NICE National Institute for Health and Care Excellence

Single Technology Appraisal

Gemtuzumab ozogamicin for untreated acute myeloid leukaemia [ID982]

Committee Papers

NICE National Institute for Health and Care Excellence

NATIONAL INSTITUTE FOR HEALTH AND CARE EXCELLENCE

SINGLE TECHNOLOGY APPRAISAL

Gemtuzumab ozogamicin for untreated acute myeloid leukaemia [ID982]

Final Scope and Final Matrix of consultees and commentators (updated March 2018)

Contents:

Any information supplied to NICE which has been marked as confidential, has been redacted. All personal information has also been redacted.

Pre-meeting briefing

Gemtuzumab ozogamicin for untreated acute myeloid leukaemia (AML)

This slide set is the pre-meeting briefing for this appraisal. It has been prepared by the technical team with input from the committee lead team and the committee chair. It is sent to the appraisal committee before the committee meeting as part of the committee papers. It summarises:

- the key evidence and views submitted by the company, the consultees and their nominated clinical experts and patient experts and
- the Evidence Review Group (ERG) report

It highlights key issues for discussion at the first appraisal committee meeting and should be read with the full supporting documents for this appraisal

Please note that this document includes information from the ERG before the company has checked the ERG report for factual inaccuracies

The lead team may use, or amend, some of these slides for their presentation at the Committee meeting

1

Abbreviation	In full
AE	Adverse effect
AML	Acute myeloid leukaemia
CHMP	Committee for Medicinal Products for Human Use
CI	Confidence intervals
CR	Complete remission
CRp	CR with incomplete platelet recovery
DA	Daunorubicin and cytarabine
DFS	Disease-free survival
ECOG	Eastern Cooperative Oncology Group
EFS	Event-free survival
EQ-5D	EuroQol five dimensions questionnaire
ERG	Evidence review group
GO	Gemtuzumab Ozogamicin
GVHD	Graft versus host disease
HR	Hazard ratio
HRQoL	Health-related quality of life
HSCT	Haematopoietic stem cell transplantation
ICER	Incremental cost-effectiveness ratio
IPD	individual patient data
KM	Kaplan-Meier
MCM	mixture cure models
mITT	Modified Intention to treat
OS	Overall survival
QALY	Quality-adjusted life year

2

Abbreviation	In full
QOL	Quality of life
RFS	Relapse-free survival
SAE	Serious adverse event
SOC	Standard of care
SMR	Standardised mortality ratio
TTO	Time trade off
VOD	Veno-occlusive disease

3

Key Issues - clinical (1)

Decision problem:

- The company's decision problem restricts the use of GO to patients not known to have unfavourable cytogenetics. This differs from the anticipated marketing authorisation, which does not restrict patients based on their cytogenetic profile.
  - Is it appropriate for the company to focus its submission to patients not known to have unfavourable cytogenetics?
  - Would the requirement for cytogenetic test results to be available before the start of treatment with GO potentially delay the start of treatment?

Clinical Pathway:

- Is it routine practice in the NHS in England to undertake cytogenetic testing before the start of treatment?
- How long does it take for cytogenetic results to be reported in routine clinical practice?
- Is molecular testing undertaken routinely? If so in what circumstances?

4

Key Issues – clinical (2)

ALFA-0701 trial:

- The dose used in the trial is different to that being used in clinical practice through the AML18 and AML19 protocols. Does this have any implications for dosing in clinical practice?
- The KM curves for EFS and OS plateau at around 36 to 48 months, suggesting no further relapse, events or mortality after around 3 to 4 years. Therefore the company claims that patients are functionally cured if there are no events within three years of treatment response. Is such an assumption appropriate?
- Patients with an unfavourable cytogenetic profile, OS and RFS outcomes appeared to be worse in the GO + DA treatment arm, compared with the DA treatment arm. How robust are these data?
- Subgroup analyses by molecular risk profile showed that the benefit seen in patients with an intermediate-1 cytogenetic and molecular risk profile was not found in patients with an intermediate-2 cytogenetic and molecular risk profile. How important is this heterogeneity in the broader 'intermediate' cytogenetic subgroup?

IPD meta-analysis

- How generalisable are the results from the IPD meta-analysis (presented as supporting evidence) to patients eligible for GO + DA in clinical practice in England?

Key issues – cost effectiveness

- Given the complexity of the company's state transition model, is it appropriate for decision-making ?
- Is the absence of an explicit structural link between relapse and HSCT appropriate given the cost-offsets assumed by the company for HSCT?
- Should the initial treatment costs of the induction and consolidation therapies be based on the IA response outcomes or on an adjustment of the IRC response outcomes as proposed by the company?
- Is it appropriate to pool response data?
- What is the most appropriate HR for long-term morbidity and survival for functionally cured patients?
- Is it appropriate to assume that functionally cured patients experience the same HRQoL as the general population?
- Are the costs for HSCT included in the company's model appropriate?
- Is it appropriate to include patients with VOD in the DA treatment arm in the model?
- Has the additional inpatient treatment costs associated with VOD been adequately captured in the company's model?

6

Key issues - cost effectiveness and other

Cost effectiveness

- What are the implications of the heterogeneity in the subgroup of patients with unknown cytogenetics and within the intermediate population on the most plausible ICER?
- What is the most plausible ICER?

Other

- Is gentuzumab ozogamicin an innovative treatment?
- Are there any equality issues?

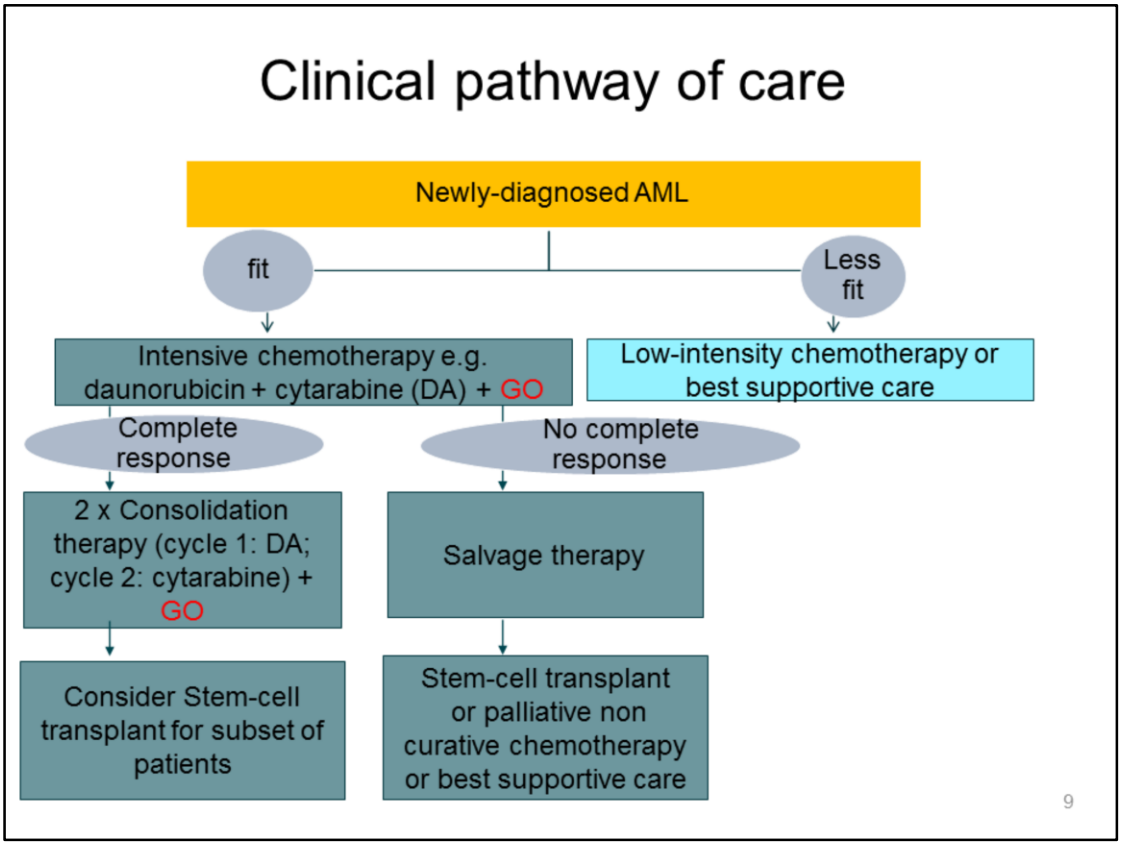
7

Disease Background

- Acute myeloid leukaemia has one of the lowest survival rates among leukaemias.
- The incidence of acute myeloid leukaemia in England is about 3,000 people per year.
- Around 55% of all cases occur in people over 70 years.
- There were 2471 new diagnoses of AML in England.
- AML is primarily a disease in older people, with incidence rising gradually from 40–44 years of age and then more steeply from 55–69 years of age.
- In 2014, there were around 2,516 deaths from AML in the UK.

Source: company submission pages 18-25

8



Adapted from company submission, Figure 4, page 31.

Related NICE Guidance

Published

TA399

Azacitidine is not recommended, within its marketing authorisation, for treating acute myeloid leukaemia with more than 30% bone marrow blasts in people of 65 years or older who are not eligible for haematopoietic stem cell transplant.

TA218

Azacitidine is recommended as a treatment option for adults who are not eligible for haematopoietic stem cell transplantation and have acute myeloid leukaemia with 20–30% blasts and multilineage dysplasia, according to the World Health Organization classification, and if the manufacturer provides azacitidine with the discount agreed as part of the patient access scheme.

In development

ID894

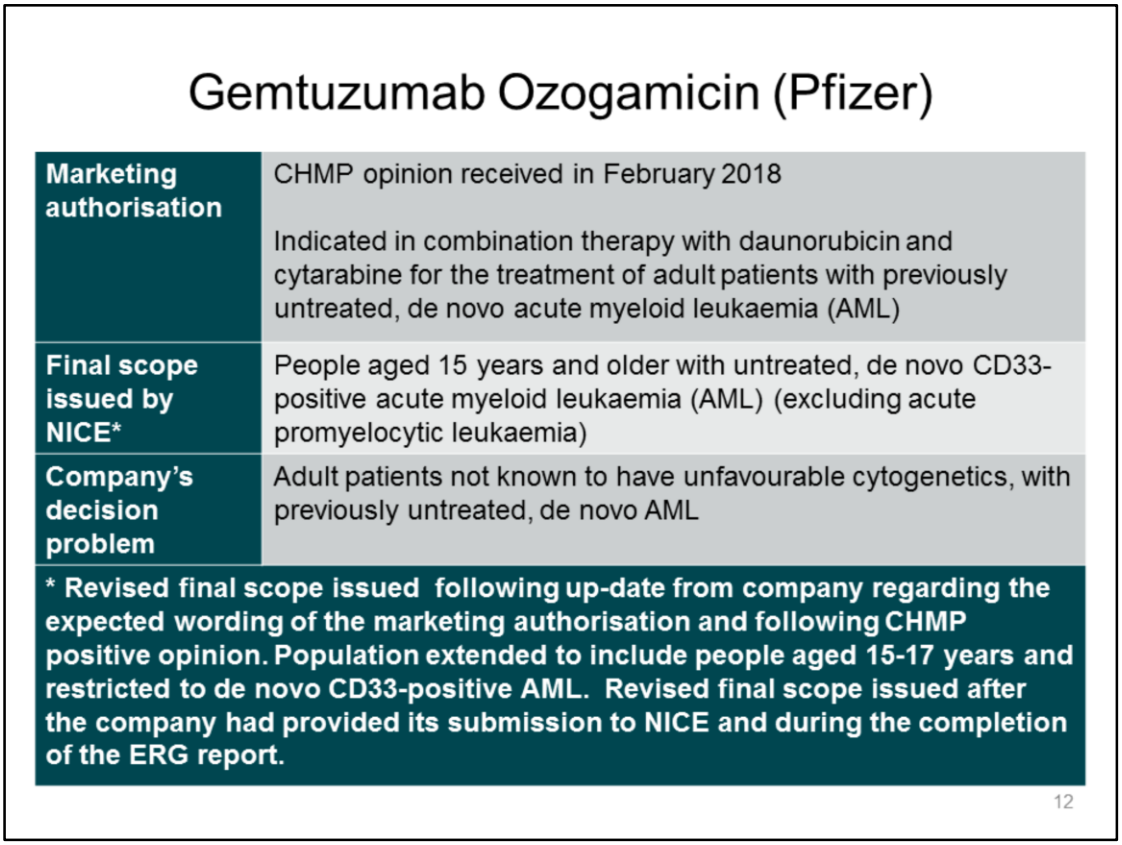
Midostaurin is being considered as a treatment option for adult patients with newly diagnosed acute myeloid leukaemia (AML) who are FLT3 mutation positive.

10

Comments from patient and professional groups

- Patient groups
  - AML has extremely poor outcomes and high unmet need.
  - AML places huge emotional strain on patients, families and carers.
  - The most common symptoms encountered by patients since their diagnosis are fatigue, infections, weakness or breathlessness, bruising or bleeding.
  - AML has a practical impact on people, causing pain, difficulty moving around and performing daily routines.
- Professional groups
  - Gemtuzumab ozogamicin appears to improve overall survival when added to induction chemotherapy for patients with favourable and intermediate risk disease karyotype.
  - There is need to access cytogenetic results very promptly at diagnosis before starting treatment, which is not a standard practice.
  - Gemtuzumab ozogamicin would be added to standard induction chemotherapy in newly-diagnosed AML.
  - Highly innovative, it will be the first routine application of antibody-directed chemotherapy in the treatment of AML.

11



The price of gemtuzumab ozogamicin is commercial-in-confidence. For the list price and average cost of a course of treatment, see Table 2, page 17 of company submission.

ERG's comments: Population in company's decision problem

- Population addressed in the company's decision problem is a subpopulation of the anticipated marketing authorisation for gemtuzumab ozogamicin and the final scope issued by NICE.
- Company's decision problem excludes patients known to have unfavourable cytogenetics as they would not receive treatment with gemtuzumab ozogamacin plus intensive chemotherapy in NHS clinical practice.
  - The clinical advisor to the ERG supported the company's rationale for excluding patients with unfavourable cytogenetics.
  - In view of the very short timeframe between diagnosis and treatment in patients with AML, the requirement for cytogenetic test results before treatment could potentially delay the start of treatment with gemtuzumab ozogamicin.
- The restriction to CD33-positive AML in the CHMP positive opinion is a narrower population than that addressed in the company's decision problem.

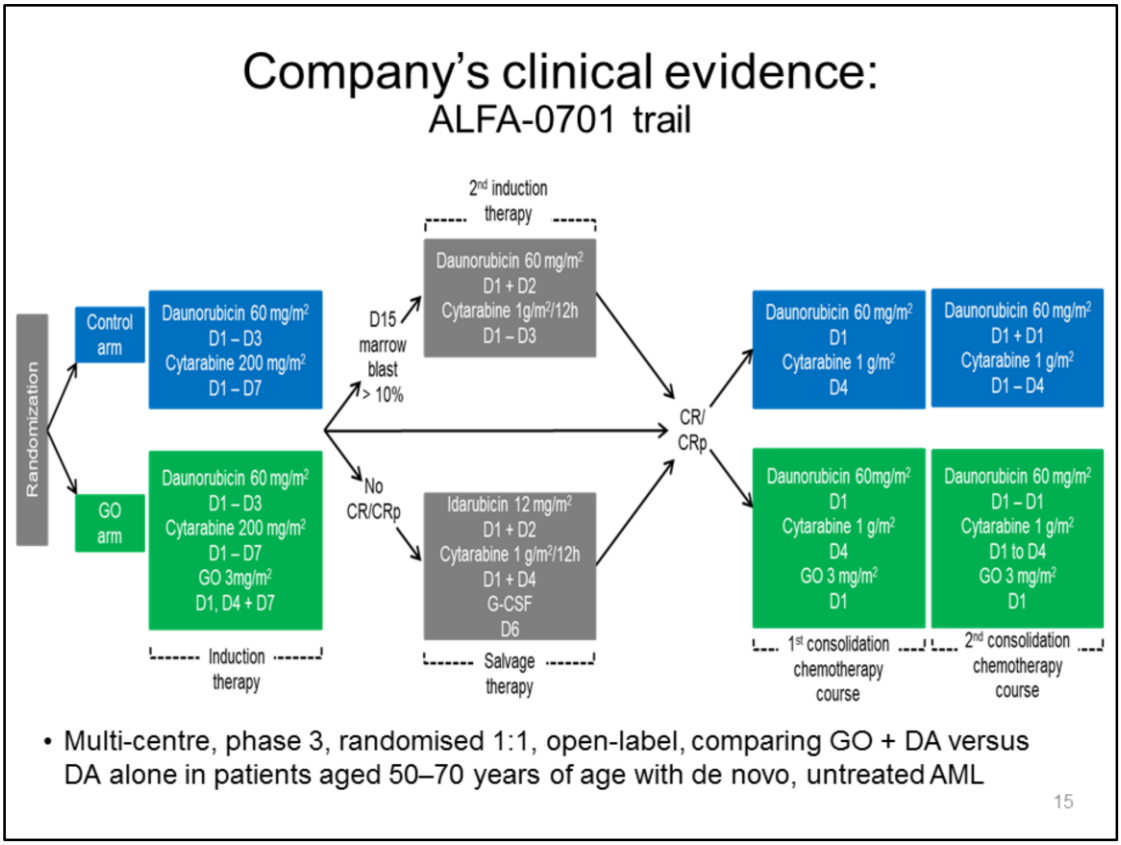
Source: ERG report, page 25

13

Company's evidence of clinical effectiveness

Evidence	Source	Used in clinical effectiveness	Used in cost effectiveness
ALFA-0701 trial	CS pages 38-79; ERG report pages 30-45	main evidence	Yes
Additional analysis by IRC assessment for EFS, RFS, OS and molecular status	Company's response to clarification; ERG report pages 44-47	Yes	Yes
IPD meta-analysis (uses evidence from ALFA-0701 and 4 other trials).	CS page 74 and, also Appendix D.3.1; ERG report pages.54-60	supportive evidence	No (different and unlicensed dosing regimens in the other trials)
Published meta-analysis	CS (Appendix D.3.2); ERG report ages 60-61	supportive evidence	No

14



Source: Company submission Figure 5, page 47.

- For further details of the trial methodology, see pages 42 to 47 of the company submission.
- The company provided results according to the IRC assessment and the assessment made by the original investigators (IA). All efficacy outcomes were pre-specified. Outcomes according to IRC and investigator analyses were available at the 1 August 2011 and 30 April 2013 data cut-offs.
- The results presented in the clinical effectiveness section of this pre-meeting document focus on the IRC assessment as the ERG considers these results to be less likely to be influenced by lack of blinding, and were similar, but generally slightly more conservative than those for the IA.

Baseline characteristics in ALFA-0701(1)

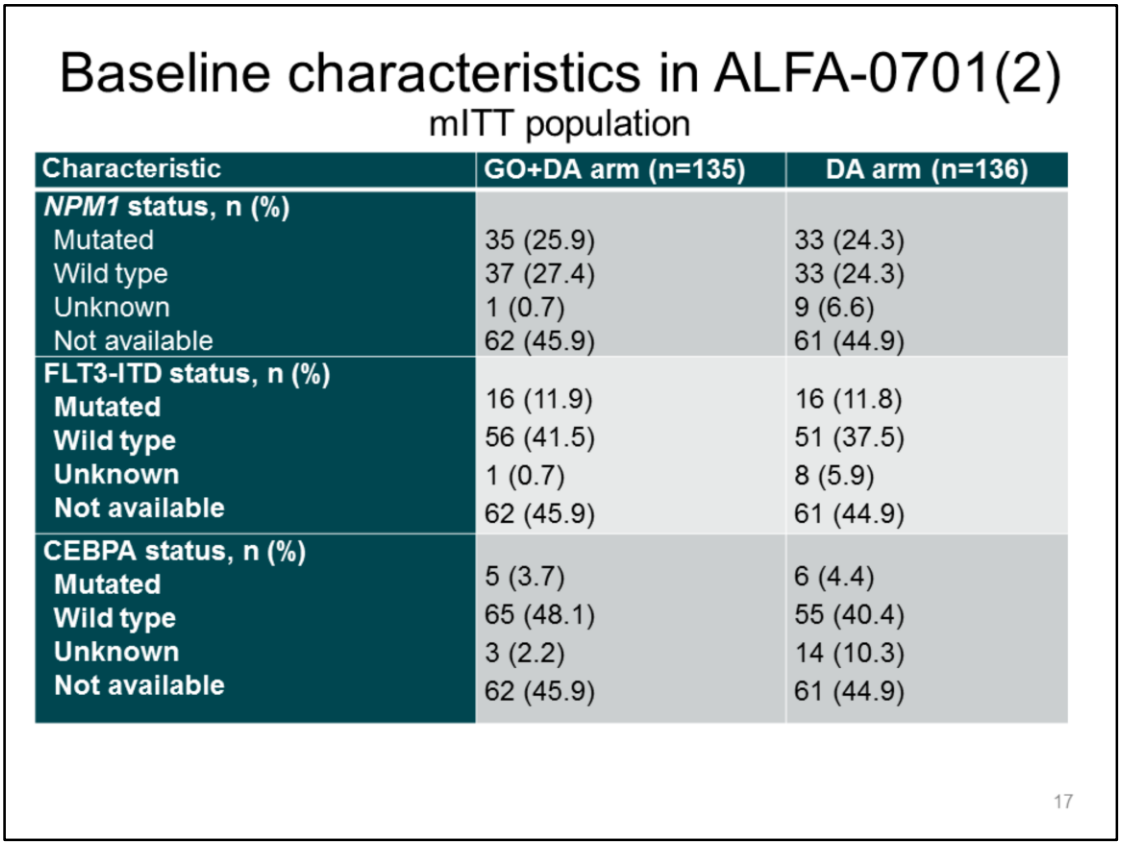
mITT population*

Characteristic	GO+DA (n=135)	DA (n=136)
Age, years, median (range)	62.0 (50–70)	61.0 (50–70)
< 60, n (%)	38 (28.1)	52 (38.2)
≥ 60, n (%)	97 (71.9)	84 (61.8)
Male, n (%)	74 (54.8)	60 (44.1)
ECOG PS, n (%)		
0–1	121 (89.6)	117 (86.0)
≥ 2	14 (10.4)	18 (13.2)
Missing	0 (0.0)	1 (0.7)
WBC count,×109/L, median (IQR)	5.8 (0.5–151.0)	4.1 (0.1-180.5)
Cytogenetics, n (%)		
Favourable	3 (2.2)	6 (4.4)
Intermediate	91 (67.4)	89 (65.4)
Unfavourable	27 (20.0)	30 (22.1)
Not available	14 (10.4)	11 (8.1)
CD33 expression, positivity		
< 30%	17 (12.6)	20 (14.7)
≥ 30%	83 (61.5)	74 (54.4)
< 70%	37 (27.4)	31 (22.8)
≥ 70%	63 (46.7)	63 (46.3)

16

Source: ERG report Table 3, page 38.

*Originally the trial included 280 patients, however informed consent forms were not transferred for nine patients when data were transferred to Pfizer. The data were therefore analysed for 271 patients in a modified intention-to-treat (mITT) population (see section 4.2.2.2, page 36 company submission).



Baseline characteristics in ALFA-0701(2)

mITT population

Characteristic	GO+DA arm (n=135)	DA arm (n=136)
***NPM1* status, n (%)**		
Mutated	35 (25.9)	33 (24.3)
Wild type	37 (27.4)	33 (24.3)
Unknown	1 (0.7)	9 (6.6)
Not available	62 (45.9)	61 (44.9)
FLT3-ITD status, n (%)		
Mutated	16 (11.9)	16 (11.8)
Wild type	56 (41.5)	51 (37.5)
Unknown	1 (0.7)	8 (5.9)
Not available	62 (45.9)	61 (44.9)
CEBPA status, n (%)		
Mutated	5 (3.7)	6 (4.4)
Wild type	65 (48.1)	55 (40.4)
Unknown	3 (2.2)	14 (10.3)
Not available	62 (45.9)	61 (44.9)

17

Adapted: from Table 71, page 77 and 78 of company's appendices.

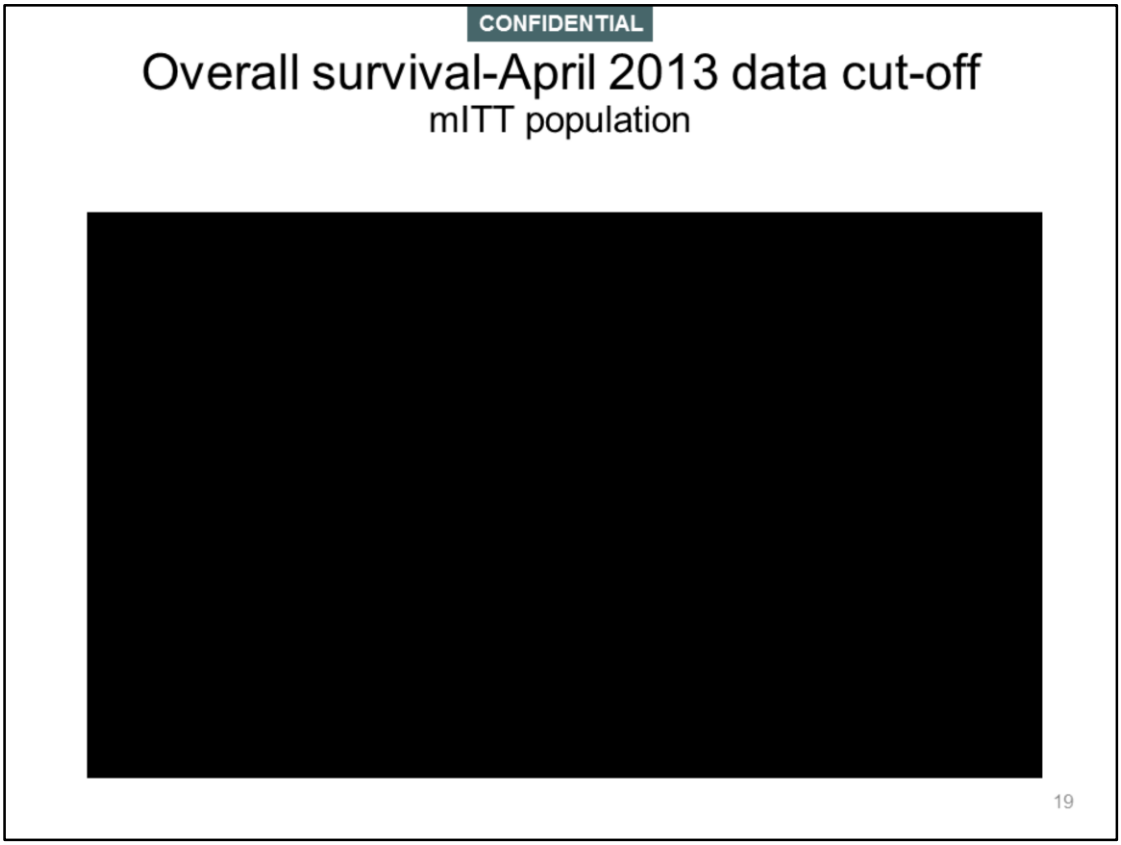
ERG's comments: ALFA-0701 trial

Area	ERG's comments
Study population	The anticipated marketing authorisation for GO specifies patients with CD33-positive AML, whilst the trial also included AML patients who were not CD33-positive, therefore, a small proportion of patients in the trial would not be eligible for GO + DA in clinical practice.
	The trial included only patients aged 50-70 years, whilst the anticipated marketing authorisation includes patients age 15 years and above; however, the majority of patients diagnosed with AML are over 50 years of age, therefore the population included in the trial is likely to be reflective of the majority of patients eligible for GO in clinical practice.
Outcome	• HRQoL was not assessed in the ALFA-0701 trial. • Patients may relapse later than 5 years, longer term events may not have been captured in the ALFA-0701 trial data. • Cytogenetic test results could potentially delay start of the treatment.
Dosing schedule	Dosing schedule in the ALFA-0701 trial is in line with the anticipated marketing authorisation, however two ongoing trials which include UK treatment centres use different dosing regime.

Overall the trial was well conducted and has a low risk of bias, up to the limits of its open-label design.

18

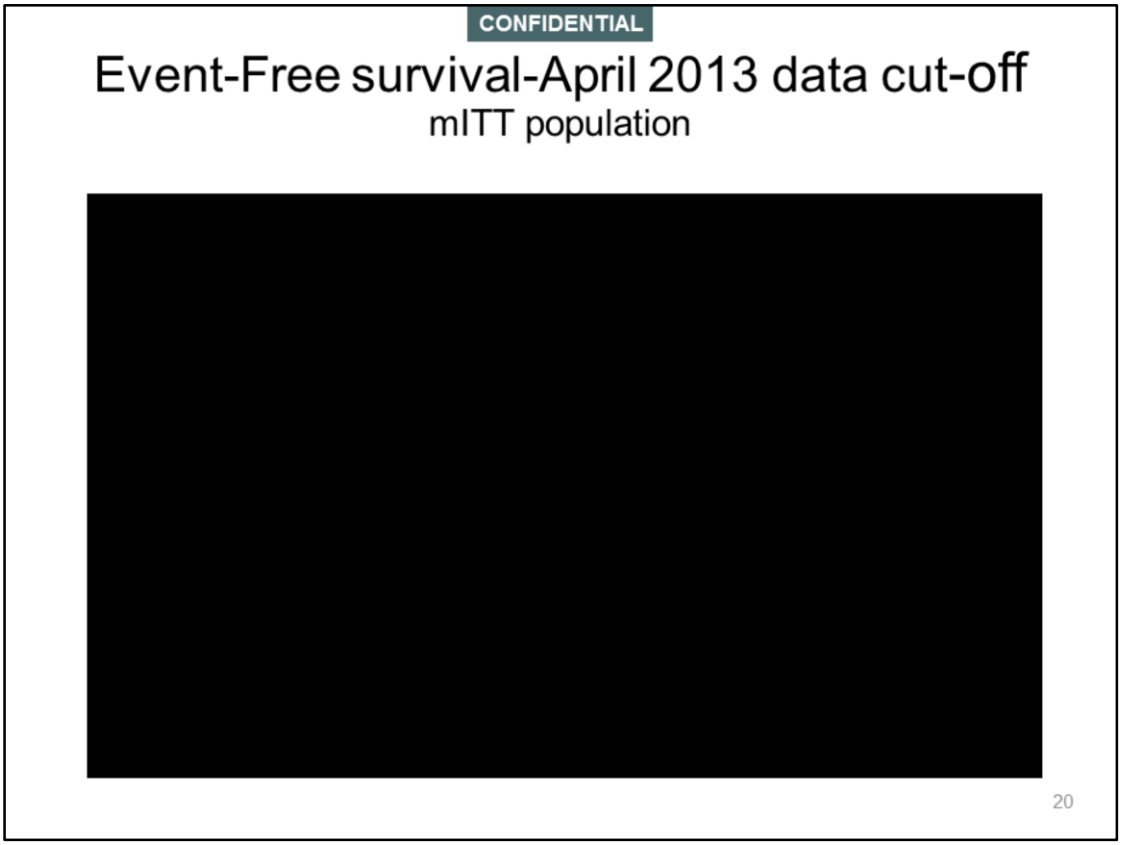
Adapted from ERG report, section 3.



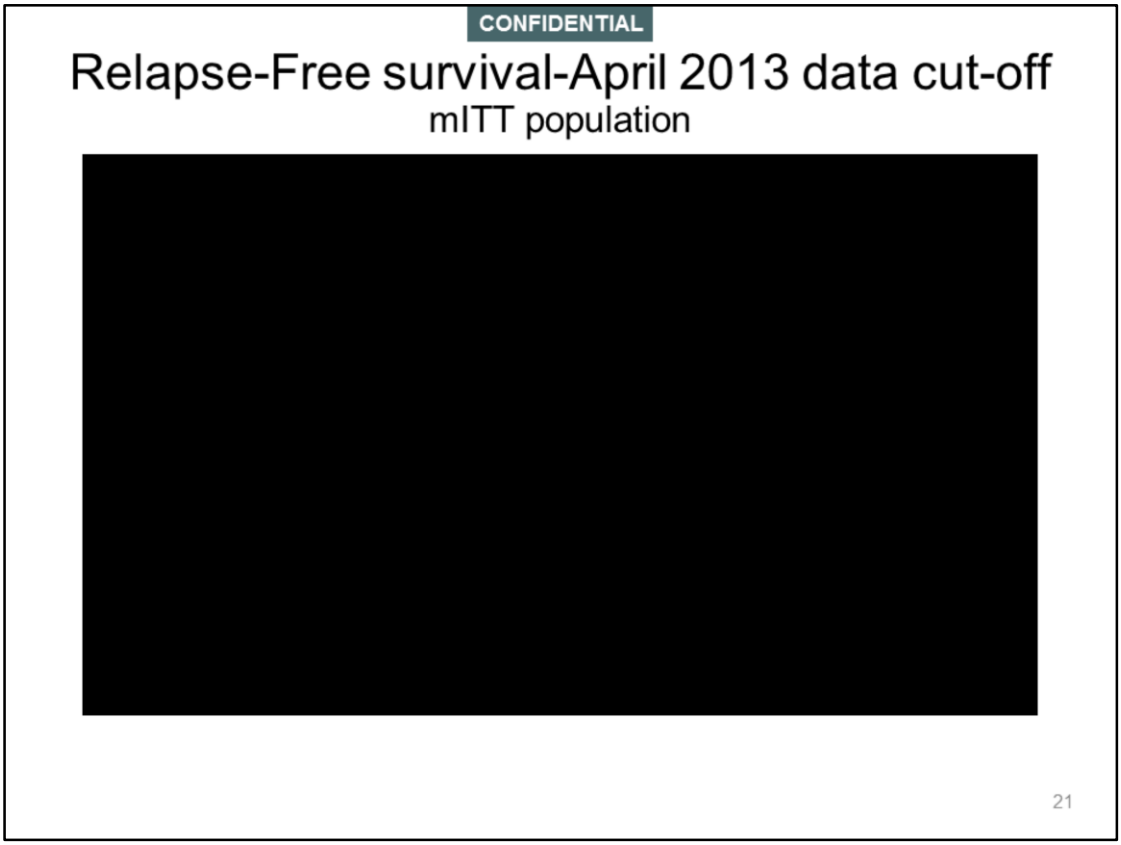
Source: Company submission, Figure 10, page 71 or ERG report Figure 4, page 41.

ERG's comments:

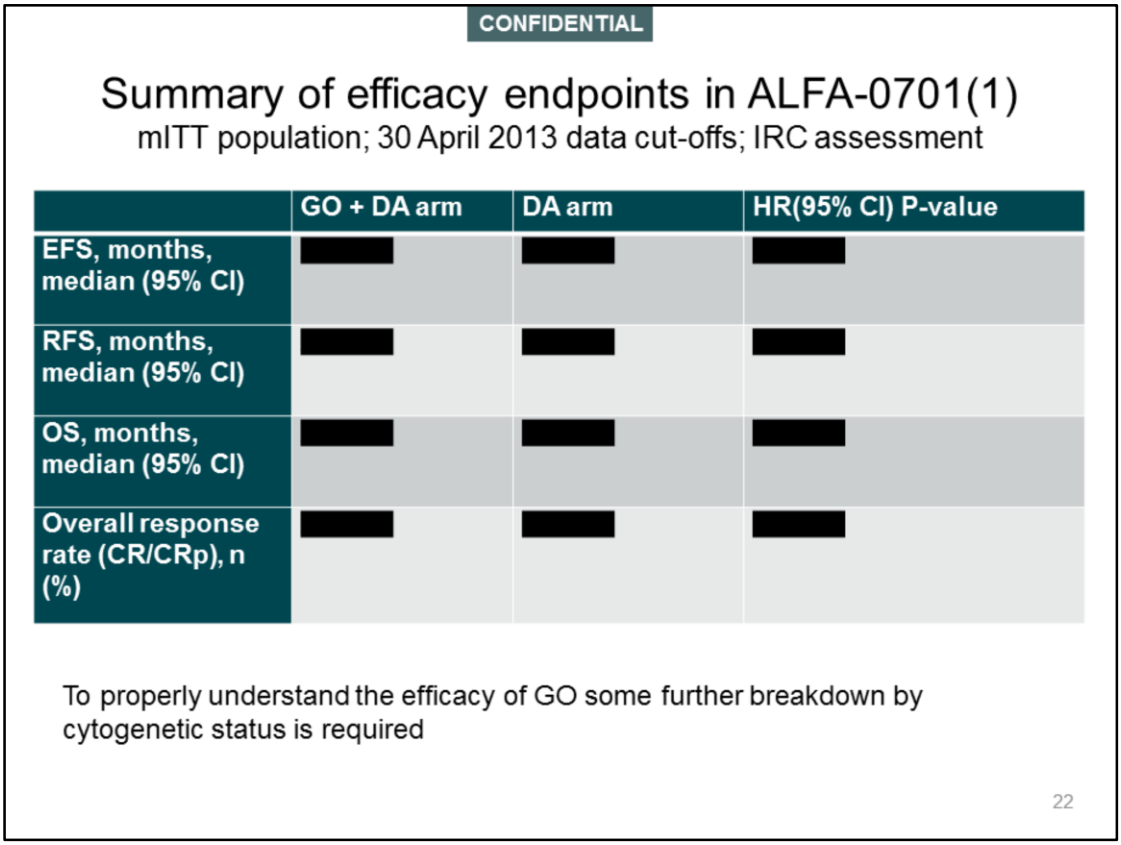
- Overall survival and response rate appeared better in the GO + DA treatment arm, these results did not reach statistical significance. This suggests that duration of remission is extended with GO, rather than proportion of patients achieving a remission.
- Both KM curves plateau, suggesting no further relapse, events or mortality after around 3 to 4 years. This supports the company's claim that patients are functionally cured if there are no events within three years of treatment response. The ERG noted that there were few patients with follow-up extending beyond 3 years.



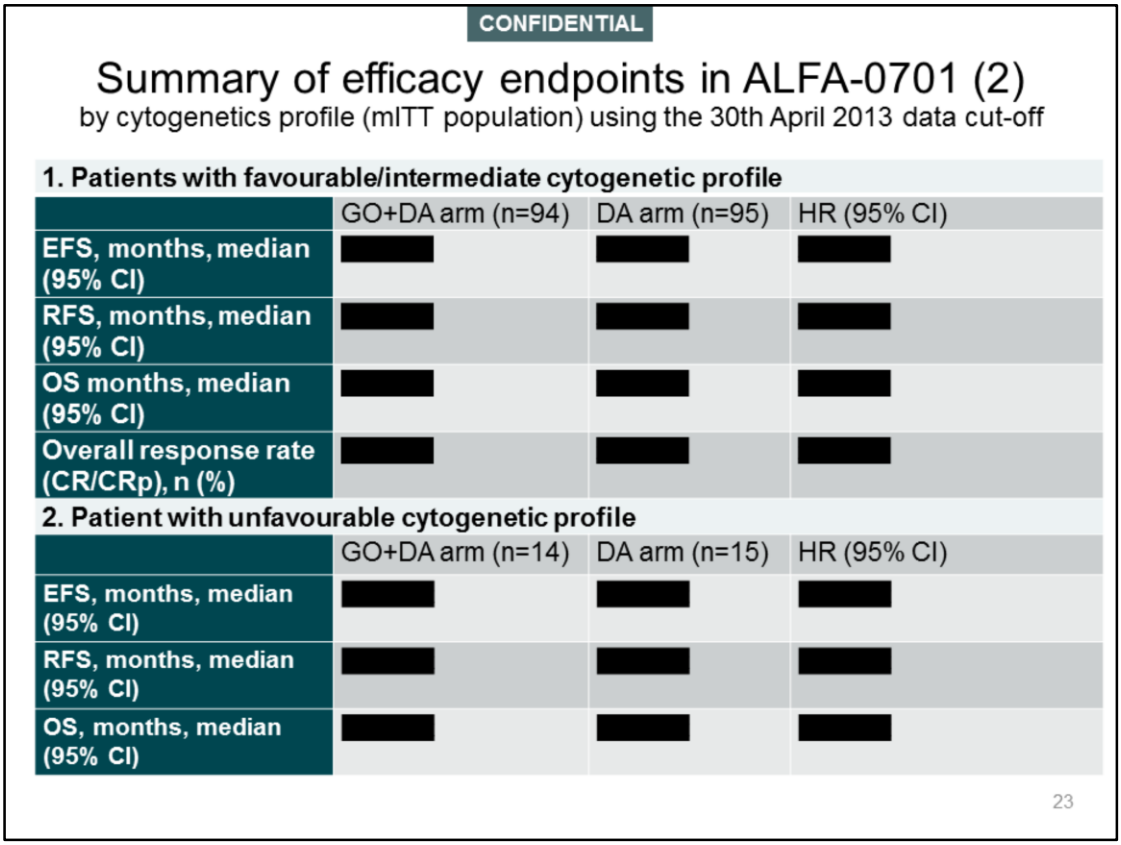
Source: Company submission Figure 6, page 63 or ERG report Figure 3, page 40.



