 Scenario analyses

Following amendments made to the original base-case at the request of the ERG, the company provided resubmissions of original scenario analyses in their response to clarification questions. The scenarios assessed the robustness of the model results and the impact of the assumptions included in the base-case analysis. The results of the scenario analyses performed on the company's updated base case are presented in Table 18.

Table 18 Company scenario analysis results

Parameter	Value	Incremental Costs	Incremental QALYs	ICER (£/QALY)
Baseline age and gender	Demographic characteristics taken from HES data	[redacted]	12.00	[redacted]
Mortality	No impairment of survival due to myeloablative conditioning (SMR = 1)	[redacted]	13.85	[redacted]
Utility	Values from vignette study only	[redacted]	12.10	[redacted]
These results were provided by the company on 20th December 2019				

The results of the company's scenario analyses show the result were sensitive to a change in the baseline demographics.

5.2.9.5 Additional scenarios requested by the ERG at points for clarification

Additional scenarios were requested by the ERG and were provided by the company at the clarification questions stage. The scenarios requested were:

i) Using age- and gender-appropriate distributions of baseline iron load and patient weight i.e. for each age and gender profile, Chart Review data to be analysed to provide the mean patient weight and baseline iron levels,

ii) Recalculate the baseline iron distribution data from the subset of patients aged 12-35 without comorbidities that would preclude treatment with Zynteglo,

iii) Recalculate baseline iron distribution data from the subset of all patients aged ≤ 18 without comorbidities that would preclude treatment,

iv) Proportion of patients receiving each type of chelation therapy and mean number of transfusions for the transfusion-dependent population (both of which are sourced from the Chart Review) to be re-analysed limited to patients aged 12-35 years,

v) Proportion of patients receiving each type of chelation therapy and mean number of transfusions for the transfusion-dependent population to be re-analysed limited to patients aged ≤18 years,

vi) Reanalyse the HRQoL data from the Chart Review to limit the included patients strictly to those who would be eligible for Zynteglo, i.e. exclude patients with 'high' T2* iron (≥15 mg/g),

vii) Limit the HRQoL data to those included in scenario (vi) and in addition only include those patients aged 12-35 years,

viii) Limit the HRQoL data to those included in scenario (vi) and in addition only include all patients aged ≤ 18,

ix) Limit the HRQoL data to those included in scenario (vi) and average the HRQoL weighted according to the age distribution in the three Zynteglo trials.

Table 19 ERG requested scenario analyses

Parameter		Value	Incremental Costs	Incremental QALYs	ICER (£/QALY)
(i)	Baseline iron and weight	Reanalysed Chart Review data	[redacted]	13.14	[redacted]
(ii)	Baseline iron	Limited to age 12-35 years	[redacted]	13.16	[redacted]
(iii)	Baseline iron	Limited to age ≤ 18 years	[redacted]	13.18	[redacted]
(iv)	Chelation therapy	Limited to age 12-35 years	[redacted]	13.14	[redacted]
(v)	Chelation therapy	Limited to age ≤ 18 years	[redacted]	13.14	[redacted]
(vi)	HRQoL	Limited to eligible patients i.e. removal of patients with T2* iron ≥15 mg/g	[redacted]	11.16	[redacted]
(vii)	HRQoL	Scenario (vi) and limited to age 12-35 years	[redacted]	8.93	[redacted]
(viii)	HRQoL	Scenario (vi) and limited to age ≤ 18 years	[redacted]	9.42	[redacted]
(ix)	HRQoL	Scenario (vi) and reweight Chart Review according to trial age distribution	[redacted]	10.41	[redacted]

For further details of the ERG's description of the limitations of the baseline characteristics, chelation therapy and the HRQoL see Sections 5.2.3, 5.2.4 and 5.2.7, respectively.

The scenario analysis with the largest impact on the results is the reanalysis of the HRQoL data to match patients included in the Zynteglo trials, i.e. through exclusion of patients with high T2* and limiting the age to 12-35 years. This reanalysis resulted in an increase in the company's base case ICER to [redacted] per QALY.

5.2.10 Model validation and face validity check

5.2.10.1 Use of the DICE framework

Transparency is one of the key benefits that has been ascribed to the DICE method. The DICE model can be implemented in a spreadsheet, containing a set of tables that specify the conditions with their values and the events with their consequences. However, the transparency of the model relies upon accurate and thorough reporting of how assumptions are implemented and outcomes are modelled. If the modeller uses inconsistent labelling and does not adhere to the prescribed table and reporting structure, the model can very quickly lose transparency. The ERG found that the model originally submitted by the company did not contain sufficient annotation, which made it challenging to understand and validate. It was also not accompanied by the model technical report, user guide and model blueprint, all of which are of key importance for understanding a DICE model. Given that this is an emerging and evolving method, the ERG would like to strongly emphasise the importance of providing the supporting documentation to review groups, and reporting standards should be formalised for NICE submissions.

Since the DICE model is a relatively recent development compared to other model structures that have traditionally been applied within the context of HTA, there are few applications of the method published to date. The ability of the DICE model to predict and replicate the results of a non-DICE model in a NICE context have been published in two analyses (although there may be others that the ERG is not aware of).[78, 79] In both examples, the double-programmed model performed very similarly to the DICE model; however, these were based on model structures that are a lot simpler (containing fewer events) than that implemented in the current appraisal of Zynteglo. The company also stated that an independent health economics agency produced a simplified cohort (Markov) version of the DICE model; however, specific details were not provided.

The company stated that the panel of independent health economists that were consulted throughout the model development process were initially sceptical regarding the use of a DICE model, given previous challenges in NICE appraisals, but were reassured by the company stating that NICE and ERGs have been provided extensive training. However, to our knowledge it is mostly the NICE guideline group and not the technology appraisal groups that have experience with using DICE models prior to this appraisal and no specific ERG training has been implemented.

The ERG also noted the significant run time of the model. To run the PSA took between 18 and 23 hours, depending on the machine used.

5.2.10.2 Model errors

The ERG conducted a validation of the functions and input parameters contained in the executable model. A validation of the underlying code that implements the DICE framework was not undertaken, since this has been reported on by prior groups.

The ERG identified an error in how age-related utilities were estimated in the originally submitted model. This was corrected in an updated version of the model provided by the company at the clarification stage. In the original model, the age characteristic used to calculate the age-related HRQoL of a patient was not updated at regular intervals beyond the first year of the modelled time-horizon. Therefore, when a patient's utility was drawn to calculate accrued QALYs, the model used the utility associated with the last time this age characteristic had been updated, which was often several decades prior, thus the health state utilities were not appropriately adjusted for age.

The impact of this correction was to reduce the number of QALYs gained across both the Zynteglo and SoC arms. As Zynteglo patients lived to a greater age in the model, and thereby incurred larger age-related utility decrements, the magnitude of the effect of the correction was greatest in these patients. The ICER of Zynteglo versus SoC, therefore, increased.

The updated model contained a number of inconsistencies regarding the input values when compared to the original model; these were also amended by the company in a subsequent version of the model.

In the final version of the model submitted by the company, the ERG noted that the estimation of the 95% confidence interval around the proportion of patients achieving transfusion independence differed to that of the original model. The 95% CI in the updated model was narrower, which had the impact of making the value appear to be less influential in the DSA. The ERG considered that the original model contained the correct calculations, and updated the analysis accordingly.

5.2.10.3 Stability of results

The model, a discrete event simulation (DES), estimated the mean cost-effectiveness results from the total costs and QALYs generated by a number of simulated patients. Stochastic models, such as DES, typically require a large number of iterations to generate stable results, as some variation will appear as a result of 'random noise', or first-order uncertainty.[80] The ERG noted that the company model used a total number of 600 patient profiles to estimate their base case deterministic results. To assess

the stability of results using different numbers of iterations, the ERG re-ran the deterministic model with up to 20,000 profiles.

Table 20 Results of the company base case analysis based on different numbers of patients profiles

Number of profiles in the model	ICER	Difference from company model
100	[redacted]	[redacted]
500	[redacted]	[redacted]
600 (company model)	[redacted]	-
1,000	[redacted]	[redacted]
5,000	[redacted]	[redacted]
10,000	[redacted]	[redacted]
20,000	[redacted]	[redacted]
50,000	[redacted]	[redacted]

As illustrated in Table 20, there is a large amount of variation in the ICER, particularly when small numbers of patient profiles were generated. Results appeared to stabilise when higher number of patient profiles were generated. To determine the driver of the differences of the results, the ERG examined the outcomes in the models of 600 and 5000 patient profiles (Table 21). With a greater number of patient profiles, Zynteglo total costs increased while SoC total costs decreased, and the number of QALYs in each arm decreased. From inspection of the breakdown of results in both models, it appears that the difference in costs was mostly due to iron chelation. The difference in life years between the two versions of the models was small, but potentially contributed towards the difference in lifetime chelation costs.

The ERG remained unclear on why the model estimated a higher ICER with a greater number of profiles, and which was the most appropriate method to generate results.

Table 21 Comparison of results of the model run with 600 and 5,000 patient profiles

Technologies	Total costs	Total QALYs	Incremental costs	Incremental QALYs	ICER (£/QALY)
Company base case (with 600 patient profiles)					
SoC	[redacted]	17.20			
Zynteglo	[redacted]	30.34	[redacted]	13.14	[redacted]
Model with 5,000 patient profiles					
SoC	[redacted]	17.01			
Zynteglo	[redacted]	29.90	[redacted]	12.90	[redacted]

5.2.10.4 Consistency of results with published economic models

To validate the prediction of life years in the standard care arm of the model, the company provided a comparison to those estimated by Weidlich,[41] a recent UK-based analysis estimating the total costs, QALYs and life years for patients on blood transfusions and iron chelation therapy with TDT.[41] Survival was predicted in this study using a meta-analysis, with a focus on patients born after 1970 and European-based studies, as thalassaemia management was considered to be most similar to the UK. Their model predicted a similar number of QALYs to the company model (11.5 in the Weidlich analysis compared with 12.55 in the company model using an equal discount rate). Higher costs predicted by the company model were due to a greater proportion of patients on the more expensive chelation agent deferasirox. However, this study was subject to the same limitations that are present within this appraisal. The survival data used in this analysis were based on patients born between 1970 and the early 2000s, and not those born more recently who will probably achieve a better life expectancy due to advancements in clinical management. The clinicians interviewed by the study authors expected disease patterns to change in the future and patients' life expectancy to improve.