Source: Company submission, Figure 8, page 67.



Adapted from company submission Table 13 page 61 or ERG report Table 4 page 40.



CONFIDENTIAL

Summary of efficacy endpoints in ALFA-0701 (2)

by cytogenetics profile (mITT population) using the 30th April 2013 data cut-off

1. Patients with favourable/intermediate cytogenetic profile

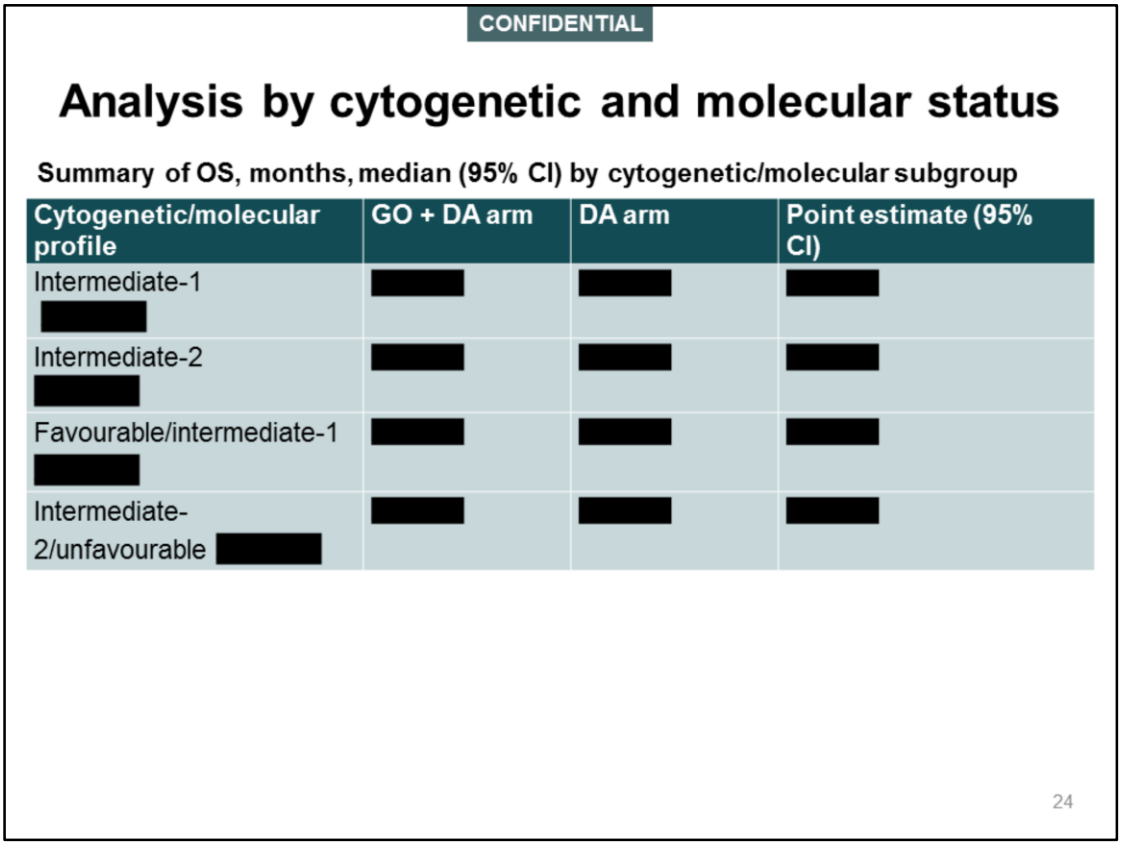
	GO+DA arm (n=94)	DA arm (n=95)	HR (95% CI)
EFS, months, median (95% CI)	[redacted]	[redacted]	[redacted]
RFS, months, median (95% CI)	[redacted]	[redacted]	[redacted]
OS months, median (95% CI)	[redacted]	[redacted]	[redacted]
Overall response rate (CR/CRp), n (%)	[redacted]	[redacted]	[redacted]

2. Patient with unfavourable cytogenetic profile

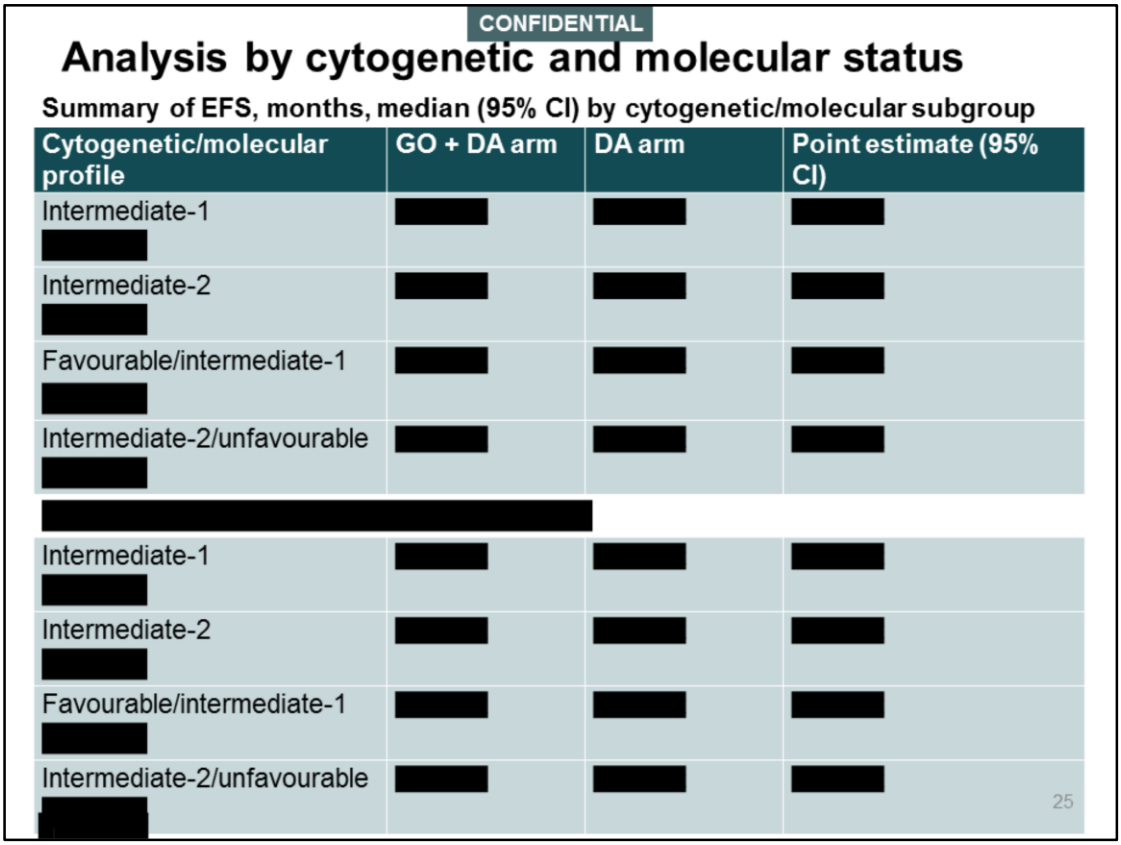
	GO+DA arm (n=14)	DA arm (n=15)	HR (95% CI)
EFS, months, median (95% CI)	[redacted]	[redacted]	[redacted]
RFS, months, median (95% CI)	[redacted]	[redacted]	[redacted]
OS, months, median (95% CI)	[redacted]	[redacted]	[redacted]

23

Adapted from ERG report Tables 5 and 6 pages 41 and 42.



Source: ERG report, Table 12, page 44.



Source: ERG report Tables 10 and 11, page 44.

CONFIDENTIAL

Summary of AEs and SAEs

	GO + DA, (N = 131) n (%)		DA, (N = 137) n (%)	
	All-causality AEs	Related AEs	All-causality AEs	Related AEs
Patients with AEs	[redacted]	[redacted]	[redacted]	[redacted]
Patients with SAEs	[redacted]	[redacted]	[redacted]	[redacted]
Patients with grade 3 or 4 or severe infection AEs	[redacted]	[redacted]	[redacted]	[redacted]
Patients with fatal events	[redacted]	[redacted]	[redacted]	[redacted]
Patients who permanently discontinued study treatment owing to AEs	[redacted]	[redacted]	[redacted]	[redacted]

Veno-occlusive disease (VOD)	GO+DA, (N=131), n (%)	DA, (N=137), n (%)
Proportion of patients with VOD	[redacted]	[redacted]

26

Source:

- **Summary of AEs and SAEs, company submission Table 25 page 78.**
- **Veno-occlusive disease (VOD), company submission Table 26 page 80.**

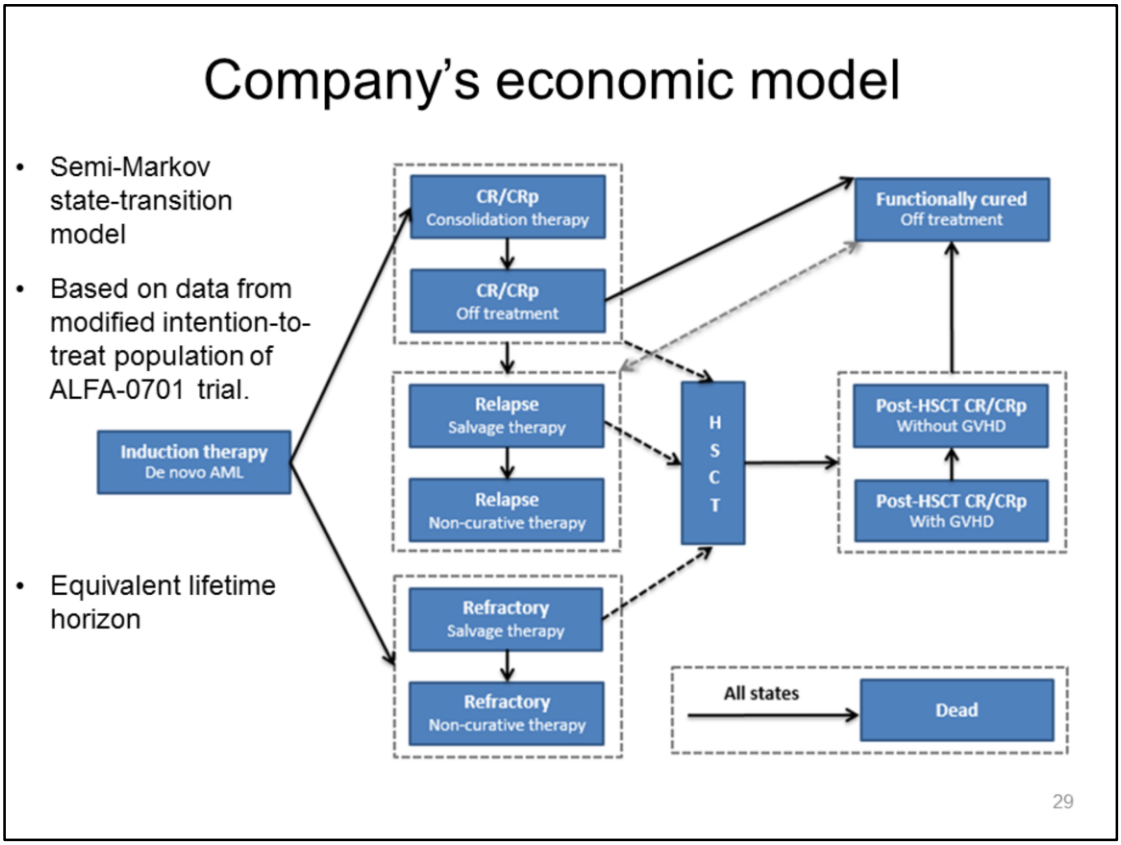
ERG's additional comments on company's clinical evidence

Company subgroup analysis by cytogenetic status
Results were better for the subgroup of patients with favourable/intermediate cytogenetic risk, than the overall population. However, for patients with unfavourable cytogenetics, outcomes appeared to be worse in the GO + DA treatment arm, compared with the DA treatment arm.
Additional subgroup analysis by cytogenetic and molecular status requested by ERG
Overall there was evidence that GO improves survival in favourable and intermediate-1 patients, but not in intermediate-2 and unfavourable patients, suggesting potentially important heterogeneity in the broader 'intermediate' cytogenetic subgroup.
The IPD meta-analysis
Included patients aged 15 years or older with newly diagnosed AML (either de novo or secondary), or high-risk myelodysplastic syndrome (MDS), which is a broader population than that defined in the decision problem or the anticipated marketing authorisation. Therefore, the results may not be entirely generalisable to patients eligible for GO+DA in clinical practice.

Adapted from ERG report.

Cost effectiveness

28



Adapted from Figure 11 p. 101 company submission.

Company's comments:

- ALFA-0701 is the pivotal trial on which regulatory submissions have been based. ALFA-0701 uses a dose and dosing schedule that is consistent with the expected EMA marketing authorisation (see section B.1.3.4. page 32 of company submission).

ERG's comments:

- There is a lack of an explicit structural link between a number of key model parameters, most importantly between relapse and HSCT. The absence of a structural link restricts the ability of the model to explore alternative scenarios in an appropriate manner, and therefore, to fully capture the uncertainty in the modelled results (see section 1.6.2, page 19 of ERG report).
- The proposed model structure is complex and was challenging to critique given the difficulties in determining the actual flow of patients through the model (see section 5.2.1, page 69 of ERG report).

Company's model: Summary (1)

Input	Source/assumption
Population	The base-case is based on a subgroup of the population in the positive opinion given by the CHMP - patients not known to have unfavourable cytogenetics. Results were also presented separately for the entire licensed population,
Intervention/ comparator	GO in combination with standard intensive chemotherapy, consisting of DA , compared with DA alone.
Treatment effectiveness	Clinical outcomes included were response (CR/CRp), RFS and OS, cure fraction, probability of HSCT, post HSCT survival. Patients alive or relapse-free at 60 months post induction therapy or HSCT assumed to be cured and experienced general population mortality adjusted to reflect excess mortality in AML survivors.
	Data taken from ALFA-0701. OS stratified by response status and parametric models fitted to extrapolate beyond the end of the trial follow-up. Parametric models fitted to RFS (CR/CRp only).
	Response and RFS endpoints based on the blinded IRC assessment, RFS and OS based on reference data 30 April 2013.

30

Source: ERG report Table 16, page 62.

ERG's comments on population (pages 73-76 ERG report)

- The licence amendments regarding age and the restriction to CD33-positivity were proposed by EMA after the company submitted to NICE.
- Company additionally included adolescent patients in the decision problem, and stated that majority of patients with de novo AML express CD33, therefore no changes were made to the model. ERG agrees that both assumptions are reasonable and changes made would not have large impact on the ICER.
- ERG agreed with the company's decision to exclude patients with known unfavourable cytogenetics .
- ERG does not consider that the company have adequately explored any remaining heterogeneity within the intermediate population and possible implications for clinical and cost-effectiveness. Further data was requested on the intermediate-1 and intermediate-2 patients.
- ERG considers that the company did not sufficiently justify the inclusion of the subgroup with unknown cytogenetics. Additional analysis showed that excluding this population from cost-effectiveness resulted in increased ICER.

Company's model: Summary (2)

Input	Source/assumption
Adverse Events	Adverse event rates taken from ALFA-0701, Grade 3 and 4 treatment related events that occurred in at least 1% of patients. GVHD as a consequence of HSCT also included. Incidence of GVHD sourced from external literature.
HRQoL	No HRQoL data collected in ALFA-0701. Health state utility values sourced from a systematic literature review: • Functionally cured patients assumed to have QoL equal to that of the aged-matched general population. • Remaining utilities sourced From NICE TA399 • Adverse event disutilities sourced from external literature. Company used an alternative set of utility values in a scenario analysis which were based on a vignette study that it had undertaken.

31

Source: ERG report Table 16, page 63.

CONFIDENTIAL

Company's model: Predicted OS (CR)

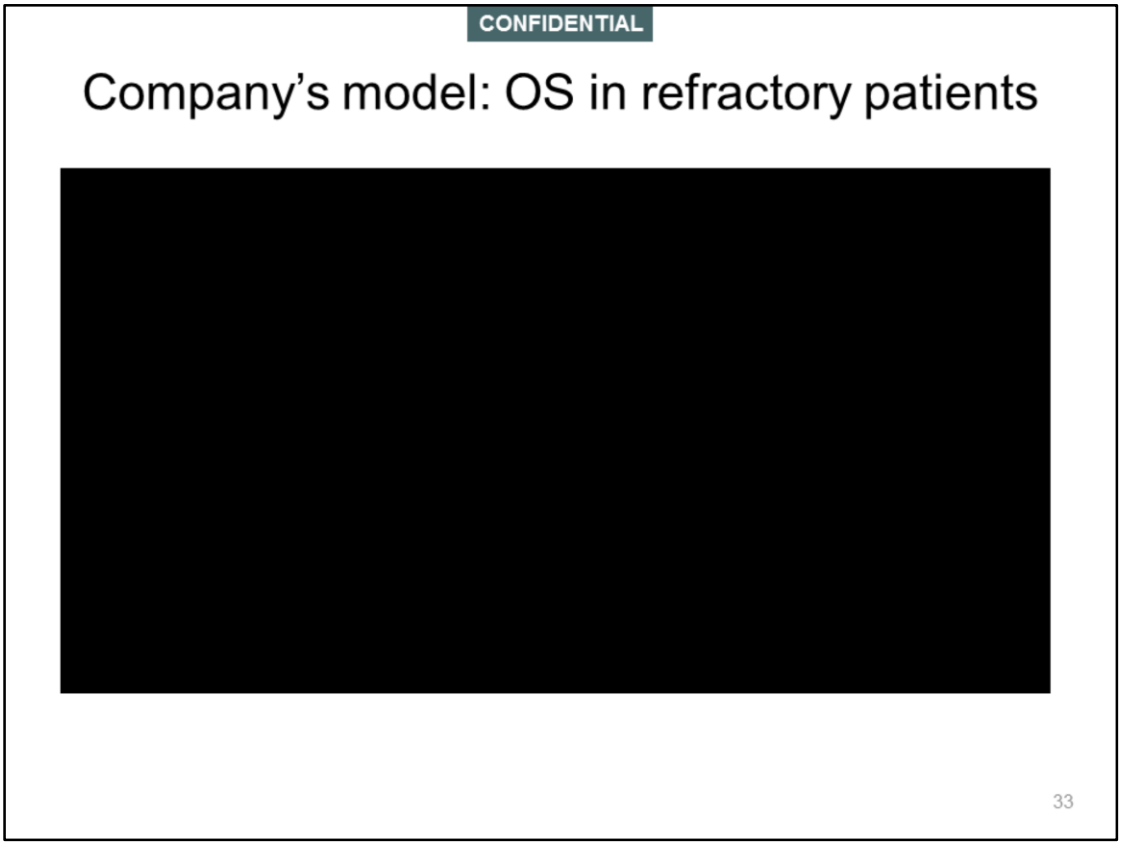
Cure fractions – overall survival in CR/CRp patients		
MCM distribution	GO+DA group	DA group
MCM Weibull	[redacted]	[redacted]
MCM lognormal	[redacted]	[redacted]
MCM generalised gamma	[redacted]	[redacted]

32

Source: ERG report Figure 11, page 84.

ERG's comments:

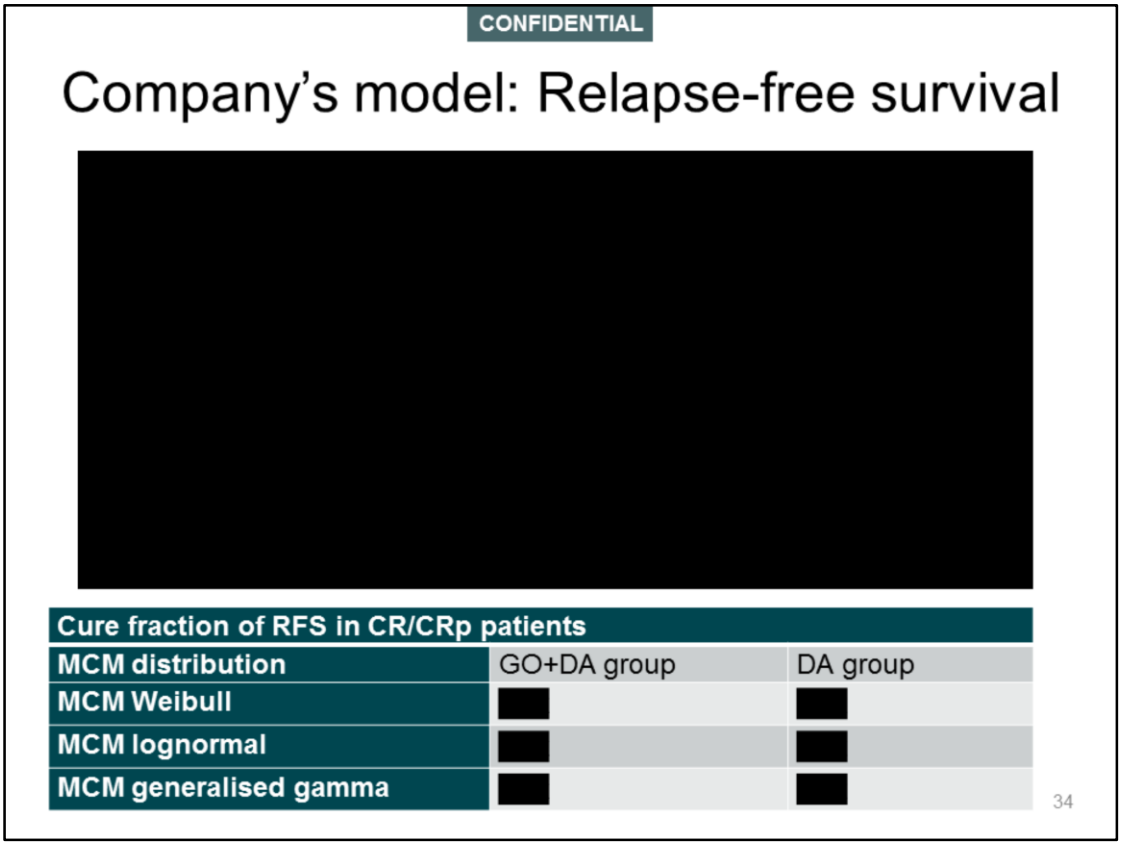
- The curve selected by the company for its base-case analysis, the MCM lognormal curve, had the best fit according to AIC/BIC statistics and provides the most conservative estimates of overall survival.



Source: ERG report Figure 12, page 85.

ERG's comments:

- The curve selected by the company for its base-case analysis, the Gompertz curve, had the best fit according to AIC/BIC statistics, and the company also considered that it had the best visual fit, stating that the spline-based models resulted in late-occurring plateaus.



Source: ERG report Figure 13, page 86.

ERG's comments:

- The company considered that the MCM Weibull and MCM lognormal provided a similar visual fit, but that the lognormal "provides the best fit to the plateau" and was considered by the company to best capture the benefits of GO+DA.

CONFIDENTIAL

Summary of survival functions in company's base-case analysis

End point	GO+DA	DA	Pooled
RFS (CR or CRp)	MCM log-normal Cure rate: [redacted]	MCM log-normal Cure rate: [redacted]	
OS (CR or CRp)	MCM log-normal Cure rate: [redacted]	MCM log-normal Cure rate: [redacted]	
OS (refractory)	-	-	Gompertz Cure rate: n/a

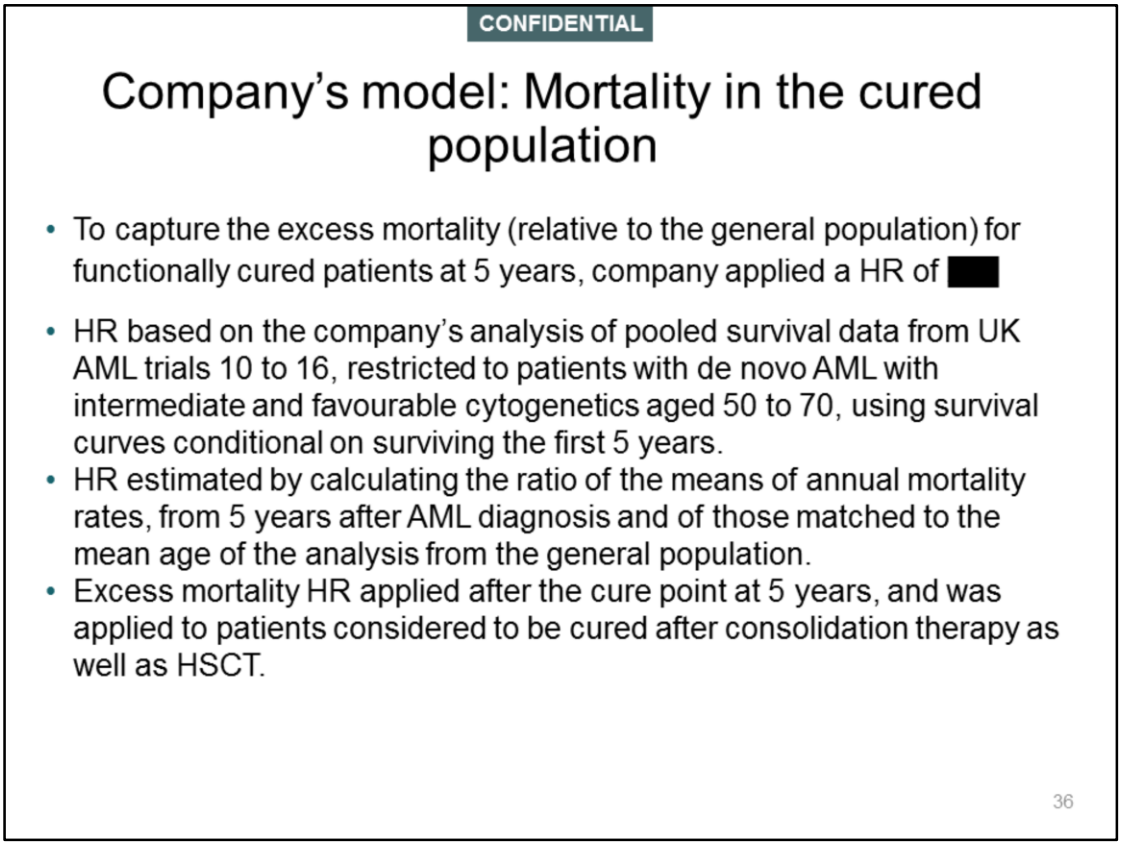
CR, complete remission; CRp: complete remission with incomplete platelet recovery; GO, gemtuzumab ozogamicin; OS, overall survival; RFS, relapse-free survival; MCM , mixed cure model

35

Source: ERG report Table 24, page 83.

ERG's comments:

- Overall, the ERG considered the company's approach to curve fitting and the rationale for selecting distributions to be appropriately justified. Uncertainties surrounding the choice of survival functions were also explored using a range of alternative functions within separate scenarios.
- Although the alternative MCM distributions reported different estimates of the absolute cure fraction for each group, the difference in the cure fraction between the groups was broadly similar for both the MCM lognormal and Weibull functions for both EFS and OS. This is important because it is the difference between the groups in the probability of long-term survival which is the main driver of QALY differences and the ICER estimates.
- For the base-case population, the ERG considers that the choice of survival function appears less critical than the assumptions which are subsequently applied to long-term survivors regarding potential excess morbidity (i.e. HRQoL assumptions) and mortality.



Source: ERG report page 87-89.

ERG's comments:

- ERG was generally satisfied with the manner in which it was implemented.
- However uncertainties remain regarding the estimation of the adjustment factor (HR):
    - The number of patients at risk in the analysis of AML10-16 trial data was not reported and therefore it was difficult to determine how robust the estimates of mortality are in later years. The values may be based on small patient numbers.
    - The HR per cycle appears higher in the years immediately following year-5 before settling into a more consistent pattern. This may indicate that surviving patients are still at risk of AML related relapse and associated mortality, suggesting that 5-years may be too early to establish that patients are functionally cured. Equally the dataset incudes all patients who are still alive at year-5, which may include patients who have relapsed and still at risk of AML related mortality.
    - In some years, the probability of death was higher in the general population than in survivors with AML, which does not seem plausible. The ERG considered that further adjustments appear appropriate, such that the mortality rate was set equal to the general population mortality rate in instances when the observed mortality rate was reported to be lower than the general population.

Company's model and ERG critique: HSCT

Company's model

- Patients were able to receive HSCT from 3 health states: CR/CRp, refractory, and relapsed.
- Probabilities of receiving HSCT estimated from data of patients receiving HSCT in ALFA-0701, excluding those with unfavourable cytogenetics.

ERG's comments

- The company limited the complexity of the model by including additional structural assumptions for the HSCT state and using calendar time (i.e. time from randomisation) rather than time in state (i.e. time from relapse) as well as absolute probabilities at fixed times. These assumptions ensured that the model predicted identical HSCT rates as observed in the trial.
- The main uncertainty was whether the data from the trial was sufficiently mature to provide an accurate estimate of the long-term difference in HSCT rates between the 2 treatment arms. The additional data and KM curves provided in response to clarification suggested no obvious bias or differences in the time at risk.
- However, the ERG noted that some of the cost-off sets for HSCT are predicted on the functional assumption. The absence of any structural link to HSCT rates limited the ERG's ability to further assess the potential impact of this source of uncertainty.

37

CONFIDENTIAL

Company's model: Health-related quality of life

- HRQoL data was not collected in ALFA-0701 so health state utilities were sourced from a systematic literature review.

Utility state	Values used in base–case analysis	Values used in scenario analysis		
Source	TA399	TA399	Pfizer TTO	Pfizer VAS
Chemotherapy treatment*	0.66	0.72	[redacted]	[redacted]
Consolidation treatment	0.66	0.72	[redacted]	[redacted]
HSCT	0.66	0.72	[redacted]	[redacted]
GVHD (post-HSCT)	0.67*	0.67*	[redacted]	[redacted]
CR/CRp off-treatment	0.74	0.77	[redacted]	[redacted]
Relapse	0.57	0.62	[redacted]	[redacted]
Refractory	0.57	0.62	[redacted]	[redacted]
Functionally cured	0.82*	0.82*	[redacted]	[redacted]

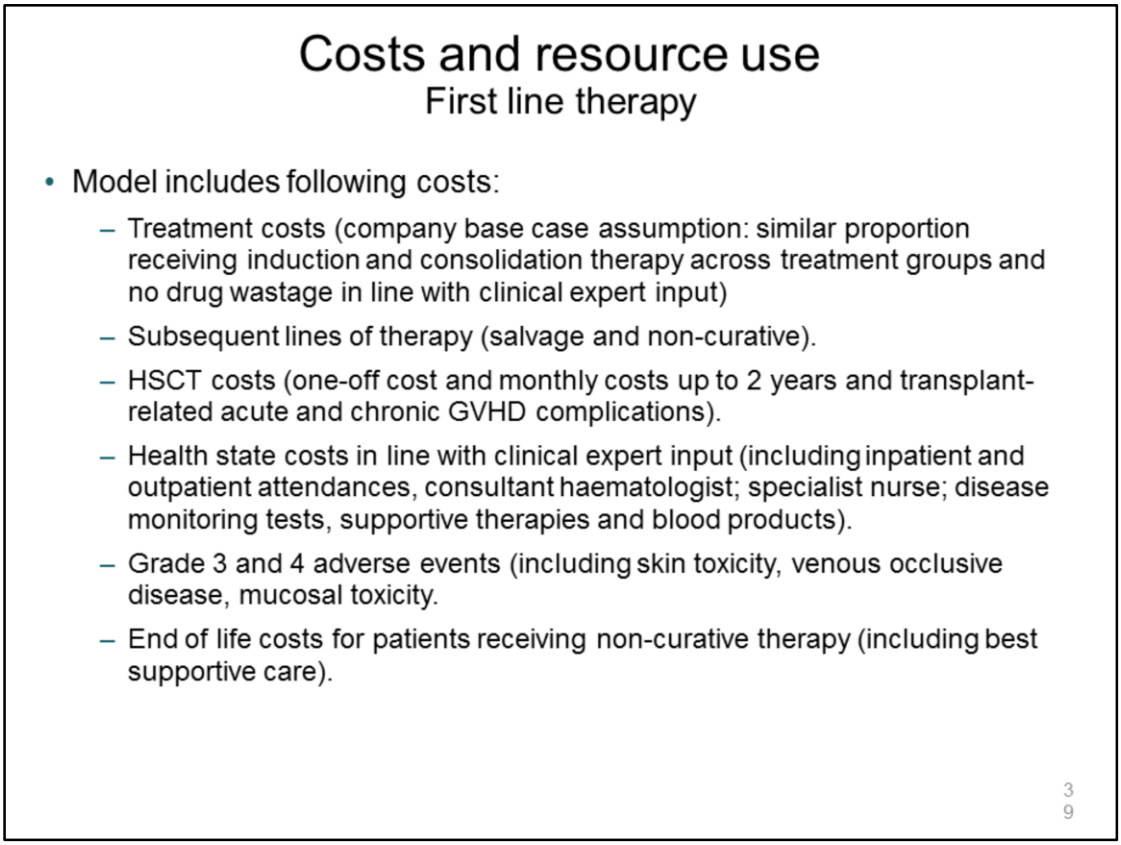
*Applied to patients in induction, salvage and non-curative
*Varied per cycle, based on mean patient age at each time point, from Ara & Brazier
*Source Kurosawa 2014

38

Adapted from ERG report Table 31 page 93.

ERG's comments:

- The ERG considered the approach used by the company to be reasonable and appropriately justified.
- The ERG commented that company's assumption that functionally cured patients experience the same HRQoL as the general population results in a marked jump in the HRQoL estimates at 5-years for functionally cured patients. The use of general population quality of life was not considered internally consistent with the excess mortality applied for functionally cured patients to OS. Given that functionally cured patients are assumed to be at higher mortality risk than the general population, the ERG concluded that it would appear reasonable to assume that functionally cured patients would also have lower quality of life than that of the general population (see ERG report section 5.2.7.1 page 94).



Adapted from ERG report, section 5.2.8 pages 95-103.

- For details of the disaggregated cost results (commercial-in-confidence) used in the company's base-case, see Table 123 page 267 of the company appendices.

CONFIDENTIAL

Company's base-case results

- Company excludes patients with known unfavourable cytogenetics based on the subgroup analysis results of ALFA-0701 and clinical advice that these patients would not be treated with GO plus intensive chemotherapy in clinical practice.
- Company included in the base-case population patients with unknown cytogenetics.

Company base- case					
Technologies	Total costs (£)	Total QALYs	Incremental costs (£)	Incremental QALYs	ICER incremental (£/QALY)
Deterministic results					
GO + DA	[redacted]	[redacted]	[redacted]	[redacted]	£12,251
DA	[redacted]	[redacted]			
Probabilistic results					
GO + DA	[redacted]	[redacted]	[redacted]	[redacted]	£13,600
DA	[redacted]	[redacted]			

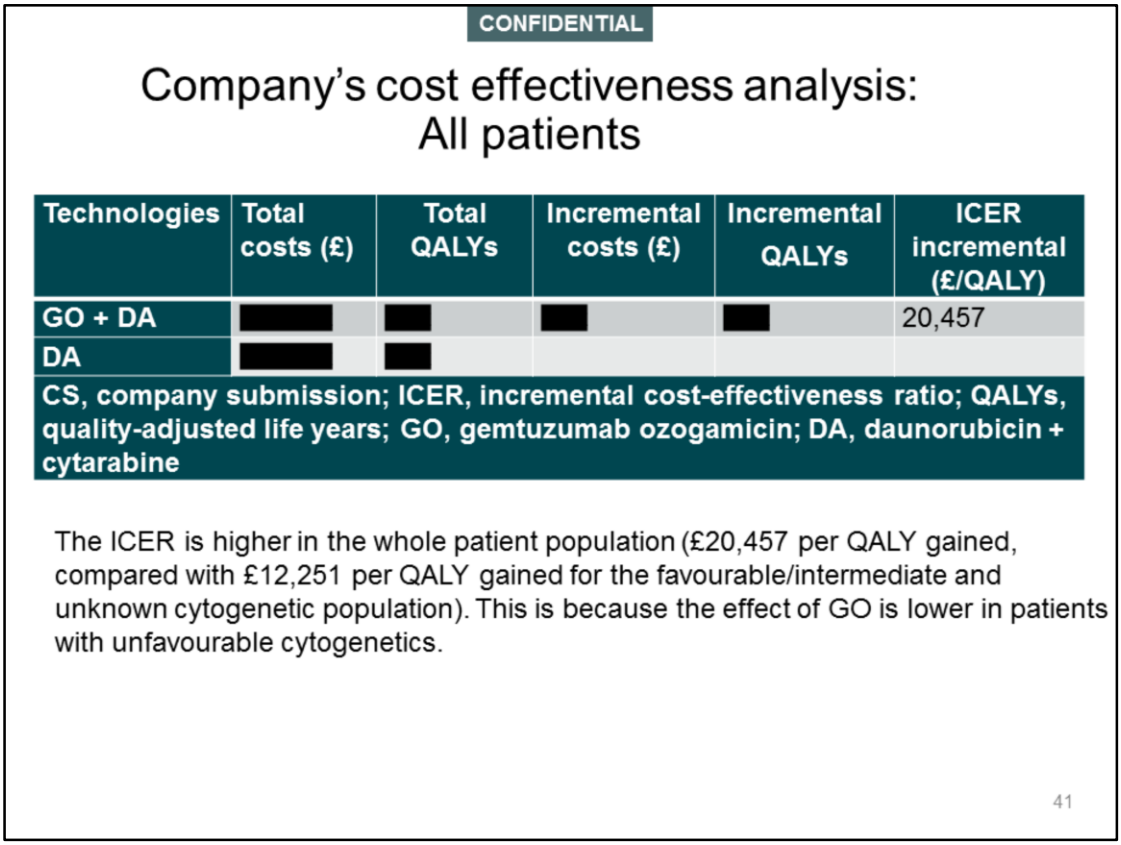
The company's sensitivity analyses showed that parameters with the largest impact on the ICER were HSCT probabilities from relapse in years 1 and 2 for the DA group, and the restricted mean survival time for relapsed patients.

40

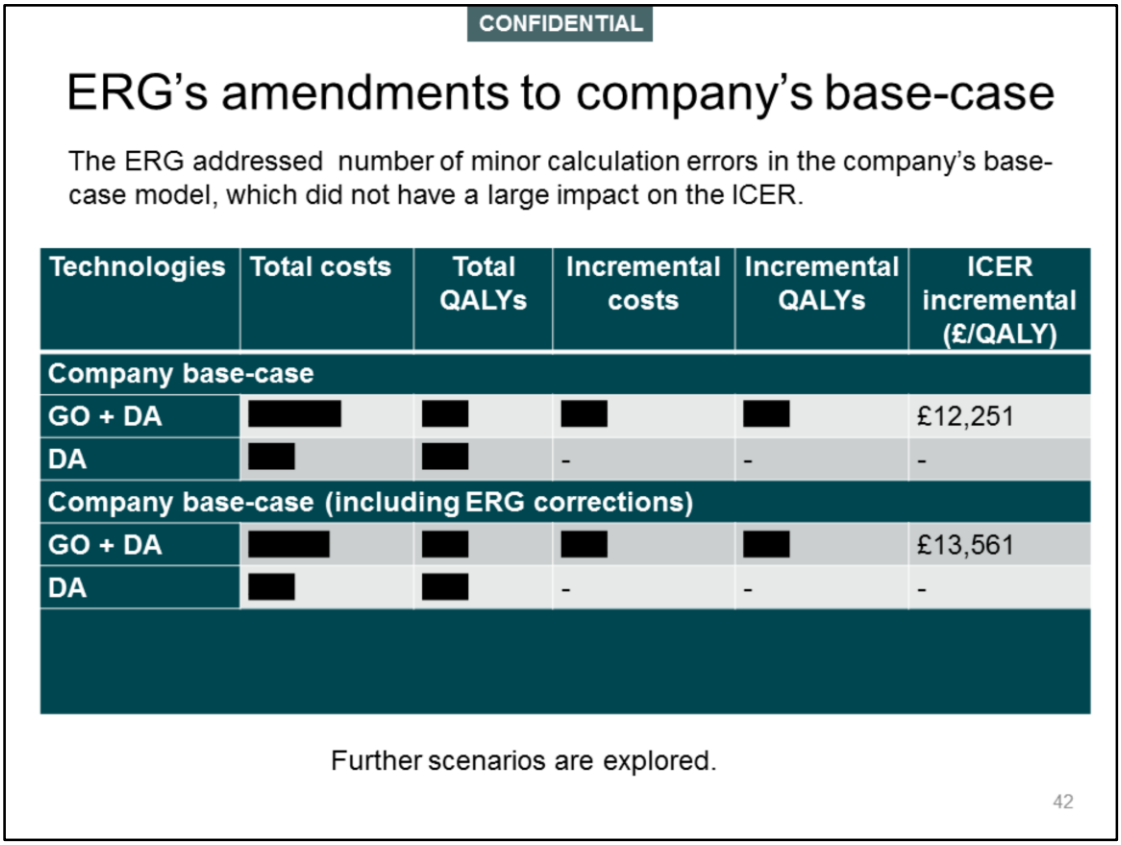
Adapted from ERG report, table 37 pages 104.

ERG's comments

- The ERG considers that the company did not sufficiently justify the inclusion of the subgroup with unknown cytogenetics and/or attempt to fully explore the implications of alternative assumptions.



Source: ERG report Table 38, page 106.



Source: ERG report Table 39, page 111.

Amendments made by the ERG:

- Inconsistencies in the data source for mortality: The ERG incorporated the more recently published mortality data for England & Wales for the survival analysis, and the mortality data for the UK for the mortality HR calculations.
- Discrepancy for HSCT probabilities after relapse: The ERG amendment involved changing the calculations to reflect the actual number of patients achieved CR/CRp in the model.
- Patients who did not receive the second cycle of induction therapy in the second cycle of the model were considered equivalent to those off-treatment for HRQoL, purpose and did not have any associated costs that cycle. The ERG applied the cost associated with the off-treatment health state to these patients in that cycle.
- Estimation of the proportion of patients with refractory disease receiving salvage therapy: patients were double adjusted. This was corrected by the ERG so that all patients with refractory disease receiving the first cycle of salvage therapy also received the subsequent cycles of salvage therapy.

CONFIDENTIAL

ERG's exploratory analyses:
courses of treatment

The impact of modelling individual rates of courses of treatment, which did not include the company adjustment to align with the IRC-assessed clinical data. Resulted in slight increase of ICER.

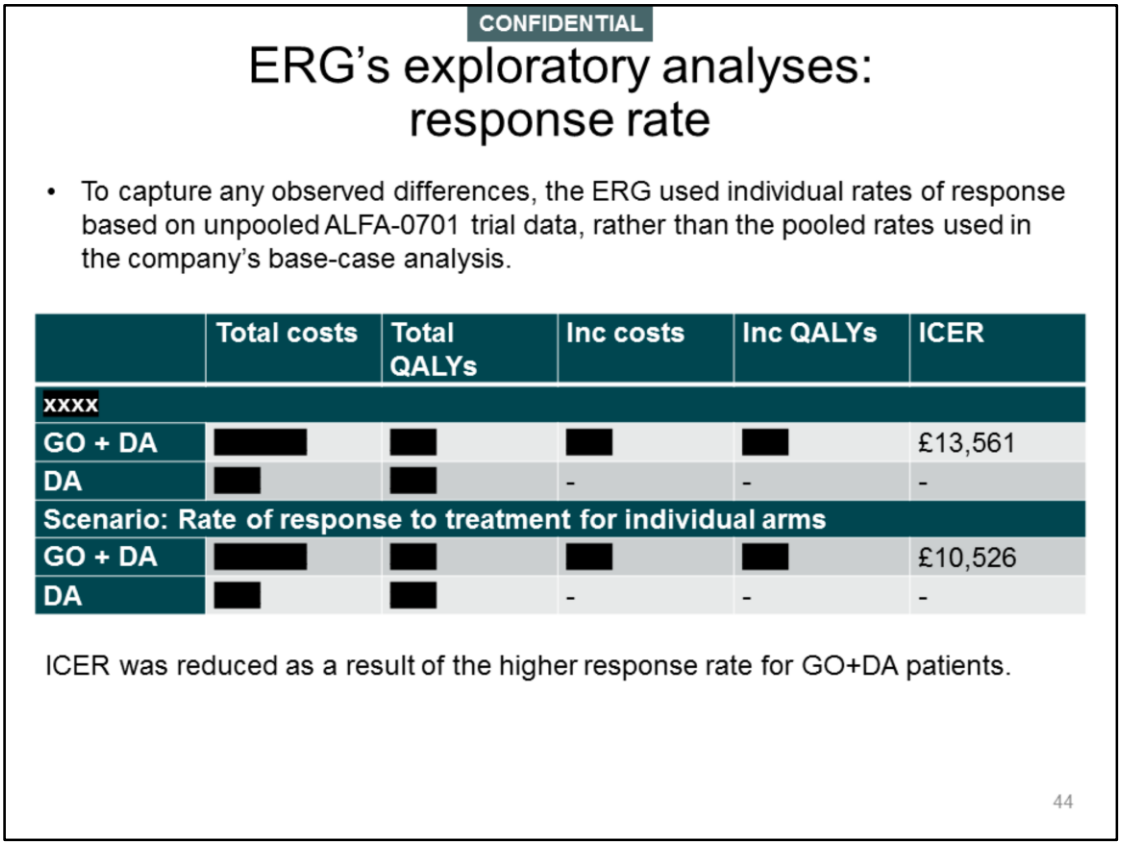
	Total costs	Total QALYs	Inc costs	Inc QALYs	ICER
Company base-case (including ERG corrections)					
GO + DA	[redacted]	[redacted]	[redacted]	[redacted]	£13,561
DA	[redacted]	[redacted]	-	-	-
Scenario: Courses of treatment based on unpooled investigator-assessed data					
GO + DA	[redacted]	[redacted]	[redacted]	[redacted]	£14,249
DA	[redacted]	[redacted]	-	-	-

	Company modelled values		ERG modelled values	
Proportion of patients	GO+DA group	DA group	GO+DA group	DA group
Induction course 1	[redacted]	[redacted]	[redacted]	[redacted]
Induction course 2	[redacted]	[redacted]	[redacted]	[redacted]
Consolidation course 1	[redacted]	[redacted]	[redacted]	[redacted]
Consolidation course 2	[redacted]	[redacted]	[redacted]	[redacted]

Source: ERG report Table 41, page 113.

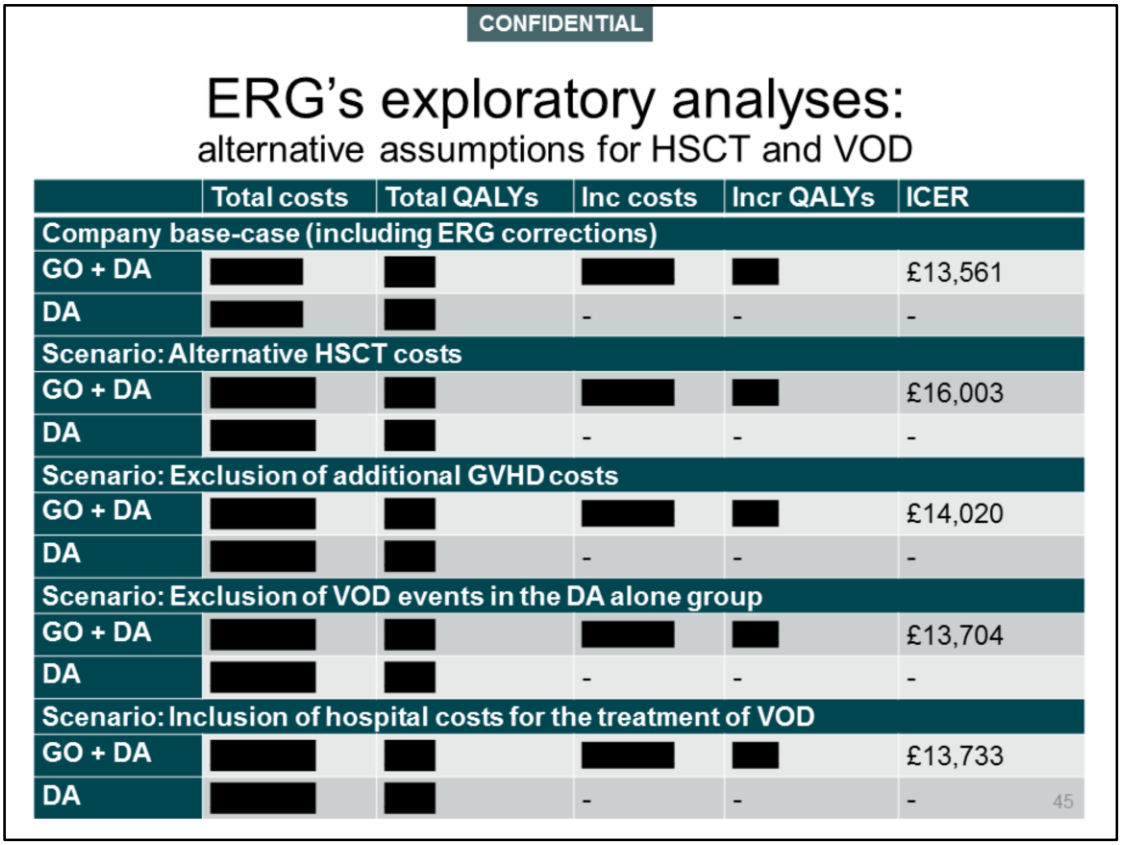
ERG's comments:

- While the ERG acknowledged the arguments made by the company to use IRC as opposed to IA analyses, the ERG considered that that the initial treatment costs of the induction and consolidation therapies should be based on the IA response outcomes. IA outcomes more appropriately reflect the actual treatment decisions and resource use incurred within the trial (see ERG report section 5.2.6.1, page 81).



	Total costs	Total QALYs	Inc costs	Inc QALYs	ICER
xxxx					
GO + DA	[redacted]	[redacted]	[redacted]	[redacted]	£13,561
DA	[redacted]	[redacted]	-	-	-
Scenario: Rate of response to treatment for individual arms					
GO + DA	[redacted]	[redacted]	[redacted]	[redacted]	£10,526
DA	[redacted]	[redacted]	-	-	-

Source: ERG report Table 42, page 114.



Source: ERG report Table 43, page 115.

ERG's comments:

- The ERG noted that HSCT costs in the company submission were obtained from a costing study conducted in the Netherlands between 1994 and 1999 and since HSCT costs changed substantially. Inflating these costs to 2017 may not accurately reflect the current costs. In the NHS reference costs of HSCT vary (from £17,344 for an autologous transplant to £38,336 for an allogeneic transplant from an unspecified donor) but they also are substantially lower than the unit cost used by the company (see ERG report section 5.2.8.4, page 101).
- Overestimating HSCT costs would bias the model in favour of GO+DA as fewer of these patients had an HSCT (see ERG report section 5.2.8.4, page 101).
- There is some uncertainty whether the additional inpatient treatment is already captured in the length of stay assumptions (see ERG report section 5.2.8.5, page 102).
- The ERG was generally satisfied with the approach to implement the AE-related costs for first-line therapy. However, the ERG considered that patients experiencing VOD would also require inpatient treatment extending beyond the standard stay for treatment with GO because of the associated high mortality risk. There is some uncertainty whether the additional inpatient treatment is already captured in the length of stay

assumptions (see ERG report section 5.2.8.5, page 102).

CONFIDENTIAL

ERG's exploratory analyses:

alternative values for functionally cured patients

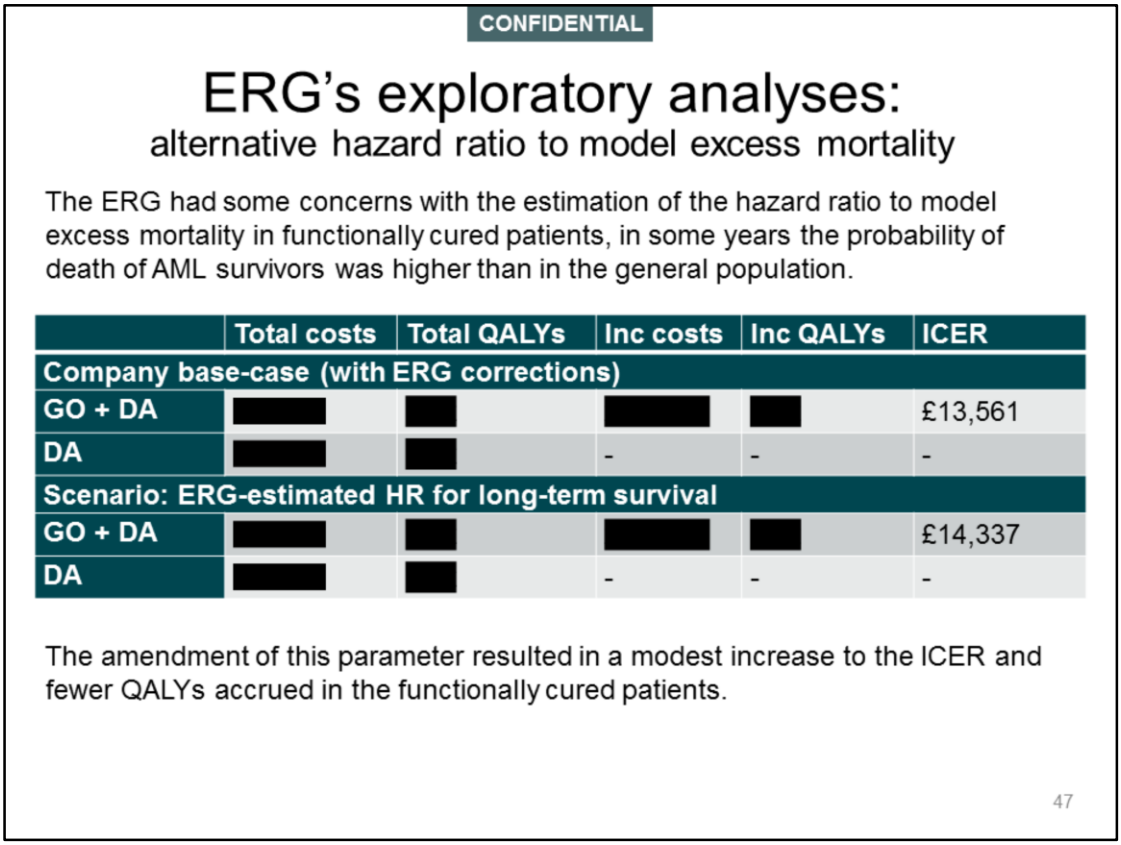
- The ERG considered the assumption that patients who are functionally cured, experience the same HRQoL as the general population, as not sufficiently justified.
- ERG explored a scenario where functionally cured patients would have lower quality of life than that of the general population.

	Total costs	Total QALYs	Incremental costs	Incremental QALYs	ICER
Company base-case (with ERG corrections)					
GO + DA	[redacted]	[redacted]	[redacted]	[redacted]	£13,561
DA	[redacted]	[redacted]	-	-	-
Scenario: Alternative utility values for functionally cured patients					
GO + DA	[redacted]	[redacted]	[redacted]	[redacted]	£13,878
DA	[redacted]	[redacted]	-	-	-
Scenario: Alternative utility values for functionally cured patients, adjusted for aging					
GO + DA	[redacted]	[redacted]	[redacted]	[redacted]	£15,279
DA	[redacted]	[redacted]	-	-	-

Both scenarios were associated with lower QALYs and higher ICERs.

46

Source: ERG report Table 44, page 116.



Source: ERG report Table 45, page 117.

ERG's comments:

- The ERG considered that a further adjustment appears appropriate, such that the mortality rate is set equal to the general population mortality rate in instances when the observed mortality rate is reported to be lower than the general population (see ERG report section 5.2.6.4, page 89).

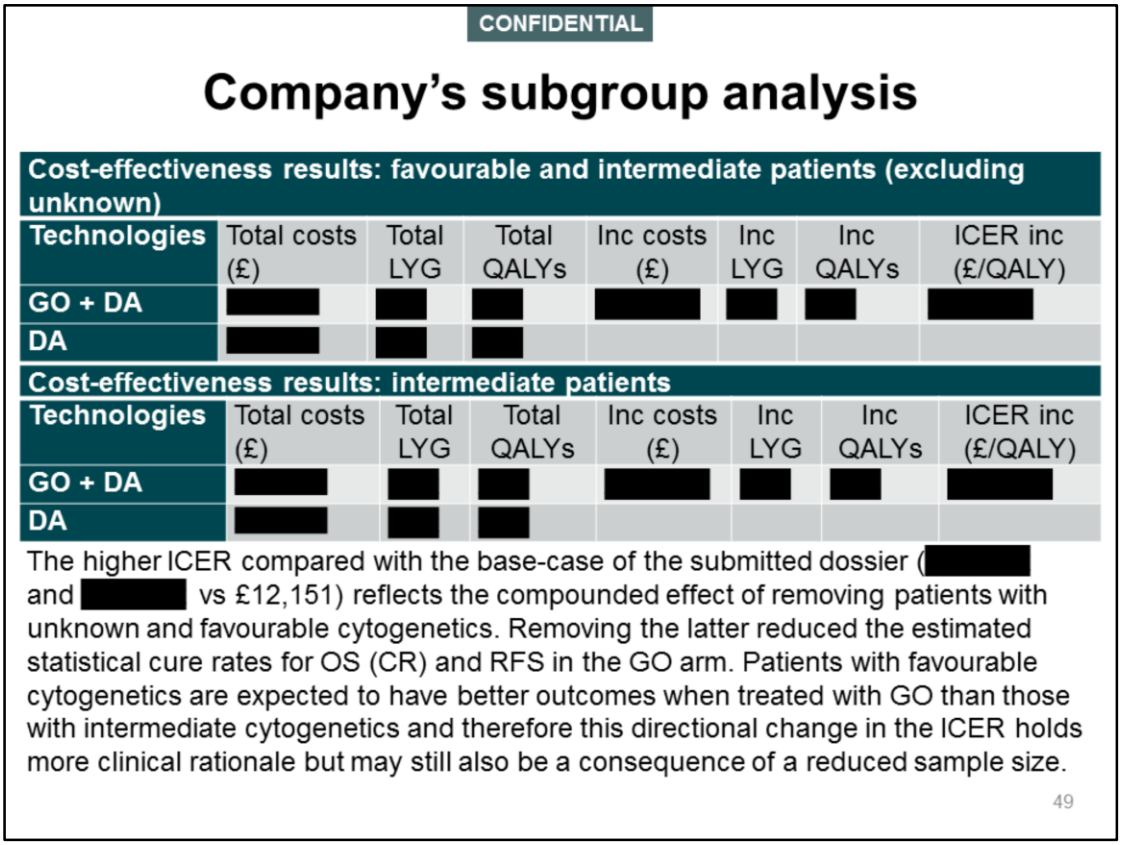
ERG's alternative base case

ERG's changes	ICER
Company base-case	£12,251
Including ERG correct minor calculation errors	£13,561
Number of induction and consolidation therapy (from ALFA-0701 trial)	£14,249
Arm-specific rate of response to treatment	£10,526
The initial cost of HSCT estimated from NHS Reference Costs	£16,003
Removal of VOD events in the DA treatment group	£13,704
Exclusion of GVHD-specific costs	£14,020
Inclusion of hospital costs for the treatment of VOD	£13,733
Quality of life in functionally cured patients based on the utility value for off-treatment CR patients, and further adjusted for age	£15,279
Long-term mortality in functionally cured patients adjusted for excess mortality using the ERG-calculated hazard ratio	£14,337
ERG's alternative base-case analysis	
Deterministic results	£16,910
Probabilistic results (preferred by ERG)	£17,956

Adapted from ERG report, section 6.4 page 117.

ERG's comments:

- The ERG noted that the probabilistic ICER was the most relevant to inform decisions based on cost-effectiveness, and is referred to as the key ICER for the ERG alternative base-case analysis elsewhere in this report.
- The scenarios were not associated with substantial differences to the ICER.



CONFIDENTIAL

Company's subgroup analysis

Cost-effectiveness results: favourable and intermediate patients (excluding unknown)

Technologies	Total costs (£)	Total LYG	Total QALYs	Inc costs (£)	Inc LYG	Inc QALYs	ICER inc (£/QALY)
GO + DA	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
DA	[redacted]	[redacted]	[redacted]				

Cost-effectiveness results: intermediate patients

Technologies	Total costs (£)	Total LYG	Total QALYs	Inc costs (£)	Inc LYG	Inc QALYs	ICER inc (£/QALY)
GO + DA	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
DA	[redacted]	[redacted]	[redacted]				

The higher ICER compared with the base-case of the submitted dossier ([redacted] and [redacted] vs £12,151) reflects the compounded effect of removing patients with unknown and favourable cytogenetics. Removing the latter reduced the estimated statistical cure rates for OS (CR) and RFS in the GO arm. Patients with favourable cytogenetics are expected to have better outcomes when treated with GO than those with intermediate cytogenetics and therefore this directional change in the ICER holds more clinical rationale but may still also be a consequence of a reduced sample size.

49

Adapted from clarification response Tables 1 and 2 page 8.

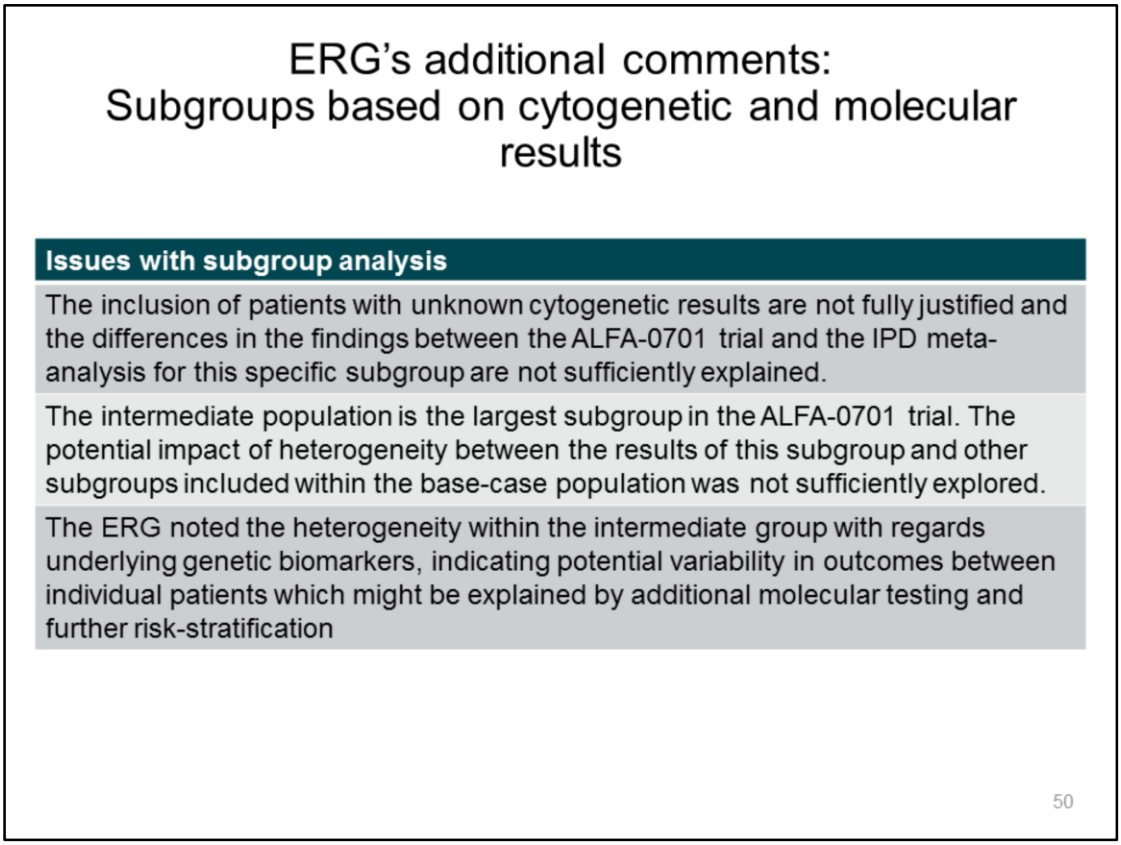
Company's rationale for including unknown cytogenetics group (clarification response Question B3 page 7)

- According to UK clinical expert opinion less than 10% of patients with de novo AML in the UK present with unknown cytogenetics (in line with the 9.2% included in ALFA-0701). An unknown classification may be a consequence of inadequate specimens or non-dividing cells making cytogenetic risk classification impossible. Depending on the severity of their symptoms these patients may need to be treated immediately rather than waiting for further confirmatory tests therefore it was considered to be appropriate to include these patients in our base-case population.

ERG's comments:

- While the ERG agreed with the company's decision to exclude patients with known unfavourable cytogenetics, the ERG did not believe that the company has sufficiently addressed the heterogeneity in the subgroup of patients with unknown cytogenetics and within the intermediate population. The ERG considered that there remains significant heterogeneity within the base-case population which may have important implications concerning the difference in the cure fraction for further subgroups within the overall

population (see ERG report section 1.5 page 17).



Source: ERG report

CONFIDENTIAL

ERG's exploratory analysis: subgroups by cytogenetic and molecular results (1)

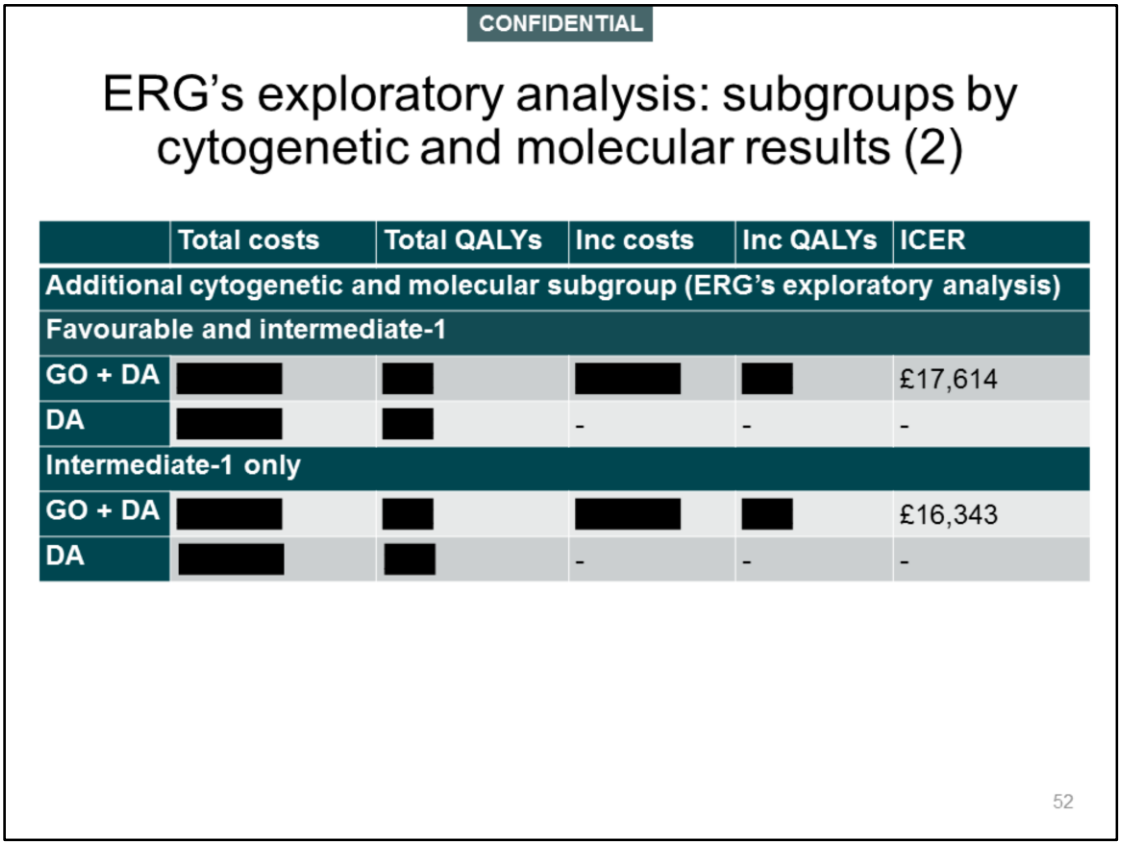
	Total costs	Total QALYs	Inc costs	Inc QALYs	ICER
Populations considered by company					
Base-case (Favourable, intermediate and unknown)					
GO + DA	[redacted]	[redacted]	[redacted]	[redacted]	£16,910
DA	[redacted]	[redacted]	-	-	-
All patients					
GO + DA	[redacted]	[redacted]	[redacted]	[redacted]	£25,941
DA	[redacted]	[redacted]	-	-	-
Additional cytogenetic subgroups (ERG's exploratory analysis)					
Favourable and intermediate					
GO + DA	[redacted]	[redacted]	[redacted]	[redacted]	£24,581
DA	[redacted]	[redacted]	-	-	-
Intermediate					
GO + DA	[redacted]	[redacted]	[redacted]	[redacted]	£31,709
DA	[redacted]	[redacted]	-	-	-

51

Adapted from ERG report Table 52 page 124.

ERG's comments

- The results suggest that clinical and economic value of GO+DA appears largely confined to the favourable and intermediate-1 population, defined by cytogenetic and molecular tests (see ERG report section 6.5.2 page 124).
- The company reported that they were unable to fit MCM models to some subgroups (see ERG report page 119).
- These findings can only be considered indicative because of data limitations. Uncertainties also remain concerning the practicality and feasibility of introducing additional risk stratification within routine clinical practice (see ERG report page 126).



	Total costs	Total QALYs	Inc costs	Inc QALYs	ICER
Additional cytogenetic and molecular subgroup (ERG's exploratory analysis)					
Favourable and intermediate-1					
GO + DA	[redacted]	[redacted]	[redacted]	[redacted]	£17,614
DA	[redacted]	[redacted]	-	-	-
Intermediate-1 only					
GO + DA	[redacted]	[redacted]	[redacted]	[redacted]	£16,343
DA	[redacted]	[redacted]	-	-	-

Adapted from ERG report Table 52 page 124.

End of life considerations

The company and ERG agree that this intervention does not meet the end of life criteria.

53

Innovation and equality

- Company considers gemtuzumab ozogamicin to be innovative:
  - gemtuzumab ozogamicin targets AML and in combination with DA is able to extend the duration of remission.
  - Gemtuzumab ozogamicin as an add-on to DA therefore represents a step-change in the management of adult patients with de novo AML.
  - The clinical benefit of adding gemtuzumab ozogamacin to DA are particularly apparent in patients with favourable/intermediate cytogenetics profile, but not in patients with unfavourable cytogenetics profile.
  - Gemtuzumab ozogamacin is able to directly target CD33-positive AML blasts in order to induce death of leukaemic cells.
  - Gemtuzumab ozogamacin reduces relapses in patients which can impact on patients' HRQoL, and therefore reduces associated increased costs owing to the need for hospitalization and chemotherapy to induce a second remission.
  - No issues equality issues raised during scoping or company submission/ patient professional statements.

54

Authors

- **Julia Sus**
  Technical Lead
- **Nicola Hay**
  Technical Adviser
- with input from the Lead Team (Derek Ward, Nigel Langford and David Chandler)

55

NATIONAL INSTITUTE FOR HEALTH AND CARE EXCELLENCE

Single technology appraisal

Gemtuzumab ozogamicin for treating acute myeloid leukaemia [ID982]

https://www.nice.org.uk/guidance/indevelopment/gid-ta10142

Document B

Company evidence submission

December 2017

File name	Version	Contains confidential information	Date
ID982_Pfizer_Document B_(8Dec2017)	1	Yes	8 December 2017
ID982_Pfizer_Document B_revised redacting_(10Jan18)	2	Yes	10 January 2018

Contents

List of tables

List of figures

Abbreviations

AE	adverse event
AIC	Akaike information criterion
ALL	acute lymphoblastic leukaemia
ALT	alanine aminotransferase
AML	acute myeloid leukaemia
ANC	absolute neutrophil count
APL	acute promyelocytic leukaemia
AST	aspartate aminotransferase
BCSH	British Committee for Standards in Haematology
BFU-E	burst-forming unit-erythroid
BIC	Bayesian information criterion
BMA	bone marrow aspirate
BNF	British National Formulary
BSA	body surface area
BSC	best supportive care
CCR	conventional care regimen
CEBPA	CCAAT/enhancer-binding protein a gene
CFU-GM	colony-forming unit granulocyte, monocyte
CFU-meg	colony-forming unit megakaryocyte
CHV	Centre Hospitalier de Versailles
CI	confidence interval
CLL	chronic lymphocytic leukaemia
CML	chronic myeloid leukaemia
CNS	central nervous system
CONSORT	Consolidated Standards of Reporting Trials
CR	complete remission
CRF	case report form
CRp	CR with incomplete platelet recovery
CSR	clinical study report
CTCAE	Common Terminology Criteria for Adverse Events
D	Day
DA	daunorubicin and cytarabine
DNA	deoxyribose nucleic acid
DoH	Department of Health
DSU	Decision Support Unit
ECOG PS	Eastern Cooperative Oncology Group performance status
EFS	Event-free survival
ELN	European LeukemiaNet
EMA	European Medicine Agency
eMiT	electronic market information tool

EORTC QLQ-C30	European Organisation for Research and Treatment of Cancer QoL questionnaire
EQ-5D	5-dimension quality of life questionnaire
EQ-VAS	European QoL visual analogue scale
ERG	evidence review group
ESMO	European Society for Medical Oncology
FAB	French–American–British
FLAG-Ida	Fludarabine, cytarabine, G-CSF and idarubicin
FLT3	*fms-like tyrosine kinase 3*
G-CSF	granulocyte colony-stimulating factor
GO	Gemtuzumab ozogamicin
GVHD	graft versus host disease
HR	hazard ratio
HRQoL	Health-related quality of life
HS	health state
HSCT	haematopoietic stem cell transplantation
ICD	informed consent document
ICER	incremental cost-effectiveness ratios
IPD	individual patient data
IRC	independent review committee
ITD	internal tandem duplication
ITT	intention-to-treat population
KM	Kaplan–Meier
LYG	life years gained
MCMs	mixture cure models
MDS	myelodysplastic syndrome
MedDRA	Medical Dictionary for Regulatory Activities
MFI	mean fluorescence intensity
mITT	modified intention-to-treat
MLL	myeloid/lymphoid leukaemia gene
MRC UK	Medical Research Council UK
NA	not applicable
NCCN	National Comprehensive Cancer Network
NCRI	National Cancer Research Institute
NE	not estimable
NHS	National Health Service
NHSBT	NHS Blood and Transplant
NICE	The National Institute for Health and Care Excellence
No. at risk	number at risk
NPM1	nucleophosmin-1 gene
NR	not reached
NZ	New Zealand
OR	odds ratio

OS	overall survival
PS	performance status
PSA	probabilistic sensitivity analysis
PSS	Personal Social Services
PSSRU	Personal Social Services Research
QALYs	quality-adjusted life years
RCT	randomized controlled trial
RFS	Relapse-free survival
RMST	restricted mean survival time
RS	relative survival
RU	resource use
SAE	serious adverse event
SD	standard deviation
SE	standard error
SF-12	12-item Short-Form Health Survey
SLR	systematic literature review
SmPC	summary of product characteristics
SOC	system organ class
SoC	standard of care
TEAEs	treatment-emergent AEs
TTO	time trade-off
ULN	upper limit of normal
v	version
VAS	visual analogue scale
VOD	veno-occlusive disease
WBC	white blood cell
WT1	Wilms' tumour suppressor gene

B.1 Decision problem, description of the technology and clinical care pathway

B.1.1 Decision problem

Gemtuzumab ozogamicin (GO) in combination with daunorubicin and cytarabine (DA) will not have received marketing at the time of this submission. The expected date for the European Medicine Agency (EMA) marketing authorization is Q2 2018.

The EMA marketing authorisation application proposes the following indication: GO in combination with DA for the treatment of adult patients with previously untreated, de novo acute myeloid leukaemia (AML).

This submission focuses on an optimized subpopulation of the expected marketing authorization for GO + DA, which is adult patients with no known unfavourable cytogenetic profile risk with previously untreated de novo AML, for the reasons presented below:

- This subpopulation reflects where GO + DA provides clinical benefit and, as a consequence, optimises the cost effectiveness of GO + DA versus DA alone.
- The proposed subpopulation is consistent with NHS clinical practice; GO + DA would not be used in patients known to have unfavourable cytogenetics profile risk, as confirmed by clinicians treating patients with AML in NHS England.

Table 1. The decision problem

	Final scope issued by NICE	Decision problem addressed in the company submission	Rationale if different from the final NICE scope
Population	Adults with untreated acute myeloid leukaemia (AML)	Adult patients not known to have unfavourable cytogenetics, with previously untreated, de novo AML	The population stated in the expected marketing authorisation is adult patients with previously untreated, de novo AML. Clinicians in England have advised that patients who are known to have unfavourable cytogenetics would not be treated with GO plus intensive chemotherapy in the NHS owing to limited evidence demonstrating superior clinical outcomes versus intensive chemotherapy alone in this patient population.
Intervention	GO in combination with chemotherapy	GO in combination with daunorubicin plus cytarabine (DA; intensive chemotherapy) **Induction course:** GO 3 mg/m^2/dose (up to a maximum of 5 mg/dose) infused on days 1, 4 and 7of induction therapy course. **Consolidation course 1 and 2:** GO 3 mg/m^2/dose (up to a maximum of 5 mg/dose) infused on day 1 of each course of consolidation therapy	In line with the expected marketing authorisation.