5.3 Conclusions of the cost effectiveness section

The ERG considered the company's economic submission to meet the requirements of the NICE reference case with a few notable exceptions, which may have a significant impact upon the cost-effectiveness results. Additionally, the ERG identified a number of uncertainties regarding key model inputs. The main concerns identified by the ERG include:

1. The complexity of the modelling approach

The ERG considers that the modelling approach employed introduced unnecessary complexity and reduced transparency, and the key benefits of a DICE modelling framework were not fully exploited. Despite this model complexity, a number of simplifying assumptions were made that undermined its face validity. A limited number of baseline characteristics were used to define individual patients, and their interaction with other outcomes was limited. Also, iron overload-related impacts to the three organ systems included in the analysis were modelled independently with no interaction between them, and the model did not account for patient history to determine the rate of future events. The ERG accepts that in many cases, this was due to limited data to inform these structural assumptions, but as a result, the model complexity created significant challenges for the ERG in terms of identifying and following key assumptions without much benefit.

2. The use of iron chelation agents in current practice is subject to uncertainty

Clinical practice regarding the use of iron chelation agents is evolving, due to improvements in evidence and confidence around the use of newer chelation agents. This is particularly true regarding to the use of chelation therapies used in combination, an option which may be becoming more common. The company's review and analysis of medical records of TDT patients in the UK showed that combination therapy was more common in those aged 12-35 than in the older age groups.

3. Differences between the modelled and the eligible population

The mean body weight used in the model was based on an average for the population. Considering that the model fails to account for an interaction between weight and age, the value was thought to be too high for younger patients. This has implications for the cost of chelating agents, as dose is calculated by patient weight.

Presence of cardiac and liver complications preclude patients from eligibility with Zynteglo. However, the model also excludes patients with endocrine complications such as hypogonadism when they would, in fact, be eligible for Zynteglo.

Patients with certain genotypes may be underrepresented in the Zynteglo trials. This limits the generalisability of the modelled population to the eligible population in the UK, and has a potential impact on the estimation of the treatment effect of Zynteglo since the underrepresented mutations are generally associated with poorer outcomes.

4. Long-term benefits of Zynteglo are uncertain

While late graft rejections are generally considered to be rare events in HSCT, the long-term benefits of Zynteglo remain uncertain. There is currently insufficient available follow-up data in the trials to determine whether there is a permanent treatment effect, although a loss of transfusion independence has not yet been recorded in any patients.

[redacted], so there is a concern regarding the consequences of a low transduction percentage upon the long-term persistence of a sufficient proportion of cells capable of producing healthy haemoglobin.

5. The iron normalisation period is potentially too short

There is a lack of representative data on how long it takes for iron levels to normalise in patients who achieve transfusion independence. The ERG considers the Zynteglo trials and the studies identified by

the company to provide insufficient support for their assumption of iron normalisation in all patients achieving transfusion independence, and that the assumed four year time to normalisation may be too optimistic. No evidence source appears to identify a time point by which all patients achieve normalised levels of iron. The studies themselves are associated with a number of limitations, and are of limited generalisability to Zynteglo since they included much younger patients who received allogeneic HSCT.

6. Mortality of transfusion-dependent patients is likely to be over-estimated

The mortality of transfusion dependent patients was likely to be over-estimated. Survival rates were estimated from an outdated study of patients performed when standard practice comprised subcutaneous chelation agents. With greater clinician experience alongside improved monitoring and iron chelation practices observed over the last decade, iron levels in TDT are more likely to be well controlled, and survival in well-chelated patients is thought to be approaching a normal pattern. Further still, patients in the current decision problem would have lower iron levels and fewer iron overload-related complications than the unrestricted TDT population followed in the study, with eligibility for Zynteglo requiring patients to be sufficiently fit to undergo the procedure.

7. The use of a discount rate of 1.5% is unjustified

The company applied a discount rate of 1.5% to costs and benefits, which is inconsistent with the NICE reference case. This significantly impacts the results of the analysis, since the costs of Zynteglo are incurred upfront while the potential benefits are generated over the patient's lifetime. The ERG believes there to be little justification for the use of the non-reference case discount rate of 1.5%, and disputes all arguments put forward by the company in support of this assumption.

The company asserted that Zynteglo restores people who would otherwise die or have a very severely impaired life to full or near full health. While HRQoL and survival of patients who are successfully treated with Zynteglo could reasonably be expected to approach that of the general population, the ERG consider that the impact of TDT on the survival and HRQoL of optimally managed patients has been overstated, particularly using modern monitoring and treatment strategies. The ERG considers that the evidence collected to date is insufficient to determine whether the benefits of Zynteglo persist into the long term. Finally, the company stated that Zynteglo will not commit the NHS to significant irrecoverable costs. However, this was on the basis of a proposed outcomes-based commercial agreement, which remained under negotiation at the time of the submission. Under most pricing scenarios, the potential cost to the NHS associated with anything but an indefinite treatment effect in all patients could be very large.

8. HRQoL of patients on standard care was underestimated

The HRQoL for patients who are transfusion dependent was estimated from a review and analysis of medical records of UK TDT patients. However, this Chart Review population is older than the modelled patients. Due to the long-term effects of transfusions and chelation therapy, age related decline in HRQoL, and historic differences in disease management, it might be expected that the HRQoL of older patients to be lower than that of optimally managed younger patients.

The population in the Chart Review also included patients with existing complications who incurred separate decrements in the model (i.e. the impact of these was double counted). The ERG was concerned that due to both of these issues the utility value used by the company was underestimated, as it is much lower than reported in other studies.[46, 47]

A re-analysis of the medical records, limiting the population to those aged from 12-35 years and without iron overload-related complications that were already accounted for in the model, produced a much higher utility estimate. This may support the idea that current management is acceptable to most patients, rather than resulting in a 'very severely impaired life'.

9. Inappropriate source of age-based HRQoL values was used

An inappropriate value set was selected as the basis for the general population utility estimates. The company used a subset of the general population, which excludes all individuals with a history of a health condition, and with any ongoing health conditions. This is inappropriate because it means that patients in the model will essentially remain in perfect health until death, excepting any treatment- or complication-related decrements associated with their TDT explicitly accounted for in the model. As a result, age-based utilities were much higher than genuine general population values, leading to an overestimation of QALYs.

The utility decrement associated with transfusion dependence, estimated from the difference between the mean value observed in the Chart Review and the "perfect health" population utility value, was subsequently estimated to be greater than it should have been. This undermines the company's argument that patients on standard care experience a 'very severely impaired life', comparable to patients with advanced cancer, which was used to justify the lower discount rate.

10. Disutility for patients who become infertile following Zynteglo is uncertain

The impact of infertility upon HRQoL is poorly understood and not typically well captured using EQ-5D, and the value was estimated from a source that was not generalisable to the present context. It

may not be appropriate to capture this impact within this appraisal, as it has previously only been used in models of assisted reproductive technologies for patients actively seeking fertility treatment. It was also uncertain whether the duration for which the decrement was applied was appropriate.

6 Impact on the ICER of additional clinical and economic analyses undertaken by the ERG

6.1 Overview

The following sections provide details of the additional analyses undertaken by the ERG to explore the key issues and uncertainties raised in the review and critique of the cost-effectiveness evidence submitted by the company. Section 6.2 describes the impact of errors identified in the ERG's validation of the company's executable model.

Section 6.3 presents the results of a series of exploratory analyses based upon uncertainties identified by the ERG. These scenario analyses examine the impact of a number of alternative assumptions upon the robustness of the cost-effectiveness results, focusing on the following:

- Baseline population characteristics and presence of co-morbidities adjusted to better reflect the eligible population,
- Distribution of chelation agents adjusted to better reflect the eligible population, using a re-analysis of the Chart Review data,
- Clinical effectiveness data adjusted for the underrepresentation of the IVS-I-5 and IVS-I-110 genotypes,
- Alternative assumptions around the rate of early engraftment failure and the long-term persistence of transfusion independence,
- Alternative assumptions on the mortality of transfusion dependent patients,
- Alternative assumptions around iron normalisation and the development of iron overload-related complications,
- Impact of using a 3.5% discount rate on costs and effects,
- Use of a number of different sources and assumptions around health-related quality of life,
- Drug acquisition costs from eMIT.

In Section 6.4, the ERG presents an alternative base-case which combines a number of the exploratory analyses presented in Section 6.3, which the ERG considers to better reflect the cost-effectiveness of Zynteglo were this technology to be approved for use on the NHS. A number of scenario analyses on the ERG alternative base-case analysis are also presented.

The results in this section are presented inclusive of the confidential PAS discount for Zynteglo. At the time of the submission, this comprised [redacted]. For brevity, the comparator of transfusions and iron chelation

is referred to as standard of care, or SoC, throughout. Life years presented are undiscounted. Deterministic results are based on 600 patient profiles, to be consistent with the company base-case analysis: results of the ERG alternative base-case analysis, based on 5,000 patient profiles, are also presented. Due to time constraints and the nature of MS Excel-based DICE models, it was not possible to produce probabilistic results for each of the following scenarios, therefore all results in the following section are deterministic unless otherwise stated.

6.2 ERG corrections and adjustments to the company's base case model

The correct and latest iteration of the company base-case results are presented in Table 22, and all additional ERG analyses that follow use this model as a basis. An error in the company's executable model was identified by the ERG, which was highlighted in the ERG's clarification letter to the company, who later provided a corrected version of the model. Further details of the error are provided in Section 5.2.10.

Table 22 Deterministic results of the corrected company base-case analysis (Table 2, clarification response)

Intervention	Total			Incremental		
	Costs	LYs	QALYs	Costs	QALYs	ICER
Company base-case analysis						
SoC	[redacted]	37.79	17.20			
Zynteglo	[redacted]	53.63	30.34	[redacted]	13.13	[redacted]

* undiscounted life years

6.3 Additional ERG analyses

6.3.1 Changes to baseline population characteristics

As discussed in Section 5.2.3, the ERG did not consider the baseline characteristics of the modelled population to adequately represent those of the population eligible to receive Zynteglo in practice. Firstly, the mean patient weight ([redacted]), based on the average weight of patients in the Chart Review, is likely to be higher than that of much of the eligible population at baseline, many of whom are aged <18.

The ERG requested that the company add into the model the functionality to link patient weight to their age at baseline. In this scenario, patients aged 12 to 17 had a mean body weight of [redacted], and the weight of those aged over 18 was [redacted]. When paediatric patients reach the age of 18, any weight-based drug costs are thereafter based on the adult weight. The results of this scenario are presented in Table 23, which show a small increase in the ICER of Zynteglo versus SoC (transfusions and iron chelation). This is attributable to the reduction in accrued costs of the chelation agents.

Table 23 ERG Scenario 1: Age category-specific body weight (paediatric and adult)

Intervention	Total			Incremental			Change from company base case ICER
	Costs	LYs*	QALYs	Costs	QALYs	ICER	
SoC	[redacted]	37.79	17.20				
Zynteglo	[redacted]	53.63	30.34	[redacted]	13.13	[redacted]	[redacted]

* undiscounted life years

Section 5.2.3 also details the ERG's concern that the company's model did not account for presence of endocrine disorders at baseline. The company stated in response to a clarification question from the ERG that individuals with TDT would not be precluded from receiving Zynteglo due to hypogonadotrophic hypogonadism or diabetes. The ERG therefore presents a scenario in which 20% of patients have hypogonadism at baseline (per the Chart Review population), and thus incur additional costs of treatment and the effects upon HRQoL (Table 24). The impact on the ICER is relatively small, since this assumption applies to patients in both treatment arms.

Table 24 ERG Scenario 2: Hypogonadism present in 20% of patients at baseline

Intervention	Total			Incremental			Change from company base case ICER
	Costs	LYs*	QALYs	Costs	QALYs	ICER	
SoC	[redacted]	37.79	17.10				
Zynteglo	[redacted]	53.63	30.08	[redacted]	12.98	[redacted]	[redacted]

* undiscounted life years

6.3.2 Clinical effectiveness data adjusted for the underrepresentation of the IVS-I-5 and IVS-I-110 genotypes

As discussed in Section 5.2.3, the ERG considers the modelled population to underrepresent the proportion of severe non-β^0/β^0 mutations covered by the marketing authorisation, such as IVS-I-5 and IVS-I-110. The base case modelled population includes [redacted] patients with severe non-β^0/β^0 mutations; however, evidence suggest this could be as high as 28% of UK patients (see Section 3.1). Limited data collected to date suggests that patients with these mutations may be less likely to achieve transfusion independence, due to their reduced ability to produce haemoglobin.

To reflect this, the ERG presents an exploratory scenario in which the proportion of patients with severe non-β^0/β^0 mutations is increased from [redacted] to 28%, which decreases the modelled probability of transplant success from 83.3% to [redacted]

The results of this scenario are presented in Table 25 which shows an increase in the ICER of Zynteglo versus SoC.

Table 25 ERG Scenario 3: Adjusting clinical effectiveness data for underrepresented genotypes

Intervention	Total			Incremental			Change from company base case ICER
	Costs	LYs*	QALYs	Costs	QALYs	ICER	
SoC	[redacted]	37.79	17.20				
Zynteglo	[redacted]	52.71	29.65	[redacted]	12.45	[redacted]	[redacted]

* undiscounted life years

6.3.3 Adjustment to distribution of chelating agents

In Section 5.2.4, the ERG discussed the appropriateness of the distribution of different chelating agents used in the model. This was originally based upon the whole Chart Review population, but the ERG received advice that chelation practices have evolved over the years, and that the regimens used in patients treated currently may not reflect historic practices. The ERG requested that the company provide the distribution of iron chelation agents in patients aged 12-35, to better reflect the current standard of care in patients who would be eligible to receive Zynteglo (Table 9). The results of this scenario are presented in Table 26 below, which show a reduction of [redacted] (-2.3%) per QALY gained in the ICER of Zynteglo versus SoC. The re-analysis of the Chart Review data showed that there was a greater proportion of patients in this age category who received combination therapy, which increased the accrued cost of chelation agents.

Table 26 ERG Scenario 4: Distribution of chelating agents based on 12-35s in Chart Review

Intervention	Total			Incremental			Change from company base case ICER
	Costs	LYs*	QALYs	Costs	QALYs	ICER	
SoC	[redacted]	37.79	17.20				
Zynteglo	[redacted]	53.63	30.34	[redacted]	13.13	[redacted]	[redacted]

* undiscounted life years

6.3.4 Early engraftment failure

The ERG also identified early engraftment failure as a potential issue in Sections 4.2.1 and 5.2.6.2. While engraftment failure has not yet occurred to date, it is possible that engraftment may not be achieved in a small number of patients. As only 42 patients have received Zynteglo as per the product license, (Section 4.2) the rate of engraftment failure may not have yet been captured in existing trial data if it occurred in sufficiently small patient numbers (e.g. less than 1 in 42 patients, or lower than 2.3%).

Table 27 illustrates the effect upon the ICER of 1% and 5% of patients failing to successfully achieve engraftment, and thus requiring rescue therapy, i.e. the reserve of the patient's non-transduced cells are re-infused, with a 54% chance of mortality, as per the company's assumptions around this

scenario. When engraftment failure occurs in only 1% of patients, the impact to the ICER remains relatively small.

Table 27 ERG Scenarios 5 and 6: Engraftment failure following Zynteglo treatment

Intervention	Total			Incremental			Change from company base case ICER
	Costs	LYs	QALYs	Costs	QALYs	ICER	
ERG Scenario 5: Engraftment failure occurs in 1% of Zynteglo patients							
SoC	[redacted]	37.79	17.20				
Zynteglo	[redacted]	53.45	30.23	[redacted]	13.03	[redacted]	[redacted]
ERG Scenario 6: Engraftment failure occurs in 5% of Zynteglo patients							
SoC	[redacted]	37.79	17.20				
Zynteglo	[redacted]	51.39	29.01	[redacted]	11.81	[redacted]	[redacted]

* undiscounted life years

6.3.5 The persistence of treatment effect

Section 5.2.6.2 highlights a number of reasons why the ERG believe that the therapeutic effect of Zynteglo may not be life-long in all patients. To explore the impact of late graft failure in a small proportion of patients upon the ICER, the ERG present two exploratory scenarios to highlight the relative importance of this assumption. Note that these scenarios are not based upon clinical data, and these are intended to be illustrative and exploratory only. They assume that all patients who relapse undergo graft failure sufficient to return them to a transfusion dependent state, regardless of their starting point. Figure 7 represents the long-term persistence of the treatment effect in ERG Scenarios 7 and 8 compared to the base case, assuming that the time at which graft rejections can occur is capped at 50 years.

In the first of these, it is assumed that every 10 years, 5% of transfusion independent and transfusion-reduced patients 'relapse' and once again become dependent upon transfusions and iron chelation. This increases the ICER by [redacted] (+26.9%) per QALY gained on Zynteglo. The second scenario assumes that 10% of patients relapse every 10 years. This results in an increase of [redacted] (+62.1%) to the ICER, illustrating the potentially significant effect uncertainty around the long-term persistence of treatment effect could have on the cost-effectiveness of Zynteglo.

Figure 7 Persistence of Zynteglo treatment effect over time in ERG Scenarios

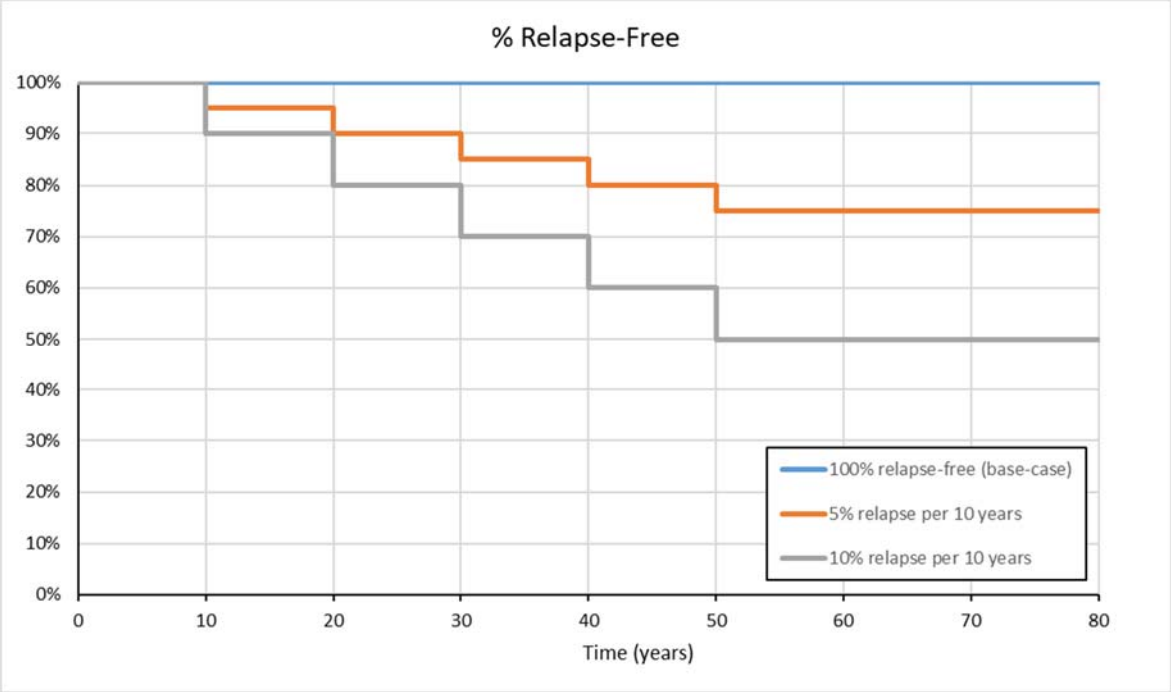


Table 28 ERG Scenarios 7 and 8: Alternative assumptions around long-term loss of treatment effect

Intervention	Total			Incremental			Change from company base case ICER
	Costs	LYs*	QALYs	Costs	QALYs	ICER	
ERG Scenario 7: 5% of patients relapse every 10 years							
SoC	[redacted]	37.79	17.20				
Zynteglo	[redacted]	50.69	28.46	[redacted]	11.26	[redacted]	[redacted]
ERG Scenario 8: 10% of patients relapse every 10 years							
SoC	[redacted]	37.79	17.20				
Zynteglo	[redacted]	48.10	26.71	[redacted]	9.51	[redacted]	[redacted]

* undiscounted life years

6.3.6 Mortality of transfusion dependent patients

The ERG questioned the appropriateness of the SMR the company applied to transfusion dependent patients in Section 5.2.6.5. The ERG considers the source that it was derived from to be obsolete in terms of its relevance to current NHS practice, due to improvements in iron chelation and patient monitoring which have led to more favourable mortality rates in TDT patients. The patients in the study of TDT survival are likely to have less favourable baseline characteristics: in order to be eligible for Zynteglo treatment, the patient must be sufficiently fit to undergo the procedure. Additionally, the ERG was particularly concerned that the company had selected the most pessimistic estimate from the literature for these patients.

Scenario 6 explores the impact of using an alternative assumption for the mortality rate associated with the transfusion dependent health state. Note that mortality due to cardiac complications is still separately accounted for in these patients: this specifically models the survival of TDT patients who have not developed cardiac complications.