	Final scope issued by NICE	Decision problem addressed in the company submission	Rationale if different from the final NICE scope
Comparator(s)	Established clinical management without GO, including but not limited to amsacrine, cytarabine, daunorubicin, etoposide, fludarabine, idarubicin, mitoxantrone, thioguanine and midostaurin (subject to ongoing NICE appraisal)	**Induction course:** 60 mg/m^2/day of daunorubicin on days 1–3 and cytarabine 200 mg/m^2/day on days 1–7 **Consolidation course 1:** 60 mg/m^2 of daunorubicin on day 1 and cytarabine 1 g/m^2/every 12 hours on days 1–4 **Consolidation, course 2:** 60 mg/m^2/day of daunorubicin on days 1–2 and cytarabine 1 g/m^2/every 12 hours on days 1–4	The most well established approach to achieving remission in untreated AML is with intensive chemotherapy for patients who are fit enough to receive it. GO is expected to be licenced as an add-on therapy to DA (intensive chemotherapy). Therefore, in order for patients to be eligible for GO +DA they must also be eligible to receive intensive chemotherapy. NHS England clinical expert opinion is that outside of clinical trials the standard of care for intensive chemotherapy is DA. Furthermore, the British Committee for Standards in Haematology (BCSH), that have developed the only published clinical guidelines for AML in the UK recommend initial therapy with daunorubicin (or anthracycline or an anthracycline-like drug) with cytarabine.[1] The evidence being presented in this dossier is therefore in line with UK clinical practice. The following drugs cannot be considered comparators for the reasons specified below:

	Final scope issued by NICE	Decision problem addressed in the company submission	Rationale if different from the final NICE scope
			Fludarabine, cytarabine, G-CSF and idarubicin (FLAG-Ida) may be used in combination (not individually as listed in the final scope) as part of FLAG-Ida induction therapy in the UK. However according to UK clinical experts, it is unlikely that this combination would be used for the patient population in this appraisal as this approach is generally reserved only for younger patients who have an unfavourable cytogenetics or therapy-related AML. This is because FLAG-Ida has similar efficacy to DA, but is not as well tolerated. Amsacrine, etoposide and mitoxantrone may be used as part of post-remission therapy by the Medical Research Council UK (as stated in the BCSH guidelines) following two courses of induction chemotherapy only. Therefore, these medicines are not comparators to GO + DA, which is used in the induction phase to induce remission. Midostaurin is targeted at patients with *FLT3* mutation-positive AML; these patients would form a subgroup of those

	Final scope issued by NICE	Decision problem addressed in the company submission	Rationale if different from the final NICE scope
			who are eligible to receive GO. Midostaurin is currently under NICE review and will not be the standard of care at the time of this submission and therefore cannot be considered as a comparator. Thioguanine is not an intensive chemotherapy and is not considered standard of care by UK clinicians in patients eligible for intensive chemotherapy.
Outcomes	• Event-free survival (EFS) • Overall survival (OS) • Disease-free survival • Adverse effects of treatment • Health-related quality of life (HRQoL)	• EFS • OS • Relapse-free survival (RFS; disease-free survival) • Adverse effects of treatment • Response rate • Health-related quality of life (HRQoL)	The pivotal study presented in this submission (ALFA-0701) did not collect HRQoL or utility data. In order to estimate HRQoL in patients with AML, Pfizer conducted a vignette study to assess utility associated with AML health states according to the preferences of the UK general population. In addition a systematic literature review was conducted to identify studies that assessed HRQoL in the relevant patient population. Furthermore, past NICE appraisals and utilizing values from NICE TA399.

	Final scope issued by NICE	Decision problem addressed in the company submission	Rationale if different from the final NICE scope
Economic analysis	The reference case stipulates that the cost effectiveness of treatments should be expressed in terms of incremental cost per quality-adjusted life-year. The reference case stipulates that the time horizon for estimating clinical and cost effectiveness should be sufficiently long to reflect any differences in costs or outcomes between the technologies being compared. Costs will be considered from an NHS and Personal Social Services perspective.	As per scope	NA
Subgroups to be considered	If evidence allows, consideration will be given to subgroups based on cytogenetics profile risk. If the evidence allows, a scenario analysis will be considered whereby stem cell transplantation is included as a subsequent treatment for people who are fit enough to undergo the procedure and whose disease remitted after standard high-dose chemotherapy with or without GO. This should reflect the proportion of people who proceed to stem cell transplantation after each treatment regimen, as well as the costs	The subgroups considered in this submission have also been considered for the key outcomes (EFS, RFS, OS). These subgroups are: 1) based on cytogenetic risk (i.e. intermediate/favourable vs unfavourable) 2) stem cell transplantation as a subsequent treatment for patients who are fit enough to undergo transplantation and whose disease remitted after	The submission takes into account cytogenetic risk as the main population to be considered in this submission: adult patients not known to have unfavourable cytogenetics, with previously untreated, de novo AML. This population excludes those patients with unfavourable risk. Clinical data demonstrate that GO in combination with DA is more clinically effective than DA alone. This is particularly the case in patients with favourable/intermediate cytogenetics profile risk. In patients with unfavourable cytogenetics profile risk, clinical data

	Final scope issued by NICE	Decision problem addressed in the company submission	Rationale if different from the final NICE scope
	and quality-adjusted life-year benefits of the procedure.	therapy	suggest that GO in combination with DA is no more effective than DA alone. Stem cell transplantation subgroup analysed in line with final scope in the economic model.
Special considerations including issues related to equity or equality	NA	NA	NA

AML, acute myeloid leukaemia; BCSH, British Committee for Standards in Haematology; DA, daunorubicin plus cytarabine; EFS, event-free survival; FLAG-Ida, fludarabine + cytarabine + granulocyte colony-stimulating factor + idarubicin; FLT3, FMS-like tyrosine kinase 3; GO, gemtuzumab ozogamicin; HRQoL, health-related quality of life; NA, not applicable; NHS, National Health Service; NICE, The National Institute for Health and Care Excellence; OS, overall survival; RFS, relapse-free survival.

B.1.2 Description of the technology being appraised

A draft version of the summary of product characteristics (SmPC) has been included in Appendix C. This document is subject to being updated until publication of the European public assessment report.

Table 2. Technology being appraised

UK approved name and brand name	Brand name: Mylotarg® UK approved name: Gemtuzumab ozogamicin (GO)
Mechanism of action	GO is an antibody–drug conjugate that combines a humanized, anti-CD33, monoclonal antibody with calicheamicin, a potent cytotoxic agent that causes DNA damage.[2-4] CD33 is a sialic acid-dependent adhesion protein that is highly expressed on the surface of acute myeloid leukaemia (AML) blast cells and on some leukaemic stem cells.[5,6] Approximately 85–90% of patients with AML are considered to be CD33 positive.[2,7] The CD33 component of GO enables targeted delivery of calicheamicin to CD33-positive blast cells.[3,4] Once GO is bound to CD33 on the cell surface it is internalized; inside the cell, calicheamicin is released from the antibody–drug conjugate complex and is able to enter the nucleus. Once activated, calicheamicin can bind to DNA to cause DNA double-strand breaks, resulting in cell-cycle arrest and apoptotic cell death.[3,8,9] *In vitro* studies showed that after a 3 mg/m^2 dose of GO, re-expression of CD33 to nearly pre-treatment levels occurred after 72 hours.[8] This led to the hypothesis that repeated administration of lower, near-saturating, fractionated doses of GO may enable increased drug internalization while improving safety versus a higher, unfractionated dosing regimen.[8,10,11] This fractionated 3 mg/m^2 dose is used in the pivotal ALFA-0701 clinical study.
Marketing authorisation/CE mark status	GO does not currently have European Medicines Agency (EMA) marketing authorization for the indication in this submission. A new marketing authorization application for GO was submitted to EMA in December 2016 for the treatment of adult patients with previously untreated, de novo

	AML in combination with daunorubicin and cytarabine (DA). A positive opinion from the Committee for Medicinal Products for Human Use is anticipated in Q1 2018 and the anticipated date of EMA approval is in Q2 2018. The EMA granted GO orphan drug status in 2000 for the treatment of AML.[12]
Indications and any restriction(s) as described in the summary of product characteristics (SmPC)	The anticipated indication of GO in the SmPC is as a combination therapy with DA for the treatment of adult patients with previously untreated, de novo AML. At the time of submission, the SmPC for GO was not available. The SmPC will be available in Q2 2018.
Method of administration and dosage	GO is administered as an intravenous infusion in combination with DA (intensive chemotherapy). In patients with previously untreated AML, the recommended dose of GO is as follows: • induction – 3 mg/m^2/day (maximum 5 mg/day) infused over 2 hours on days 1, 4 and 7 as part of therapy with DA • consolidation therapy in patients experiencing complete remission – 3 mg/m^2/day (maximum 5 mg/day) infused over 2 hours on day 1 as part of therapy with DA, for up to two cycles.
Additional tests or investigations	No additional tests would be required before patients are prescribed GO, other than those that are already routine in the diagnosis of AML (e.g. cytogenetic profile testing). After initiating GO, patients would require no tests or investigations additional to those that would already be performed following treatment with standard intensive chemotherapy.
List price and average cost of a course of treatment	List price of GO: [redacted] Average cost for a course of treatment (excluding wastage) is [redacted]
Patient access scheme (if applicable)	Not applicable

AML, acute myeloid leukaemia; DA, daunorubicin + cytarabine; DNA, deoxyribose nucleic acid; EMA, European Medicines Agency; GO, gemtuzumab ozogamicin; SmPC, summary of product characteristics.

B.1.3 Health condition and position of the technology in the treatment pathway

Acute myeloid leukaemia (AML) is a highly symptomatic, rare disease with high unmet need

- AML is characterized by the overproduction of blasts, which by overcrowding the bone marrow also prevent normal production of red blood cells and platelets.
- AML is a rare disease and has orphan status as per the European Medicines Agency (EMA) designation, with an incidence of 5.2 per 100 000 in England. Based on data from 2015, there were 2471 new cases of AML in England. Owing to the fact that 73.6% of patients are likely to have de novo AML, of these new cases, 1819 would have been expected to have de novo AML in 2015.
- AML is primarily a disease of the elderly, with incidence rising gradually from 40–44 years of age and then more steeply from 55–69 years of age.
- There is high unmet need in AML: patients with AML have poor prognosis with regard to overall survival, and survival is strongly related to age, being worse in those aged ≥ 65 years than those aged 25–64 years (5-year overall survival: 6% vs 41%).
- Patients with AML have worse health-related quality of life (HRQoL) compared with the general population. This may relate to its high symptomatic burden and poor prognosis, which can impact on patients' lives.
- AML imposes a high economic burden owing to treatments, hospitalisations and management of adverse events from treatment and symptoms. The cost of AML increases in those who relapse owing to the cost of reinduction therapy and further hospitalisations.

The aim of treatment in AML is to achieve and maintain complete remission (CR). The standard approach in the UK in order to achieve CR is intensive chemotherapy with daunorubicin plus cytarabine (DA)

- The duration of first CR is positively correlated with survival. Patients who relapse have poor prognosis with regard to achieving a second CR and survival. Patients who have been in CR for 3 years have little risk of relapsing. UK clinical experts consider patients to be 'functionally cured' and to have a risk of death similar to the general population when they have been in CR for 3–5 years. Additionally, patients who achieve CR have better HRQoL than those who do not achieve CR.
- Cytogenetics profile is the most powerful prognostic factor for predicting response to treatment and the durability of response. Patients with favourable/intermediate cytogenetics profile have a better prognosis than those with unfavourable cytogenetics profile with regard to treatment response, risk of relapse and survival.
- There have been no new approved therapies for patients with AML in the past 40 years; as a consequence, more than 80% of patients in the UK are treated in the context of clinical trials so that they can access more innovative therapies, such as gemtuzumab ozogamicin (GO) that can improve their long-term outcomes.

There is no NICE guidance informing clinical pathway of care in AML. British

Committee for Standards in Haematology (BCSH) guidelines recommend that most patients should be enrolled in ongoing clinical trials. This is a clear indication of the unmet need in this disease area and the need for innovative medicines such as GO to be recommended as part of routine care.

- There is good agreement between BCSH guidelines and UK clinical practice (as described by UK clinical experts). If patients are not enrolled in trials, they should receive intensive chemotherapy with DA. HSCT should only be used in patients who are at high risk of relapses (e.g. those with unfavourable cytogenetics profile) once they have achieved CR, or can be used in patients who relapse or who do not achieve CR following salvage therapy.
- Use of GO has evolved, informed by clinical trials. ALFA-0701 is the pivotal trial on which regulatory submissions have been based. The dose and dosing schedule of ALFA-0701, which forms the primary evidence of this submission, is consistent with the expected EMA marketing authorization for GO.
- The proposed position of GO is alongside DA. GO + DA is therefore not expected to displace first-line treatments. However, greater use of GO +DA may displace second-line treatments by avoiding relapses.

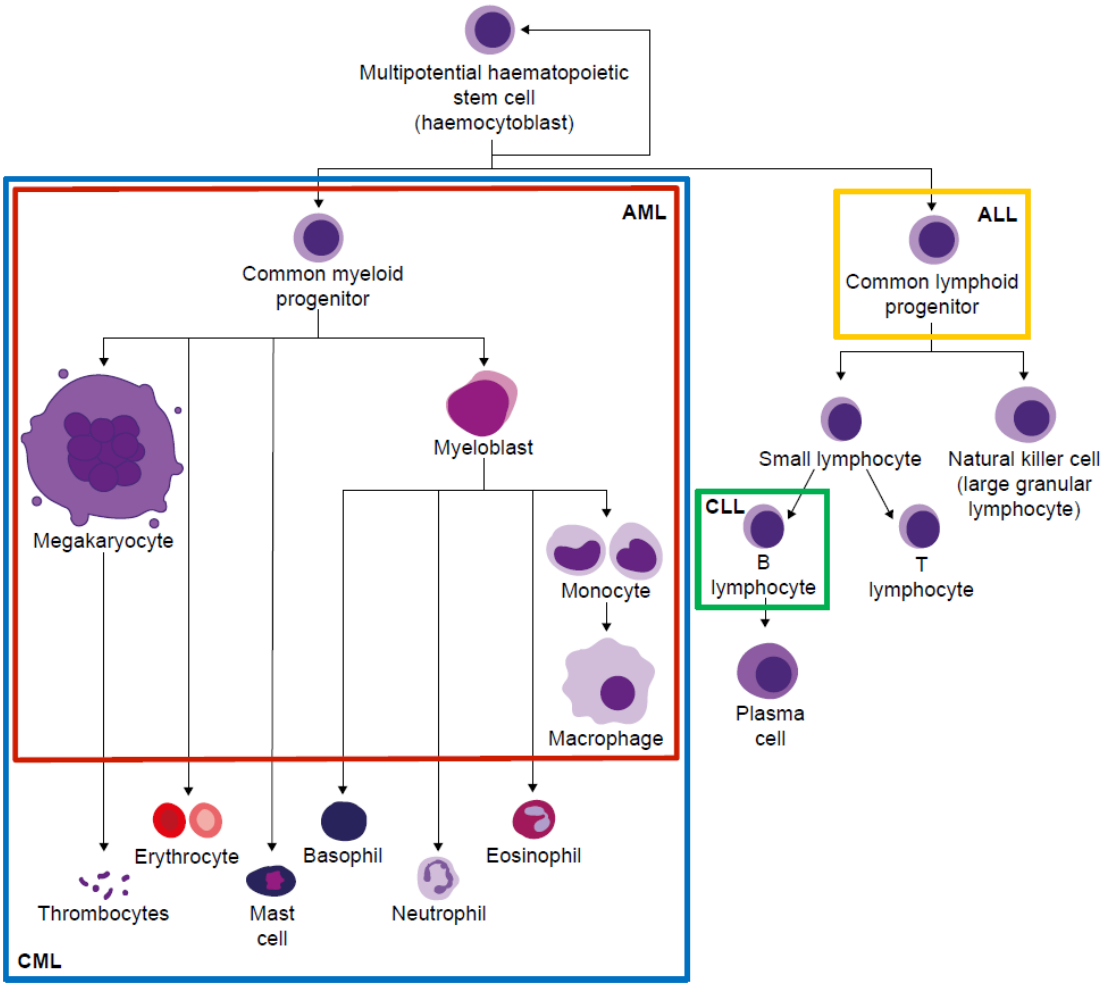
B.1.3.1. Disease overview

AML is a haematopoietic malignancy characterized by the rapid proliferation of immature, non-functional myeloblasts

Leukaemia is a type of blood cancer that originates in the bone marrow.[13,14] It is characterized by abnormal differentiation of haematopoietic stem cells and subsequent clonal over-proliferation of blood cells that cannot properly mature.[13,15] This submission is concerned with acute myeloid leukaemia (AML), which involves overproduction of immature granulocytic and monocytic white blood cells, known as blasts.[4,13,14] Blasts cannot function as mature, healthy white blood cells, and by overcrowding the bone marrow they prevent normal production of red blood cells and platelets.[13,14,16] Cells affected in AML and in other leukaemias are shown in Figure 1.

As AML progresses, blasts can also infiltrate other organs, including the spleen, liver, skin, lymph nodes, bones and central nervous system.[16] The most common causes of death in AML are bone marrow failure and fatal infiltration of the disease into organs, commonly the lungs and the brain.[16]

Figure 1. Haematopoietic cells affected in AML, CML, ALL and CLL



Source: Hjelle *et al*. 2010;[17] CRUK.[18,19]

ALL, acute lymphoblastic leukaemia; AML, acute myeloid leukaemia; CLL, chronic lymphocytic leukaemia; CML, chronic myeloid leukaemia.

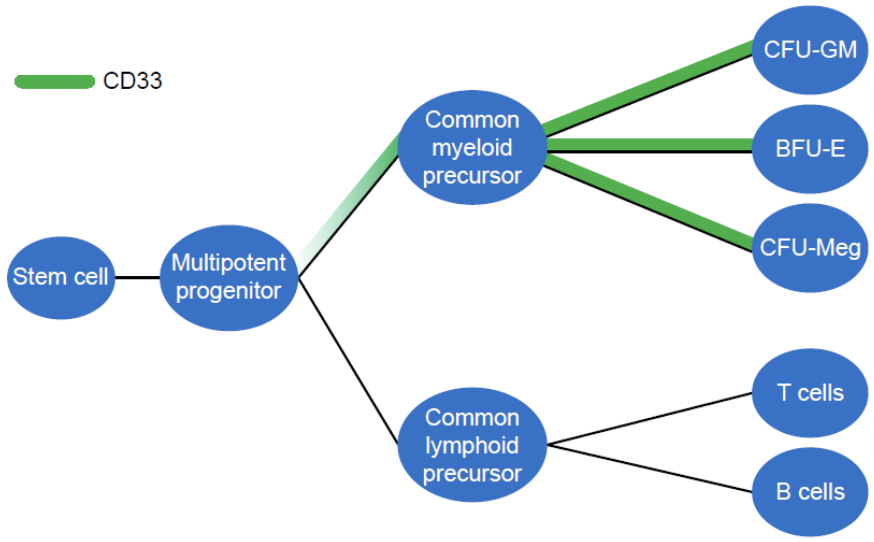
Based on a large population-based study conducted in Sweden, AML was found to arise de novo in 73.6% of patients.[20] AML can also arise secondarily, either owing to the progression of pre-existing haematological disease or following therapy used to treat unrelated malignancies (therapy-related AML).[16]

In the UK, AML is diagnosed using blood tests, bone marrow examination (to determine cell morphology and degree of bone marrow infiltration with disease), immunophenotyping (to determine cell lineage, e.g. whether AML blasts are CD33 positive), cytogenetics profile and molecular genetics.[1,14,21] For an AML diagnosis, patients are required to have a marrow blast count of ≥ 20% (World Health Organization definition).[21,22] In UK clinical practice, patients are classified based on blast morphology (French–American–British [FAB] system), which is considered

alongside results of immunophenotyping and cytogenetic tests.[4,16,22,23] The advantage of the FAB system is its ease of use and the speed with which a diagnosis can be made.

CD33 is a sialic acid-dependent adhesion protein that is highly expressed on the surface of AML blasts and on some leukaemic stem cells.[5,6] CD33 is also expressed to a lower level on multilineage haematopoietic progenitors, myelocytes and some white blood cells, but not on pluripotent haematopoietic stem cells or non-haematopoietic tissue (Figure 2).[2,5,6,10] The function of CD33 is poorly understood.[8] Approximately 85–90% of patients with AML are CD33 positive.[2,7] Treatments directed at CD33-positive cells would target CD33-positive AML blasts while sparing cells and tissues not expressing CD33 (e.g. haematopoietic stem cells and non-haematopoietic cells).[5,9] True leukaemia-specific epitopes have yet to be identified.[8]

Figure 2. CD33 expression on haematological cells[a]



Source: Walter *et al*. 2012.[10]

[a]Green shading shows the level of CD33 expression; darker shading indicates higher CD33 expression.

BFU-E, burst-forming unit-erythroid; CFU-GM, colony-forming unit granulocyte, monocyte; CFU-meg, colony-forming unit megakaryocyte.

AML is a rare disease that predominantly affects older patients

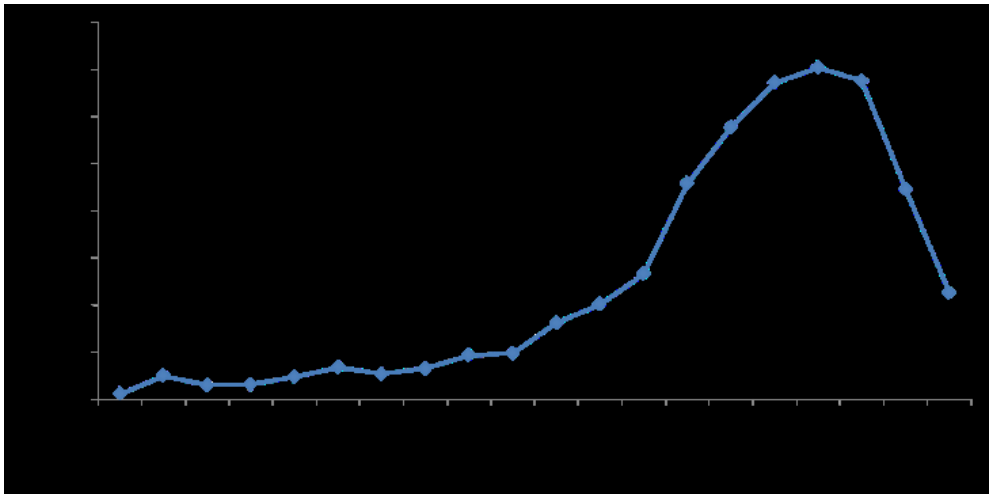
AML is a rare disease according to the European Medicines Agency (EMA) definition (incidence of < 5 per 10 000 people in the EU).[24,25] The total estimated prevalence of AML in the UK based on the observed number of newly diagnosed cases of AML

between 2004 and 2011 and patient survival is 9.6 per 100 000.[26] In the UK in 2014, there were 3100 new cases of AML.[27,28]

The 2014 incidence of AML (European age-standardized) in England was 5.2 per 100 000 population.[27] According to the most recent data from England, in 2015 there were 2471 new cases of AML.[29,30] Owing to the fact that 73.6% of patients have de novo AML, of these new cases, 1819 patients would have been expected to have de novo AML.[20]

AML is primarily a disease of the elderly; according to clinical guidelines and AML trials, older patients are typically defined as those over the age of 60 years.[1,31] Using data from England from 2015, the number of newly diagnosed cases increased with patient age, rising gradually from 40–44 years of age and then more steeply from 55–69 years of age (Figure 3).[30]

Figure 3. Number of newly diagnosed cases of AML in England in 2015, stratified by age



Source: Office for National Statistics.[30]

Survival is worse among older patients with AML (≥ 60 years) than younger patients

In England in 2014 there were 2127 deaths from AML, corresponding to a mortality rate of 4.3 per 100 000.[29,32]

Patients with AML have poor survival. In a study of data from general practices across the UK (collected between 1987 and 2006) patients with AML had a median

survival of 9.5 months.[33] This study did not restrict patients by age, fitness to receive intensive chemotherapy or any other prognostic factors.

When patients in England were stratified by age, 5-year overall survival (OS) was worse in patients aged ≥ 65 years than in those aged 25–64 years (Table 3).[34]

Table 3. Overall survival up to 2010 among patients in England diagnosed with leukaemia in 2008–2010, stratified by age

Age group	Survival (years)	OS (%)
25–64 years	**1**	64
	5	41
65+ years	**1**	20
	5	6

Source: National Cancer Intelligence Network, 2014.[34]
CI, confidence interval; OS, overall survival.

Patients who are eligible to receive intensive chemotherapy regimens may have better prognosis than the overall population of patients with AML, some of whom may not be fit enough to receive intensive chemotherapy. However, age remains the key prognostic factor. In studies of patients receiving intensive chemotherapy, including those in the UK, rates ranged from 41% for 5-year OS in patients aged > 15 years (median age: 49 years) to 20% for 3-year OS in patients aged > 60 years (median age: 67 years) and median OS was 27.5 and 12 months, respectively.[31,35,36]

AML has a high unmet need, with the worst survival of all leukaemias

AML has the poorest survival of all leukaemias: in a study conducted in England, 5-year OS in patients aged ≥ 65 years was lowest in those with AML (Table 4).[34]

Table 4. Five-year overall survival up to 2010 in patients in England aged ≥ 65 years diagnosed with leukaemia from 2008–2010.

Leukaemia	5 year OS (%)
ALL	11%
AML	6%
CLL	53%
CML	35%

Source: National Cancer Intelligence Network, 2014.[34]
ALL, acute lymphoblastic leukaemia; AML, acute myeloid leukaemia; CI, confidence interval; CLL, chronic lymphocytic leukaemia; CML, chronic myeloid leukaemia; OS, overall survival.

B1.3.2. Burden of disease

AML is associated with a high symptomatic burden

Initial presenting symptoms of AML include fever, fatigue, difficulty breathing, weight loss, bruising, bleeding, and aches and pains in the bones and joints.[13,16,37-39] Symptoms can be non-specific or a consequence of bone marrow failure (shortage of red blood cells, white blood cells and platelets) or the presence of AML blasts in the bone marrow, peripheral blood or, infrequently, other organs.[4,13,37,38] Patients who experience a relapse have a return of symptoms caused by the emergence of blasts in the bone marrow, the recirculation of AML blasts in the blood or the development of extramedullary disease.

With regard to the symptomatic burden of AML, a study using the MD Anderson Symptom Inventory scale showed that the most severely rated symptoms in patients were fatigue, disturbed sleep, drowsiness, muscle weakness and dry mouth. Although most symptoms were assessed as mild, patients still reported that AML interfered with their lives (general activities, mood, work and relations with others).[40]

AML can negatively impact on many aspects of patients' QoL and its management can put an emotional strain on caregivers

The high symptomatic burden, poor prognosis and intensive treatment associated with AML have a negative impact on patients' HRQoL, as described in a number of studies presented below.[40-45]

There is evidence that HRQoL in patients with AML is worse than in the general population according to the 30-item European Organisation for Research and Treatment of Cancer QoL questionnaire (EORTC QLQ-C30; $p < 0.05$),[41] the European QoL visual analogue scale (EQ-VAS)[41] and the 12-item Short-Form Health Survey (SF-12) instrument.[42] Based on a study that used these instruments, patients with AML had significantly worse HRQoL scores overall (75 versus 79), and on the physical functioning (80 versus 91), role functioning (75 versus 89), emotional functioning (83 versus 89), cognitive functioning (78 versus 93) and social functioning (82 versus 95) scales versus the general population (all $p < 0.05$).[41]

The goal of treatment in patients with AML is to achieve remission.[16] In two studies, patients who achieved CR had better HRQoL outcomes than those who did not achieve CR.[43,44] In another study, based on responses during interviews, patients reported that uncertainty about long-term remission and fear of relapse were threats to their psychosocial wellbeing.[45] These studies highlight the importance of treating patients with effective therapies that allow them to achieve and maintain remission.

Apart from the impact of AML on patients, the disease can also impact on caregivers. One study has demonstrated the impact of AML on caregivers. In this study, caregivers had significantly higher total scores for post-traumatic stress compared with the patients with acute leukaemia (including those with AML) for whom they were caring ($p < 0.001$); a significantly higher proportion also met the criteria for post-traumatic stress disorder (36.8% versus 18.4%; $p < 0.001$).[46] Post-traumatic stress disorder symptoms in carers were found to be positively correlated with those of patients.[46] The humanistic burden of AML and treatment on caregivers is not captured in the quality-adjusted life-year (QALY) calculation, and therefore the benefit of treatments that improve patient outcomes is potentially underestimated.

AML is associated with a high economic burden, related to the disease itself as well as the cost of treatments and managing side effects

AML has a high economic burden owing to treatment with intensive chemotherapy, long hospitalizations, management of adverse events from treatment, and the symptoms of AML.[47,48]

A study conducted in the UK (based on 2012 costs) demonstrated that during the first 6 months of therapy, per patient, the most costly treatment pathway for AML was haematopoietic stem cell transplantation (HSCT; consisting of one cycle of induction therapy and two cycles of consolidation chemotherapy leading to HSCT). The next most expensive was intensive chemotherapy alone (consisting of a cycle of induction therapy then four cycles of consolidation therapy).[47] Cost components that contribute to the economic burden of intensive chemotherapy are in-hospital stays (including stays in the intensive care unit), medical staff costs and use of expensive haemato-oncology resources.[49-51]

The use of an effective intensive chemotherapy regimen early in the care pathway can lead to resource savings for several reasons. The first is by reducing the need for HSCT, which as mentioned above can be expensive. Furthermore, maintaining a durable remission and avoiding relapses confers economic benefit by avoiding the costs of reinduction therapy to achieve a second remission.[49] Indeed, a Swedish study (1995 costs) of patients with newly diagnosed AML aged < 65 years found that the costs associated with treating patients who relapsed were more than those incurred achieving the first remission.[48] In this study, the major cost of treating a relapse was for reinduction therapy.[48] The cost premium of treating a relapse has also been reported in a recent US study of patients with newly diagnosed AML.[49]

Indirect costs associated with AML remain poorly studied. A Swedish study (1992 costs) reported that the biggest contributor to indirect costs in AML (91% of indirect costs), was production losses owing to premature mortality.[52] Indirect costs are not captured in the QALY calculation. Therefore the value of treatments that improve patient outcomes will be underestimated.

B.1.3.3. Treatment overview

The aim of treatment in AML is to achieve and maintain remission

The aim of treatment in AML is to achieve and maintain complete remission (CR), which is the only outcome that leads to an extension of survival.[16]

The primary goal of treatment in AML is not typically to bridge patients to HSCT, unless they are at high risk of relapse. This is because patients who can achieve CR with standard treatment can maintain long-term disease-free survival without HSCT.

Treatment in AML comprises two phases: induction followed by consolidation in those who achieve CR

AML is treated in two phases: the induction phase followed by the consolidation phase for those who achieve CR.[16,53] The goal of the induction phase in AML is to clear the bone marrow of all haematopoietic cells, regardless of whether they are normal cells or AML blasts.[4,15,16] This allows the bone marrow to repopulate with healthy cells in order to establish normal haematopoiesis and achieve CR.[1,15,16,21] The goal of the consolidation phase is to increase the durability of remission by

eliminating all remaining disease in patients who have achieved remission following induction therapy.[4,15,16] Patients who do not receive treatment during the consolidation phase will most likely relapse within 4–9 months.[1,16]

The duration of first remission is positively correlated with length of survival

The duration of first remission is positively correlated with survival: 5-year OS was 5% in those whose first CR was ≤ 6 months versus 26% when the duration of first CR was > 18 months.[54] Patients who have been in CR for 3 consecutive years have little risk of relapsing.[16] UK clinical experts consider patients to be 'functionally cured' when they have been in CR for 3–5 years, meaning that these patients' mortality risk can be considered equivalent to that of the general population.

The standard approach to treatment in AML is intensive chemotherapy with DA, which has not changed for the past 40 years

For the past 40 years, the standard approach to treating AML has been with an intensive chemotherapy regimen in patients who are able to tolerate and are considered eligible for this form of treatment.[1,55,56] According to the UK British Committee for Standards in Haematology (BCSH) guidelines, patients over the age of 60–65 years should be considered for intensive chemotherapy if they have a good performance status score (World Health Organization grade 0–2), white cell count < 100 × 10^9/L, normal organ function, lack of unfavourable cytogenetics profile and a lack of multidrug resistant gene expression.[1]

As recommended in the BCSH guidelines, and validated by UK clinical experts, the intensive chemotherapy regimen used in the UK consists of daunorubicin (or another anthracyclin) and cytarabine (DA), which are administered over 3 and 10 days, respectively (known as the 3 + 10 regimen).[1,36] According to European AML guidelines, each respective treatment can also be administered over 3 and 7 days, respectively (3 + 7 regimen).[4,16,55,57] There are no known trials comparing the 3 + 7 and 3 + 10 regimens, but according to UK clinical experts no difference in outcomes between the 3 + 7 and 3 + 10 treatment regimens is expected.[1,36]

For patients not able to receive intensive chemotherapy, treatment with non-curative, low intensity chemotherapy can be considered. In the UK this comprises low-dose cytarabine (recommended by the BCSH), hydroxycarbamide or azacitidine.[1]

According to the BCSH guidelines, consolidation therapy can comprise the same treatments used during intensive induction therapy.[1] HSCT can also be used as part of consolidation therapy.[16,55]

HSCT is not a standard goal of treatment in AML in patients who achieve CR unless they are at high risk of relapse (e.g. those with unfavourable cytogenetics profile)

Although HSCT has curative potential, it is associated with increased risk of initial mortality and transplantation-related morbidity.[15,16,29,58] There is evidence to suggest that patients with a high pre-transplantation comorbidity burden are at higher risk of HSCT-related morbidity and mortality than those with a low comorbidity burden.[59] Older patients, who make up the majority of those diagnosed with AML, are at particularly high risk of transplantation-related morbidity and mortality with standard myeloablative conditioning HSCT.[1,57,60] This may relate to the fact that these patients are more likely to have comorbidities that may make them fundamentally unsuitable for HSCT.[4,61,62] However, older patients may still be eligible for and benefit from consolidation with intensive chemotherapy. Therefore, for those who can maintain remission with and are receiving chemotherapy, HSCT offers no additional survival benefit and consolidation with intensive chemotherapy alone remains the mainstay of treatment.[16,47,57,60,63,64] In contrast, for those at high risk of relapse (e.g. those with unfavourable cytogenetics profile) an early decision for transplantation should be made, preferably as soon as first CR has been achieved.[22,60,64]

Patients who do not achieve CR or who relapse can receive salvage therapy with the intention of bridging to HSCT, or palliative non-curative chemotherapy

Patients with AML who do not achieve remission following induction therapy, or who relapse following remission, may be able to receive salvage therapy.[57] Salvage therapy is given with the intention of bridging patients to HSCT, with the overall goal of achieving CR and then performing HSCT.[57] For patients who do not achieve CR after salvage therapy, or if HSCT/salvage therapy is not an option, which can be for various reasons (e.g. a high comorbidity burden),[59] palliative, non-curative

chemotherapy can be offered in order to lessen symptoms and improve quality of life (QoL).[65] In the UK, this can be in the form of azacitidine for a subgroup of patients who cannot receive HSCT, or best supportive care with hydroxcarbamide.[1,66]

Details of treatments and the pathway of care are provided in Section B.1.3.4.

Cytogenetics profile is a powerful prognostic factor in predicting response to treatment; patients with favourable/intermediate cytogenetics profile have better outcomes than those with unfavourable cytogenetics profile

AML is a heterogeneous disease, being accompanied by a diverse range of chromosome abnormalities, gene mutations and changes in gene and micro-ribonucleic acid expression. These AML-related factors may contribute to the pathophysiology of AML and can be used to categorize patients into prognostic groups with regard to response to induction therapy (remission, relapse, survival) and treatment-related mortality.[4,13,16,21,62]

Chromosomal abnormalities, as detected by cytogenetics profile, are the most powerful prognostic factor in predicting response to induction therapy.[16,21] Cytogenetic abnormalities have been identified in approximately half of all patients with newly diagnosed AML.[16,67] The majority of cases are associated with non-random chromosomal translocations that can result in gene rearrangements.[16] Based on diagnostic karyotyping, patients can be characterized as having favourable, intermediate or unfavourable cytogenetics profile according to the types of abnormalities that are present.[16] The incidence of unfavourable cytogenetic abnormalities increases with increasing age.[16] It should be noted that not all patients will receive a cytogenetic profile classification. These patients are classified as having unknown cytogenetics.

Patients with favourable/intermediate cytogenetics profile have a better prognosis than those with unfavourable cytogenetics profile with regard to CR, incidence of relapse, and survival.[4,62,68] A study of 1213 patients (aged 15–86 years) with de novo AML showed that compared with those with favourable/intermediate cytogenetics profile, patients with unfavourable cytogenetics profile had a lower probability of achieving CR (11.9- and 4-fold, respectively) and a higher risk of relapse (4.4- and 3.0-fold, respectively) and death (4.3- and 2.1-fold, respectively).[68] In another study

of 1344 previously untreated patients (aged 16–88 years), those with favourable/intermediate cytogenetics profile had higher rates of CR (favourable: 71%; intermediate: 46%; unfavourable: 42%) and 4-year OS (49%, 27%, 9%, respectively) than those with unfavourable cytogenetics profile.[69]

The particular cytogenetic profile classification that a patient receives is based on the system that is used to define their karyotype (Appendix L1.1).[67] The revised MRC classification is based on cytogenetic abnormalities, and stratifies patients into favourable, intermediate or adverse (in this submission, referred to as 'unfavourable') cytogenetics profile.[70] In UK studies investigating GO +DA in patients with AML, between 59% and 69% of patients in the trials had favourable or intermediate cytogenetics profile according to the revised MRC cytogenetic classification.[31,35,36] The European LeukemiaNet (ELN) classification recommends that both cytogenetic and genetic abnormalities are taken into account when assigning patients to prognostic groups.[21,22] According to the ELN classification, patients can be categorized as having favourable, intermediate or adverse (in this submission, referred to as "unfavourable") cytogenetics profile depending on the combination of cytogenetic and genetic alterations that are present.

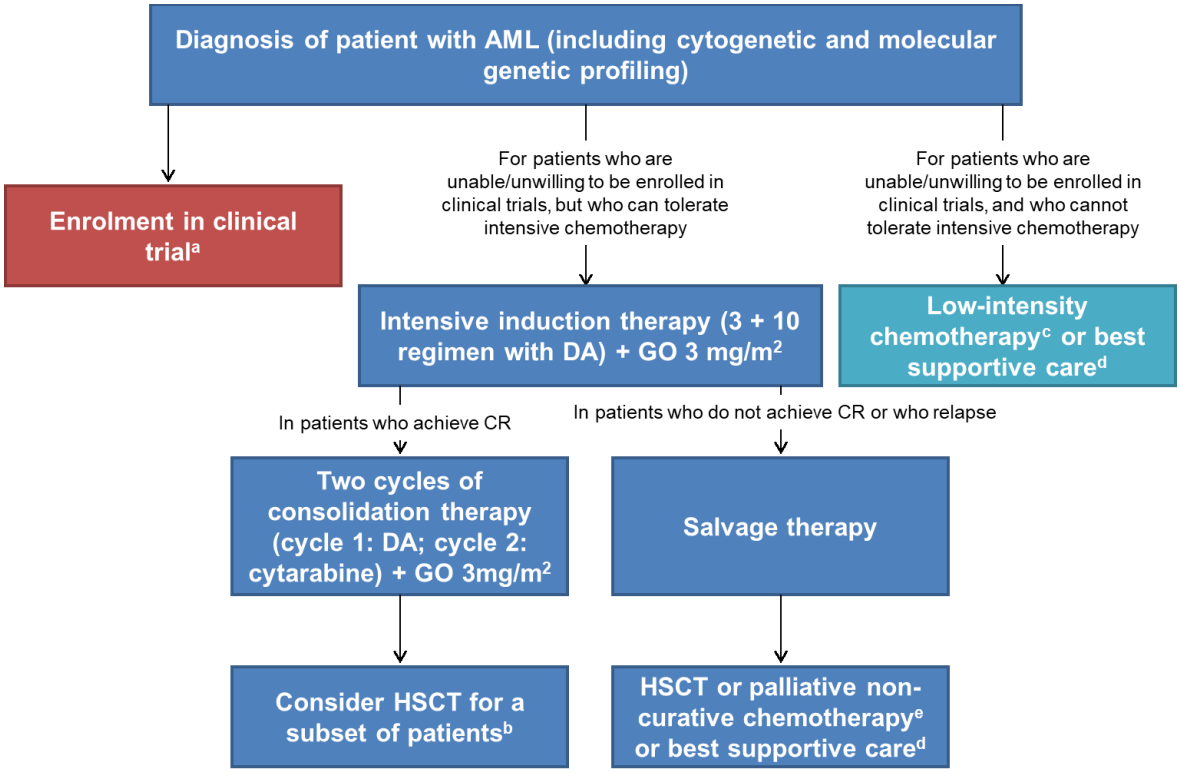
One such genetic alteration that has prognostic impact is the *FMS-like tyrosine kinase 3* (*FLT3*) internal tandem duplication (ITD), which is associated with unfavourable prognosis, particularly when the mutant-to-wild type ITD allelic ratio is high (≥ 0.5).[22,55] The prognostic impact of *FLT3* ITD should also be considered in the context of other mutations that may be present.[22,55] For example, *FLT3-ITD* often co-occurs with *nucleophosmin-1* (*NPM1*) mutations and the extent of their prognostic impact is determined by the expression level of *FLT3* ITD and the mutation status of *NPM1*.[22,55,71] Midostaurin, a kinase inhibitor under appraisal by NICE (section B.1.3.4), is active in patients with the *FLT3* mutation, but does not distinguish between mutations in other genes that are required for the *FLT3* mutation to be prognostic.