In ERG Scenario 9 (Table 29), a lower SMR value of 2.0 was assumed for transfusion dependent patients. This results in a small decrease in the ICER of Zynteglo versus SoC. This is driven by the increased cost of SoC due to more of these patients remaining alive and requiring treatment, which outweighs the increase in QALYs due to a reduction in mortality.

Table 29 ERG Scenario 9: SMR of 2.0 applied to transfusion-dependent patients

Intervention	Total			Incremental			Change from company base case ICER
	Costs	LYs*	QALYs	Costs	QALYs	ICER	
SoC	[redacted]	42.41	18.50				
Zynteglo	[redacted]	54.09	30.49	[redacted]	11.99	[redacted]	[redacted]

* undiscounted life years

6.3.7 Iron normalisation and iron-related complications

The ERG identified a number of issues with the company's assumptions regarding the time to iron normalisation, and the risk of complications faced by patients who achieve normalised levels of iron after achieving transfusion independence. The following subsections address each of the issues raised in Section 5.2.7.

6.3.7.1 Iron normalisation period

As discussed in Section 5.2.6.3, the ERG did not concur with the company's conclusion that identified evidence supported an assumption of normalised iron levels in all patients after four years. Indeed, the limited data available from the Zynteglo trials was not sufficient to support this assumption; the iron levels of some TI patients remained elevated after 48 months of follow-up, and many trial patients remained on iron removal therapy (chelation or therapeutic phlebotomy) at the latest follow-up. It, therefore, appears that the assumption that all patients have normal iron levels four years after treatment may be not be accurate.

The ERG therefore presents scenarios in which the effect of an iron normalisation period of 5, 7, and 10 years is explored. During this time, patients incur the additional costs of iron chelation therapy, and are at a higher risk of certain complications associated with their elevated iron levels.

The results of ERG Scenarios 10, 11 and 12 are shown in Table 30, illustrating the potentially significant increase in the ICER of Zynteglo if iron levels are not normalised at 4 years, as assumed in the company's analysis. If iron levels were to take 7 years to normalise following treatment with Zynteglo, the resulting ICER would be [redacted] (17.4%) higher per QALY gained.

Table 30 ERG Scenarios 10-12: Alternative time to iron normalisation in transfusion independent patients

Intervention	Total			Incremental			Change from company base case ICER
	Costs	LYs*	QALYs	Costs	QALYs	ICER	
ERG Scenario 10: 5 year iron normalisation period							
SoC	[redacted]	37.79	17.20				
Zynteglo	[redacted]	53.40	30.05	[redacted]	12.85	[redacted]	[redacted]
ERG Scenario 11: 7 year iron normalisation period							
SoC	[redacted]	37.79	17.20				
Zynteglo	[redacted]	52.28	29.21	[redacted]	12.01	[redacted]	[redacted]
ERG Scenario 12: 10 year iron normalisation period							
SoC	[redacted]	37.79	17.20				
Zynteglo	[redacted]	51.26	28.49	[redacted]	11.28	[redacted]	[redacted]

* undiscounted life years

6.3.7.2 Complications due to iron overload in transfusion independent patient

The long-term consequences of iron damage in patients who achieve transfusion independence is an uncertainty in this analysis, since current evidence is limited and often contradictory. The company assumed that patients who have normalised iron levels are no longer at risk of developing complications. However, clinical advice to the ERG suggested that there may be a degree of pre-existing irreversible damage in many patients prior to treatment, albeit sufficiently small to allow for eligibility, which could theoretically result in a long-term risk of developing complications. In Scenario 13, the ERG applied the rates of developing cardiac complications associated with low iron overload, to patients with normalised iron levels (as described in Section 5.2.6.4).

The results of this analysis are presented in Table 31. The ERG highlight that this scenario is illustrative and represents the most conservative scenario regarding this assumption. However, given the large increase to the ICER, this scenario demonstrates that this source of uncertainty is important and shows that the development of long-term complications has consequences to the cost-effectiveness of Zynteglo.

Table 31 ERG Scenario 13: Patients with normalised levels of iron face a residual risk of developing iron overload-related complications

Intervention	Total			Incremental			Change from company base case ICER
	Costs	LYs*	QALYs	Costs	QALYs	ICER	
SoC	[redacted]	37.79	17.20				
Zynteglo	[redacted]	45.06	26.33	[redacted]	9.13	[redacted]	[redacted]

* undiscounted life years

6.3.8 Application of a 3.5% discount rate on costs and effects

The ERG's position is that Zynteglo does not meet the criteria for a 1.5% discount rate for costs and outcomes, as was used in the economic evaluation provided by the company. Section 5.2.5 provides a detailed description of the reasons for this. The results of the company's corrected base-case analysis are presented using a 3.5% discount rate for costs and outcomes in Table 32. This has the effect of more than doubling the ICER of Zynteglo versus SoC to [redacted] (+130.5%) per QALY gained.

The reasons for this are twofold; firstly, the majority of the hypothetical QALY gains associated with Zynteglo are received long into the future. Secondly, SoC is associated with much higher long-term costs, which are now subject to a higher discount rate, while almost all of the costs associated with Zynteglo are incurred upfront and thus have no discounting. This has the effect of substantially reducing the total cost of SoC but not of Zynteglo. This is appropriate, as there may be potentially substantial irrecoverable costs to the NHS, given the uncertainties around the long-term therapeutic effect of Zynteglo, which is better understood for transfusions and iron chelation therapy.

Table 32 ERG Scenario 14: 3.5% discount rate on costs and effects

Intervention	Total			Incremental			Change from company base case ICER
	Costs	LYs*	QALYs	Costs	QALYs	ICER	
SoC	[redacted]	37.79	12.55				
Zynteglo	[redacted]	53.63	19.84	[redacted]	7.29	[redacted]	[redacted]

* undiscounted life years

6.3.9 Health related quality of life scenarios

The ERG identified a number of issues with the company's preferred utility values and the way in which they had been implemented in the economic model. The following subsections address each of the issues raised in Section 5.2.7.

6.3.9.1 Age decrements based on Ara and Brazier general population values

As discussed in Section 5.2.7, the company's base case analysis used an inappropriate value set as the basis of general population utility by age. The subgroup from Ara and Brazier[62] selected by the

company included only those individuals in perfect health with no history of health problems in each age group. The company's base-case therefore assumes that patients would remain in otherwise perfect health until death, and never develop any health condition not explicitly included in the model.

The use of the whole-population dataset from Ara and Brazier[62] has extensive precedent in cost-utility analyses, and was therefore used in Scenario 15 in Table 33 below. The resulting ICER for Zynteglo is increased by [redacted] (+16.1%) per QALY gained versus SoC.

Table 33 ERG Scenario 15: Age-related disutilities taken from full Ara and Brazier population

Intervention	Total			Incremental			Change from company base case ICER
	Costs	LYs*	QALYs	Costs	QALYs	ICER	
SoC	[redacted]	37.79	16.43				
Zynteglo	[redacted]	53.63	27.74	[redacted]	11.32	[redacted]	[redacted]

* undiscounted life years

6.3.9.2 Utility for TDT based on adjusted Chart Review population

The utility describing the HRQoL of transfusion dependent patients used in the company's economic model was 0.69. This was the mean of all patients aged ≥16 in the Chart Review, but the ERG highlighted this population was substantially older than the patients included in the trial, and had no restrictions on co-morbidities, iron load, or general baseline health. The ERG, therefore, requested that the company provide a re-analysis of HRQoL in the Chart Review dataset, limiting the population to those aged 12-35, and excluding patients with a 'high' cardiac T2* and those whose existing co-morbidities were already separately accounted for in the model (such as diabetes and hypogonadism) to avoid double counting these decrements. The mean utility of this population was [redacted] which the ERG considered to be more comparable to that of the baseline value for patients included in the Zynteglo trials, and to the population who might be eligible for treatment in practice. By removing older patients from this analysis, it also means that we are no longer adjusting down an already age-adjusted utility value for age, which resulted in an overly low utility for older patients in the company's analysis.

The use of the age-appropriate Chart Review population utility increased the ICER of Zynteglo by [redacted] (+47.2%) per QALY gained versus SoC, from the corrected company base-case analysis (Table 34). The impact upon the ICER is so significant because the majority of QALY gains on Zynteglo in the company's analysis are generated through improvement in the modelled HRQoL of patients over a long period of time, rather than by an extension to life. By aligning the HRQoL of TDT patients to be consistent with the trial population and avoiding double counting the impact of

iron-related complications and age, Zynteglo generates fewer incremental QALYs compared with the company base case analysis.

Table 34 ERG Scenario 16: Transfusion dependence utility from adjusted Chart Review population

Intervention	Total			Incremental			Change from company base case ICER
	Costs	LYs*	QALYs	Costs	QALYs	ICER	
SoC	[redacted]	37.79	21.95				
Zynteglo	[redacted]	53.63	30.87	[redacted]	8.92	[redacted]	[redacted]

* undiscounted life years

The ERG also considered the combined impact of the two scenarios explored above: use of the age-related disutilities taken from the full Ara and Brazier population, and the analysis of the adjusted Chart Review population to estimate the health state utility value for transfusion dependent patients. As presented in Table 35, the impact to the ICER is substantial, as both analyses result in a higher HRQoL for standard care patients, leading to fewer incremental QALYs for Zynteglo.

Table 35 ERG Scenario 17: Combination of Scenario 14 and 15

Intervention	Total			Incremental			Change from company base case ICER
	Costs	LYs*	QALYs	Costs	QALYs	ICER	
SoC	[redacted]	37.79	21.17				
Zynteglo	[redacted]	53.63	28.28	[redacted]	7.11	[redacted]	[redacted]

* undiscounted life years

6.3.9.3 Transfusion independent patients incur a disutility for subcutaneous chelation during iron normalisation period

The ERG did not consider it appropriate to disregard the disutilities associated with subcutaneous iron chelation therapy in transfusion independent patients who continued to receive this treatment during the iron normalisation period. In the Zynteglo trials, a proportion of patients remained on, or returned to, iron chelation therapies following their treatment with Zynteglo. The company applied utilities derived from the vignette study for SoC patients, and specifically referenced the disadvantages of using subcutaneous iron chelators throughout their submission.

The ERG examined the impact of including a utility decrement relative to the patient's baseline utility, equivalent to that captured in the vignette study, for those treated using subcutaneous iron chelators during the iron normalisation period, thus capturing the additional burden of administration and adverse effects associated with this treatment. The results of this scenario are presented in Table 36 below.

Table 36 ERG Scenario 18: Disutility for TI patients on subcutaneous chelation therapy

Intervention	Total			Incremental			Change from company base case ICER
	Costs	LYs*	QALYs	Costs	QALYs	ICER	
SoC	[redacted]	37.79	17.20				
Zynteglo	[redacted]	53.63	30.26	[redacted]	13.06	[redacted]	[redacted]

* undiscounted life years

6.3.9.4 Removal of the disutility for infertility

The ERG questioned the appropriateness of including the impact of infertility upon HRQoL as a result of the myeloablative conditioning procure given prior to Zynteglo treatment. Its impact is poorly understood and not typically well captured using EQ-5D, and the value was estimated from a source that was not generalizable to the present context. It may not be appropriate to capture this impact within the present appraisal, and it has previously only been used in models of assisted reproductive technologies for patients actively seeking fertility treatment.

The ERG examined the impact of removing this disutility from the analysis. The results of this scenario are presented in Table 37 below. A modest decrease in the ICER was observed.

Table 37 ERG Scenario 19: Removal of the disutility for infertility

Intervention	Total			Incremental			Change from company base case ICER
	Costs	LYs*	QALYs	Costs	QALYs	ICER	
SoC	[redacted]	37.79	17.20				
Zynteglo	[redacted]	53.63	30.91	[redacted]	13.71	[redacted]	[redacted]

* undiscounted life years

6.3.10 Drug acquisition costs from eMIT

Drug acquisition costs for a number of therapies used in the treatment process were obtained from the BNF. However, some of the therapies, including the subcutaneous chelation agent desferrioxamine and two of the therapies used in the Zynteglo treatment process, are generic products that are widely available to the NHS at discounted prices. The ERG considers eMIT to be a more representative estimate of drug expenditure (costs presented in Table 15 in Section 5.2.8). Unit costs from eMIT are generally considerably lower than those in the BNF, and the use of the BNF costs will overestimate drug expenditure.

Results of this scenario are presented in Table 38. The inclusion of eMIT unit costs results in a small increase to the ICER. While using eMIT unit costs impacts the total costs accrued in both arms, the

effect is more pronounced in the SoC arm where desferrioxamine is applied over a lifetime, while the two treatments associated with Zynteglo are one-off.

Table 38 ERG Scenario 20: Drug unit costs acquisition costs from eMIT

Intervention	Total			Incremental			Change from company base case ICER
	Costs	LYs*	QALYs	Costs	QALYs	ICER	
SoC	[redacted]	37.79	17.20				
Zynteglo	[redacted]	53.63	30.34	[redacted]	13.13	[redacted]	[redacted]

* undiscounted life years

6.4 ERG's alternative base-case analysis

The ERG's alternative base-case analysis combines a number of the above scenario analyses. The ERG considers this new analysis to better reflect the uncertainties around the clinical data, and to address the ERG's concerns surrounding the assumptions and data sources used company's base-case analysis.

This analysis includes the following changes from the company's base-case:

- Age category-specific body weight used in chelation dosage calculations (Scenario 1),
- 20% of the population have hypogonadism at baseline (Scenario 2),
- Distribution of iron chelation therapies based on 12-35s in Chart Review (Scenario 3),
- Iron normalisation period of 5 years (Scenario 10),
- Costs and outcomes discounted at 3.5% (Scenario 14),
- Age-adjustment of utilities based on Ara and Brazier general population values (Scenario 15),
- Age- and comorbidity-adjusted Chart Review utility values (Scenario 16),
- eMIT drug acquisition costs used (Scenario 20).

The results of the probabilistic and deterministic base-case analyses are presented in Table 39. The deterministic ICER generated using this set of assumptions is [redacted] and the probabilistic ICER is [redacted] per QALY for Zynteglo compared with standard care. These results include the confidential PAS discount available for Zynteglo.

Table 39 Results of the ERG's alternative base-case analysis

Intervention	Total			Incremental			Change from company base case ICER
	Costs	LYs*	QALYs	Costs	QALYs	ICER	
Company's corrected base-case results (deterministic)							
SoC	[redacted]	37.79	17.20				
Zynteglo	[redacted]	53.63	30.34	[redacted]	13.13	[redacted]	-
ERG deterministic base-case							
SoC	[redacted]	37.79	15.48				
Zynteglo	[redacted]	53.40	18.53	[redacted]	3.05	[redacted]	[redacted]
ERG probabilistic base-case							
SoC	[redacted]	39.03	15.62				
Zynteglo	[redacted]	53.54	18.55	[redacted]	2.93	[redacted]	[redacted] **

* undiscounted life years, ** relative to the base-case probabilistic ICER

6.4.1 Scenarios on the ERG alternative base case analysis

The selection of changes made to the ERG base-case analysis were driven by the available evidence; however, a number of important uncertainties remain. To address the remaining uncertainty, the ERG conducted a number of scenarios on their alternative base-case analysis.

The first of these scenarios included the use of a 1.5% discount rate for costs and benefits, which was originally applied by the company in their base-case analysis. The company cited the example of Strimvelis, the only other gene therapy considered by NICE, which was appraised using the NICE HST process. The NICE committee discussing the evidence for Strimvelis stated that "it was uncertain about whether Strimvelis fully met the criteria to use a discounting rate of 1.5%, and that both discount rates should be considered by the committee during its decision-making".[49] As such, the ERG hereby presents the results of their analysis using the 1.5% discount rate to enable a comparison of results, but considers that the discount rate of 3.5% remains the most appropriate for reasons discussed extensively in Section 5.2.5.

In a second scenario, the ERG considers the impact of assuming a less severe impact to the survival of transfusion dependent patients, and applies an SMR of 2 in their alternative base case analysis. This value is not driven by published evidence but rather a clinically justified assumption that current transfusion-dependent patients, who do not have any cardiac complications, would have much improved survival than that observed in previous decades (Section 5.2.6.5). Since it is subject to a high degree of uncertainty, the ERG did not consider it appropriate to use this value in their

alternative base case scenario, but considers it an important assumption to explore, as the alternative base-case almost certainly underestimates the number of LYs generated by standard care patients.

Table 40 presents the results of these additional scenarios. Notably, with the application of the company's preferred discount rate of 1.5%, Zynteglo cannot be considered cost-effective, with an ICER of [redacted]. As demonstrated in the second scenario, the application of a lower value for the SMR for transfusion dependent patients had a much greater impact to the ICER than when it was applied to the company's base-case analysis (a decrease of [redacted] to the ICER when applied to the company analysis, versus an increase of [redacted] when applied to the ERG analysis). In the company model, the increased cost of SoC due to more of these patients remaining alive and requiring treatment outweighs the increase in QALYs due to a reduction in mortality. However, in the ERG analysis, greater QALYs are generated by patients on SoC due to higher utility estimates being estimated and applied by the ERG for transfusion dependent patients.

Table 40 Results of scenario analyses on the ERG alternative base case analysis

Intervention	Total			Incremental			Change from company base case ICER
	Costs	LYs*	QALYs	Costs	QALYs	ICER	
Scenario: 1.5% discount rate							
SoC	[redacted]	37.79	21.07				
Zynteglo	[redacted]	53.40	27.78	[redacted]	6.71	[redacted]	[redacted]
Scenario: SMR of 2 for transfusion dependent patients							
SoC	[redacted]	42.41	16.14				
Zynteglo	[redacted]	53.89	18.62	[redacted]	2.48	[redacted]	[redacted]
Scenario: 1.5% discount rate and SMR of 2 for transfusion dependent patients							
SoC	[redacted]	42.41	22.54				
Zynteglo	[redacted]	53.89	27.96	[redacted]	5.42	[redacted]	[redacted]

* undiscounted life years

Due to the uncertainty in the modelling sampling procedure and its impact on the ICER, the results of the ERG base case are presented in Table 41, in which results are based on 5,000 modelled profiles, rather than 600 as in the company base-case analysis.

The results show a considerable increase in the ICERs of all three scenarios compared to the same scenarios based on 600 profiles (presented in Table 40). The change from the company base-case ICER was estimated using the company ICER based on 5,000 patient profiles. Notably, increasing the number of 5,000 profiles increases the ICER of the ERG base case to [redacted], from [redacted] per QALY.

Table 41 Results of scenario analyses on the ERG alternative base case analysis, using 5,000 profiles

Intervention	Total			Incremental			Change from company base case ICER
	Costs	LYs*	QALYs	Costs	QALYs	ICER	
ERG base case analysis, based on 5,000 model profiles							
SOC	[redacted]	37.34	15.32				
Zynteglo	[redacted]	52.60	18.25	[redacted]	2.93	[redacted]	[redacted]
Scenario: ERG base-case analysis with 1.5% discount rate, based on 5,000 model profiles							
SoC	[redacted]	37.34	20.81				
Zynteglo	[redacted]	52.60	27.28	[redacted]	6.47	[redacted]	[redacted]
Scenario: ERG base-case analysis with SMR of 2 for transfusion dependent patients, based on 5,000 model profiles							
SoC	[redacted]	41.97	15.99				
Zynteglo	[redacted]	53.13	18.35	[redacted]	2.35	[redacted]	[redacted]
Scenario: ERG base-case analysis with 1.5% discount rate and SMR of 2 for transfusion dependent patients, based on 5,000 model profiles							
SoC	[redacted]	41.97	22.30				
Zynteglo	[redacted]	53.13	27.47	[redacted]	5.17	[redacted]	[redacted]

* undiscounted life years

6.5 Conclusions from the ERG analyses

The ERG has presented a number of additional analyses carried out in stages. These exploratory analyses were undertaken on a model provided by the company at the clarification stage, which addressed an error identified by the ERG, and included updates to a number of modelling assumptions. The impact of these changes resulted in a company base-case ICER of £[redacted] per QALY.

Using the corrected and updated model, the ERG then presented a number of "single-change" analyses considering a range of issues raised in Section 5.2. These scenario analyses addressed the following issues:

- Baseline population characteristics and presence of co-morbidities adjusted to reflect the eligible population,
- Distribution of iron chelation agents adjusted to reflect the eligible population, using a re-analysis of the Chart Review data,
- Clinical effectiveness data adjusted for the underrepresentation of the IVS-I-5 and IVS-I-110 genotypes,

- Alternative assumptions around the rate of early engraftment failure and the long-term persistence of transfusion independence,
- Alternative assumptions on the upon mortality of transfusion dependent patients,
- Alternative assumptions around iron normalisation and the development of iron overload-related complications,
- Impact of using a 3.5% discount rate on costs and effects,
- Use of a number of different sources and assumptions around health-related quality of life,
- Alternative drug acquisition unit costs.