Other factors that can be prognostic of patient outcomes include age (e.g. < 60 years versus ≥ 60 years) and Easter Cooperative Oncology Group performance status score and relapsed disease (Appendix L1.2).[16,72]

B.1.3.4. Gemtuzumab ozogamicin (GO) and the clinical pathway of care

In the absence of NICE guidance, the UK clinical pathway has been based on available guidelines and information from UK clinical experts (Figure 4).

Figure 4. Clinical pathway for the treatment of patients with de novo AML and the positioning of GO + DA in the clinical pathway



Source: BCSH 2006;[1] EBMT guidance; NICE TA399[66]

[a]Guidelines recommend that patients are first enrolled in a clinical trial, but for those who are unable or unwilling to participate in trials, intensive chemotherapy with DA should be offered. [b]HSCT is offered as a treatment option in patients with unfavourable cytogenetics profile after first CR. It is also considered in patients with intermediate cytogenetics profile who are fit for transplantation and have a suitable donor, but not in patients with favourable cytogenetics profile, for whom transplantation offers no additional benefit over chemotherapy; [c]Low-intensity chemotherapy options include low-dose cytarabine, azacitidine or hydroxycarbamide;[1,66] [d]Best supportive care options include transfusion support and hydroxycarbamide to control white cell counts;[1] [e]Palliative non-curative chemotherapy can include azacitidine[66]

AML, acute myeloid leukaemia; CR, complete remission; DA, daunorubicin + cytarabine; GO, gemtuzumab ozogamicin; HSCT, hematopoietic stem cell transplantation.

Enrolment in clinical trials is standard practice in the UK because the alternative treatment options have limited ability to keep patients in remission. This is a clear indication of the unmet need in AML and the need for innovative medicines, such as GO, to be recommended as part of routine care

The BCSH guidelines recommend that patients with de novo AML be asked to participate in clinical trials.[1] This is a clear indication of the unmet need in AML with regard to the availability of innovative medicines in routine clinical practice that can improve patient outcomes. In the absence of a suitable trial, or for patients who are unable or unwilling to participate, the BCSH guidelines recommend induction chemotherapy with DA (using a 3 + 10 or 3 + 7 regimen).[1] Patients who CR with induction therapy should receive consolidation therapy in order to maintain remission.[1] For patients who cannot receive intensive chemotherapy, low-intensity chemotherapy or best supportive care should be offered (transfusion support and hydroxycarbamide to control white cell count).[1]

Enrolment in clinical trials is standard practice in the UK because the alternative treatment option with intensive chemotherapy has only limited ability to maintain patients in remission (Appendix L1.3). As new evidence from trials becomes available, clinicians adapt their practice. Therefore, the treatment regimen used in clinical trials has the potential to impact on the clinical pathway.

Use of GO is evolving as new evidence becomes available; decisions on its use in clinical practice should be evidence based. Based on ongoing and completed trials, it is recommended that GO 3 mg/m^2 be administered alongside DA as part of induction and consolidation therapy

Recent trials open to patients with AML have investigated GO + DA as a first-line therapy in combination with standard intensive induction chemotherapy with DA, and have also investigated dosing of GO (Appendix L1.4). These trials are being conducted globally, with some also recruiting patients from the UK. Recently, a French study, ALFA-0701, was completed, which assessed fractionally dosed GO 3 mg/m^2 administered in combination with DA as part of induction and consolidation therapy (five fractionated doses in total). ALFA-0701 is the pivotal trial on which regulatory submissions have been based. Data from ALFA-0701 are presented in section B2 using a dose and dosing schedule that is consistent with the expected

EMA marketing authorization. Data from ALFA-0701, as presented in section B2, indicate that GO + DA is associated with improved efficacy and a manageable safety profile (including a reduced risk adverse events, specifically veno-occlusive disease [VOD] compared with trials that used higher doses of GO).

According to the clinical pathway of care (Figure 4), and as per AML trials, including ALFA-0701, it is recommended that GO is administered in fractionated doses as part of induction and consolidation therapy alongside DA. DA is the current first-line standard of care in the UK in patients not enrolled in trials. In line with this, patients with AML who are not eligible to receive DA would not be eligible to receive GO and therefore DA is the main comparator against which GO + DA should be assessed.

Use of GO alongside DA will displace second-line therapies by avoiding relapses

The approval of GO + DA as per the indication in this submission would be consistent with its current use and no first-line therapies would be displaced. However, greater use of GO + DA as part of routine clinical practice will displace second-line treatments for some patients, by avoiding relapses and extending the duration of remission versus DA alone. In ALFA-0701, the primary source of data for this submission, GO + DA during induction and consolidation therapy significantly improved event-free and relapse-free survival over 2 years in patients with untreated, de novo AML versus DA alone.

There is good agreement between the BCSH guidelines and UK clinical practice, as described by clinical experts in the UK

An expert panel of three UK clinicians was convened in February 2017 by Pfizer to provide clarity on existing clinical practice for the treatment of AML in the UK. They stated that approximately 80% of UK patients with AML enter clinical trials. Outside a trial, standard first-line treatment was considered to be DA in patients eligible for intensive chemotherapy. It was stated that FLAG-Ida was unlikely to be prescribed unless patients had unfavourable cytogenetics profile and were fit enough to tolerate this form of treatment. It was their opinion that patients who could maintain CR with chemotherapy would receive no additional benefit from HSCT. They estimated that approximately half of patients who achieved a first CR would receive HSCT: this includes all patients with unfavourable cytogenetics profile who are fit. Patients who

relapse and then achieve a second CR receive HSCT as standard of care. No other treatments were mentioned as being used in UK clinical practice.

According to NICE guidance, adults with blood and bone marrow cancers who receive high-intensity chemotherapy should be treated at a haematology unit and managed by a multidisciplinary team, which includes haemato-oncologists, haematopathologists, nurses, palliative specialists and support staff as core team members.[73] The team should be involved in meetings with patients, where their care is discussed, and should be responsible for initial recommendations about care and also delivery of treatment and long-term support.[73] There should be sufficient provisions and levels of staffing to rapidly assess and manage potentially life-threatening complications.[73]

The principal European guidelines for the management of AML are provided by the ELN, published in 2017, and the European Society for Medical Oncology (ESMO), published in 2013.[22,74] These are similar to the UK guidelines and, although evidence for GO is summarized, neither guidelines addresses GO in its recommendations.

For HSCT, the European Society for Blood and Bone Marrow Transplantation handbook provides some guidance for patients with AML.[60] According to the revised 2012 guidance, patients should be stratified according to molecular aberrations and/or ELN risk stratification (molecular and cytogenetic): patients with unfavourable risk should be offered HSCT and those with favourable risk should receive consolidation therapy.[60] According to the ELN guidelines, those with intermediate risk should receive HSCT; those with favourable risk can receive HSCT or consolidation therapy.[60] According to ESMO guidelines, HSCT is recommended in those with intermediate/unfavourable risk.[74]

Three other treatments are under consideration by NICE for the treatment of AML, but these are not relevant in this appraisal because they target different populations or they will not be standard of care at the time of this submission

NICE has considered azacitidine, recommending its use for adults not eligible for HSCT with 20–30% blasts and multilineage dysplasia (TA218).[75] This is not a comparator for GO + DA in this appraisal because GO + DA is positioned for patients with de novo AML, whereas azacitidine is also positioned for patients with secondary

AML. In addition, two AML appraisals are ongoing. The first is for midostaurin (ID894) for patients with untreated *FLT3* mutation- or *FLT3* ITD-positive AML. This is a subgroup of the population that would be eligible to receive GO. However, this appraisal will not be completed and midostaurin will not be a standard of care at the time of this submission.[76,77] The second is for decitabine (ID1114), which has marketing authorization for adults with newly diagnosed de novo or secondary AML, who are not candidates for standard induction chemotherapy;[78] GO + DA is indicated in the population of patients who can receive standard chemotherapy, and therefore will not overlap with the population eligible to receive decitabine.

There is high unmet need in AML: outcomes following treatment with DA remain poor, particularly with regard to long-term survival and disease-free survival

Current intensive chemotherapy options do not fully address the unmet need in AML. While rates of remission can be as high as 69% following treatment with DA in patients with newly diagnosed or previously untreated AML (Appendix L1.3),[16,79-83] rates of long-term survival[79-83] and disease-free survival remain poor.[82] For example, the rate of 5-year OS ranged from 2% in patients aged ≥ 60 years to 48% in patients aged 15–64 years. Furthermore, response to DA was not was not durable, with one study reporting an EFS of 3.4 months in patients aged ≥ 60 years and other studies reporting high rates of relapse.[79,81] Patients who relapse face a low chance of achieving a second CR with existing treatments.[84,85] This highlights the need for an AML treatment that is associated with durable, long-term remission that can increase the chance of patients being 'functionally cured'.

B.1.4 Equality considerations

There are no equality considerations expected for GO.

B.2 Clinical effectiveness

Direct head-to-head evidence from the pivotal study, ALFA-0701, demonstrates the clinical benefit of adding gemtuzumab ozogamicin (GO) to daunorubicin plus cytarabine (DA) compared with DA alone for the first-line treatment of de novo AML. Event-free survival (EFS) and relapse-free survival (RFS) were significantly improved in patients receiving GO + DA versus DA alone. These improvements in RFS translated into durable responses in a higher proportion of patients in the GO + DA arm than in the DA arm ([redacted] versus [redacted]; observed from 3 years in patients with favourable/intermediate cytogenetics).

- ALFA-0701, a phase 3 randomized controlled trial (RCT), provides evidence in a population of patients with AML that includes those with favourable/intermediate cytogenetics profile, as per the population stated in the decision problem. Efficacy outcomes were reported separately for patients with favourable/intermediate cytogenetics profile, owing to the fact that this is the population relevant for decision making. This trial provides evidence across a total of 271 patients randomly assigned either to the GO + DA arm (n = 135) or the DA arm (n = 136) in the modified intention-to-treat (mITT) population and in patients with favourable/intermediate cytogenetics profile (GO +DA arm: n = 95; DA arm: n = 94).

- Adding GO to DA led to a statistically significant and clinically meaningful [redacted] improvement in the primary endpoint of EFS in both the mITT population and in the subpopulation with favourable/intermediate cytogenetics profile, and in both populations [redacted] more patients were event free at 3 years with the addition of GO. After 3 years, the Kaplan-Meier (KM) curve started to plateau (at approximately [redacted] in the GO + DA arm and [redacted] in the DA arm), indicating the durability of the response. Results below are according to independent review committee (IRC) analysis.

    - In the mITT population median EFS was significantly longer in the GO + DA ([redacted]) versus DA ([redacted]) arm ([redacted]).
    - The EFS advantage in the GO + DA versus the DA arm was most apparent in the subpopulation of patients with favourable/intermediate cytogenetics profile ([redacted]). There was no significant difference in median EFS in patients with unfavourable cytogenetics profile.
    - At 3 years, the proportion of patients who were event free was higher in the GO + DA versus the DA arm in the mITT population [redacted]) and the subpopulation with favourable/intermediate cytogenetics profile ([redacted] vs [redacted]).
    - Results according to IRC analyses were in line with those of investigator analyses.

- In patients in complete remission (CR) or CR with incomplete platelet recovery (CRp), adding GO to DA led to a significant improvement in RFS of at least [redacted] in both

the mITT population and in those with favourable/intermediate cytogenetics profile, and in both populations [redacted] more patients were event free at 3 years with the addition of GO. After 3 years, the KM curve started to plateau (at approximately [redacted] in the GO + DA arm and [redacted] in the DA arm), indicating the durability of the response. Results below are according to IRC analysis.

  - In the mITT population, median RFS was significantly longer in the GO + DA ([redacted]) versus DA ([redacted]) arm ([redacted]).
  - The RFS advantage was most apparent in patients with favourable/intermediate cytogenetics profile ([redacted]). There was no significant difference in median RFS in patients with unfavourable cytogenetics profile.
  - At 3 years, the proportion of patients who were relapse free was higher in the GO + DA arm than in the DA arm in both the mITT population ([redacted]) and the subpopulation with favourable/intermediate cytogenetics profile ([redacted]).
  - Results according to IRC analyses were in line with investigator analyses.

- In support of the RFS results, in an ad hoc analysis, median time to subsequent anti-cancer therapy after induction failure or relapse was significantly delayed in the GO + DA arm versus the DA arm by [redacted].

  - Median time to subsequent anti-cancer therapy administered after induction failure or relapse was significantly longer in the overall GO + DA arm ([redacted] months; 95% CI: [redacted]) compared with the DA arm ([redacted] months; 95% CI: [redacted]; HR: [redacted]).

- In the mITT population, adding GO to DA led to a numerical extension in OS of [redacted] compared with DA alone that trended towards significance; the numerical extension in OS in the subpopulation with favourable/intermediate cytogenetics profile was aligned with and more apparent than that of the mITT population ([redacted]). After approximately 3 years, the OS risk started to plateau (at approximately [redacted] in the GO + DA arm and [redacted] in the DA arm)

  - In the mITT population, there was a trend towards longer median OS in the GO + DA arm ([redacted]) versus the DA arm ([redacted]). ALFA-0701 was not powered for OS and did not reach statistical significance ([redacted]).

  - A trend toward longer OS following treatment with GO + DA versus DA alone was more apparent in patients with favourable/intermediate cytogenetics profile ([redacted]) than in the mITT population. There was no difference in OS in patients with unfavourable cytogenetics profile between the

treatment arms.

- In an individual patient data (IPD) meta-analysis that included a larger population (i.e. five studies, including ALFA-0701, n = 3331 patients in total) than in ALFA-0701, patients who were randomised to GO had a significant OS advantage versus those randomized to a comparator arm without GO (no-GO) in the first-line treatment of AML (Appendix D.3)

- Adding GO to DA led to a numerical increase in the proportion of patients who achieved CR/CRp. However, this did not reach statistical significance. Results below are according to IRC analysis.
  - In the mITT population, [redacted] of patients in the GO + DA arm achieved CR/CRp after induction therapy versus [redacted] in the DA arm. In patients with favourable/intermediate cytogenetics profile, the corresponding proportions with an overall response were [redacted] and [redacted].
  - Results according to IRC analyses are in line with investigator analyses.
- In ALFA-0701, GO in combination with DA was generally associated with manageable and reversible adverse events, consistent with the known safety profile of each of the individual agents. Trends for treatment-related adverse events in the subpopulation with favourable/intermediate cytogenetics profile were consistent with the overall population.

B.2.1 Identification and selection of relevant studies

To assess the comparative efficacy and safety of GO + DA versus DA in previously untreated de novo AML, a systematic literature review (SLR) was conducted to identify RCTs and non-RCTs.

Full details of the methodology used to identify and select the RCT and non-RCT clinical evidence relevant to the technology being appraised are reported in Appendix D.1. A full summary of the included and excluded studies, including the PRISMA flow diagram and reasons for exclusion, are also provided in Appendix D.1.

B.2.2 List of relevant clinical effectiveness evidence

The SLR identified eight RCTs assessing clinical efficacy and safety of GO + DA vs DA alone in untreated patients with AML: ALFA-0701, MRC AML15, NCRI AML16, MRC AML17, AML18, AML19, SWOG S0106 and GOELAMS AML 2006 IR (Appendix D.1.1.4).[31,36,86-94]

ALFA-0701 is the pivotal study used to support the European Medicines Agency (EMA) marketing authorisation (Table 5). ALFA-0701 compares fractionated GO 3 mg/m² (days 1, 4 and 7 during induction therapy and day 1 for two cycles of consolidation therapy; up to 5 mg per dose) in combination with DA (3 + 7 regimen during induction) versus DA alone (3 + 7 regimen during induction) in patients with untreated, de novo AML.[89] It was therefore considered the most appropriate primary source of evidence for this submission and informed the cost-effectiveness analysis.

All other RCTs identified in the SLR used dosing regimens of GO that were not considered for approval by the EMA. These studies cannot therefore be considered as primary sources of evidence for this submission and are only supportive. However, four of these trials were assessed as pooled evidence in an IPD meta-analysis alongside ALFA-0701, which was used as supportive evidence for the regulatory submission (ALFA-0701, MRC AML15, NCRI AML16, MRC AML17, SWOG S0106 and GOELAMS AML 2006 IR).[35] AML18 and AML19 are still ongoing and were not included in the IPD meta-analysis.

With the exception of ALFA-0701, the dose and dosing schedule of GO in the other studies would reflect off-label use under the expected marketing authorisation for GO, and were not considered for the economic model in this appraisal.

No relevant studies presenting data on non-RCTs were identified through the SLR (Appendix D.1.2).

Table 5. Clinical effectiveness evidence for ALFA-0701

Study	ALFA-0701 (NCT00927498)
Study design	Randomized, phase 3, open-label
Population	Patients with previously untreated de novo AML (50–70 years old)
Intervention(s)	Induction treatment (in combination with DA as outlined below in comparator section) o GO (3 mg/m^2 [maximum dose 5 mg] intravenously over 2 hours, days 1, 4, 7) First and second consolidation courses (in combination with DA as outlined below in comparator section) o GO (3 mg/m^2 on day 1)
Comparator(s)	DA

	Induction treatment: DA (3 + 7) ○ Daunorubicin (60 mg/m^2 intravenously, days 1–3) ○ Cytarabine (200 mg/m^2 as continuous infusion over 7 days) Second induction course (if there were more than 10% persistent leukaemic blasts): ○ Daunorubicin (60 mg/m^2 per day for 2 days) ○ Cytarabine (1000 mg/m^2 per 12 hours, infused over 2 hours for 3 days) ○ Lenograstim 263 µg intravenously until neutrophil recovery ○ With no GO addition First and second consolidation courses: ○ Daunorubicin (60 mg/m^2 for 1 day in first course or 2 days in second course) ○ Cytarabine (1000 mg/m^2 per 12 hours, infused over 2 hours on days 1–4)		
Indicate if trial supports application for marketing authorisation	Yes	**Indicate if trial used in the economic model**	Yes
Rationale for use/non-use in the model	Pivotal study used for marketing authorisation application		
Reported outcomes specified in the decision problem	• Event-free survival [time frame: relapse or death measured from randomization] • Overall survival [time frame: survival from randomization] • Relapse-free survival (disease-free survival) • Adverse effects of treatment [time frame: duration of study]		
All other reported outcomes	• Complete remission rate [time frame: complete remission after induction] • Cumulative incidence of relapse [time frame: relapse from complete remission] • Response (event-free survival; relapse-free survival; overall survival and response rate) in the subpopulation with favourable/intermediate cytogenetics profile [time frame: duration of study]		

Source: Castaigne *et al*. 2012;[89] ALFA-0701 CSR.[24]

AML, acute myeloid leukaemia; DA, daunorubicin + cytarabine; GO, gemtuzumab ozogamicin

B.2.3 Summary of methodology of the relevant clinical effectiveness evidence

Results from ALFA-0701 were originally presented by Castaigne *et al*. 2012.[89] Following publication, the study sponsor (Centre Hospitalier de Versailles [CHV]) continued to collect data up to a later cut-off not included in Castaigne ok*et al*. 2012.

The Castaigne *et al*. 2012 publication reports on all 280 patients who were randomized (intention-to-treat population [ITT]).[89] However, when the data were transferred to Pfizer, nine patients' signed informed consent forms were not transferred. The clinical study report (CSR) therefore summarizes data for 271 randomized patients, who make up the modified intention-to-treat (mITT) population.[24]

After data transfer to Pfizer, additional analyses and data collection were conducted. These were not included in Castaigne *et al*. 2012 but are reported in the CSR:

- retrospective data collection and analysis (up to 1 November 2013) to complete the dataset for any regulatory submissions
- blinded analyses of ALFA-0701 efficacy data by an IRC to confirm the results of investigator analyses. As part of the IRC analyses, two independent reviewers evaluated the same cases separately, and a third reviewer resolved any discrepancies.

Data in the CSR are presented for the mITT population according to investigator and IRC analyses. Table 6 summarizes data cut-offs and populations in Castaigne *et al*. 2012 and the CSR.

Table 6. Summary of data availability from Castaigne et al. 2012 and the CSR

	1 August 2011 data cut-off		***30 April 2013 data cut-off***	
	Investigator analysis	***IRC analysis***	***Investigator analysis***	***IRC analysis***
Castaigne *et al.* 2012 (ITT population,	✓	×	×	×

	1 August 2011 data cut-off		*30 April 2013 data cut-off*	
	Investigator analysis	***IRC analysis***	***Investigator analysis***	***IRC analysis***
n = 280)				
ALFA-0701 CSR (mITT population, n = 271)	✓	✓	✓	✓

Source: Castaigne *et al*. 2012;[89] ALFA-0701 CSR.[24]

CSR, clinical study report; IRC, independent review committee; ITT, intention-to-treat; mITT, modified intention-to-treat.

The IRC analyses at the 30 April 2013 data cut-off (minimum length of follow-up of 3 years) are the focus of sections B.2.3–B.2.6, as presented in the CSR.[24] The IRC analyses for the mITT population for the 30 April 2013 data cut-off, as reported in the CSR, are also included in the economic model. The use of the IRC analyses is owing to the fact that these data were generated according to regulatory requirements. The investigator analysis may be more reflective of clinical practice but as shown in Appendix D.2.4 there is good agreement between the IRC and investigator analyses for all outcomes.

B.2.3.1. Trial methodology

The trial methodology of the ALFA-0701 study is summarized in Table 7. Further details of the trial are provided in the sections below.

Table 7. Summary of ALFA-0701 trial methodology

Trial number (acronym)	NCT00927498 (ALFA-0701)
Location	26 haematology centres in France
Trial design	Investigator-sponsored, randomized (1:1 randomization), open-label, phase 3 trial comparing GO + DA versus DA alone in patients aged 50–70 years of age with de novo, untreated AML
Eligibility criteria for participants	*Inclusion criteria* Previously untreated patients with a confirmed diagnosis of AML, aged 50–70 years with normal cardiac function assessed by use of radionucleotide scintigraphy or

	echography, who had an ECOG PS score of 0–3, had blood and bone marrow specimens taken for molecular assessment and had signed the informed consent document (ICD) *Exclusion criteria* Patients with acute promyelocytic leukaemia (APL), previous myeloproliferative or myelodysplastic syndrome, AML secondary to exposure to chemotherapy and/or radiotherapy, AML with central nervous system (CNS) involvement, severe uncontrolled infection, liver or renal dysfunction, other malignant diseases, seropositivity to HIV or hepatitis B or C (except post-vaccination), prior antileukaemia treatment (except hydroxyurea in case of hyperleukocytic leukaemia) or a positive pregnancy test in women of childbearing age
Trial drugs (the interventions for each group with sufficient details to allow replication, including how and when they were administered) **Intervention(s) (n = 135) and comparator(s) (n = 136)[a]** **Permitted and disallowed concomitant medication**	**Induction therapy:** induction treatment with intravenous daunorubicin (60 mg/m² on days 1–3) + cytarabine (200 mg/m² as continuous infusion for 7 days; also known as the 3 + 7 regimen) without (DA arm) or with intravenous fractionally dosed GO (3 mg/m² [maximum dose 5 mg] infused over 2 h on days 1, 4 and 7; GO + DA arm) Patients with > 5% leukaemic blasts in the bone marrow (or > 10%, depending on the protocol amendment) on day 15 were given a second round of induction therapy without GO, irrespective of randomization. If the day 15 bone marrow assessment was indeterminate, assessment could be repeated 7 days later; however, the second induction course could not be initiated later than day 22 **Salvage therapy:** patients who did not achieve a CR after the first course of induction therapy and who did not receive the second course of induction therapy could receive a salvage course, comprising idarubicin (12 mg/m^2; days 1 and 2), cytarabine (1 g/m^2 twice a day; days 1 and 4) and granulocyte colony-stimulating factor (G-CSF) (day 6). To be eligible for salvage therapy, patients needed an ECOG PS score of < 3 and creatinine clearance > 30 mL/min. Patients who did not respond to induction therapy (including salvage course) discontinued study treatment. Patients who achieved CR/CRp after induction or salvage therapy went on to receive two courses of consolidation therapy according to their initial randomization.

	Consolidation therapy (two courses): intravenous daunorubicin (60 mg/m² for 1 day [first course] or 2 days [second course]) in combination with intravenous cytarabine (1000 mg/m² per 12 h, infused over 2 h on days 1–4), without (DA arm) or with intravenous GO (3 mg/m² on day 1 of each course of consolidation therapy; GO + DA arm) **HSCT:** Patients who experienced CR could be considered for allogeneic transplant according to PS, age, the existence or not of a related donor, and cytogenetic and molecular risk categories. Patients with favourable and intermediate-1 cytogenetic and molecular risk categories were not to be sent for transplant in first CR; patients with intermediate-2 or unfavourable cytogenetic and molecular risk categories who experienced a CR were considered for transplant if qualified by other criteria. Permitted and disallowed concomitant medications were not specified
Duration of study	First patient, first visit: 8 January 2008 Last patient, last visit: 30 April 2013 Data cut-off for retrospective data collection: 1 November 2013 Minimum duration of follow-up: 3 years
Protocol amendments	February 2008 (resulting from discussions with study investigators): additional clarity on day 15 bone marrow aspirate (BMA); addition of biological sampling for residual disease assessment, and amending conditions for re-inducing patients (revised from day 15 BMA blasts > 10% to day 15 BMA blasts > 5%); and changing modification of daunorubicin dosing from 60 mg/m^2/day to 35 mg/m^2/day May 2009 (resulting from discussions with study investigators): inclusion of a salvage course and clarifying bone marrow transplantation with respect to the last dose of GO December 2009 (resulting from safety signals observed during course of study): GO discontinuation in cases of persistent thrombocytopenia (platelet count < 1 × 10^{11}/L within 14 days after the planned date of the next treatment

	course)
Method of randomization	After providing informed consent, patients were randomly allocated to one of two treatment arms in a 1:1 ratio using a computer-generated random allocation sequence (R software, version 2.10.1). Randomization was undertaken centrally and communicated via telephone.
Patient stratification	By centre; the 1:1 allocation ratio had block sizes of four
Blinding	Open-label: the study treatment was not blinded to patients or investigators After data were transferred, an IRC at Pfizer performed retrospective analyses to confirm investigator-collected data and to collect additional data as per regulatory requirements. The IRC was blinded to the treatment arms.
Primary outcomes (including scoring methods and timings of assessments)	EFS: time from date of randomisation to date of induction failure, relapse or death due to any causes, whichever came first. Remission was assessed after induction, before each course of consolidation and at 1–2 months after haematological recovery from the last cycle of consolidation therapy. Patients still in remission after haematological recovery were followed every 3 months for 2 years from the start of therapy. For patients still in remission at 2 years, evaluations were extended to every 6 months. Patients were followed until death. EFS was determined by IRC or investigator analyses using definitions provided in Table 4.
Secondary/tertiary outcomes (including scoring methods and timings of assessments)	OS: time from randomization to the date of death due to any cause. A reference date of 30 April was applied to the OS data for final analyses and all deaths occurring after this date were not included. Haematological remission: assessed as the proportion of patients achieving CR/CRp; assessed after induction, before each course of consolidation therapy and 1–2 months after haematological recovery from the last cycle of consolidation therapy. Remission was determined by IRC or investigator analyses using definitions provided in Table 4. RFS: defined for patients experiencing remission as time from the date of remission to the date of relapse or death from any cause, whichever came first. The timings of assessments are described above. RFS was determined by IRC or investigator analyses using definitions provided

	in Table 4. Safety: A retrospective safety data assessment was conducted to collect treatment-emergent AEs (TEAEs) of special interest, based on review of patient medical files. For veno-occlusive disease (VOD), data were collected until the patient's death or the retrospective data cut-off of 1 November 2013, whichever occurred first, in order to identify any late drug toxicity associated with VOD. During retrospective data collection, all events meeting the definition of SAE within 28 days after the last dose of study drug were recorded.
Pre-planned subgroups	Age, WBC count, ECOG PS, CD33 expression, CD33 MFI ratio, NCCN risk classification, ELN risk classification, FLT3-ITD status, NPM1 status, CEBPA status, MLL status, WT1 status, cytogenetics profile as classified by CHV, genotype Subgroups analysis according to IRC, was evaluated in all subgroups with the exception of CD33 MFI ratio.

[a]In Castaigne *et al*. 2012 the size of the patient population was n = 139; comparator, n = 139.

Source: ALFA-0701 CSR;[24] Castaigne *et al.* 2012.[89]

AE, adverse event; AML, acute myeloid leukaemia; APL, acute promyelocytic leukaemia; BMA, bone marrow aspirate; CEBPA, CCAAT/enhancer-binding protein α gene; CHV, Centre Hospitalier de Versailles; CNS, central nervous system; CR, complete remission; CRp; complete remission with incomplete platelet recovery; DA, daunorubicin + cytarabine; ECOG PS, Eastern Cooperative Oncology Group performance status; EFS, event-free survival; ELN, European LeukemiaNet; FLT3-ITD, internal tandem duplication of the FMS-like tyrosine kinase 3 gene; G-CSF, granulocyte colony-stimulating factor; GO, gemtuzumab ozogamicin; HIV, human immunodeficiency virus; HSCT, hematopoietic stem cell transplantation; ICD, informed consent document; IRC, independent review committee; MFI, mean fluorescence intensity; MLL, myeloid/lymphoid leukaemia gene; NCCN, National Comprehensive Cancer Network; NPM1, nucleophosmin-1 gene; OS, overall survival; RFS, relapse-free survival; SAE, serious adverse event; VOD, veno-occlusive disease; WBC, white blood cell; WT1, Wilms' tumour suppressor gene.

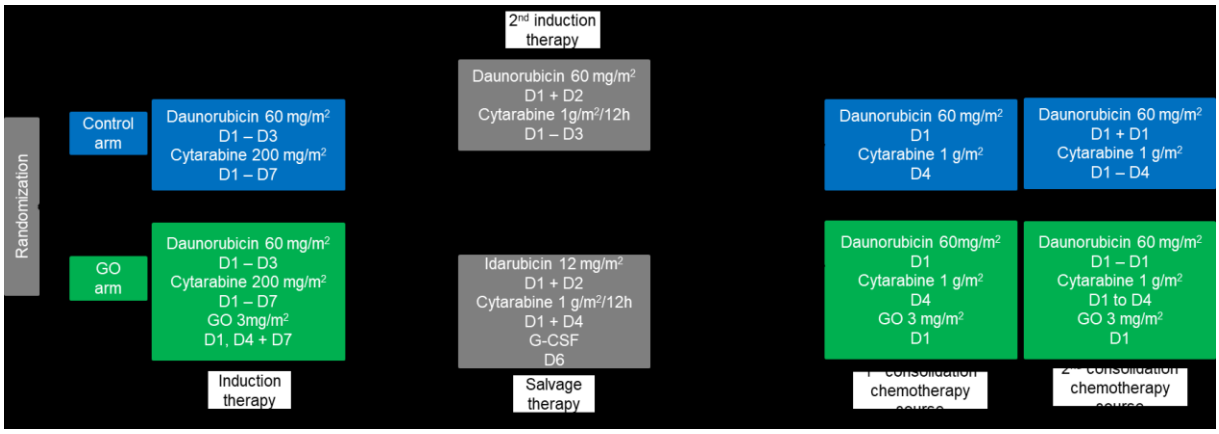
B.2.3.2. Treatment in ALFA-0701

Treatment was divided into two phases: induction and consolidation (Figure 5), with the option of salvage therapy if CR/CRp was not achieved after induction.

If patients experienced a leukocyte count greater than 30 000/mm^3, aspartate aminotransferase (AST) or alanine aminotransferase (ALT) greater than 2.5 × the upper limit of normal (ULN) and/or total bilirubin greater than 2 × the ULN, administration of GO was postponed. At the consolidation stage, patients were also required to have a platelet count of at least 100 000/mm^3. If liver function tests were abnormal at the consolidation stage, consolidation therapy was not postponed, but

only DA was administered. If patients had thrombocytopenia (platelet count < 100 000/mm^3), consolidation therapy was postponed until it had corrected. If moderate thrombocytopenia persisted (platelet count ≥ 50 000/mm^3 and < 100 000/mm^3), consolidation therapy was administered without GO.

Figure 5. ALFA-0701 study design



Source: figure adapted from Castaigne *et al*. 2012.[95]

CR, complete remission; CRp, complete remission with incomplete platelet recovery; D, day; G-CSF, granulocyte colony-stimulating factor; GO, gemtuzumab ozogamicin.

B.2.3.3. Efficacy outcomes

Primary and secondary efficacy outcomes have been defined in Table 8 according to IRC and investigator analysis. All efficacy outcomes were pre-specified.

Outcomes according to IRC and investigator analyses were available at the 1 August 2011 and 30 April 2013 data cut-offs as summarized in Table 9.

Table 8. Primary and secondary efficacy outcomes

Efficacy outcome	*Description for IRC analysis*	*Description for investigator analysis*
Primary endpoint		
EFS	Time from date of randomisation to date of induction failure, relapse or death due to any causes, whichever came first. EFS is not confounded by therapies subsequent to relapse	
	• Determined programmatically, using definition of CR/CRp detailed below, and by comparing relapse date indicated by the investigator to any supporting peripheral blood or BMA data within 1 day of the relapse, if available • Remission: CR (absence of blast cells in peripheral blood and no investigator report of extramedullary, molecular or cytogenetic disease; blast cell count < 5% in BMA with no Auer rods; neutrophil count > 1×10^9/L; platelet count > 1×10^{11}/L in the absence of transfusion) or CRp (CR with platelet count ≤ 1×10^{11}/L) • Relapse: assessed in patients achieving CR as any of presence of ≥ 5% blast cells in BMA or presence of Auer rods; blast cells in the CBC not attributable to bone marrow recovery following chemotherapy or G-CSF therapy; investigator report of extramedullary disease, molecular or cytogenetic disease	• Determined by the individual investigators for each patient treated at their site • Remission: CR (< 5% blasts in a normocellular marrow and ANC > 1×10^9 /L with a platelet count ≥ 100×10^9 /L in the peripheral blood in the absence of transfusion) or CRp (CR with residual thrombocytopenia [platelet count < 100×10^9 /L]) • Date of induction failure: date of evaluation of bone marrow response after the last induction cycle if a CR (by investigator assessment) had not been achieved • Disease progression: assessed/classified according to the International Working Group Criteria[96] • Relapse: recurrence of circulating blasts or cytopenia that led to a BMA revealing excess blasts; this was not specified in the protocol but is the usual definition of haematological relapse after CR/CRp

Efficacy outcome	*Description for IRC analysis*	*Description for investigator analysis*
Secondary endpoint		
OS	Time from date of randomization to date of death due to any cause. Patients known to be alive were censored at the last follow-up date or the reference date, and all deaths occurring after this date were not included in the OS analysis	
Response rate	Proportion of patients achieving an overall response	
	• CR + CRp	• The eCRF did not distinguish between CR and CRp; thus all responders were assessed as CR/CRp
RFS	Assessed in patients achieving CR/CRp as time from the date of remission to the date of relapse or death from any cause, whichever came first	
	• Determined programmatically, using definition of CR/CRp detailed above, and by comparing relapse date indicated by the investigator to any supporting peripheral blood or BMA data within 1 day of the relapse, if available	• Determined by individual investigators • Relapse: see above • Remission: categories at induction were CR, CRp, alive but treatment failure, early death before day 15, death in non-blastic aplasia (day 15 BMA), death after day 15 (day 15 BMA not carried out); categories in consolidation were CR, CRp, relapse, died, dropped out of the study, lost to follow-up and not assessed

Source: ALFA-0701 CSR.[24]

ANC, absolute neutrophil count; BMA, bone marrow assessment; CBC, complete blood count; CR, complete remission; CRp, complete remission with incomplete platelet recovery; eCRF, electronic case report form; EFS, event-free survival; G-CSF, granulocyte colony-stimulating factor; IRC, independent review committee; NA, not applicable; OS, overall survival; RFS, relapse-free survival.

Table 9. IRC- or investigator-assessed outcome measures available from the CSR by date of data cut-off (mITT population)

Outcome measure	*Data available from 1 August 2011 data cut-off*	*Data available from 30 April 2013 data cut-off*	*Used in economic analysis*
Primary endpoint			
EFS – IRC	✓	✓	
EFS – investigator	✓[a]	✓	✓ (2013 data cut)
Secondary endpoint			
OS	✗	✓	✓ (2013 data cut)
RFS – IRC	✓	✓	✓ (2013 data cut)
RFS – investigator	✓	✓	
Response rate – IRC[b,c]	NA	NA	✓ (2013 data cut)
Response rate – investigator[c,d]	NA	NA	

[a] Anticipated number of events observed for EFS. [b]CR/CRp. [c]Response rate was determined during the treatment period and these data were therefore available prior to any data cut-off. [d]CR + CRp.
Source: ALFA-0701 CSR.[24]

CR, complete remission; CRp, complete remission with incomplete platelet recovery; CSR, clinical study report; EFS, event-free survival; IRC, independent review committee; mITT, modified intention-to-treat; OS, overall survival; RFS, relapse-free survival.

B.2.3.4. Patient characteristics

In ALFA-0701, patient characteristics in the mITT population were well balanced between the treatment arms, but a higher proportion in the GO + DA versus DA arm were male (GO + DA: 54.8%; DA: 44.1%) and aged ≥ 60 years (GO + DA: 71.9%; DA: 61.8%). Most patients in the mITT population (189/271 patients) had favourable/intermediate cytogenetics profile according to the CHV classification, with a similar proportion having favourable/intermediate cytogenetics profile in each arm (GO + DA: 69.6%; DA: 69.9%). Only a small proportion of patients had < 30% CD33 expression (GO + DA: 12.6%; DA: 14.7%) (appendix D2.2).

Baseline characteristics were largely similar between the ITT and mITT populations, with the exception of white blood cell count, which was higher in both treatment arms in the ITT (GO + DA vs DA: 5.8 × 10^9/L vs 4.1 × 10^9/L) and mITT (GO + DA vs DA:

6.9 × 10^9/L vs 5.0 × 10^9/L) population (appendix D2.2). There were also differences in nucleophosmin 1 (NPM1) status, FMS-like tyrosine kinase 3 (FLT3-ITD) status and CCAAT/enhancer-binding protein α (CEBPA) status between the treatment arms.

The patient population in ALFA-0701 is generalizable to the UK population. This is demonstrated by patient baseline characteristics for ALFA-0701 being similar to those in UK AML studies (e.g. AML15 and AML16; appendix D.3.1.10). In particular, a similar proportion in AML15 had favourable/intermediate cytogenetics profile (MRC classification) as in ALFA-0701 (63.5%), but the proportion was lower than ALFA-0701 in AML16 (55%).