The scenarios associated with the greatest impact on cost-effectiveness outcomes involved the use of a 3.5% discount rate for costs and benefits, as per the NICE reference case. The discount rate significantly impacts the analysis due to the difference in the way Zynteglo and standard care patients accrue costs and QALYs. With Zynteglo, the majority of costs are allocated upfront while benefits are generated over the lifetime. Aligning the discount rate in the model to that specified in the reference case resulted in the ICER increasing from £[redacted] to £[redacted]. Scenario analyses that explored the impact of an alternative source of general population age-related utility values and a re-estimate of the utility values for transfusion dependent patients demonstrated that the results of the model are sensitive to these assumptions, as the majority of the benefit attributed to Zynteglo was assumed to be due to reductions in morbidity as well as mortality. The company's original assumptions over-estimated HRQoL following Zynteglo treatment and under-estimated HRQoL on standard care.

The ERG alternative base-case implemented a number of the assumptions that were included in the company exploratory analyses. This analysis estimated Zynteglo to be more costly (cost difference [redacted]) and more effective (3.05 QALY gain) compared with standard care, and suggests that the ICER for Zynteglo compared with standard care is £[redacted] per QALY. Further analyses undertaken by the ERG on their alternative base-case suggested that the mortality rate for transfusion-dependent patients was also an influential parameter in the analysis.

7 End of life

This intervention does not meet the end of life criteria published by NICE.

8 Overall conclusions

Results from the three Zynteglo trials demonstrate that most patients respond well to Zynteglo and achieve transfusion independence. However, the Zynteglo trial population appears not to be a particularly good representation of the UK TDT population and consequently there is uncertainty about how effective Zynteglo is in the subgroup of patients with severe non-β^0/β^0 genotypes. It is possible that [redacted] may increase the probability of achieving TI in patients with severe non-β^0/β^0 genotypes (compared to the previous processes used in the trial programme) but only results from the study HGB-212 and further data from HGB-207 can resolve this uncertainty. The trials results are still quite immature, and the number of patients treated is small, so uncertainty exists regarding the longevity of Zynteglo and the possibility of adverse events in the medium-to-long term.

The applicability of the comparator data selected by the company is not optimal since some studies did not reflect the improvements in TDT patient treatment and management achieved over the last 10-20 years. This was exacerbated by the lack of transparency regarding the methods used to identify many studies and the limited consideration of the implication of study limitations.

[redacted]

The ERG's critique of the economic evaluation presented by the company focused upon a number of key challenges, and performed a range of scenarios to explore the impact of alternative assumptions upon the cost-effectiveness of Zynteglo.

The ERG proposed an alternative base-case analysis to address several of the key uncertainties identified. The main changes implemented by the ERG include adjustment of the baseline characteristics of the modelled population based on a reanalysis of the Chart Review data, increasing the iron normalisation period to 5 years, discounting of costs and outcomes at 3.5% per annum, use of a more appropriate value set to calculate age-related decline in HRQoL, and using utility estimates for TDT patients based on reanalysis of the Chart Review population adjusted for age and co-morbidities. The probabilistic ICER estimated using the ERG's preferred assumptions was [redacted], which was [redacted] higher than the company's base-case probabilistic ICER. The single most significant reason for this increase was the use of a 3.5% discount rate over the company's 1.5% rate, since for Zynteglo the majority of costs are incurred upfront and are therefore mostly unaffected by the higher discount rate, while benefits are generated over the patient's lifetime and are thus subject to more discounting.

A number of uncertainties remain unexplored in the absence of more appropriate evidence. This could mean that the results of the ERG's preferred base-case analysis may underestimate the true ICER for Zynteglo. Firstly, the long-term benefits of Zynteglo are uncertain due to the immaturity of the evidence base. The ERG is concerned that anything other than an indefinite treatment effect across all patients could have a substantial impact upon the cost-effectiveness of Zynteglo, and may subject the NHS to large irrecoverable costs. Heterogeneity of treatment effect also remains an area of uncertainty, as heterogeneity based on genotype and manufacturing process is not addressed in the evidence. Further uncertainty also surrounds the resolution of elevated iron levels following successful treatment with Zynteglo, and the associated reduction in risk of iron-related complications.

8.1 Implications for research

As Zynteglo has a conditional license via the EMA's adaptive pathways programme, further follow-up data on efficacy and safety will continue to be routinely collected for all the Zynteglo trials. Of particular interest would be data on the iron normalisation period following Zynteglo treatment, and persistence of the therapeutic effect in the long term.

More contemporary evidence on survival and HRQoL of TDT patients on current therapies, including the use of combination chelation therapy, would be beneficial.

9 References

1. Zynteglo SmPC. bluebird bio, Zynteglo 1.2 20 × 106 cells/mL dispersion for infusion. Draft Summary of Product Characteristics. 2019.
2. Henderson S, Timbs A, McCarthy J, Gallienne A, Van Mourik M, Masters G, et al. Incidence of haemoglobinopathies in various populations - the impact of immigration. *Clin Biochem* 2009;**42**:1745-56.
3. Kulozik AE, Locatelli F, Yannaki E, Porter JB, Thuret I, Sauer MG, et al. *Results from the phase 3 Northstar-3 study evaluating lentiglobin gene therapy in patients with transfusion-dependent β-thalassaemia and a β0 or IVS-I-110 mutation at both alleles of the HBB gene*. In: 24th European Hematology Association Congress. Amsterdam, June 13-16; 2019.
4. Lal A, Locatelli F, Kwiatkowski JL, Kulozik AE, Yannaki E, Porter JB, et al. Northstar-3: Interim results from a phase 3 study evaluating lentiglobin gene therapy in patients with transfusion-dependent β-thalassemia and either a β0 or IVS-I-110 mutation at both alleles of the HBB gene. *Blood* 2019;**134**:815.
5. Hongeng S, Thompson AA, Kwiatkowski JL, Rasko JE, Schiller GJ, Deary B, et al. *Clinical outcomes following lentiglobin gene therapy for transfusion-dependent β-thalassemia in the northstar HGB-204 study*. In: 23rd Annual Congress of 2018 Asia-Pacific Blood and Marrow Transplantation Group (APBMT). Taipei, Taiwan, 2-4 Nov; 2018.
6. Jackson R, Ameratunga S, Broad J, Connor J, Lethaby A, Robb G, et al. The GATE frame: critical appraisal with pictures. *Evid Based Med* 2006;**11**:35-8.
7. National Institute for Health and Care Excellence (NICE). *Methods for the development of NICE public health guidance (third edition)*. London: NICE; 2012.
8. Foster M. *National Haemoglobinopathy Registry annual report - 2018/19*. Manchester: Medical Data Solutions and Services (MDSAS); 2019.
9. bluebird bio Inc. *Clinical Study Report HGB 205: A phase 1/2 open label study evaluating the safety and efficacy of gene therapy of the β-hemoglobinopathies (sickle cell anemia and β-thalassemia major) by transplantation of autologous CD34+ stem cells transduced ex vivo with a lentiviral β^{A-T87Q}- globin vector (Lentiglobin BB305 drug product). [Unpublished, confidential document]*. Cambridge, MA: bluebird bio Inc; 2018.
10. bluebird bio Inc. *Clinical Study Report HGB-204. A phase 1/2, open label study evaluating the safety and efficacy of gene therapy in subjects with β-thalassemia major by transplantation of autologous CD34+ stem cells transduced ex vivo with a lentiviral $\beta^{A\ T87Q}$-globin vector (LentiGlobin BB305 drug product). [Draft, unpublished confidential document]*. . Cambridge, MA: bluebird bio Inc; 2018.
11. Thompson AA, Walters MC, Kwiatkowski J, Rasko JEJ, Ribeil J-A, Hongeng S, et al. Gene therapy in patients with transfusion-dependent β-thalassemia. *N Engl J Med* 2018;**378**:1479-93.
12. bluebird bio Inc. *Interim Clinical Study Report HGB-207. A phase 3 single arm study evaluating the efficacy and safety of gene therapy in subjects with transfusion-dependent β-thalassemia, who do not have a β^0/β^0 genotype, by transplantation of autologous CD34+ stem cells transduced ex vivo with a lentiviral $\beta^{a\text{-}t87q}$-globin vector in subjects ≤50 years of age. [Unpublished, confidential document]*. Cambridge, MA: bluebird bio Inc; 2018.
13. bluebird bio Inc. *June 2019 integrated TLFs for HGB-204, HGB-205, HGB-207 and LTF-303*. In: bluebird bio Inc; 2019.
14. Locatelli F, Wlater MC, Kwiatkowski JL, Porter JB, Sauer MG, Thuret I, et al. LentiGlobin gene therapy for patients with transfusion-dependent β-thalassemia (TDT): results from the phase 3 Northstar-2 and Northstar-3 studies. In: *American Society of Hematology*; 2018; San Diego, CA. 2018.
15. Rasko JE, Thompson AA, Kwiatkowski JL, Hongeng S, Schiller GJ, Anurathapan U, et al. *Clinical outcomes of LentiGlobin gene therapy for transfusion-dependent β-thalassemia following completion of the Northstar HGB-204 study*. In: American Society of Hematology. San Diego, CA; 2018.