B.2.4 Statistical analysis and definition of study groups in the relevant clinical effectiveness evidence

B.2.4.1. Sample size calculation

The primary analysis of ALFA-0701 was to compare EFS in the GO + DA arm versus the DA arm in patients with de novo, untreated AML. EFS at 3 years was assumed to be 40% in the GO + DA arm and 25% in the DA arm (hazard ratio [HR]: 0.66, based on an underlying exponential distribution). ALFA-0701 aimed to randomize 280 patients 1:1 to the two treatment groups (140 per treatment arm). Sample size calculation showed that with 140 patients per treatment arm, 184 events were needed to give a statistical power of 80%, assuming a type-I error rate of 5% (two-sided) and a type-II error rate of 20%. For OS, 187 events would give sufficient power for testing the treatment effect based on a two-sided 0.05 level log-rank test, if the same HR for the sample size calculation is assumed as for the EFS in the protocol (HR: 0.66).

B.2.4.2. Populations

The mITT population included all patients randomized for whom informed consent forms were transferred (n = 271). The mITT was analysed according to initial randomization, regardless of whether patients received the study drug to which they were randomized or a different drug. Patients referred to HSCT at any time were

included in the mITT population. In ALFA-0701, the mITT population presented in the CSR was the primary population for evaluating outcomes and patient characteristics.

Safety analyses were based on the as-treated population, which included all patients who received at least one dose of the study medication. In the case of treatment misallocation, patients in the as-treated population were reported according to whether they received GO or not.

B.2.4.3. Interim and final analyses

A summary of when interim analyses were conducted for ALFA-0701 are summarized in Table 10.

B.2.4.4. Statistical tests

A summary of statistical analyses used in ALFA-0701 is shown in Table 11.

B.2.4.5. Participant flow in ALFA-0701

A Consolidated Standards of Reporting Trials (CONSORT) image depicting patient flow in the ALFA-0701 trial is presented in appendix D.2.1.

In total, 280 patients were enrolled across 26 centres in France between January 2008 and November 2010 and randomized (1:1) to the GO + DA arm or the DA arm (appendix D.2.). Of these patients, two withdrew consent and were excluded (one patient from each treatment arm). When data were transferred to Pfizer, results were available for 271 randomized patients (mITT population): 135 were assigned to receive GO + DA and 136 were assigned to the DA arm.

Of patients randomly assigned to receive treatment, [redacted] received study treatment ([redacted] in each arm). In the GO + DA arm, [redacted] completed treatment with chemotherapy and of those, [redacted] completed treatment with GO + DA. In the DA arm, [redacted] completed treatment with chemotherapy. The most common reasons ([redacted]) for permanent discontinuation of GO + DA were [redacted] ([redacted]) and resistant disease (i.e., [redacted]). In the GO + DA arm, the most common ([redacted]) reasons for permanent discontinuation of chemotherapy were [redacted]

[redacted]) and [redacted] ([redacted]); in the DA arm the most common reason was [redacted] ([redacted]).

Table 10. Interim and final analyses available for the mITT population in ALFA-0701

Outcome measure	*Unplanned interim analyses*	*Planned interim analysis*	*Final analysis*
Primary endpoint			
EFS – IRC	• NA	• 1 August 2011	• 30 April 2013
EFS – investigator	• July 2010 (requested by French regulatory agency) • July 2011 (provide results for a congress submission)[97]	• 1 August 2011 (all anticipated events for EFS observed)	• 30 April 2013
Secondary endpoints			
OS	• January 2010 (requested by French regulatory agency based on safety data[a]) • July 2010 • July 2011 • February 2012 (provide results or a primary publication)[89]	• 1 August 2011	• 30 April 2013
Response rate – IRC	NA[b]		
Response rate – investigator	NA[b]		
RFS – IRC	NA	1 August 2011	30 April 2013
RFS – investigator	NA	1 August 2011	30 April 2013

[a]Requested by the French Regulatory Agency in December 2009 in response to toxic deaths. There was no evidence of an increase in number of deaths and so the study was continued. [b]Response rate was determined during the treatment period and these data were therefore available prior to any data cut-off.

Source: ALFA-0701 CSR.[24]

EFS, event-free survival; IRC, independent review committee; mITT, modified intention-to-treat; NA, not applicable; OS, overall survival; RFS, relapse-free survival.

Table 11. Summary of statistical analyses used in ALFA-0701

Outcome	*Statistical analysis*
EFS	• The Kaplan–Meier method was used with comparisons made between the two treatment arms using two-sided log-rank tests • Two-sided 95% CIs for median time to event were estimated using the Brookmeyer–Crowley method with log-log transformation • HRs and associated two-sided 95% CIs were estimated using the Cox proportional hazards model • In subgroup analyses to assess the impact of baseline covariates on outcomes, a Cox proportional hazards model was used to calculate HRs, and 95% CIs and *p* values were calculated using log-rank tests
OS	• The Kaplan–Meier method was used with comparisons made between the two treatment arms using two-sided log-rank tests • Two-sided 95% CIs for median time to event were estimated using the Brookmeyer–Crowley method with log-log transformation • HRs and associated two-sided 95% CIs were estimated using the Cox proportional hazards model • Survival probabilities at 1, 2 and 3 years were estimated based on the Kaplan–Meir method, with their two-sided 95% CIs • The number and percentage of deaths, and the number and percentage were censored • In subgroup analyses to assess the impact of baseline covariates on outcomes, a Cox proportional hazards model was used to calculate HRs, and 95% CIs and *p* values were calculated using log-rank tests
CR, CRp, CR/CRp	• The number and percentage of patients in each category were calculated, as well as the risk difference, odds ratio comparing the two treatment arms and their two-sided 95% CIs, and *p* value from Fisher's exact test • Hematological response per investigator was cross-classified with hematological response as determined by independent review
RFS	• The Kaplan–Meier method was used with comparisons made between the two treatment arms using two-sided log-rank tests • Two-sided 95% CIs for median time to event were estimated using the Brookmeyer–Crowley method with log-log transformation • HRs and associated two-sided 95% CIs were estimated using the Cox proportional hazards model • Survival probabilities at 1, 2 and 3 years were estimated based on the Kaplan–Meier method, with their two-sided 95% CIs • The number and percentage of deaths, and the number and percentage were censored • In subgroup analyses to assess the impact of baseline covariates on outcomes, a Cox proportional hazards model was used to calculate HRs, and 95% CIs and *p* values were calculated using log-rank tests
Imputation for missing data	• Only partially missing dates were imputed • Missing start dates: for all analyses, only partially missing dates were imputed. If the day of the month is missing for any date used in a

Outcome	Statistical analysis
	calculation of time-to-event endpoints, the first of the month will be used to replace the missing date unless the calculation results in a negative time duration (e.g. the date of resolution cannot be before the date of onset; if replacing the resolution date with the first of the month results in a negative duration, the resolution date will be set to the onset date). If both month and day are missing, the date will be imputed to 1 January unless this results in a negative time duration. For OS, EFS and RFS if conventions result in a negative duration, duration will be reset to 1 day • Missing start dates of follow-up therapies were imputed using the general rules. If this resulted in a date earlier than the last date of the study drug, then the last date of the study drug was used as the start date • Missing stop dates were replaced using the last day of the month to replace the missing date. If both the day and month were missing, 31 December of the non-missing year replaced the missing date • Patients with missing response were treated as non-responders. All patients in the mITT population were included in the denominator in calculating the response rate • For time-to-event data, patients without an event were censored. In general, all events that occurred after a reference date will not be considered in EFS derivations, so are not counted as events. There are two cases for the censoring date for patients last known to be event-free: (1) if there is information indicating a patient has been followed until after the reference date (e.g. there is a disease assessment confirming that the patient is free of an event after the reference date), then the censoring date for this patient will be the reference date; (2) otherwise the censoring date will be the last assessment date before the reference date • In covariate analyses, patients with unknown data were treated as missing and excluded

Source: ALFA-0701 CSR.[24]

CI, confidence interval; CR, complete remission; CRp, complete remission with incomplete platelet recovery; EFS, event-free survival; HR, hazard ratio; mITT, modified intention-to-treat; OS, overall survival; RFS, relapse-free survival.

B.2.5 Quality assessment of the relevant clinical effectiveness evidence

The quality assessment of ALFA-0701 is shown in Table 12.

Table 12. Quality assessment results for ALFA-0701

Trial number (acronym)	NCT00927498 (ALFA-0701)
Was randomisation carried out appropriately?	Yes Patients were randomized in a 1:1 ratio. Randomization was performed using a computer-generated sequence
Was the concealment of	ALFA-0701 was an open-label study

treatment allocation adequate?	
Were the groups similar at the outset of the study in terms of prognostic factors?	Yes Baseline characteristics between the arms in the mITT population were well balanced in terms of median age, ECOG PS, cytogenetics profile (favourable/intermediate) and CD33 expression. However, a higher proportion in the GO + DA versus DA arm were male and aged ≥ 60 years
Were the care providers, participants and outcome assessors blind to treatment allocation?	Although the study treatment was not blinded to patients or investigators during the study period, after data were transferred to Pfizer, a retrospective analysis verified investigator-collected data. The IRC was blinded to the treatment arms to ensure unbiased assessment of data
Were there any unexpected imbalances in drop-outs between groups?	No A similar proportion of patients completed chemotherapy treatment in the GO + DA arm and the DA arm: in the GO + DA arm 48.9% of patients completed GO + DA treatment. The most common reason for discontinuation of GO + DA was AEs and resistant disease. The most common reasons for discontinuation of chemotherapy in the GO + DA arm were AEs and resistant disease; in the DA arm, the most common reason was resistant disease
Is there any evidence to suggest that the authors measured more outcomes than they reported?	No All outcomes were reported
Did the analysis include an intention-to-treat analysis? If so, was this appropriate and were appropriate methods used to account for missing data?	The analysis included a mITT population, which was the most appropriate population as it included all randomized patients for whom data were transferred to Pfizer, with the exception of those patients who withdrew consent prior to the start of treatment Appropriate methods were used to account for missing or incomplete data

Source: ALFA-0701 CSR.[24]

AE, adverse event; mITT, modified intention-to-treat; ECOG PS, Eastern Cooperative Oncology Group performance status; GO, gemtuzumab ozogamicin.

B.2.6 Clinical effectiveness results of the relevant trials

- ALFA-0701 provides strong evidence to demonstrate the clinical benefit of adding GO to DA compared with DA alone for the first-line treatment of patients with de novo AML with a favourable/intermediate cytogenetics profile. EFS and RFS were significantly improved in patients receiving GO + DA versus DA alone. Increasing the duration of remission can lead to long-term durable survival (or a 'functional cure'). A durable response was observed from 3 years in approximately [redacted] of patients receiving GO + DA and [redacted] of patients receiving DA alone.

- Adding GO to DA led to a statistically significant and clinically meaningful [redacted] improvement in the primary endpoint of event-free survival (EFS) in both the mITT population and in the subpopulation with favourable/intermediate cytogenetics profile, and in both populations [redacted] more patients were event free at 3 years with the addition of GO. After 3 years, the KM curve started to plateau (approximately [redacted] in the GO + DA arm and [redacted] in the DA arm), indicating the durability of the response. Results presented below are according to IRC analysis

  - In the mITT population median EFS was significantly longer in the GO + DA ([redacted]) versus DA ([redacted]) arm ([redacted]). The EFS advantage in the GO + DA versus the DA arm was most apparent in the subpopulation of patients with favourable/intermediate cytogenetics profile ([redacted]). There was no significant difference in median EFS in patients with unfavourable cytogenetics profile.
  - At 3 years, the proportion of patients who were event free was higher in the GO + DA arm than the DA arm in the mITT population ([redacted]; [redacted] versus [redacted]) and the subpopulation with favourable/intermediate cytogenetics profile ([redacted] versus [redacted]).
  - Results according to IRC analyses are in line with investigator analyses.

- In patients in complete remission (CR) or CR with incomplete platelet recovery (CRp), adding GO to DA led to a significant improvement in relapse-free survival (RFS) of at least [redacted] in both the mITT population and in those with favourable/intermediate cytogenetics profile, and in both populations [redacted] more patients were event free at 3 years with the addition of GO. After 3 years, the KM curve started to plateau (at approximately [redacted] in the GO + DA arm and [redacted] in the DA arm), indicating the durability of the response. Results presented below are according to IRC analysis

  - In the mITT population, median RFS was significantly longer in the GO + DA ([redacted]) versus DA ([redacted]) arm [redacted]). The RFS advantage was most apparent in patients with favourable/intermediate cytogenetics profile ([redacted]). There was no significant difference in median RFS in patients with unfavourable cytogenetics profile.

- At 3 years, the proportion of patients who were relapse free was higher in the GO + DA arm than in the DA arm in both the mITT population (vs) and the subpopulation with favourable/intermediate cytogenetics profile (; vs).
- In patients who experienced an event following CR/CRp, a numerically lower proportion in the GO + DA arm than in the DA arm relapsed in both the mITT population (vs) and in the subpopulation with favourable/intermediate cytogenetics (vs).
- Results according to IRC analyses are in line with those by investigator analyses.

- In support of the RFS results, in an ad hoc analysis, median time to subsequent anti-cancer therapy after induction failure or relapse was significantly delayed in the GO + DA arm versus the DA arm by .

  - Median time to subsequent anti-cancer therapy administered after induction failure or relapse was significantly longer in the GO + DA arm (;) versus the DA arm (; ;).

- In the mITT population, adding GO to DA led to a numerical extension in OS of versus DA alone that trended towards significance; the numerical extension in OS was more apparent in the subpopulation with favourable/intermediate cytogenetics profile ( ). After 3 years, the OS risk started to plateau (at approximately in the GO + DA arm and in the DA arm)

  - In the mITT population, there was a trend towards longer median OS in the GO + DA ( ) versus the DA ( ) arm ( ). There was a more apparent trend toward longer OS following treatment with GO + DA vs DA alone in patients with favourable/intermediate cytogenetics profile than in the mITT population (vs). In patients with unfavourable cytogenetics profile there was no numerical difference in OS between the GO + DA and DA arm. ALFA-0701 was not powered for OS and did not reach statistical significance in the mITT population or the subpopulation with favourable/intermediate cytogenetics profile.

- Adding GO to DA led to a numerical increase in the proportion of patients who achieved CR/CRp in the mITT population, as well as in the intermediate/favourable cytogenetic subpopulation, but this did not reach statistical significance. Results presented below are according to IRC analysis

  - In the mITT population, of patients in the GO + DA arm achieved CR/CRp after induction therapy versus in the DA arm. In patients with favourable/intermediate cytogenetics profile, the corresponding proportions were and . There was no difference in response rate between the GO + DA arm and DA arm in patient with unfavourable cytogenetics profile.
  - Results according to IRC analyses are in line with those of investigator analyses.

The ALFA-0701 study provides strong evidence to demonstrate the clinical benefit of adding GO to DA compared with DA alone for the first-line treatment of de novo AML. The clinical benefit observed is most apparent in the subpopulation with favourable/intermediate cytogenetics profile. This is therefore the key subpopulation that is central to this submission and should be the focus for decision making.

The focus of this section is the mITT population and the subpopulation with favourable/intermediate cytogenetics profile according to IRC analyses; these data are used in the economic model. This is because IRC analyses were conducted as per regulatory standards. IRC analyses are consistent with investigator analyses, which may reflect of clinical practice (Appendix D.2.4). In summary, data in this section focus on:

- **EFS and RFS:** IRC analyses for the mITT population and the subpopulation with favourable/intermediate cytogenetics profile; 30 April 2013 data cut-off
- **Response rate:** IRC analyses for the mITT population and the subpopulation with favourable/intermediate cytogenetics profile; data were available prior to any data cut-off because this was evaluated during the treatment period
- **OS:** Data not categorized according to IRC or investigator analysis, because no difference between the analyses would be expected. OS was evaluated for the mITT population and the subpopulation with favourable/intermediate cytogenetics profile at the 30 April 2013 cut-off

Table 13 summarises efficacy data for the mITT population at the 30 April 2013 cut-off, according to IRC and investigator analyses.[24] Similar data at the 1 August 2011 cut-off are summarized in Appendix D.2.3.[24] Table 14 summarises efficacy data for the favourable/intermediate cytogenetic subpopulation: EFS and RFS data are presented at the 30 April 2013 cut-off for the IRC analysis and 1 August 2011 cut-off for the investigator analysis. OS data are presented for the 30 April 2013 cut-off. Response rate data are presented for IRC and investigator analysis. The methods for subgroup analyses are described in section B.2.7 and detailed results for the unfavourable cytogenetic subpopulation and other subgroups are reported in Appendix E. A summary of data from Castaigne *et al*. 2012 publication for the ITT population at the 1 August 2011 data cut-off are presented in Appendix D2.3.

Table 13. Summary of efficacy endpoints in ALFA-0701 (mITT population; 30 April 2013 data cut-offs)[a]

	GO + DA arm	*DA arm*	*Point estimate (95% CI)*	*p value*	*Economic model*
Randomized (n)	135	136	–	–	–
Efficacy (n) (mITT population)	135	136	–	–	–
EFS, months, median (95% CI)					
IRC assessment	[redacted]	[redacted]	[redacted]	[redacted]	✓
Investigator assessment	[redacted]	[redacted]	[redacted]	[redacted]	✗
RFS, months, median (95% CI)					
IRC assessment	[redacted]	[redacted]	[redacted]	[redacted]	✓
Investigator assessment	[redacted]	[redacted]	[redacted]	[redacted]	✗
OS, months, median (95% CI)	[redacted]	[redacted]	[redacted]	[redacted]	✓
Overall response rate (CR + CRp), n (%)[a]					
IRC assessment	[redacted]	[redacted]	[redacted]	[redacted]	✓
Investigator assessment	[redacted]	[redacted]	[redacted]	[redacted]	✗

Source: ALFA-0701 CSR.[24]

[a] Response rate was determined during the treatment period and these data were therefore available prior to any data cut-off.

CI, confidence interval; CR, complete remission; CRp, complete remission with incomplete platelet recovery; DA, daunorubicin + cytarabine; EFS, event-free survival; GO, gemtuzumab ozogamicin; HR, hazard ratio; IRC, independent review committee; mITT, modified intention-to-treat; NE, not estimable; OR, odds ratio; OS, overall survival; RFS, relapse-free survival.

Table 14. Summary of efficacy endpoints in ALFA-0701 for patients with favourable/intermediate cytogenetics profile (mITT population)

	GO + DA arm (n = 94)	*DA arm (n = 95)*	*Point estimate (95% CI)*	*p value*	*Economic model*
EFS, months, median (95% CI)[a]					
IRC assessment (30 April 2013 data cut-off)	[redacted]	[redacted]	[redacted]	[redacted]	✓
RFS, months, median (95% CI)[a]					
IRC assessment (30 April 2013 data cut-off)	[redacted]	[redacted]	[redacted]	[redacted]	✓
OS (30 April 2013 data cut-off), months, median (95% CI)	[redacted]	[redacted]	[redacted]	[redacted]	✓
Overall response rate (CR/CRp), n (%)[b]					
IRC assessment	[redacted]	[redacted]	[redacted]	[redacted]	✓
Investigator assessment	[redacted]	[redacted]	[redacted]	[redacted]	✖

Source: ALFA-0701 CSR;[24] Favourable/intermediate cytogenetics profile subpopulation (IRC data)[98]

[a]At the 30 April 2013 data cut-off, data were not available for EFS and RFS according to investigator analysis in the favourable/intermediate cytogenetic subpopulation. [b]CR/CRp. Response rate was determined during the treatment period and these data were therefore available prior to any data cut-off.

CI, confidence interval; CR, complete remission; CRp, complete remission with incomplete platelet recovery; DA, daunorubicin + cytarabine; EFS, event-free survival; GO, gemtuzumab ozogamicin; HR, hazard ratio; IRC, independent review committee; mITT, modified intention-to-treat; NE, not estimable; NR, not reached; OR, odds ratio; OS, overall survival; RFS, relapse-free survival.

B.2.6.1. Event-free survival

Adding GO to DA led to a statistically significant and clinically meaningful [redacted] improvement in EFS, with [redacted] more patients being event free; the clinical benefit on EFS was most apparent in the subpopulation with favourable/intermediate cytogenetics profile ([redacted] improvement; [redacted] more patients being event free)

EFS was the primary outcome of ALFA-0701. IRC analyses at the 30 April 2013 cut-off found median EFS to be significantly longer in the GO + DA arm ([redacted]) than in the DA arm ([redacted]; Figure 6 and Table 15). This corresponds to a [redacted] reduction in the risk of an event for patients in the GO + DA arm compared with the DA arm. From approximately 36 months, the KM curve plateaued in both treatment arms and remained stable for the remaining follow-up (Figure 6; approximately [redacted] in the GO + DA arm and [redacted] in the DA arm), reflecting a decrease in hazard. The rate of EFS at 2 years was [redacted] ([redacted]) in the GO + DA arm and [redacted] ([redacted]) in the DA arm; at 3 years, the corresponding proportions were [redacted] ([redacted]) in the GO + DA arm and [redacted] ([redacted]) in the DA arm.

Figure 6. Kaplan–Meier plot of EFS (mITT population; 30 April 2013 cut-off; IRC analysis)

[redacted]

Source: ALFA-0701 CSR.[24]

No. at risk, number at risk; DA, daunorubicin + cytarabine; GO, gemtuzumab ozogamicin; mITT; modified intention-to-treat. Circles indicate censoring observations.

Table 15. Analysis of EFS conducted according to IRC analysis (mITT population; 30 April 2013 cut-off)

	GO + DA arm (n = 135)	*DA arm (n = 136)*
EFS events, n (%) Induction failure Relapse Death	[redacted]	[redacted]
Censored, n (%)[b]	[redacted]	[redacted]
Median time to event, months (95% CI)	[redacted]	[redacted]
HR (95% CI); *p* value	[redacted]	

[a] [redacted]

Source: ALFA-0701 CSR.[24]

CI, confidence interval; DA, daunorubicin + cytarabine; EFS, event-free survival; GO, gemtuzumab ozogamicin; HR, hazard ratio.

The benefit of GO + DA was most apparent in those with favourable/intermediate cytogenetics profile (Figure 7 and Table 16). According to IRC analysis at the 30 April 2013 cut-off, median EFS was longer in the GO + DA ([redacted]) versus DA ([redacted]) arm.[24] From approximately 36 months, the KM curve started to plateau in both treatment arms and remained stable for the remaining follow-up (Figure 6; approximately [redacted] in the GO + DA arm and [redacted] in the DA arm). The rate of EFS at 2 years was [redacted] in the GO + DA arm and [redacted] in the DA arm; at 3 years, the corresponding proportions were [redacted] and [redacted]. In patients with unfavourable cytogenetics profile there was no difference in median EFS between the arms (Appendix E.2. 1).

Data for EFS according to IRC analyses at the 30 April 2013 cut-off for the mITT population were consistent with those conducted according to investigator analyses for the mITT population at the 30 April 2013 cut-off (Appendix D.2.4.1) and the favourable/intermediate cytogenetics profile subpopulation at the 1 August 2011 cut-off (Appendix E.2.2).

Figure 7. Kaplan–Meier plot of EFS (favourable/intermediate cytogenetics profile subpopulation; 30 April 2013 cut-off; IRC analysis)

Source: Favourable/intermediate cytogenetics profile subpopulation (IRC data)[98]

No at risk, number at risk; DA, daunorubicin + cytarabine; GO, gemtuzumab ozogamicin; mITT, modified intention-to-treat; RFS, relapse-free survival.

Circles indicate censoring observations.

Table 16. Analysis of EFS by IRC analysis (favourable/intermediate cytogenetic subpopulation; 30 April 2013 cut-off)

	GO + DA arm (n = 95)	*DA arm (n = 94)*
EFS events, n (%) Induction failure Relapse Death	[redacted]	[redacted]
Censored, n (%)[a]	[redacted]	[redacted]
Median time to event, months (95% CI)	[redacted]	[redacted]
HR (95% CI); *p* value	[redacted]	

[a] [redacted]

Source: Favourable/intermediate cytogenetics profile subpopulation (IRC data)[98]

CI, confidence interval; DA, daunorubicin + cytarabine; EFS, event-free survival; GO, gemtuzumab ozogamicin; HR, hazard ratio.

B.2.6.2. Relapse-free survival

In patients who achieved CR/CRp in the mITT population, RFS was improved by 9 months in the GO + DA arm versus the DA arm, and [redacted] more patients were event free at 3 years; the RFS benefit was most apparent in the favourable/intermediate cytogenetic subpopulation ([redacted] improvement; [redacted] more event free at 3 years)

RFS was a secondary outcome in ALFA-0701.[24] It was measured in patients who achieved a CR/CRp. According to IRC analysis at the 30 April 2013 cut-off, RFS was significantly longer in the GO + DA arm ([redacted]) than in the DA arm ([redacted]; Table 17 and Figure 8). This corresponds to a [redacted] reduction in the risk of an event for patients in the GO + DA arm compared with those in the DA arm. The rate of RFS at 2 years according to IRC analyses was [redacted] ([redacted]) in the GO + DA arm and [redacted] ([redacted]) in the DA arm; at 3 years, the corresponding proportions were [redacted] ([redacted]) and [redacted] ([redacted]) in the DA arm.

Table 17. Analysis of RFS by IRC analysis (mITT population; 30 April 2013 cut-off)

	GO + DA arm (n = [redacted])	*DA arm (n = [redacted])*
RFS events, n (%) Relapse Death before relapse	[redacted]	[redacted]
Censored, n (%)[a]	[redacted]	[redacted]
Median time to event, months (95% CI)	[redacted]	[redacted]
HR (95% CI); *p* value	[redacted]	

[a]The main cause of censoring was patients being event-free at the reference date (GO + DA arm: [redacted]; DA arm: [redacted]).

Source: ALFA-0701 CSR.[24]

CI, confidence interval; CR, complete remission; CRp, complete remission with incomplete platelet recovery; DA, daunorubicin + cytarabine; GO, gemtuzumab ozogamicin; HR, hazard ratio; IRC, independent review committee; RFS, relapse-free survival.

Figure 8. Kaplan–Meier plot of RFS (mITT population; 30 April 2013 cut-off; IRC analysis)

Source: ALFA-0701 CSR.[24]

No at risk, number at risk; DA, daunorubicin + cytarabine; GO, gemtuzumab ozogamicin; mITT, modified intention-to-treat; RFS, relapse-free survival.

Circles indicate censoring observations.

The benefit of adding GO to DA on RFS was most pronounced in those with favourable/intermediate cytogenetics profile (Figure 9 and Table 18). According to IRC analysis at the 30 April 2013 cut-off, median RFS was significantly longer in the GO + DA ( ) versus DA arm (; Figure 9 and Table 18). The rate of RFS at 2 years according to IRC analyses was ( ) in the GO + DA arm and ( ) in the DA arm; at 3 years, the corresponding proportions were ( ) and ( ) in the DA arm. In patients with unfavourable cytogenetics profile there was no difference in median RFS between the GO + DA arm and the DA arm (Appendix E.3.1).

Figure 9. Kaplan–Meier plot of RFS (favourable/intermediate cytogenetic subpopulation; 30 April 2013 cut-off; IRC analysis)

Source: Favourable/intermediate cytogenetics profile subpopulation (IRC data)[98]

No at risk, number at risk; DA, daunorubicin + cytarabine; GO, gemtuzumab ozogamicin; mITT, modified intention-to-treat; RFS, relapse-free survival.

Circles indicate censoring observations.

Table 18. Analysis of RFS by IRC analysis (favourable/intermediate cytogenetic subpopulation; 30 April 2013 cut-off)

	GO + DA (N = [redacted])	DA (N = [redacted])
RFS events, n (%) Relapse Death before relapse	[redacted]	[redacted]
Censored, n (%)[a]	[redacted]	[redacted]
Median time to event, months (95% CI)	[redacted]	[redacted]
HR (95% CI); *p* value	[redacted]	

[a]The main cause of censoring was patients being event-free at the reference date (GO + DA arm: [redacted]; DA arm: [redacted]).

Source: Favourable/intermediate cytogenetics profile subpopulation (IRC data)[98]

CI, confidence interval; DA, daunorubicin + cytarabine; GO, gemtuzumab ozogamicin; HR, hazard ratio; KM, Kaplan–Meier; mITT, modified intention-to-treat; NE, not estimable; NR, not reached; RFS, relapse-free survival.

Data for RFS according to IRC analyses at the 30 April 2013 cut-off for the mITT population were consistent with those conducted according to investigator analyses for the mITT population at the 30 April 2013 cut-off (Appendix D.2.4.2) and the favourable/intermediate cytogenetics profile subpopulation at the 1 August 2011 cut-off (Appendix E.3.2).

Adding GO to DA can significantly delay the time to subsequent anti-cancer therapy after induction failure or relapse by

In support of the RFS results, an *ad hoc* analysis demonstrated that median time to subsequent anti-cancer therapy administered after induction failure or relapse was significantly longer in the overall GO + DA arm ( ) compared with the DA arm ( ).

B.2.6.3. Overall survival

In the mITT population, there was a trend for longer OS in the GO + DA arm than in the DA arm by ; the OS trend was most apparent in the favourable/intermediate cytogenetic subpopulation, being extended by with the addition of GO to DA

OS was a secondary outcome in the ALFA-0701 trial. Unlike EFS, OS analysis is potentially confounded by subsequent treatments that patients received for AML.[24]

At the data cut-off of 30 April 2013, deaths occurred in the GO + DA arm and deaths occurred in the DA arm (Table 19). There was a trend towards longer median OS in the GO + DA arm ( ) than in the DA arm ( ), although the difference was not significant (; Figure 10). The lack of significance for OS is likely because the study was not powered to detect statistically significant differences between the arms for this outcome. Furthermore, the majority of patients in ALFA-0701 received at least one follow-up therapy for AML (including HSCT-conditioning regimens), which would have confounded the results.

The proportions of patients receiving follow-up therapy were similar between the GO + DA arm ( ) and the DA arm ( ). patients in the DA arm received GO + DA as a part of follow-up therapy. patients in the

GO + DA arm and [redacted] in the DA arm underwent HSCT at any time during the study (as part of consolidation therapy or following relapse or induction failure). After HSCT, clinicians assume similar survival benefit, irrespective of the type of induction therapy received.

A significant difference in OS has, however, been observed in meta-analyses evaluating a larger sample of patients who received GO in combination with intensive chemotherapy with those who were in the comparator arm and did not receive GO (see Appendix D3).

In patients with favourable/intermediate cytogenetics profile, at the data cut-off of 30 April 2013, [redacted] deaths occurred in the GO + DA arm and [redacted] deaths occurred in the DA arm. The OS benefit in the mITT population was more apparent in the favourable/intermediate cytogenetics profile subpopulation in the GO + DA arm ([redacted]; Table 20) versus the DA arm ([redacted]), which trended towards significance ([redacted]). In patients with unfavourable cytogenetics profile the numerical difference in OS between patients receiving GO + DA and those receiving DA was less apparent than in patients with favourable/intermediate cytogenetics profile (Appendix E.1).

Table 19. Analysis of OS (mITT population; 30 April 2013 cut-off)

	GO + DA arm (n = 135)	***DA arm (n = 136)***
Number of deaths, n (%)	[redacted]	[redacted]
Censored, n (%)	[redacted]	[redacted]
Median time to event, months (95% CI)	[redacted]	[redacted]
HR (95% CI); *p* value[a]	[redacted]	

[a]Based on two-sided *p* value from log-rank test.

Source: ALFA-0701 CSR.[24]

CI, confidence interval; DA, daunorubicin + cytarabine; GO, gemtuzumab ozogamicin; HR, hazard ratio; OS, overall survival.

Figure 10. Kaplan–Meier plot of OS (mITT population; 30 April 2013 data cut-off)

Source: ALFA-0701 CSR.[24]

No at risk, number at risk; DA, daunorubicin + cytarabine; GO, gemtuzumab ozogamicin; mITT, modified intention-to-treat; OS, overall survival.

Table 20. Analysis of OS by investigator analysis (favourable/intermediate cytogenetic subpopulation; 30 April 2013 cut-off)

N	*GO + DA (N = 94)*	*DA (N = 95)*	*Hazard ratio (95% CI) p value*
Number of deaths, n (%)	[redacted]	[redacted]	-
KM estimate of median time to event versus control, months (95% CI)	[redacted]	[redacted]	[redacted]

Source: ALFA-0701 CSR.[24]

CI, confidence interval; DA, daunorubicin + cytarabine; GO, gemtuzumab ozogamicin; KM, Kaplan–Meier; mITT, modified intention-to-treat; NE, not estimable; OS, overall survival.

B.2.6.4. Response rate

There was a slight increase in proportion of patients achieving CR or CRp in the induction phase in the GO + DA arm than in the DA arm

Response rate was a secondary outcome in ALFA-0701. According to IRC analysis the overall response rate (ORR) following induction therapy was [redacted] in the GO + DA arm and [redacted] in the DA arm ([redacted]).

In patients with favourable/intermediate cytogenetics profile, response rates according to IRC analysis were numerically higher in the GO + DA ([redacted]) versus DA ([redacted]) arm (Table 22), but the difference was not statistically significant. In patients with unfavourable cytogenetics profile, there was no difference in response rate between patients receiving GO + DA versus DA.

Although response rate was similar between the treatment arms, it is important to note that based on data for RFS, the relapse-free period was more durable in patients who received GO + DA than DA alone.

The overall response rate in the mITT population and the subpopulation with favourable/intermediate cytogenetics profile according to IRC analyses were consistent with those conducted according to investigator analyses (appendix D.2.4.3 and appendix E.4).

Table 21. Response rate by IRC analysis (mITT population)[a]

	GO + DA arm (n = 135) n (%)	***DA arm (n = 136) n (%)***	***OR[b] (95% CI)[c]***	***p value[d]***
Overall response rate (CR/CRp)[e]	[redacted]	[redacted]	[redacted]	[redacted]

[a]Response rate was determined during the treatment period and these data were therefore available prior to any data cut-off; [b]Based on Wald statistic. [c]Based on Clopper–Pearson interval method. [d]Obtained using Fisher's exact test. [e]No distinction was made between CR and CRp.

Source: ALFA-0701 CSR.[24]

CI, confidence interval; CR, complete remission; CRp, complete remission with incomplete platelet recovery; DA, daunorubicin + cytarabine; GO, gemtuzumab ozogamicin; IRC, independent review committee; mITT, modified intention-to-treat; OR, odds ratio.

Table 22. Analysis of response rate by IRC analysis (favourable/intermediate cytogenetic profile subpopulation)[a]

	GO + DA arm (n = 94) n (%)	**DA arm (n = 95) n (%)**	**OR[b] (95% CI)[c]**	***p* value[d]**
Overall response rate (CR + CRp)	[redacted]	[redacted]	[redacted]	[redacted]

[a]Response rate was determined during the treatment period and these data were therefore available prior to any data cut-off. [b]Based on Wald statistic. [c]Based on Clopper–Pearson interval method. [d]Obtained using Fisher's exact test.

Source: Favourable/intermediate cytogenetics profile subpopulation (IRC data)[98]

CI, confidence interval; CR, complete remission; CRp, complete remission with incomplete platelet recovery; DA, daunorubicin + cytarabine; GO, gemtuzumab ozogamicin; IRC, independent review committee; mITT, modified intention-to-treat; OR, odds ratio.

B.2.7 Subgroup analysis

B.2.7.1. ALFA-0701 stratification factors

Owing to the fact that AML is a heterogeneous disease, it is important to assess the impact of patient baseline characteristics on treatment outcomes in subgroups that are known to have a clinical impact on prognosis.

The following pre-planned baseline subgroup analyses were performed in the mITT population for EFS, OS, RFS and response rate in the ALFA-0701 trial according to the following baseline covariate stratifications:[24] cytogenetic risk (according to CHV classification), age (at time of randomization), Eastern Cooperative Oncology Group performance status score, CD33 expression (by percentage of CD33-positive cells), CD33 MFI ratio, risk classification based on National Comprehensive Cancer Network (NCCN) and ELN guidelines, FLT3-ITD status, NPM1 status, CEBPA status, myeloid/lymphoid leukaemia gene status, Wilms' tumour suppressor gene status, and genotype (according to CHV classification).

Subgroup analyses address the decision problem by investigating the proportion of patients with favourable/intermediate cytogenetics profile for whom adding GO to DA has the greatest clinical benefit. Outcomes in other subgroups were consistent with the overall analysis (see Appendix E for full details).

B.2.7.2. ALFA-0701 statistical methods for subgroup analysis

Subgroup analyses were based on the mITT population.[24] Outcome definitions and statistical methods that are specific to the subgroup analysis are summarized in Table 23. Other statistical methods are summarized in Table 11. The results of all subgroup analyses are shown in Appendix E.

Table 23. Summary of ALFA-0701 statistical methodology for subgroup analysis

Outcome	Definition and statistical analysis
EFS	• For IRC analysis the censoring date was the reference date of 30 April 2013, or the date of the last disease assessment before the reference date for all subgroups except CD33 MFI ratio • Evaluated according to the primary definition of EFS with event rates being determined by investigator assessment with the censoring date being the reference date of 1 August 2011, or the date of the last disease

	assessment before the reference date. • HRs and 95% CIs were calculated using the unstratified Cox model • *p* values were calculated using the log-rank test
OS	• The reference date for assessment was 30 April 2013 • HRs and 95% CIs were calculated using the unstratified Cox model • *p* values were calculated using the log-rank test
RFS	• For IRC analysis, the censoring date was the reference date of 30 April 2013, or the date of last disease assessment before the reference date. • Analysis was based on the primary definition of RFS with event rates being determined by investigator assessment with the censoring date being the reference date of 1 August 2011, or the date of the last disease assessment before the reference date. HRs and 95% CIs were calculated using the unstratified Cox model • *p* values were calculated using the log-rank test

Source: ALFA-0701 CSR.[24]

CI, confidence interval; EFS, event-free survival; HR, hazard ratio; IRC, independent review committee; MFI, mean fluorescence intensity; RFS, relapse-free survival.

B.2.8 Meta-analysis

A meta-analysis was conducted by Pfizer, which updated a published meta-analysis (Hills *et al.* 2014) by using later data cut-offs and individual patient data (IPD), where available.[35] This IPD meta-analysis includes ALFA-0701, which is the pivotal trial in the submission, and four other trials that use different dosing regimens of GO that will not be approved by the EMA. With the exception of ALFA-0701, the trials included in the IPD meta-analysis were therefore not considered relevant in this submission. The IPD meta-analysis comparing GO with the comparator arm that does not include GO (no-GO) therefore only provides supportive evidence, and was the only meta-analysis used for the regulatory submission to the EMA. However, data from the IPD meta-analysis were not used to populate the economic model in the present submission.

The methodology and key results of the IPD meta-analysis are reported in Appendix D.3.1 for completeness. Four other meta-analysis were identified as part of an SLR, and these have been summarized in Appendix D.3.2.

B.2.9 Indirect and mixed treatment comparisons

No indirect or mixed treatment comparisons were performed.

B.2.10 Adverse reactions

- The ALFA-0701 study shows that GO in combination with DA is generally associated with manageable and reversible adverse events (AEs), consistent with the known safety profile of each of the individual agents. The most common AEs were haematological, which can be managed through supportive measures and the withholding of GO during consolidation therapy. Treatment-emergent AEs (TEAEs) in the subpopulation of as-treated patients with favourable/intermediate cytogenetics profile was consistent with the overall as-treated population.