16. *Summary of product characterstics for Busilvex 6mg/ml concentrate for solution for infusion*: European Medicines Agency; 2017.
17. Evidera, bluebird bio Inc. *Systematic literature review to support lentiglobin® in transfusion-dependent β-thalassemia for NICE submission. Study Report. EVA-20726-04 (version 1.0) [confidential]*: Evidera; 2019.
18. bluebird bio Inc. *An observational study to evaluate the routine management, healthcare resource use and outcomes for patients with transfusion-dependent β-thalassaemia treated in the United Kingdom (TDT Chart Review) [draft manuscript, confidential]*. In: Open Vie; 2019.
19. Evidera, bluebird bio Inc. *Systematic review of the burden of disease and treatment for transfusion dependent β-thalassemia EVA-20726 (version 3.0) [confidential]*: Evidera; 2018.
20. Higgins JPT, Sterne JAC, Savovic J, Page MJ, Hróbjartsson A, Boutron I, et al. A revised tool for assessing risk of bias in randomized trials. *Cochrane Database of Systematic Reviews* 2016;**10**:29-31.
21. Delea TE, Sofrygin O, Thomas SK, Baladi JF, Phatak PD, Coates TD. Cost effectiveness of once-daily oral chelation therapy with deferasirox versus infusional deferoxamine in transfusion-dependent thalassaemia patients: US healthcare system perspective. *Pharmacoeconomics* 2007;**25**:329-42.
22. Gabutti V, Piga A. Results of long-term iron-chelating therapy. *Acta Haematol* 1996;**95**:26-36.
23. Piga A, Gaglioti C, Fogliacco E, Tricta F. Comparative effects of deferiprone and deferoxamine on survival and cardiac disease in patients with thalassemia major: a retrospective analysis. *Haematologica* 2003;**88**:489-96.
24. Anderson LJ, Holden S, Davis B, Prescott E, Charrier CC, Bunce NH, et al. Cardiovascular T2-star (T2*) magnetic resonance for the early diagnosis of myocardial iron overload. *Eur Heart J* 2001;**22**:2171-9.
25. Kremastinos DT, Tsetsos GA, Tsiapras DP, Karavolias GK, Ladis VA, Kattamis CA. Heart failure in beta thalassemia: a 5-year follow-up study. *Am J Med* 2001;**111**:349-54.
26. Casale M, Filosa A, Ragozzino A, Amendola G, Roberti D, Tartaglione I, et al. Long-term improvement in cardiac magnetic resonance in beta-thalassemia major patients treated with deferasirox extends to patients with abnormal baseline cardiac function. *Am J Hematol* 2019;**94**:312-8.
27. Modell B, Khan M, Darlison M, Westwood MA, Ingram D, Pennell DJ. Improved survival of thalassaemia major in the UK and relation to T2* cardiovascular magnetic resonance. *J Cardiovasc Magn Reson* 2008;**10**:42.
28. Pepe A, Rossi G, Bentley A, Putti MC, Frizziero L, D'Ascola DG, et al. Cost-utility analysis of three iron chelators used in monotherapy for the treatment of chronic iron overload in beta-thalassaemia major patients: an Italian perspective. *Clin Drug Investig* 2017;**37**:453-64.
29. Angelucci E, Muretto P, Nicolucci A, Baronciani D, Erer B, Gaziev J, et al. Effects of iron overload and hepatitis C virus positivity in determining progression of liver fibrosis in thalassemia following bone marrow transplantation. *Blood* 2002;**100**:17-21.
30. Ang AL, Tzoulis P, Prescott E, Davis BA, Barnard M, Shah FT. History of myocardial iron loading is a strong risk factor for diabetes mellitus and hypogonadism in adults with beta thalassemia major. *Eur J Haematol* 2014;**92**:229-36.
31. Novartis Pharmaceuticals Corporation. *Prescribing information for JADENU® (deferasirox) tablets, for oral use and JADENU® Sprinkle (deferasirox) granules, for oral use*. 2019. Available from: https://www.pharma.us.novartis.com/sites/www.pharma.us.novartis.com/files/jadenu.pdf. [accessed 14th January 2020].
32. Apotex. *Prescribing information for FERRIPROX® (deferiprone) tablets, for oral use*. 2011. Available from: https://www.accessdata.fda.gov/drugsatfda_docs/label/2011/021825lbl.pdf. [accessed 14th January 2020].
33. Fisher SA, Brunskill SJ, Doree C, Gooding S, Chowdhury O, Roberts DJ. Desferrioxamine mesylate for managing transfusional iron overload in people with transfusion-dependent thalassaemia. *Cochrane Database Syst Rev* 2013:CD004450.
34. Pistoia L, Meloni A, Salvadori S, Spasiano A, Lisi R, Rosso R, et al. Cardiac involvement by CMR in different genotypic groups of thalassemia major patients. *Blood Cells Mol Dis* 2019;**77**:1-7.

35. Poomthavorn P, Chawalitdamrong P, Hongeng S, Mahachoklertwattana P, Pakakasama S, Khlairit P, et al. Gonadal function of beta-thalassemics following stem cell transplantation conditioned with myeloablative and reduced intensity regimens. *J Pediatr Endocrinol Metab* 2013;**26**:925-32.
36. Chaudhury S, Ayas M, Rosen C, Ma M, Viqaruddin M, Parikh S, et al. A multicenter retrospective analysis stressing the importance of long-term follow-up after hematopoietic cell transplantation for beta-thalassemia. *Biol Blood Marrow Transplant* 2017;**23**:1695-700.
37. John MJ, Jyani G, Jindal A, Mashon RS, Mathew A, Kakkar S, et al. Cost Effectiveness of Hematopoietic Stem Cell Transplantation Compared with Transfusion Chelation for Treatment of Thalassemia Major. *Biol Blood Marrow Transplant* 2018;**24**:2119-26.
38. Leelahavarong P, Chaikledkaew U, Hongeng S, Kasemsup V, Lubell Y, Teerawattananon Y. A cost-utility and budget impact analysis of allogeneic hematopoietic stem cell transplantation for severe thalassemic patients in Thailand. *BMC Health Serv Res* 2010;**10**:209.
39. Karnon J, Tolley K, Vieira J, Chandiwana D. Lifetime cost-utility analyses of deferasirox in beta-thalassaemia patients with chronic iron overload: a UK perspective. *Clin Drug Investig* 2012;**32**:805-15.
40. Caro JJ. Discretely Integrated Condition Event (DICE) Simulation for Pharmacoeconomics. *Pharmacoeconomics* 2016;**34**:665-72.
41. Weidlich D, Kefalas P, Guest JF. Healthcare costs and outcomes of managing β-thalassemia major over 50 years in the United Kingdom. *Transfusion* 2016;**56**:1038-45.
42. Kwiatkowski JL, Kim HY, Thompson AA, Quinn CT, Mueller BU, Odame I, et al. Chelation use and iron burden in North American and British thalassemia patients: a report from the Thalassemia Longitudinal Cohort. *Blood* 2012;**119**:2746-53.
43. *Girls UK growth chart 2-18 years*. Royal College of Paediatrics and Child Health; 2012. Available from: https://www.rcpch.ac.uk/sites/default/files/Girls_2-18_years_growth_chart.pdf [accessed 10th January 2019].
44. *Boys UK growth chart 2-18 years*. Royal College of Paediatrics and Child Health; 2012. Available from: https://www.rcpch.ac.uk/sites/default/files/Boys_2-18_years_growth_chart.pdf [accessed 10th January 2019].
45. United Kingdom Thalassaemia Society. *Standards for the Clinical Care of Children and Adults with Thalassaemia in the UK, 3rd Edition*. London: United Kingdom Thalassaemia Society; 2016.
46. Javanbakht M, Keshtkaran A, Shabaninejad H, Karami H, Zakerinia M, Delavari S. Comparison of Blood Transfusion Plus Chelation Therapy and Bone Marrow Transplantation in Patients with β-Thalassemia: Application of SF-36, EQ-5D, and Visual Analogue Scale Measures. *Int J Health Policy Manag* 2015;**4**:733-40.
47. Seyedifar M, Dorkoosh FA, Hamidieh AA, Naderi M, Karami H, Karimi M, et al. Health-Related Quality of Life and Health Utility Values in Beta Thalassemia Major Patients Receiving Different Types of Iron Chelators in Iran. *Int J Hematol Oncol Stem Cell Res* 2016;**10**:224-31.
48. bluebird bio Inc. *Zynteglo TDT: UK Patient Preference report*. In; 2019.
49. National Institute for Health and Care Excellence (NICE). *Strimvelis for treating adenosine deaminase deficiency-severe combined immunodeficiency [HST7]*. NICE; 2018. Available from: https://www.nice.org.uk/guidance/hst7 [accessed 13th January 2020].
50. Inati A, Kahale M, Sbeiti N, Cappellini MD, Taher AT, Koussa S, et al. One-year results from a prospective randomized trial comparing phlebotomy with deferasirox for the treatment of iron overload in pediatric patients with thalassemia major following curative stem cell transplantation. *Pediatr Blood Cancer* 2017;**64**:188-96.
51. Angelucci E. Complication free survival long-term after hemopoietic cell transplantation in thalassemia. *Haematologica* 2018;**103**:1094-6.
52. Pepe A, Meloni A, Rossi G, Midiri M, Missere M, Valeri G, et al. Prediction of cardiac complications for thalassemia major in the widespread cardiac magnetic resonance era: a prospective multicentre study by a multi-parametric approach. *Eur Heart J Cardiovasc Imaging* 2018;**19**:299-309.
53. Angelucci E, Burrows N, Losi S, Bartiromo C, Hu XH. Beta-Thalassemia (BT) Prevalence and Treatment Patterns in Italy: A Survey of Treating Physicians. 2016;**128**:3533-.