- Adding GO to DA did not greatly increase the proportion of patients who experienced a treatment-related AE, but a higher proportion of patients in the GO + DA arm than the DA arm experienced a treatment-related serious AE (SAE)

    - In total, [redacted] of patients in the GO + DA arm and [redacted] in the DA arm experienced a treatment-related AE. For treatment-related SAEs, the corresponding proportions were [redacted] and [redacted], respectively. The most common SAE in the GO + DA arm was thrombocytopenia ([redacted] versus [redacted]). However, treatment-related severe infections (grade ≥ 3) were experienced by [redacted] of patients in the GO + DA arm and [redacted] of patients in the DA arm

- Adding GO to DA resulted in higher proportion of patients experiencing veno-occlusive disease (VOD) than with DA alone ([redacted] vs [redacted]). For patients in the DA arm, VOD occurred within 28 days of the last dose of GO. Expert opinion of clinicians in the UK is that the observed rates of VOD for GO are acceptably low and importantly, are lower than previous studies with GO, which can be attributed to its fractionated dosing. Adding GO to DA did not lead to an increase in treatment-related deaths

    - A similar proportion of patients in both the GO + DA arm ([redacted]) and the DA arm ([redacted]) had treated-related deaths and a similar proportion of patients died owing to study treatment toxicity ([redacted] versus [redacted], respectively)

- Rates of TEAEs in the subpopulation of as-treated patients with favourable/intermediate cytogenetics profile were aligned with the overall as-treated population.

B.2.10.1. Methodology of safety in ALFA-0701

A retrospective safety data assessment was conducted to collect treatment-emergent AEs (TEAEs) of special interest, based on review of patient medical files. The events assessed retrospectively were:

- haemorrhage (all grades)
- veno-occlusive disease (VOD; all grades)

- severe infections
- any other AE leading to permanent discontinuation of treatment
- hospitalizations
- transfusions performed up to 28 days after the last dose of study drug
- cardiac function assessments as per the planned protocol, even if they occurred more than 28 days after the last study drug.

Details on the methodology for collecting safety data for ALFA-0701 are presented in appendix F.

Data presented in this submission are for the as-treated population (all patients who received at least one dose of study medication; GO + DA: n = 137 [97%]; DA: n = 137 [100.7%]; because some patients randomized to the GO arm received only DA) in ALFA-0701, and are based on retrospective data collection unless stated otherwise. TEAEs are also presented for a subpopulation of the as-treated population who had a favourable/intermediate cytogenetics profile (GO + DA: n = 91; DA: n = 95).

B.2.10.2. Safety results in ALFA-0701

In ALFA-0701, safety analyses were based on a total of 268 (98.9%) patients who comprised the AT population, defined as all patients who received at least 1 dose of study medication and who were analysed according to the actual treatment received. Within the AT population, 131 patients received GO + DA, 137 patients received DA.

This section covers the safety profile related to treatment in the GO + DA arm compared with treatment in the DA arm in ALFA-0701, as presented in the CSR.[24] Data on treatment-related deaths are summarized in Table 24, while AEs and SAEs are summarized in Table 25; a list of predefined TEAEs is shown in Table 26.

B.2.10.3. Drug exposure

The median overall durations of study treatment, from first dose until last dose (including recovery periods without treatment), were similar between treatment groups, at [redacted] (range: [redacted]) in the GO + DA arm and [redacted] (range: [redacted]) in the DA arm.[24] The median durations of study treatment

for patients who experienced CR/CRp were also similar between treatment groups and by treatment phase.

A total of [redacted] discontinued study drug owing to TEAEs. The proportion of patients who permanently discontinued treatment owing to TEAEs was higher in the GO + DA arm ([redacted]) than in the DA arm ([redacted]; [redacted]). This difference between arms in TEAEs was accounted for mainly by thrombocytopenia ([redacted]) in the GO + DA arm (no discontinuations for this reason in the DA arm) and hepatobiliary disorders (GO + DA arm: [redacted]; DA arm: [redacted]).

Over [redacted] of patients were able to receive all three fractionated doses of GO during induction therapy, and just under half were able to receive GO during the two courses of consolidation therapy (Appendix D2.1).

B.2.10.4. Treatment-related deaths

In total, [redacted] of patients had treatment-related deaths: [redacted] patients in the GO + DA arm and [redacted] patients in the DA arm (Table 24).[24] Of patients who died, a similar number in both treatment arms died owing to study treatment toxicity (GO + DA arm: [redacted]; DA arm: [redacted]).

Table 24. Summary of treatment-related deaths

	GO + DA (N = 131) n (%)	DA (N = 137) n (%)
Treatment-related deaths	[redacted]	[redacted]
Cause of death		
Disease under study	[redacted]	[redacted]
Study treatment toxicity	[redacted]	[redacted]
Unknown	[redacted]	[redacted]
Other	[redacted]	[redacted]
Mechanism of death		
Disease progression or relapse	[redacted]	[redacted]
Septic shock	[redacted]	[redacted]
Infection	[redacted]	[redacted]
GVHD	[redacted]	[redacted]
Liver toxicity	[redacted]	[redacted]
Haemorrhage	[redacted]	[redacted]

Other	[redacted]	[redacted]
Treatment related death during CR/CRp[a]	[redacted]	[redacted]

More than one mechanism of death could have been selected; therefore row totals may exceed the total number of deaths. Treatment-related deaths were fatal SAEs classified by the company as related to any study treatment as reported in the SAE report, grade 5 events (or infection events indicating the patient withdrew from the study) considered related to any study treatment by the investigator on the CRF, or deaths classified as study treatment toxicity by the investigator on the CRF.

[a]Deaths in CR/CRp are defined as patients who experienced CR/CRp by investigator assessment and died without relapse by investigator assessment and did not undergo transplantation.

Source: ALFA-0701 CSR.[24]

CR, complete remission; CRF, case report form; CRp, complete remission with incomplete platelet recovery; DA, daunorubicin + cytarabine; GO, gemtuzumab ozogamicin; GVHD, graft versus host disease; SAE, serious adverse event.

B.2.10.5. Serious adverse events

A high proportion of patients experienced AEs in both the GO + DA ([redacted]) and DA ([redacted]) arms (Table 25). For SAEs, the corresponding proportions were [redacted] and [redacted], respectively. The most common SAE in the GO + DA arm was thrombocytopenia reported in [redacted] ([redacted]) and in [redacted] patients ([redacted]) in the control group. Other common treatment-related SAEs (MedDRA preferred terms) experienced by over 5% of patients in the GO + DA and control treatment arms respectively were: bronchopulmonary aspergillosis in [redacted] ([redacted]) and [redacted] ([redacted]) patients; febrile bone marrow aplasia in [redacted] ([redacted]) and [redacted] ([redacted]) patients; and septic shock in [redacted] ([redacted]) and [redacted] ([redacted]) patients.

Table 25. Summary of AEs and SAEs

	GO + DA (N = 131) n (%)		**DA (N = 137) n (%)**	
	All-causality AEs	**Related AEs**	**All-causality AEs**	**Related AEs**
Patients with AEs	[redacted]	[redacted]	[redacted]	[redacted]
Patients with SAEs	[redacted]	[redacted]	[redacted]	[redacted]
Patients with grade 3 or 4 or severe infection AEs	[redacted]	[redacted]	[redacted]	[redacted]
Patients with fatal events	[redacted]	[redacted]	[redacted]	[redacted]
Patients who permanently discontinued study treatment owing to AEs	[redacted]	[redacted]	[redacted]	[redacted]

Source: ALFA-0701 CSR.[24]

AE, adverse event; DA, daunorubicin + cytarabine; GO, gemtuzumab ozogamicin; SAE, serious adverse event.

B.2.10.6. Pre-defined TEAEs

TEAEs are summarized in Table 26 for the full population of as-treated patients. TEAEs that occurred in the highest proportion of patients in the GO + DA arm were haemorrhage (GO + DA: [redacted]; DA: [redacted]).

Table 26. Predefined TEAEs (all causalities) by maximum CTCAE grade (as-treated population)

	GO + DA (N = 131) n (%)	DA (N = 137) n (%)
CHV CRF[a] predefined CRF term		
Skin toxicity, total	[redacted]	[redacted]
Grade 3	[redacted]	[redacted]
Grade 4	[redacted]	[redacted]
Mucosal toxicity, total	[redacted]	[redacted]
Grade 3	[redacted]	[redacted]
Grade 4	[redacted]	[redacted]
Pain, total	[redacted]	[redacted]
Grade 3	[redacted]	[redacted]
Grade 4	[redacted]	[redacted]
Nausea, vomiting, diarrhoea, total	[redacted]	[redacted]
Grade 3	[redacted]	[redacted]
Grade 4	[redacted]	[redacted]
Constipation, total	[redacted]	[redacted]
Grade 3	[redacted]	[redacted]
Grade 4	[redacted]	[redacted]
Pulmonary toxicity, total	[redacted]	[redacted]
Grade 3	[redacted]	[redacted]
Grade 4	[redacted]	[redacted]
Cardiac rhythm disorder, total	[redacted]	[redacted]
Grade 3	[redacted]	[redacted]
Grade 4	[redacted]	[redacted]
Other cardiac toxicity, total	[redacted]	[redacted]
Grade 3	[redacted]	[redacted]

	GO + DA (N = 131) n (%)	DA (N = 137) n (%)
Grade 4	[redacted]	[redacted]
Central neurological toxicity, total	[redacted]	[redacted]
Grade 3	[redacted]	[redacted]
Grade 4	[redacted]	[redacted]
Peripheral neurological toxicity, total	[redacted]	[redacted]
Grade 3	[redacted]	[redacted]
Grade 4	[redacted]	[redacted]
Retrospective data collection[b]		
Infections and infestations severe (≥ grade 3)	[redacted]	[redacted]
Haemorrhage all grades (≥ grade 1) cluster TEAEs, total[c]	[redacted]	[redacted]
Grade 3	[redacted]	[redacted]
Grade 4	[redacted]	[redacted]
Grade 5	[redacted]	[redacted]
VOD all grades (grade 1) cluster TEAEs, total[c,d]	[redacted]	[redacted]
Grade 3	[redacted]	[redacted]
Grade 4	[redacted]	[redacted]
Grade 5	[redacted]	[redacted]

[a]Notes for CHV CRF data: CHV CRF captured only for grade 3 or 4 events. Patients were counted once, at the highest toxicity severity, at each level (i.e. once at "any non-haematological toxicity" and once for each specific toxicity parameter). [b]Notes for retrospective data collection grade 3. Infections and all grades haemorrhage, VOD and other AEs which led to permanent discontinuation of study drugs were collected and reported. Multiple occurrences of the same AE in a patient at the preferred term level or SOC level were counted as one AE per treatment in each row. [c]In the presence of a patient who had both missing and non-missing CTCAE grades for AEs with the same preferred term, the missing CTCAE grade of the AE was treated as the lowest severity grade. [d]Includes preferred term of veno-occlusive liver disease and veno-occlusive disease.

Grade per CTCAE v3.0. MedDRA v18.0 coding dictionary applied.

Source: ALFA-0701 CSR.[24]

AE, adverse event; CHV, Centre Hospitalier de Versailles; CRF, case report form; DA, daunorubicin + cytarabine; GO, gemtuzumab ozogamicin; CTCAE, Common Terminology Criteria for Adverse Events; MedDRA, Medical Dictionary for Regulatory Activities; SOC, system organ class; TEAE, treatment-emergent adverse event; v, version; VOD, veno-occlusive disease.

Rates of TEAEs in the favourable/intermediate cytogenetics profile subpopulation of the as-treated population were consistent with the overall population (Table 27).[99]

Table 27. Predefined TEAEs (all causalities) by maximum CTCAE grade (favourable/intermediate cytogenetics profile subpopulation)

	GO + DA (N = 91) n (%)	DA (N = 95) n (%)
CHV CRF[a] predefined CRF term		
Skin toxicity, total	[redacted]	[redacted]
Grade 3	[redacted]	[redacted]
Grade 4	[redacted]	[redacted]
Mucosal toxicity, total	[redacted]	[redacted]
Grade 3	[redacted]	[redacted]
Grade 4	[redacted]	[redacted]
Pain, total	[redacted]	[redacted]
Grade 3	[redacted]	[redacted]
Grade 4	[redacted]	[redacted]
Nausea, vomiting, diarrhoea, total	[redacted]	[redacted]
Grade 3	[redacted]	[redacted]
Grade 4	[redacted]	[redacted]
Constipation, total	[redacted]	[redacted]
Grade 3	[redacted]	[redacted]
Grade 4	[redacted]	[redacted]
Pulmonary toxicity, total	[redacted]	[redacted]
Grade 3	[redacted]	[redacted]
Grade 4	[redacted]	[redacted]
Cardiac rhythm disorder, total	[redacted]	[redacted]
Grade 3	[redacted]	[redacted]
Grade 4	[redacted]	[redacted]
Other cardiac toxicity, total	[redacted]	[redacted]
Grade 3	[redacted]	[redacted]
Grade 4	[redacted]	[redacted]
Central neurological toxicity, total	[redacted]	[redacted]
Grade 3	[redacted]	[redacted]
Grade 4	[redacted]	[redacted]
Peripheral neurological toxicity, total	[redacted]	[redacted]
Grade 3	[redacted]	[redacted]
Grade 4	[redacted]	[redacted]
Retrospective data collection[b]		

	GO + DA (N = 91) n (%)	DA (N = 95) n (%)
Infections and infestations severe (≥ grade 3)	[redacted]	[redacted]
Haemorrhage all grades (≥ grade 1) cluster TEAEs, total[c]	[redacted]	[redacted]
Grade 3	[redacted]	[redacted]
Grade 4	[redacted]	[redacted]
Grade 5	[redacted]	[redacted]
VOD all grades (grade 1) cluster TEAEs, total[c,d]	[redacted]	[redacted]
Grade 3	[redacted]	[redacted]
Grade 4	[redacted]	[redacted]
Grade 5	[redacted]	[redacted]

[a]Notes for CHV CRF data: CHV CRF captured only for grade 3 or 4 events. Patients were counted once, at the highest toxicity severity, at each level (i.e. once at "any non-haematological toxicity" and once for each specific toxicity parameter). [b]Notes for retrospective data collection grade 3. Infections and all grades haemorrhage, VOD and other AEs which led to permanent discontinuation of study drugs were collected and reported. Multiple occurrences of the same AE in a patient at the preferred term level or SOC level were counted as one AE per treatment in each row. [c]In the presence of a patient who had both missing and non-missing CTCAE grades for AEs with the same preferred term, the missing CTCAE grade of the AE was treated as the lowest severity grade. [d]Includes preferred term of veno-occlusive liver disease and veno-occlusive disease.

Grade per CTCAE v3.0. MedDRA v18.0 coding dictionary applied.

Source: Treatment-emergent AEs in patients with favourable/intermediate cytogenetics profile[99]

AE, adverse event; CHV, Centre Hospitalier de Versailles; CRF, case report form; DA, daunorubicin + cytarabine; GO, gemtuzumab ozogamicin; CTCAE, Common Terminology Criteria for Adverse Events; MedDRA, Medical Dictionary for Regulatory Activities; SOC, system organ class; TEAE, treatment-emergent adverse event; v, version; VOD, veno-occlusive disease.

B.2.10.7. Haematological toxicity

Treatment-induced cytopenia was observed in both treatment arms (Table 28).[24] Grade 3/4 persistent thrombocytopenia was defined as the non-recovery of a platelet count > 50 000/mm^3 at day 45 after the start of a treatment course. During the induction phase, [redacted] of patients in the GO + DA arm and [redacted] of patients in the DA arm had platelet counts recover to > 50 000/mm^3. The corresponding proportions were [redacted] and [redacted] during the first consolidation phase and [redacted] and [redacted] during the second consolidation phase. At each treatment phase, the median time to platelet recovery was [redacted] in the GO + DA arm than the DA arm.

Table 28. Haematological toxicity

	GO + DA arm (n = 131)		DA arm (n = 137)	
Treatment-induced thrombocytopenia (< 50 000 cells/mm^3)				
	Patients with platelet recovery, n (%)	**Duration of thrombocytopenia, days[a]**	**Number of patients with platelet recovery**	**Duration of thrombocytopenia, days[a]**
After induction	[redacted]	[redacted]	[redacted]	[redacted]
After first consolidation	[redacted]	[redacted]	[redacted]	[redacted]
After second consolidation	[redacted]	[redacted]	[redacted]	[redacted]

Source: ALFA-0701 CSR.[24]
DA, daunorubicin + cytarabine; GO, gemtuzumab ozogamicin.

B.2.10.8. Infections

Treatment-related severe infections (grade ≥ 3) were experienced by a similar proportion of patients in the GO + DA arm ([redacted]) and the DA arm ([redacted]; Table 26), and the total times for which patients experienced severe infections were similar between the GO + DA and DA arms across all phases of treatment (median: [redacted] versus [redacted]).[24] In the subpopulation of patients with favourable/intermediate cytogenetics, the corresponding proportions of patients with treatment-related severe infections (grade ≥ 3) in the GO + DA and DA arms were [redacted] and [redacted], respectively.[99]

B.2.10.9. Hepatotoxicity

The most commonly occurring hepatoxicity TEAE was VOD, which was reported in [redacted] of patients in the GO + DA arm and [redacted] of patients in the DA arm (in both arms, all cases of VOD were identified as being treatment-related).[24] Both patients who experienced VOD in the DA arm experienced VOD 28 days after the last dose of GO + DA (Table 29).

Table 29. Timing of veno-occlusive disease (as-treated population)

	GO + DA (N = 131)	DA (N = 137)
Patients with VOD, n (%)	[redacted]	[redacted]

	GO + DA (N = 131)	DA (N = 137)
Days from last dose of study treatment to VOD		
Mean (SD)	[redacted]	[redacted]
Range	[redacted]	[redacted]
Time from HSCT to VOD (days), N	[redacted]	[redacted]
Mean (SD)	NA	[redacted]
Range	NA	[redacted]
Timing of VOD in relation to study drug,[a,b,c] n (%)		
Within 28 days after receiving any GO[a]	[redacted]	[redacted]
After 28 days after last dose of any GO[a]	[redacted]	[redacted]
With no prior GO received[a]	[redacted]	[redacted]
Timing of VOD with regard to HSCT,[b] n (%)		
With no prior HSCT	[redacted]	[redacted]
Within 28 days after receiving an allogeneic HSCT	[redacted]	[redacted]
After 28 days after receiving an allogeneic HSCT	[redacted]	[redacted]
Within 28 days after receiving an autologous HSCT	[redacted]	[redacted]
Within 28 days after receiving an autologous HSCT	[redacted]	[redacted]

Last dose of study treatment was determined from the GO, daunorubicin and cytarabine dosing records and also included idarubicin (a component of salvage therapy).

[a]Follow-up treatments (after study treatments) were included when determining any GO treatment. [b]One patient experienced two episodes of VOD, but only the first episode was taken into consideration. [c]One patient was not included as VOD occurred 28 days before a GO dose and the patient was rechallenged with GO after VOD.

Source: ALFA-0701 CSR.[24]

DA, daunorubicin + cytarabine; GO, gemtuzumab ozogamicin; HSCT, haematopoietic stem cell transplantation; NA, not applicable; NE, not estimable; SD, standard deviation; VOD, veno-occlusive disease.

In the subpopulation of patients with favourable/intermediate cytogenetics, the corresponding proportions of patients with VOD in the GO + DA and DA arms were [redacted] and [redacted], respectively.[99]

B.2.10.10. Hospitalisations

All patients were hospitalized during the study for administration of study treatment.[24] However, more patients in the GO + DA arm ([redacted]) were readmitted or had planned hospitalization prolonged owing to reasons related to AEs, compared with patients in the DA arm ([redacted]). The rates of admission to intensive care units were similar between the treatment groups (GO + DA arm: [redacted]; DA arm: [redacted]).

B.2.11 Ongoing studies

There are no ongoing trials assessing GO + DA in line with the expected marketing authorisation dosing schedule. Details of all ongoing studies are summarized in Appendix L1.4.

B.2.12 Innovation

The primary goal of treatment in de novo AML is to achieve and maintain CR/CRp for as long as possible, since relapse in AML is associated with poor survival and high healthcare costs.[15,16,49,61] However, despite 40–70% of patients with newly diagnosed, previously untreated AML achieving an initial CR with DA, a high proportion of patients relapse and face a low chance of achieving a second CR.[16,79-83,85,100] Indeed, fewer than 25% of people who go into remission following treatment with currently available therapies will survive beyond 3 years.[79] There is a need for an approved therapy that can increase the durability of remission in patients with AML and reduce mortality rates.

B.2.12.1. GO + DA delivers a more durable response compared with DA alone, with a more marked clinical benefit in patients with favourable/intermediate cytogenetics profile

GO targets AML blasts to induce cell death, and in combination with DA is able to extend the duration of remission.[2,3,24,89] In ALFA-0701, GO + DA significantly improved EFS and RFS when added to DA, compared with DA alone.[89] The clinical benefit of adding GO to DA was particularly apparent in patients with favourable/intermediate cytogenetics profile, but not in patients with unfavourable cytogenetics profile.

GO as an add-on to DA therefore represents a step-change in the management of adult patients with de novo AML. This improvement is meaningful to patients, particularly those with favourable/intermediate cytogenetics profile and to the NHS. This is by reducing the relapses, which can impact on patients' HRQoL, and are associated increased costs owing to the need for hospitalization and chemotherapy to induce a second remission.

B.2.12.2. GO specifically targets CD33-positive blast cells in AML

As described in section B.1, GO is an anti-CD33 antibody that is conjugated to calicheamicin, a potent cytotoxic agent.[2,3] GO is able to directly target CD33-positive AML blasts in order to induce death of leukaemic cells;[2,3] worldwide approximately 85–90% of patients with AML are CD33 positive.[2,7] The selectivity of GO minimizes the impact of calicheamicin on healthy cells and tissues.

B2.13 Interpretation of clinical effectiveness and safety evidence

B.2.13.1. Efficacy

GO + DA prolonged the time to relapse or death in patients with de novo AML versus DA alone, particularly in those with favourable/intermediate cytogenetics profile

In ALFA-0701, according to IRC analysis, GO + DA improved EFS ([redacted]) and reduced the risk of an event by [redacted]. In patients achieving CR/CRp, GO + DA significantly improved RFS ([redacted]) and reduced the relapse risk by [redacted]. In patients with favourable/intermediate cytogenetics profile had an EFS ([redacted]) and RFS ([redacted]) benefit when they received GO + DA, compared with DA alone. But this clinical benefit was not observed in patients with unfavourable cytogenetics profile.[24] These trends are aligned with those in the IPD meta-analysis, which provided supportive evidence (Appendix D.3.1).

The response to GO + DA in the mITT population and in the subpopulation with favourable/intermediate cytogenetics profile was durable, as indicated by the fact that the KM curves for EFS and RFS plateaued from [redacted] onwards, indicating a decrease in hazard. The patients that are represented by this portion of the KM curves likely represent those who are ‘functionally cured’. This is aligned with UK clinical expert opinion that patients who achieve EFS or RFS for 3–5 years are regarded as functionally cured and have a risk of death that is similar to the general population (as described in Section B1). Therefore, despite the fact that response rates were similar irrespective of whether patients received GO + DA or DA alone, adding GO to DA increases the durability of the response versus DA, and makes it more likely that patients will remain in remission for long enough (3–5 years) to achieve the status of being ‘functionally cured’.

An extension in time to relapse or death represents a clinically important benefit to patients. Patients who experience CR/CRp experience improvements in their QoL over time compared with patients who do not experience CR (see section B.1), in part by delaying the inconvenience and toxicities associated with subsequent therapies.[43,44] Extending EFS and RFS is also likely to benefit the NHS by increasing the time before patients need medical intervention to manage relapses, thereby reducing healthcare resource use and costs. In line with this, an *ad hoc* analysis in ALFA-0701 demonstrated that the median time to patients requiring subsequent anti-cancer therapy following induction failure or relapse was significantly improved by 9 months in patients receiving GO + DA compared with DA ([redacted]).

Although GO + DA was not associated with a significant OS benefit in the mITT population of ALFA-0701 or the subpopulation with favourable/intermediate cytogenetics profile, this may have been because the trial was not powered to measure changes in OS (i.e. pre-specified events for primary analysis at a specific statistical power). Furthermore, OS would have been confounded by follow-up therapies that patients received, including HSCT and control arm cross-over to GO. However, in larger UK studies, a significant improvement in 3-year survival has been observed in patients over 60 years of age following the addition of GO to induction chemotherapy.[31,36] Furthermore, a significant OS advantage was observed in the GO versus no-GO arm of the IPD meta-analysis.[35]

UK clinical experts anticipate that the findings of ALFA-0701 will be generalizable to the UK population. This is owing to the fact that baseline characteristics of ALFA-0701 are similar to baseline characteristics of in trials conducted in the UK, MRC AML15 and NCRI AML16. Furthermore, In ALFA-0701, nearly 70% of patients had favourable/intermediate cytogenetics profile.[24] In UK studies (AML15 and AML16), the proportions of patients with AML who had favourable/intermediate cytogenetic risk were in the range 55–64%.[31,36] These results indicate that a large proportion of with AML have the potential to benefit from treatment with GO + DA in the UK.

In other subgroups evaluated, the benefit of GO + DA was consistent with outcomes for the overall population

The effects of GO + DA compared with DA alone were generally consistent with the overall results across most of the other subgroups. This indicates that the benefits of GO + DA are likely to be observed across the range of patients in clinical practice, but are particularly apparent in patients with favourable/intermediate cytogenetics profile.

B.2.13.2. Safety

GO + DA has a predictable and manageable safety profile

In ALFA-0701, GO + DA had a safety profile that was consistent with that of the individual therapies.[24] Over [redacted] of patients were able to receive all three fractionated doses of GO during induction therapy, and [redacted] were able to receive GO during the two courses of consolidation therapy.

A similar proportion of patients in both arms died (GO + DA arm: [redacted]; DA arm: [redacted]), and a similar proportion died owing to study treatment toxicity (GO + DA arm: [redacted]; DA arm: [redacted]).[24] When TEAEs were considered, the proportions of patients with platelet recovery were similar between the arms, but the time for recovery was longer in the GO + DA arm than the DA arm.

VOD is a well-recognized complication following treatment with GO and clinicians are better able to manage this risk owing to their ongoing clinical experience with GO. In the present study, a higher proportion in the GO + DA versus DA arm had VOD ([redacted] and [redacted], respectively); notably, for patients with VOD in the DA arm, this occurred within 28 days after the last dose of GO. However, these rates are regarded by clinical experts in the UK as being acceptably low, and importantly are lower than in previous studies, which can be attributed to the fractionated dosing of GO. Other TEAEs such as severe infections were experienced by similar proportions of patients in the GO + DA and DA arms and consistent with this, similar proportions of patients had neutrophil recovery after each treatment phase.

For SAEs, a higher proportion of patients in the GO + DA arm than the DA arm experienced SAEs, with the most frequently occurring SAE in the GO + DA arm being thrombocytopenia (GO + DA arm: [redacted]; DA arm: [redacted]).[24] However, this can

be managed through supportive measures and the withholding of GO during consolidation therapy.

Results in the subpopulation of patients with favourable/intermediate cytogenetics profile were consistent with the overall as-treated population.

B.2.13.3. Strengths of the evidence base

Internal validity of the evidence base

The primary evidence for the efficacy and safety of GO + DA comes from a high-quality multicentre RCT involving 271 patients randomized 1:1 to GO in combination with DA or to DA alone.[24,89] This study included patients with previously untreated, de novo AML who were aged 50–70 years. AML is a heterogeneous disease, and the eligibility criteria of ALFA-0701 allowed for a broad range of predefined patient subgroups to be included, from across the cytogenetic, genetic and performance status spectrum. Furthermore, data collected according to regulatory standards that were used in the IRC analyses were in good agreement with data collected during investigator analyses, which may be reflective of clinical practice.

External validity of the evidence base

This study also has high external validity: according to UK clinical experts, the patient population in ALFA-0701 is likely to be representative of the UK population. This is supported by the fact that patient baseline characteristics in ALFA-0701 (appendix D2.2) were consistent with those in UK clinical trial, MRC AML15 and NCRI AML16 (included in the IPD meta-analysis). Outcomes in ALFA-0701 are therefore likely to be generalizable to the UK population.

Data from ALFA-0701 are also likely to be generalizable to patients outside the 50–70 years age range because eligibility for intensive chemotherapy is based mostly on fitness not on age, although age and fitness may be correlated in some cases.

The effects of GO + DA were most apparent in the subpopulation of patients with favourable/intermediate cytogenetics profile.

Evidence-based dosing regimen

The dosing of GO was based on a large body of pre-existing clinical data. Phase 2 studies administering two doses of GO at 9 mg/m^2 14 days apart demonstrated that at this dose and schedule, GO was associated with haematological toxicity and frequent liver toxicity, including VOD.[24] In SWOG S0106, GO was given at a dose of 6 mg/m^2, and was associated with excess mortality not counterbalanced by later benefit.[86] Subsequently, two phase 2 studies using GO (3 × 3 mg/m^2) with DA demonstrated benefit without excessive haematological toxicity.[11,24,101] On this basis, the ALFA-0701 study dosing regimen and schedule were chosen.

In ALFA-0701, GO was administered alongside DA in the GO + DA arm and patients in DA arm did not receive GO.[24] DA is the most commonly used UK standard of care at present for patients who are not eligible for enrolment in clinical trials, but who are able to receive intensive chemotherapy.[1] In patients who are eligible for enrolment in clinical trials, currently ongoing trials administer GO in combination with DA during induction therapy; in these trials, GO can be administered either with DA as a single dose or, in newer studies, as a fractionated dose (GO + DA arm), or patients can receive DA alone (DA arm).

Clinical practice for treating AML in the UK is constantly evolving and being adapted in light of emerging evidence, and completed or ongoing clinical trials exploring treatment with fractionated GO in combination with DA.

Clinically relevant endpoints

The primary endpoint of the ALFA-0701 study was EFS.[24] EFS is recognized by haematologists to represent clinical benefit by assessing the duration of delay before the inconvenience and toxicity associated with subsequent therapies and morbidities that accompany recurrent AML. EFS is not confounded by therapies subsequent to relapse in AML. Secondary outcomes were OS, haematological remission and RFS. EFS, OS and RFS have all been noted as recognized endpoints in AML clinical trials by international expert panels, and in clinical guidance documents, including from the National Cancer Institute.[24] Extending the duration of RFS is a key outcome in AML, which reduces costs and resource use for the NHS. All outcomes were based on laboratory data and not subjective criteria that are likely to be influenced by a lack of

investigator blinding. In line with this, data analysed by the IRC were consistent with those from investigator-analysed data.

B.2.13.4. Weaknesses of the evidence base

Study design

ALFA-0701 is an open label, investigator-sponsored trial that was not originally designed to meet regulatory requirements.[24] However, a retrospective data collection was conducted to ensure that the data met regulatory requirements. These data were included in an IRC analysis, which reported outcomes that were in agreement with the original investigator analyses.

ALFA-0701 was not powered to measure OS

ALFA-0701 was powered to measure EFS, but not OS.[24] OS is an impractical endpoint in AML owing to confounding by post-induction treatment, including other chemotherapy regimens, HSCT and control arm cross-over to GO. No appropriately censored data were available to enable robust conclusions around benefits of GO in OS.[24]

Quality-of-life data

No QoL data were generated as part of ALFA-0701, and no QoL data were available from the meta-analyses. To address the limited QoL data available for AML, Pfizer have conducted a vignette study to measure the utilities associated with different health states in AML based on preferences of the UK general population – such evidence has been accepted by NICE for other submissions, when alternative data are lacking.[102] Pfizer have also conducted SLRs to identify EQ-5D QoL data in the literature to match the NICE reference case.

B.2.13.5. End-of-life criteria

GO + DA in adult patients with de novo AML is not expected to meet end-of-life criteria.

B.2.13.6. Life expectancy and potential patient population

A diagnosis of AML is associated with poor survival

AML is a rare, incurable, malignant haematological disease that progresses rapidly and is typically fatal within weeks or months if left untreated.[16]

As described in section B.1, AML is a heterogeneous disease and patient age at diagnosis is a key modifier of survival probability. In England, 5-year relative survival was worse in patients aged ≥ 65 years than in those aged 25–64 years (20.1% versus 62.6%; using data from 2008–2010).[34] In the UK, in a study which did not restrict by age, the median survival of patients with AML was estimated to be 9.5 months (1987–2006 data; Table 30).[33] However, the population of patients in this study may include those with poor outcomes or those who may not have been able to tolerate intensive chemotherapy, which may not reflect the population of patients who would be eligible to receive GO + DA in the UK.

Table 30. Survival of patients with AML

Country data	N	Age range (median), years	Median OS	Survival	
			Months	Time period	%
UK[33]	NR	All	9.5	2001–2010	5-year: 20.0
England[34]	NR	25–64 vs 65+	NR	2008–2010	5-year: 37.9 vs 4.5%
Survival data from RCTs					
MRC AML15 (UK/Denmark/NZ)[35, 36]	1113	15–71 (50)	27.5[a]	2002–2009	5-year: 41.0[a]
NCRI AML16 (UK/Denmark)[31,35]	1115	51–84 (67)	12.0[a]	2006–2011	4-year: 15.0[b]

[a]Chemotherapy arm only.

Source: Bhayat *et al*. 2009;[33] Burnett *et al*. 2011;[36] Burnett *et al*. 2012;[31] National Cancer Intelligence Network;[34] Pfizer meta-analysis CSR[35]

AML, acute myeloid leukaemia; CI, confidence interval; MRC, Medical Research Council; N/A, not applicable; NCRI, National Cancer Research Institute; NR, not reported; NZ, New Zealand; OS, overall survival; RCT, randomized controlled trial; UK, United Kingdom.

In line with this, median OS in UK clinical trials (AML15, AML16) in patients who received intensive chemotherapy ranged from 12 months in those with a median age of 67 years to 27.5 years in those with a median age of 50 years. In UK clinical practice, UK clinical experts have estimated that approximately 80% of patients will enter a trial. Patients able to tolerate intensive chemotherapy are likely to be enrolled

in trials. Patients who can tolerate intensive chemotherapy are also likely to be eligible to receive GO + DA. Therefore, survival estimates from trials are likely to be representative of survival in patients who can receive GO + DA.

In the UK, GO is anticipated to be indicated as a first-line treatment in combination with DA for adult patients with previously untreated, de novo AML not known to have unfavourable cytogenetics profile. The estimated population in England who would be expected to be eligible to receive GO + DA is 715–738 patients per year for the next 5 years (see Budget impact analysis section).

B.3 Cost effectiveness

***De novo* cost-effectiveness model**

- A *de novo* lifetime cohort state-transition model (Microsoft Excel®) with an NHS and PSS perspective was constructed to evaluate the cost-effectiveness of GO as first line treatment in combination with intensive chemotherapy (DA) for de novo AML patients. This submission investigates two district populations, one aligned with anticipated EU MA and a subpopulation excluding patients with adverse cytogenetics that addresses the decision problem of this submission and reflects UK clinical practice. The latter is the focus of this section and the former is included in the appendix for completeness.
- Clinical outcomes used to inform this economic analysis including response rates, relapse free survival, overall survival, probability of HSCTs, occurrence of adverse events were sourced from the ALFA-0701 study. Health-related quality of life data, costs, medical resource use data were sourced from published literature, past NICE technology appraisal TA399, clinical opinion and the ALFA-0701 study.
- The key driver of relative cost-effectiveness was the difference in relapse-free survival (RFS), which meant more patients in the GO arm remained relapse-free for longer than in the comparator arm and accrued more QALYs and less of the high costs associated with relapse. Average accrued lifetime costs and QALYs in the refractory health state (failed induction) were identical between arms, but almost [redacted] of patients in the ALFA-0701 study achieved complete remission where these advantages of GO are prominent.

Survival

- Mixture cure models were used to model both treatment arms for RFS and overall survival in remission states (Log-normal distribution) to properly fit the tail of the KM data; standard parametric curves were deemed inappropriate because they cut the plateau and so did not reflect the benefit GO provided in extending RFS that was shown in the ALFA-0701 trial. Flexible spine models were also investigated. Overall survival for patients in the refractory state used standard parametric models (Gompertz distribution) in the base case.
- Clinical outcome projections were validated by UK clinical experts who were aligned in their expectation that the addition of GO significantly improves outcomes for those patients in relapse free survival in a way consistent with long term model predictions.

Base case (cytogenetic subpopulation) results

- For the cytogenetic subpopulation, the lifetime incremental QALYs and life-years were [redacted], respectively, yielding an ICER of £12,251 for the average patient. Therefore GO represents a cost-effective use of NHS resources.

Sensitivity analyses

- Parameter uncertainty was explored through probabilistic sensitivity analysis with structural uncertainty and key assumptions explored through scenario analyses and deterministic one-way sensitivity analyses. Probabilistic ICERs were similar to deterministic ICERs.
- The cytogenetic subpopulation results were most sensitive to the probabilities of HSCT and pooled restricted mean survival time for those patients that relapsed.
- A variety of scenario analyses were tested and showed that ICERs were relatively stable to changes in the methods for survival analysis, ranging from £6,821 to £12,233. The mixture cure model base case gave the highest ICERs.

B.3.1 Published cost-effectiveness studies

B.3.1.1 Systematic Literature Review

An SLR was conducted to identify relevant economic evaluations. The SLR identified nine cost-effectiveness evaluations of treatments for AML (detailed in Appendix G), none of which evaluated GO plus DA. Of the nine analyses, only four were conducted from a UK payer perspective. Of these four analyses, only two provided sufficient detail on model methodology and reported actual incremental cost-effectiveness ratios (ICERs); both were conducted as part of previous submissions to NICE, Table 31. NICE TA399 was used as a source for adverse event and health state utilities.[66,103]

Table 31 Summary list of published cost-effectiveness studies

Study	Year	Summary of model	Patient population (average age in years)	QALYs (intervention, comparator)	Costs (currency) (intervention, comparator)	ICER (per QALY gained)
Edlin, 2011[104] Full-text NICE TA218	2011	Critical review of cost-effectiveness analysis (state transition model) comparing azacitidine and a conventional care regimen.	Patients with AML, 20%-30% bone marrow blasts, and multi-lineage dysplasia who were not eligible for HSCT.	QALY gain (company submission after appeal following ERG input and suggested amendments) Azacitidine vs. BSC: 1.01 Azacitidine vs. low-dose chemo: 1.34 Azacitidine vs. usual care: 1.09	NR	Manufacturer's base-case: Azacitidine vs. BSC: £47,432/QALY Azacitidine vs. low-dose chemo: £40,754/QALY Azacitidine vs. usual care: £37,105/QALY Base-case plus vial sharing scenario: Azacitidine vs. BSC: £44,440/QALY Azacitidine vs. low-dose chemo: £37,929/QALY Azacitidine vs. usual care: £34,366/QALY Company submission after appeal (following ERG input and suggested amendments): Azacitidine vs. BSC: £63,177/QALY Azacitidine vs. low-dose chemo: £49,030/QALY Azacitidine vs. usual care: £56,945/QALY Committee best available estimate: £47,200/QALY The probabilistic ICERs all £47,000-£48,000.
Tikhonova, 2016[103] Full-text	2016	Critical review of cost-effectiveness analysis (partition-	Patients with AML with more than 30% bone marrow blasts	NR	NR, but drug acquisition and administration costs were	Company submission: Base-case: £20,648/QALY Patients with poor cytogenetics: £20,227/QALY

Study	Year	Summary of model	Patient population (average age in years)	QALYs (intervention, comparator)	Costs (currency) (intervention, comparator)	ICER (per QALY gained)
NICE TA399		survival model) comparing azacitidine with conventional care regimen (comprised of three individual comparators: intensive chemotherapy followed by BSC upon disease relapse or progression, non-intensive chemotherapy followed by BSC, and BSC alone).	who were not eligible for hematopoietic stem-cell transplantation		the largest cost components in the azacitidine and CCR arms	Patients with MDS-related changes: £19,175/QALY ERG amendments: After model errors corrected: £20,648/QALY After all amendments: £273,308/QALY Company submission: Deterministic sensitivity analysis showed that the ICER was most sensitive to administration costs in the CCR arm, the HR for OS, the remission rates in the CCR arm, and the acquisition and administration costs in the azacitidine arm. The mean ICER in this analysis was £17,423 per QALY. ERG amendments: Deterministic sensitivity analysis resulted in ICERs of above £200,000 per QALY. The mean ICER of £277,123 per QALY was obtained in a probabilistic sensitivity analysis conducted by the ERG.