54. Shenoy S, Angelucci E, Arnold SD, Baker KS, Bhatia M, Bresters D, et al. Current results and future research priorities in late effects after hematopoietic stem cell transplantation for children with sickle cell disease and thalassemia: a consensus statement from the Second Pediatric Blood and Marrow Transplant Consortium International Conference on late effects after pediatric hematopoietic stem cell transplantation. *Biol Blood Marrow Transplant* 2017;**23**:552-61.
55. Rahal I, Galambrun C, Bertrand Y, Garnier N, Paillard C, Frange P, et al. Late effects after hematopoietic stem cell transplantation for beta-thalassemia major: the French national experience. *Haematologica* 2018;**103**:1143-9.
56. ONS. *National life tables, UK: 2015 to 2017*. 2018. Available from: https://www.ons.gov.uk/releases/nationallifetablesuk2015to2017 [accessed
57. Jobanputra M, Paramore C, Laird SG, McGahan M, Telfer P. Co-morbidities and mortality associated with Transfusion Dependent Beta-thalassaemia Patients in England: A 10-Year Retrospective Cohort Analysis - Draft manuscript. Draft manuscript.
58. *Summary of product characteristics for Zynteglo*: European Medicines Agency; 2019.
59. Matza LS, Paramore LC, Stewart KD, Karn H, Jobanputra M, Dietz AC. Health state utilities associated with treatment for transfusion-dependent beta-thalassemia. *Eur J Health Econ* 2019.
60. Busnelli A, Somigliana E, Benaglia L, Leonardi M, Ragni G, Fedele L. In vitro fertilization outcomes in treated hypothyroidism. *Thyroid* 2013;**23**:1319-25.
61. Scotland G, McLernon D, Kurinczuk J, McNamee P, Harrild K, Lyall H, et al. Minimising twins in *in vitro* fertilisation: a modelling study assessing the costs, consequences and cost–utility of elective single versus double embryo transfer over a 20-year time horizon. *Br J Obstet Gynaecol* 2011;**118**:1073-83.
62. Ara R, Brazier JE. Using health state utility values from the general population to approximate baselines in decision analytic models when condition-specific data are not available. *Value Health* 2011;**14**:539-45.
63. Fryback DG, Dasbach EJ, Klein R, Klein BE, Dom N, Peterson K, et al. The Beaver Dam Health Outcomes Study: initial catalog of health-state quality factors *Med Decis Making* 1993;**13**:89-102.
64. Overview of analytic approach and results. In: Institute of Medicine, Stratton KR, Durch JS, Lawrence RS, editors. *Vaccines for the 21st century: a tool for decisionmaking*. Washington, DC: The National Academies Press; 2000. p. 53-92.
65. Appendix 17 Neisseria gonorrhea. In: Institute of Medicine, Stratton KR, Durch JS, Lawrence RS, editors. *Vaccines for the 21st century: a tool for decisionmaking*. Washington, D C: The National Academies Press; 2000. p. 257-66.
66. Jalkanen K, Aarnio E, Lavikainen P, Jauhonen H-M, Enlund H, Martikainen J. Impact of type 2 diabetes treated with non-insulin medication and number of diabetes-coexisting diseases on EQ-5D-5 L index scores in the Finnish population. *Health Qual Life Outcomes* 2019;**17**:117.
67. Solli O, Stavem K, Kristiansen IS. Health-related quality of life in diabetes: The associations of complications with EQ-5D scores. *Health Qual Life Outcomes* 2010;**8**:18.
68. NHS Improvement. *National schedule of reference costs*. Available from: https://improvement.nhs.uk/resources/reference-costs/#rc1718 [accessed 17th January 2020].
69. Joint Formulary Committee. *British National Formulary (online)* BMJ Group and Pharmaceutical Press; 2019. Available from: https://bnf.nice.org.uk/ [accessed 17th January 2020].
70. Curtis LA, Burns A. *Unit costs of health and social care 2018*. University of Kent; 2018. Available from: https://kar.kent.ac.uk/70995/ [accessed 17th January 2020].
71. Drezet A, Touzot F, Magalon J, Caccavelli L, Lezoray M, Taupin P, et al. *Budget impact analysis of gene therapy of hematopoietic stem cells for the management of beta-Thalassemia major*. In; 2019.
72. National Institute for Health and Care Excellence (NICE. *Costing statement: Blood transfusion. Implementing the NICE guideline on blood transfusion (NG24)*: NICE; 2015. Available from: https://www.nice.org.uk/guidance/ng24/resources/costing-statement-2177158141
73. Agrawal S, Davidson N, Walker M, Gibson S, Lim C, Morgan CL, et al. Assessing the total costs of blood delivery to hospital oncology and haematology patients. *Curr Med Res Opin* 2006;**22**:1903-9.

74. Bentley A, Gillard S, Spino M, Connelly J, Tricta F. Cost-utility analysis of deferiprone for the treatment of beta-thalassaemia patients with chronic iron overload: a UK perspective. *Pharmacoeconomics* 2013;**31**:807-22.
75. Karnon J, Tolley K, Oyee J, Jewitt K, Ossa D, Akehurst R. Cost-utility analysis of deferasirox compared to standard therapy with desferrioxamine for patients requiring iron chelation therapy in the United Kingdom. *Curr Med Res Opin* 2008;**24**:1609-21.
76. NHS England. *National Genomic Test Directory.Testing criteria for rare and inherited disease*: NHS England; 2019.
77. Department of Health and Social Care. *Drugs and pharmacuetical electronic market information (eMit)*. 2011. Available from: https://www.gov.uk/government/publications/drugs-and-pharmaceutical-electronic-market-information-emit [accessed 10th December 2019].
78. Moller J, Davis S, Stevenson M, Caro JJ. Validation of a DICE simulation against a discrete event simulation implemented entirely in code. *Pharmacoeconomics* 2017;**35**:1103-9.
79. National Institute for Health and Care Excellence (NICE). *Cabozantinib for treating medullary thyroid cancer [TA516]*. NICE; 2018. Available from: https://www.nice.org.uk/guidance/ta516 [accessed 16th January 2019].
80. Davis S, Stevenson M, Tappenden P, Wailoo AJ. *NICE DSU Technical Support Document 15: Cost-effectiveness modelling using patient-level simulation* Sheffield: Decision Support Unit, School of Health and Related Research, University of Sheffield; 2014.

10 Appendices

10.1 Appendix 1: Drummond Checklist

Table 42 Quality assessment of included CEA study using Drummond et al. checklist completed by the ERG

	CEA quality assessment questions	Answer (Yes/No/Unclear)	Notes/Explanation for No or Unclear
1	Was the research question stated?	Yes	-
2	Was the economic importance of the research question stated?	Yes	-
3	Was/were the viewpoint(s) of the analysis clearly stated and justified?	Yes	-
4	Was a rationale reported for the choice of the alternative programmes or interventions compared?	Yes	-
5	Were the alternatives being compared clearly described?	Yes	-
6	Was the form of economic evaluation stated?	Yes	-
7	Was the choice of form of economic evaluation justified in relation to the questions addressed?	Yes	-
8	Was/were the source(s) of effectiveness estimates used stated?	Yes	-
9	Were details of the design and results of the effectiveness study given (if based on a single study)?	Yes	-

10	Were details of the methods of synthesis or meta-analysis of estimates given (if based on an overview of a number of effectiveness studies)?	Yes	-
11	Were the primary outcome measure(s) for the economic evaluation clearly stated?	Yes	-
12	Were the methods used to value health states and other benefits stated?	Yes	-
13	Were the details of the subjects from whom valuations were obtained given?	Yes	-
14	Were productivity changes (if included) reported separately?	Yes	-
15	Was the relevance of productivity changes to the study question discussed?	Yes	-
16	Were quantities of resources reported separately from their unit cost?	Yes	-
17	Were the methods for the estimation of quantities and unit costs described?	Yes	-
18	Were currency and price data recorded?	Yes	-
19	Were details of price adjustments for inflation or currency conversion given?	Yes	-

20	Were details of any model used given?	Yes	-
21	Was there a justification for the choice of model used and the key parameters on which it was based?	Partly	The company provided justification for using a discrete event simulation structure, however there was insufficient transparency in the original discretely integrated condition event simulation framework to allow validation.
22	Was the time horizon of cost and benefits stated?	Yes	-
23	Was the discount rate stated?	Yes	-
24	Was the choice of rate justified?	No	The company used a non-reference case discount rate.
25	Was an explanation given if cost or benefits were not discounted?	N/A	-
26	Were the details of statistical test(s) and confidence intervals given for stochastic data?	Yes	-
27	Was the approach to sensitivity analysis described?	Yes	-
28	Was the choice of variables for sensitivity analysis justified?	Partly	Many variables were tested in sensitivity analyses, however some were implemented incorrectly (% transplant success) and some were not included despite being a driver of cost effectiveness (discount rate).
29	Were the ranges over which the parameters were varied stated?	Yes	-

30	Were relevant alternatives compared? (That is, were appropriate comparisons made when conducting the incremental analysis?)	Yes	-
31	Was an incremental analysis reported?	Yes	-
32	Were major outcomes presented in a disaggregated as well as aggregated form?	Yes	-
33	Was the answer to the study question given?	Yes	-
34	Did conclusions follow from the data reported?	Yes	-
35	Were conclusions accompanied by the appropriate caveats?	No	-
36	Were generalisability issues addressed?	No	There were many generalisability issues not addressed by the company.

10.2 Appendix 2: Critique of the company's search strategies for cost-effectiveness evidence

The company confirmed in their response to the points for clarification that the search strategies included in these two reports by Evidera were used to identify evidence for the SLRs included within the submission. The 2018 report by Evidera contained the description of the searches and full search strategies in Section 3.1.1, p. 14-15.[19] In the 2019 report by Evidera, the original searches were updated, with a description of the searches on p. 5 and the full search strategies included in Appendix D, p. 111-115.[17]

The following databases were searched on 19th May 2017: MEDLINE (via PubMed.com and Embase.com) and Embase (via Embase.com). Retrieval was limited to English language studies with an abstract, published from 2007 to 31st December 2017. The searches of MEDLINE and Embase were updated on 1st April 2019 along with searches of the following additional databases: EconLit

(via Ovid) and PsycINFO (via EbscoHost). All additional databases were searched from 2007 to 1st April 2019. Retrieval of records from MEDLINE, Embase and PsycINFO was limited to English language publications.

Specific conferences taking place from 2015 onwards were searched via Embase.com to identify relevant conference abstracts or posters: American Society of Hematology (ASH), European Hematology Association (EHA), British Blood Transfusion Society (BBTS), European Society for Blood and Marrow Transplantation (EBMT) and the International Society of Blood Transfusion. In addition, the following online conference websites were searched: ASH (2018), EHA (2018) and the International Society of Blood Transfusion (2018). Further unpublished studies or grey literature were identified through searches of the HTA database and the Cost Effectiveness Analysis (CEA) Registry.

The search strategies for all databases (MEDLINE, Embase, PsycINFO and EconLit) were structured appropriately using terms for thalassaemia combined with terms for either transfusion or iron chelation therapies. The term thalassemia major has been included within the intervention terms in most of the search strategies. This appears to be a mistake, however it would not have caused relevant studies to be missed. A lack of truncation was noted throughout all database strategies for some search terms – thalassemia, thalassaemia, anemia and anaemia. These terms could have been truncated as follows: thalassemi$, thalassaemi$, anemi$, anaemi$, to allow maximal retrieval of relevant records that use the same word stem but have different endings eg: thalassaemic, thalassaemias, anaemic, anaemias etc.

Subject headings for iron chelation therapies and blood transfusion were missing from the MEDLINE and Embase search strategies. It is usual practice for systematic review searches to include both textword searches of the title and abstract fields as well as relevant subject headings to ensure all relevant studies are retrieved.

The search strategies for EMBASE were limited by publication type to records that have been assigned as articles or articles in press eg: line #7, table 2, p. 14-15 in the 2019 report by Evidera. [17] This may have omitted erratums or corrections to published articles as well as other publication types such as book chapters, short surveys, reviews and conference papers.

Some minor reporting errors were found by the ERG relating to the searches. The PRISMA flow diagram (figure 1, p. 9, 2019 report by Evidera) had typing errors in the first box – the 2017 searches of PubMed retrieved 2398 results and the 2017 searches of Embase retrieved 2372 results.[17] In the company submission report, the PRISMA flow diagrams for the original review of economic evaluation and cost and resource use studies (Figure 1, Appendix G, p. 2) and the diagram for the

original review of HRQoL studies (Figure 1, Appendix H, p. 1) had typing errors in the first box – the 2017 searches of PubMed retrieved 2398 results and the 2017 searches of Embase retrieved 2372 results. In Appendix G of the company submission on p. 1, the year of initial search is reported incorrectly – this should be 19th May 2017.