AML, acute myeloid leukaemia; BSC, best supportive care; CCR, conventional care regimen; ICER, incremental cost-effectiveness ratio; ERG, evidence review group; HR, hazard ratio; HSCT, haematopoietic stem cell transplantation; MDS, myelodysplastic syndrome; NR, not reported; QALYs, quality-adjusted life years; OS, overall survival.

B.3.2 Economic analysis

B.3.2.1. Patient population

This economic evaluation is aligned with the decision problem outlined in section B.1 (Table 1) and thus focuses on the cost-effectiveness of GO as add-on treatment to first line standard therapy (DA) in de novo AML patients eligible for intensive chemotherapy. A de novo model was developed because no published models evaluated GO plus DA. The analysis utilises ALFA-0701 individual patient level mITT data (see section B.2.3) and is presented for two cohorts due to distinct differences in expected efficacy and in particular relapse free survival:

- A subpopulation of the total trial population (mITT) that excludes patients with unfavourable cytogenetics profile, referred to as "favourable/intermediate" throughout the submission dossier. AML is a heterogeneous disease and baseline cytogenetic risk constitutes one of the most significant prognostic makers of diseases outcomes.[24] It was therefore clinically meaningful to include this subpopulation in the economic analysis. Moreover GO has no proven benefit for patients with unfavourable cytogenetic risk profile and so it was deemed appropriate to exclude these patients from our base case analyses.

- The total trial population (mITT) or "all patients" that includes patients with favourable, intermediate, unfavourable and unknown cytogenetic profiles in alignment with anticipated EU MA. Those results were presented for completeness and can be found in Appendix O; input and results tables/graphs in this appendix follow the same order as the presentation of subpopulation analyses in the main section.

A summary of the ALFA-0701 study, including baseline patient characteristics per populations included in the model is provided in section B.2 and Appendix D. Key model inputs that varied between the 2 populations included response rates (see section B.3.3.1), survival analyses of relapse-free and overall survival (section B.3.3.2), probability of HSCT (section B.3.3.8), and incidence of AEs (section B.3.4.3).

B.3.2.2. Model structure

A cohort state-transition model was developed incorporating relatively novel survival analyses compared to those previously submitted to NICE.[103,104] The design of the model structure was informed by the clinical pathway, clinical expert input, previous AML models,[105-107] and the nature of the available data. Relevant health states were identified, developed, and validated as part of a preference elicitation study commissioned by Pfizer.[108] The corresponding model health states and descriptions of what they capture are presented in Appendix M. The model structure is presented in Figure 11.

The model followed the National Health Service (NHS) and Personal Social Services (PSS) perspective to align with the NICE reference case (Table 32). In order to capture lifetime costs and benefits the time horizon is 40 years. UK clinical experts advised (see Table 50) that each cycle of active therapy (induction, consolidation, and salvage therapy) lasts for approximately 1 month therefore cycles of treatment are modelled accordingly.

De novo AML patients enter the model on commencement of systemic therapy (either GO + DA or DA alone). Patients receive one or two induction courses, depending on their initial response to treatment. At the end of induction therapy, patients are assessed and either attain CR or CRp, or fail induction therapy (i.e., are refractory). All patients leave the induction therapy health states after two model cycles. Patients who attain CR or CRp continue up to two courses of consolidation therapy based on data from the ALFA-0701 study. However, some patients in ALFA-0701 with poor risk profiles did not complete the consolidation courses due to disease progression or adverse events, and received a HSCT. This is reflected in the model.

Patients in CR or CRp who relapse, and patients who are refractory to induction therapy, move to second-line therapies (salvage or non-curative). From this group of patients, some can receive salvage therapy, with the aim of attaining second-line CR or CRp and subsequently a potentially curative HSCT. The proportions of patients who receive salvage therapy upon entering the relapse or refractory health states was based on their level fitness and clinician's discretion (e.g., their suitability for

salvage therapy). Following salvage therapy, patients receive best supportive care, which is also based on clinical opinion (Table 50). Patients who are not deemed fit enough for salvage therapy receive non-curative therapies, i.e. low-intensity chemotherapy or best supportive care. Patients who fail to attain second-line CR or CRp following salvage therapy receive best supportive care only.

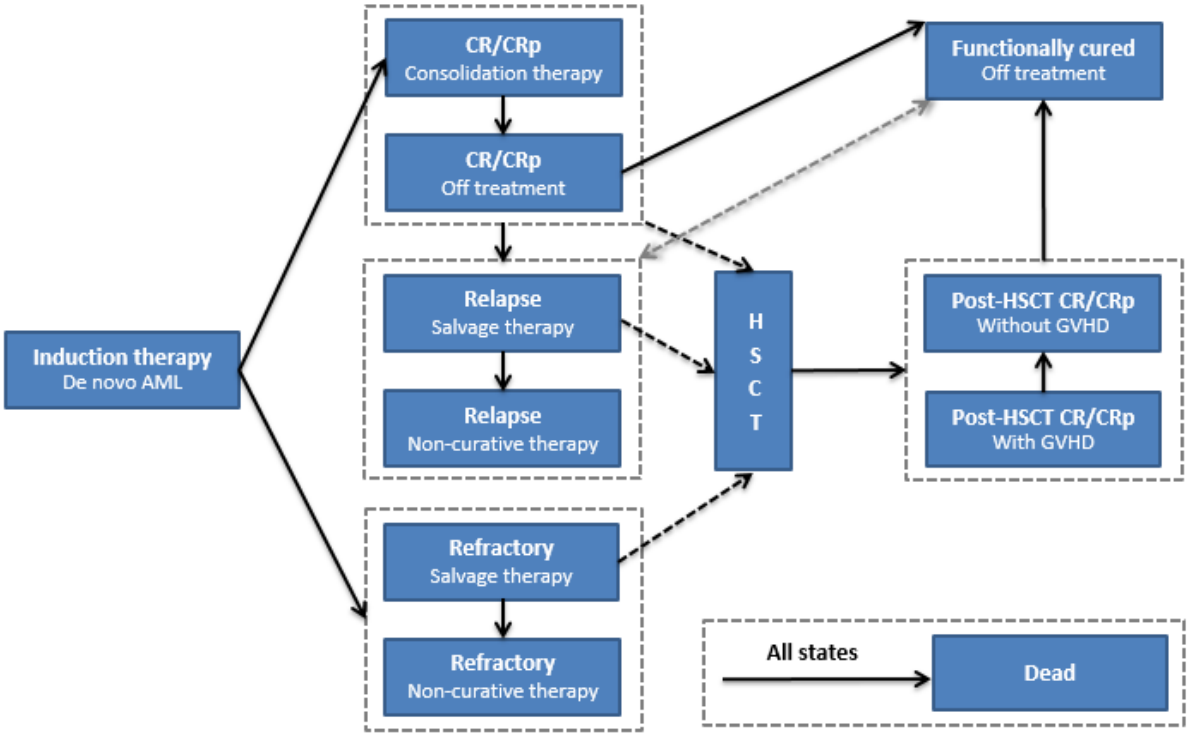
Patients in the CR or CRp, relapse, and refractory health states can receive HSCT and the calculations for this in the model are based on ALFA-0701 data. UK clinicians advised (Table 50) that patients receiving a HSCT need to be hospitalized for more than a month in an isolation room. Therefore the period of hospitalization associated with an HSCT procedure is assumed to last for one cycle (i.e. a month) in the model.

It is implicit that all patients who received a HSCT in ALFA-0701 from the relapse or refractory health state had first achieved second-line CR or CRp, and this was verified by clinical opinion. However, some relapsed patients in ALFA-0701 who did not receive a HSCT but had a prolonged survival of 5 years in the relapse state are likely to have achieved second line CR or CRp and can be considered functionally cured. This is captured in the model and was validated by clinical opinion (Table 50) and literature.[109] The proportion of the total cohort for which the base case model makes this assumption at 5 years was [redacted] and [redacted] for GO and the comparator arm, respectively. Therefore, this assumption made little impact to incremental results.

Patients then move to a post-transplant CR or CRp health state, either with or without graft versus host disease (GVHD). Patients who remain in the CR or CRp or the post-HSCT CR or CRp health states without relapsing for a fixed duration of 5 years transition to the functionally cured health state. In clinical practice, patients are classified as being functionally cured (i.e., have long-term disease-free survival) after remaining in CR or CRp for between 3 and 5 years. The model assumes a duration of 5 years which is aligned with plateau trend seen in ALFA-0701 KM data (plateau phase starting roughly at [redacted]). The cure time point of three years is tested in scenario analyses for completeness.

Deaths from the post-HSCT CR or CRp health state or from the functionally cured health state are captured by the underlying OS curves. Relapses following second-line CR or CRp were not measured in the ALFA-0701 study. Therefore, first relapse of those patients who attained induction success was captured in the post-HSCT CR or CRp health state and in the functionally cured health state by the underlying RFS curve, whereas relapses following HSCT from the relapse and refractory health states could not be captured.

Figure 11 Model Structure Diagram



AML, acute myeloid leukaemia; CR, complete remission; CRp, complete remission with incomplete platelet recovery; GVHD, graft versus host disease; HSCT, hematopoietic stem-cell transplantation.

Data from the ALFA-0701 study on the timing of disease assessments and duration of study treatments are presented in Appendix M. The model applied half cycle correction to costs and QALYs. In line with the NICE reference case an annual discount rate of 3.5% is applied to costs and outcomes.

B.3.2.3. Feature of the de novo economic analysis

Table 32 Features of the de novo economic analysis

Factor	Chosen values	Justification
Time horizon	40 years	Sufficient to ensure that all costs and benefits over a patient's life time are captured
Cycle length	1 month	Consistent with the length of the cycles of active therapy relevant to the model and short enough to accurately model costs and outcomes
Half-cycle correction	Yes	Mitigates bias due to cycle length
Were health effects measured in QALYs; if not, what was used?	Yes	NICE reference case
Discount of 3.5% for utilities and costs	Yes	NICE reference case
Perspective (NHS/PSS)	Yes	NICE reference case

NHS, National Health Service; NICE, National Institute of Health and Care Excellence; PSS, Personal Social Services; QALYs, quality-adjusted life years.

B.3.2.4. Intervention technology and comparators

The intervention considered in this economic evaluation is GO in combination with DA, the latter being the standard intensive chemotherapy regimen used in routine clinical practice. The cost-effectiveness model incorporates ALFA-0701 study[24,89] dosing for GO and DA (see section B.2, Table 7) aligned with the expected GO marketing authorisation.

It is worthwhile noting that there are some differences between dosing schedules used in UK clinical trials (see Appendix M). Despite this, experts expressed that the fractional dosing schedule used in the ALFA-0701 trial would allow more dosing flexibility in UK clinical practice. Therefore, the model uses the GO dose of 3mg/m^2 (up to one 4.5mg vial) on days 1, 4 and 7, as per expected marketing authorisation, and no scenarios exploring alternative dosing schedules are presented in this submission.

Patients who present with newly diagnosed AML have their cytogenetic profile classified by bone marrow analysis. Clinical experts (Table 50) advised that current

UK clinical practice is to obtain cytogenetic results before starting treatment with GO with time for test results assumed to be a maximum of 7 days in current practice. Experts also suggested that a small proportion of patients initiate treatment with backbone chemotherapy and GO before cytogenetic testing results are known (~ 9–10%) and so this was incorporated in scenario analyses. This corresponds, in the subpopulation analysis, to a proportion of patients with unfavourable cytogenetics profile treated with GO and was modelled appropriately to reflect UK guidelines. Therefore in the model these patients only accrue the relevant costs and QALYs during first induction treatment.

The proportion of patients who received each course of treatment in the ALFA-0701 study was applied in the model to account for treatment discontinuation (see section B.3.5.2).

B.3.3 Clinical parameters and variables

Independent review committee clinical data were obtained from the ALFA-0701 study, described in section B.2. Data from the study that are used in the model are summarised in Table 33.

Analyses were performed using the mITT analysis set. The mITT population consisted of patients who were randomized, but excluded nine patients who withdrew consent prior to the start of treatment. The projected trial enrolment period was 3 years, and the duration of follow-up was planned to be 2 years from the date of last patient enrolled. Patients known to be alive were censored at the last follow-up date or reference date. A reference date of April 30, 2013, was applied to the OS data for the final analyses; all deaths occurring after that date were not included in OS analyses. If there was information confirming that a patient was alive after the reference date, the reference date was used as the censoring date.

The mITT population comprised a total of 271 patients: 135 in the GO + DA arm and 136 in the control arm. The population of the cytogenetic subpopulation corresponded to 108 patients in the GO + DA arm and 106 in the control arm.

Table 33 Application of ALFA-0701 study data in the model for favourable/intermediate cytogenetics profile patients

Data	Application in the model	Value
Response status	Used to determine the proportion of patients who attained CR or CRp following first-line induction therapy and stratify survival outcomes based on induction success or failure.	• GO: [redacted] • DA: [redacted] • Pooled: [redacted]
RFS	Used to fit parametric survival curves to extrapolate long-term RFS estimates.	Sections B.3.3.2 and B.3.3.5
OS	Used to fit parametric survival curves to extrapolate long-term OS estimates.	Sections B.3.3.2 and B.3.3.5
Time to HSCT	Informs how many patients receive an HSCT per cycle. (CR, relapse and refractory)	Section B.3.3.8
RMST	Used to estimate the duration of second-line treatment costs for relapsed and refractory patients.	• Refractory (no HSCT): [redacted] • Relapsed (no HSCT):[redacted]
Post-HSCT cure rate	Used to adjust OS estimates for patients who receive HSCT.	[redacted]
Adverse event incidence	Informs the proportion of patients who experienced a grade 3 or 4 treatment related adverse event (>1% incidence in ALFA-0701 trial) and associated cost in each arm.	Section B.3.3.10
Age	Used to calculate background mortality rates and age-adjusted utilities	Mean baseline age: • Pooled: [redacted]
Gender	Used to calculate background mortality rates and age-adjusted utilities.	Mean % female at baseline: • Pooled: [redacted]
BSA	Used to calculate drug costs based on average dose received per cycle.	Mean baseline BSA: • Pooled: [redacted]
Weight	Used to calculate drug costs based on average dose received per cycle.	Mean baseline weight: • Pooled: [redacted]
Treatment courses	Used to calculate the cost of first-line treatments in both arms.	Section B.3.5.2
Blood products	Used to calculate the cost of blood products in both arms.	Section B.3.5.5

BSA, body surface area; CR, complete remission; CRp, complete remission with incomplete platelet recovery; OS, overall survival; RFS, relapse-free survival; HSCT, stem cell transplant; RMST, restricted mean survival time.

B.3.3.1. Response to First-Line Treatment

Response in ALFA-0701 study was measured in terms of CR or CRp. Patients whose data was insufficient to determine response were assumed to have failed induction therapy (GO: [redacted], DA: [redacted]). The model uses independent review committee-assessed response status based on the mITT population (see section B.2) and presented in Table 34. The base case assumes pooled treatment arms (GO + DA and DA) because although in ALFA-0701 trial GO demonstrated an improvement in quality of CR+CRp, differences between arms were [redacted]. This is aligned with clinical expectations that GO does not affect response status. Response to treatment that varies by arm is tested in a scenario analysis.

It is important to note that the vast majority of patients achieved CR or CRp in the ALFA-0701 study, which is where the advantages of GO in extending RFS are reflected (see sections B.3.3.2 and B.3.3.5).

Table 34 Response Status Data

IRC	GO + DA (N = [redacted])	DA (N = [redacted])	Pooled (N = [redacted])[a]	Source
CR + CRp, n (%)	[redacted]	[redacted]	[redacted]	ALFA-0701 CSR[24]
Induction failure,[b] n (%)	[redacted]	[redacted]	[redacted]	ALFA-0701 CSR[24]

[a] Pooled data were calculated from the individual treatment-arm data reported in the ALFA-0701 study.
[b] Includes patients for whom there was insufficient data to determine response.
CR, complete remission; CRp, complete remission with incomplete platelet recovery; CSR, clinical study report; DA, daunorubicin and cytarabine; GO, gemtuzumab ozogamicin; PSA, probabilistic sensitivity analysis.

B.3.3.2. Survival Analyses

Parametric survival curves were fitted to the patient-level data from the ALFA-0701 study for EFS and the OS endpoint. OS was stratified by response status because survival for patients who attained CR or CRp was expected to be longer than that for refractory patients, as demonstrated by the higher plateaus in the KM data for the former compared to the latter. Moreover, GO is known to extend relapse-free survival. Clinical advisors believed that GO would not affect OS for refractory patients, since those patients failing induction treatment tend to demonstrate poor

prognosis and survival outcomes post treatment failure. Therefore, OS for refractory patients is pooled in the base case.

EFS in the trial was defined as the time from date of randomization to date of induction failure, relapse, or death due to any cause, whichever was first. Therefore, analyses of EFS for patients with induction success included the same outcomes as RFS: relapse and death due to any cause. For the purposes of this report, EFS is hereafter referred to as RFS.

The following survival curves for the modeled health states were fitted:

- RFS of the patients entering CR or CRp, by treatment arm
- OS of the patients entering CR or CRp, by treatment arm
- OS of the refractory patients, pooled treatment arm (base case)
- OS of the refractory patients, by treatment arm (scenario analysis)

The RFS and OS Kaplan-Meier curves in the figures below for the subpopulation of favourable/intermediate cytogenetics profile (Figure 12 and Figure 13) show a plateau from approximately [redacted] to [redacted] (i.e. a functionally cured proportion). It is important to note the large visible distances between arms, particularly for the RFS endpoint (labelled below as EFS) of around [redacted] in the GO arm vs [redacted] for the comparator. The corresponding KM curves for the total population, including patients with adverse risk, demonstrate the same plateau trend and are presented in Appendix M.

Figure 12 RFS and OS (CR or CRp), Kaplan-Meier Curves, Cytogenetic subpopulation

EFS, event-free survival; tx, treatment; OS, overall survival.

Figure 13 OS (Refractory), Kaplan-Meier Curves, Cytogenetic subpopulation

OS, overall survival; tx, treatment.

In order to be aligned with recommendations published by NICE's Decision Support Unit (DSU) in Technical Support Document No. 14,[110] survival analysis methods are used to capture the visible plateau in KM data and the more complex instantaneous risk of events (i.e. hazard function) found in this disease area. This is essential if the model is to capture the benefits of add-on GO reflected in the KM data above: a higher number of patients remaining relapse free for longer with the addition of GO (i.e. a higher RFS curve for GO relative to comparator).

These methods were flexible spline modelling and mixture cure models (MCMs); methods recommended by statistical expert Professor Nicolas Latimer (Table 50). The advantage is that these fitted models do not cut the prominent plateau seen in KM curves as would be seen with standard fitted parametric curves (see appendix M) for survival outcomes in health states that involve a clinically meaningful "functionally cured" rate (i.e. RFS and OS in remission). Indeed, MCM is well established statistical practice for the AML disease area.[111-118]

In summary, MCM functions are used in the base case for RFS and OS in remission health states (section B.3.3.5); with spline models and parametric curves tested in scenario analyses. For OS in refractory the selection was made between standard parametric curves and spline models (section B.3.3.5). Detailed descriptions for the flexible spline and standard parametric functions, and the fitted functions plotted over the Kaplan-Meier curves for RFS and OS in remission are presented in Appendix M.

In all analyses the standard six distributions were explored in alignment with the NICE reference case,[110] as well as the spline modelling framework of Royston & Palmer (Flexsurv statistical package)[119] for fitting the following three distributions:

- Weibull
- Log-normal
- Log-logistic

Up to three knots were tested for all the subsets, using the base case settings from the R package with each of three knots placed at 50%, 33% and 66%, and 25%, 50% and 75% of uncensored survival, respectively. During evaluation of optimal numbers of knots and statistical fit, results using a single knot were considered the most appropriate for use as the addition of a second and third knot did not improve statistical fit.

Standard guidance for fitting and selecting survival functions was followed (NICE DSU 14).[110] Specifically, models were selected based on indicators for proportional hazards: log-cumulative hazards plots, proportional hazards test p-value and the behavior of the Kaplan-Meier curves, e.g., convergence or crossing over of the treatment arms. Measures of statistical fit were compared: Akaike information criterion (AIC) and Bayesian information criterion (BIC) results. Finally, the fit of the curve to KM data and consistency with clinical opinion were assessed.

B.3.3.3. Model selection

The MCM, flexible-spline and standard parametric functions fit to RFS and OS were extrapolated over the model time horizon (40 years) in order to assess long term outcomes. MCMs were fitted using the strsmix package in STATA (StataCorp LLC; College Station, Texas), see Lambert (2007)[120] for an explanation of the MCM approach employed in this submission. STATA was chosen over the statistical

package R (R Foundation for Statistical Computing; Vienna, Austria) - statistical package gfcure[116] - because STATA could take general population mortality into account in the MCM parametrisation.[120] Therefore analyses using R were not included in the model, but can be requested from Pfizer. [98, 99]

MCMs divide the cohort according to disease status. A group of patients are marked as the "cured" and the remaining as the "uncured" fraction. The proportion of cured patients is a parameter that the MCM model intends to fit. The key function for the MCM that was used in the base case is presented in Equation 1 Lambert (2007)[120] and simply expresses that the OS (or RFS) curve is an average of uncured and cured survival weighted by the fitted proportion that is cured.

Equation 1: MCM functional form:

$$\boldsymbol{S_{Overall}(t) = S_{General\ population}(t)\{\pi + (1-\pi)S_{Uncured}(t)\}}$$

T,time; S_Overall (t), survivor function for all-cause mortality; S_general population (t), survivor function for expected mortality (i.e., general population); S_Uncured, the survivor function for the uncured population; π, cured fraction or proportion cured.

MCMs were fitted using the following curves for the uncured fraction:

- Weibull
- Log-normal
- Generalized gamma

The exponential and log-logistic models could not be fitted using the strsmix package.

All MCM curve functions were explored with and without treatment as a covariate, but results in this section present only those functions with individual treatment arms (i.e., treatment not included as a covariate) because combined STATA MCM did not allow for treatment differences in the uncured survival. Therefore results from those analyses would fail to capture separately between treatment arms those surviving in the uncured state which is an essential distinguishing factor according to clinicians

and statistical experts. This approach was also validated with a survival analysis expert (Professor Nick Latimer).

There are several distinctive advantages in using the MCM framework. First, cure models acknowledge that there is a cured subset of patients and allow estimation of the cure fraction, i.e., the cure fraction is treated as an unknown parameter to be estimated via maximum likelihood estimation, in a similar way to the fitting of parameters that determine survival curve fit and shape. This is more statistically valid and less prone to decision bias than estimating a cure proportion based on visual inspection of Kaplan-Meier data (i.e. applying a cure point manually).

Othus et al,[112] argued that the use of MCMs is warranted for this disease area because, commonly, a long-term cure rate is seen for a proportion of patients. Such models can therefore estimate and compare, between treatments, the proportion of long-term survivors (cured) and the proportion of patients who do not survive in the long-term (uncured).[112]

Second, by explicitly modelling the cure fraction as a parameter to be fitted, the probabilistic sensitivity analysis results will reflect the stochastic uncertainty related to this variable - i.e., the probabilistic sensitivity analysis accounts for the smaller sample of patients in the plateau section of the KM curves and the associated random sampling error. Third, MCM models allow stratification of the RFS and OS curves by cured and uncured cohorts; this allows an additional layer of clinical validation to be applied.[113] For this evaluation, the MCM extrapolations as well as predicted cured proportions were validated with clinical experts (Table 50).

There is also a remarkable level of similarity between models fitted with the different statistical packages, despite these packages using slightly different parametrizations of MCM models. Visually, most of the MCMs fitted the Kaplan-Meier data well. Indeed, the MCM model base case provides more conservative cost-effectiveness results with respect to the GO arm than best fitting functions from the other models.

The log-cumulative hazards plots indicated that the proportional hazards assumption was likely to hold for RFS (complete remission) and OS (refractory), but not for OS (complete remission). Therefore, individual treatment-arm functions were selected

for the model base case, but single functions with treatment covariate were explored for OS (complete remission) and are presented in scenario analyses. Appendix M.3. presents log-cumulative hazards plots and graphs of all fitted curves.

B.3.3.4. Cure rate and survival beyond ALFA-0701 study follow-up period

A cure in AML is established for specific patient groups as seen in published literature and validated with UK clinical experts. The fixed time point of 5 years (60 months) was unanimously considered the most clinically valid point by all clinical experts and by a UK AML statistical expert (Table 50). Mortality post the 5 year cure point incorporates age- and sex-specific probabilities estimated from annual mortality rates for the England and Wales general population published in life tables by the Office for National Statistics (2016).[30] These are adjusted for baseline age and gender characteristics of the ALFA-0701 study used in the analysis.

As seen in published evidence and suggested by expert opinion, there may be some excess long term mortality risk associated with cured AML patients. This is because such patients may be more susceptible to cancer occurring or re-occurring or may experience a higher risk of mortality associated with their cancer treatments. HSCT is also known to cause an increased risk of late complications and excess mortality compared to the general population.[122-125] Literature suggests the most common cause of late mortality is relapse, followed by chronic GVHD, infections, organ toxicity, and second cancers. It is also recognized that patients who receive cardiotoxic anthracyclines (such as daunorubicin and idarubicin) have a long-term increased risk of developing cardiovascular disease.[123]

The base-case hazard ratio (subpopulation = [redacted];corresponding to a calculated excess mortality risk of [redacted] compared to that of the general population) was derived using an analysis of pooled survival data from UK AML trials 10 to 16 (subpopulation with favourable and intermediate cytogenetics and all patients) performed by professor Robert Hills (see Appendix M.4) which is ideally suited for calculating an excess mortality HR to better inform the economic evaluation of this submission compared to published literature. First, it is UK trial data for a de novo AML population. Second, the mean study entry age (subpopulation = [redacted]) of this pooled sample is very close to our model entry age (subpopulation = [redacted]). Third, we used

survival curves that were conditional on surviving the first 5 years - i.e. a sample excluding those that died within 5 years was used, which is appropriate because the model applies the HR adjusted general mortality rates to RFS and OS at a fixed time point of 5 years (60 months), when a "functionally cured" proportion was evident.

Estimates found in literature did not satisfy these key requirements.[126-131] It is important to note that AML is a heterogeneous disease and registry based estimates do not capture prognostic differences in the population that can determine mortality. Better outcomes are expected in the population relevant for this evaluation (de novo AML patients fit for intensive chemotherapy) than a population of less fit patients treated with low dose chemotherapy. Indeed, clinical expert opinion highlighted that long term AML survivors have an excess mortality risk of around [redacted], which is more optimistic than Professor Hills' data (see Appendix M.4 for more details).

In the case of MCMs, selected as the base-case modelling approach in RFS and OS remission health states the main statistical package (STATA strsmix) can take account of general population mortality rates during the fitting and parametrization of the MCMs.[120] In other words, this approach requires general population background mortality to be taken into account since this is embedded in the generated STATA parameters. Therefore, the Excel model explicitly incorporates general background mortality in relevant survival functions for valid long term extrapolations. In each model cycle, the probability of death was calculated from the OS curves and from the general population mortality rates. The excess mortality hazard ratio was applied post cure point to capture long term morality risk of those "cured" patients.

B.3.3.5. Base case survival extrapolations

This section presents results for the subpopulation with favorable/intermediate cytogenetics profile. Results for the total population can be found for completeness in Appendix O.

RFS Survival Extrapolations

It is clear throughout the following visual fit sections that the GO arm curves sit above the comparator curves (DA only). This can be seen for OS, but the biggest differences can be seen for RFS and in particularly in the cytogenetic subpopulation

which is consistent with clinical expectations. The general shape of long term extrapolations do not differ greatly, but predict that patients receiving add-on GO are more likely to remain relapse-free at each time point; at some points almost [redacted] as many patients in the GO arm are expected to be in RFS compared to the comparator arm.

Visual fit

In terms of visual fit both the MCM Weibull and MCM log-normal curves fit the data well, but on balance the Log-normal provides the best fit to the plateau (Figure 14). The MCM generalised gamma had the worst visual fit, particularly in the GO arm.

Figure 14 RFS Survival Extrapolations

DA, daunorubicin and cytarabine; GO, gemtuzumab ozogamicin; KM, Kaplan-Meier; MCM, mixture cure models; RFS, relapse free survival.

Statistical fit

According to AIC/BIC values (Table 35), the lognormal function provides the best statistical fit for all analyses.

Table 35 AIC and BIC statistics: RFS

Survival Distribution	GO + DA				DA			
	AIC	Rank	BIC	Rank	AIC	Rank	BIC	Rank
MCM Weibull	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
MCM log-normal	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
MCM generalized gamma	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]

AIC, Akaike information criterion; BIC, Bayesian information criterion; NA, not available; RFS, relapse free survival.

OS (CR or CRp) Survival Extrapolations

Visual fit

Similar to above analyses, the log-normal provides the better visual fit for both arms (Figure 15).

Figure 15 OS (CR or CRp) Survival Extrapolations

CR, complete remission; CRp, complete remission with incomplete platelet recovery; DA, daunorubicin and cytarabine; GO, gemtuzumab ozogamicin; KM, Kaplan-Meier; MCM, mixture cure models; OS, overall survival.

Statistical fit

Based on AIC/BIC values (Table 36), the log-normal functions provide the best statistical fit for the OS remission survival data.

Table 36 AIC and BIC statistics: OS (CR or CRp)

Survival distribution	GO + DA				DA			
	AIC	Rank	BIC	Rank	AIC	Rank	BIC	Rank
MCM Weibull	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
MCM log-normal	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
MCM generalized gamma	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]

AIC, Akaike information criterion; BIC, Bayesian information criterion; NA, not available; RFS, relapse free survival.

OS (refractory) Survival Extrapolations

Visual fit

The base-case pools OS for refractory patients. The standard parametric curves give a wider spread of fitted curves compared to the spline models with 1-knot in all analyses (Figure 16). However, the latter produces fitted curves that consistently plateau late compared to the KM data. The parametric Gompertz tends to provide best overall visual fit in all analyses of OS data for refractory patients.

Figure 16 OS (Refractory) Survival Extrapolations

KM, Kaplan-Meier; OS, overall survival.

Statistical fit

As seen in Table 37 below, AIC/BIC values demonstrate that the gompertz curve provides the best statistical fit across all distributions.

Table 37 AIC and BIC statistics: OS (refractory), pooled treatment arms

	Cytogenetic subpopulation			
Survival distribution	**AIC**	**Rank**	**BIC**	**Rank**
Weibull	[redacted]	[redacted]	[redacted]	[redacted]
Log-normal	[redacted]	[redacted]	[redacted]	[redacted]
Log-logistic	[redacted]	[redacted]	[redacted]	[redacted]
Exponential	[redacted]	[redacted]	[redacted]	[redacted]
Gompertz	[redacted]	[redacted]	[redacted]	[redacted]
Generalized Gamma	[redacted]	[redacted]	[redacted]	[redacted]
Spline Weibull 1-knot	[redacted]	[redacted]	[redacted]	[redacted]
Spline Log-normal 1-knot	[redacted]	[redacted]	[redacted]	[redacted]
Spline Log-logistic 1-knot	[redacted]	[redacted]	[redacted]	[redacted]

AIC, Akaike information criterion; BIC, Bayesian information criterion.

Selection of Preferable Survival Distributions –Base Case

As part of the statistical model selection process recommended by NICE DSU, a group of experts (two clinicians and a statistical expert; Table 50) examined the visual correspondence between model predicted proportions (and extrapolations) and their clinical experience. These meetings occurred on October 9th and 16th 2017 and as well as validation of RFS and OS in CR over time, the corresponding cure proportions predicted by the base case MCM curves were validated.

The base-case functions selected for each treatment arm are presented in Table 38 and based on the visual and statistical fit discussed above.

Table 38 Base-case survival functions

Endpoint	GO + DA	DA	Pooled
RFS (CR or CRp)	MCM log-normal	MCM log-normal	—
OS (CR or CRp)	MCM log-normal	MCM log-normal	—
OS (refractory)	—	—	Gompertz

CR, complete remission; CRp, complete remission with incomplete platelet recovery; GO, gemtuzumab ozogamicin; OS, overall survival; RFS, relapse-free survival.

B.3.3.6. Survival adjustments for treatment-switching

A number of patients in the chemotherapy arm of the ALFA-0701 study ([redacted] or [redacted]) received GO as follow-up therapy through a compassionate use programme. A feasibility assessment was conducted in order to evaluate whether it would be possible to adjust for this treatment switching in the model but it was not deemed possible (see Appendix M.5 for details). This is because the assumptions underlying the different adjustment methods were not considered to hold and therefore the results would not provide reliable and robust estimates of OS. It should be noted that these switchers could bias relative efficacy in favour of the comparator arm.

B.3.3.7. Subsequent Lines of Active Treatment and Non-curative Therapies

Relapsed or refractory patients can receive high-intensity salvage therapy with curative intent if they are deemed fit enough by their clinician. The aim is for patients to attain CR or CRp, at which point they are eligible for a transplant. Patients who are not deemed fit enough for high-intensity salvage therapy (for example, due to complications and comorbidities) receive non-curative therapies and palliative care.

The base case value for the proportion of patients with relapsed or refractory disease who received salvage therapy was assumed to be [redacted] based on clinical estimates.

B.3.3.8. Hematopoietic Stem-Cell Transplantation

Patients can receive HSCT from the CR or CRp, relapse, and refractory health states. HSCT probabilities were calculated from the number of patients in the ALFA-0701 study who underwent HSCT from time-to-HSCT analyses for the following cohorts: those who attained CR or CRp (without relapse), who relapsed (after attaining CR or CRp), and who were refractory. The time-to-HSCT data are presented in Appendix M. The mean time to HSCT was calculated for each cohort and was used to dictate the model cycle in which HSCTs occur.

HSCT probabilities were calculated for the CR or CRp and refractory health states from pooled data in the ALFA-0701 study for the base case analysis because clinicians advised that the probability of HSCT would not be affected by GO. This is a conservative assumption for refractory patients because there were more HSCTs in the control arm in the ALFA-0701 study. HSCTs for CR or CRp (without relapse) patients were the same between arms in the ALFA-0701 study and therefore pooling is clinically accurate. HSCTs from the relapse health-state occurred over 4 years because patients relapsed at different times. Although GO was not expected to impact the probability of HSCT for patients who relapsed, GO prevents relapses; therefore, the total number of HSCTs for relapse patients was expected to be lower for GO. Annual probabilities for each treatment arm were calculated based the time-to HSCT analyses (Appendix M). As a validation, the total number of HSCTs in the model matches the ALFA-0701 study data. The HSCT probabilities are presented in Table 39.

Relapsed patients were assumed to receive a HSCT at the midpoint of each model year, from year 2 onward. The first HSCT for relapsed patients occurred [redacted] from randomization; patients were assumed to receive a transplant at the midpoint of [redacted]. The annual mean times to HSCT in the ALFA-0701 study were not used because there were not enough transplants in some years to calculate accurate mean estimates. Table 39 show the probabilities for the subpopulation of favourable/intermediate cytogenetics that will inform the base case

analysis. Annual HSCT probabilities for the total population are shown in Appendix O.

Table 39. Annual Probability of Hematopoietic Stem-Cell Transplantation

Probability of HSCT	Subpopulation			Timing (months)[b]	Source
	GO + DA	DA	Pooled[a]		
From CR or CRp (without relapse)[c]					
Total, %	NA	NA	[redacted]	[redacted]	Calculation
From Refractory[d]					
Total, %	NA	NA	[redacted]	[redacted]	Calculation
From Relapse[c]					
Year 1, %	[redacted]	[redacted]	[redacted]	[redacted]	Calculation
Year 2, %	[redacted]	[redacted]	[redacted]	[redacted]	Calculation
Year 3, %	[redacted]	[redacted]	[redacted]	[redacted]	Calculation
Year 4, %	[redacted]	[redacted]	[redacted]	[redacted]	Calculation
Year 5, %	[redacted]	[redacted]	[redacted]	[redacted]	Calculation

DA, daunorubicin and cytarabine; GO, gemtuzumab ozogamicin; HSCT, hematopoietic stem-cell transplant; PSA, probabilistic sensitivity analysis.

[a] Pooled data were calculated from the individual treatment-arm data reported in the ALFA-0701 study.

[b] Time at which patients transition to HSCT in the model.

[c] HSCT probabilities calculated from the total number of patients who enter CR or CRp.

[d] HSCT probabilities calculated from the total number of patients who enter refractory.

B.3.3.9. Post-HSCT Survival Adjustment

Analyses of post-HSCT OS in the ALFA-0701 study were performed from the time of HSCT. Clinical experts (see Table 50) expected the OS after HSCT to be similar for all patients, regardless of whether the HSCT followed first-line or second-line CR or CRp and the chemotherapies received, or whether the patients were refractory to induction therapy or relapsed. Therefore, data were pooled for all HSCT patients and analysed from the time of HSCT (Figure 17). The curve flattens after approximately 2 years as it reaches the cured proportion. This is corroborated by previous research in which the majority of HSCT deaths occur within the first 2 years after the HSCT.[122]

Figure 17 OS, Kaplan-Meier Curve, All HSCT Patients from the Time of HSCT

HSCT, hematopoietic stem-cell transplant; OS, overall survival.

Post-HSCT mortality in the model is governed by the same underlying OS curves for patients who do not receive HSCT. Therefore, the model would under-predict mortality for relapsed and refractory patients who do not receive an HSCT and over-predict mortality for those who do receive an HSCT. [redacted] (see Appendix M).

HRQoL is higher for patients after HSCT than for relapsed and refractory patients receiving non-curative therapies, as validated by clinicians (see section B.3.4.4; Table 41). Therefore, the QALYs predicted by the model may be biased toward GO. Adjustments are included in the model to increase the predicted survival for patients after HSCT and reduce the predicted survival for patients who do not receive HSCT, proportionally, such that the total number of deaths remains the same.

The adjustment is performed based on a cured proportion of patients calculated from the ALFA-0701 study. Post-HSCT OS was [redacted] at the end of study follow-up (Figure 17) which is incorporated in the model as proportion of patients being cured after transplantation. When the survival of the HSCT patients has dropped to [redacted], as predicted by the OS curves, adjusted general population mortality rates are

applied to these patients (see section B.3.3.4). The additional deaths that would have occurred for the HSCT patients, had the cure rate not been applied, are instead taken from the patients who did not receive HSCT. This approach mitigates bias toward GO by preventing the over-prediction of deaths and under-prediction of QALYs for HSCT patients using ALFA-0701 trial data.

B.3.3.10. Adverse Events

The model incorporates the costs of managing treatment-related AEs and the disutility associated with these events. All grade 3 and 4 TEAEs that occurred in at least 1% of patients in the ALFA-0701 study were included in the model. Clinical experts (see Table 50) did not identify any important AEs missing from the model. Rates in Table 40 below are for the subpopulation of favourable/intermediate cytogenetics profile for rates of the total population, see Appendix F.

Table 40 Incidence of adverse events for first-line AML therapies

Adverse Event: n (%)	GO + DA (N = 105)[a]	DA (N = 106)[a]	Source
Skin toxicity	[redacted]	[redacted]	ALFA-0701 CSR[24]
Mucosal toxicity	[redacted]	[redacted]	ALFA-0701 CSR[24]
Pain	[redacted]	[redacted]	ALFA-0701 CSR[24]
Nausea, vomiting, and diarrhea	[redacted]	[redacted]	ALFA-0701 CSR[24]
Pulmonary toxicity	[redacted]	[redacted]	ALFA-0701 CSR[24]
Cardiac rhythm disorder	[redacted]	[redacted]	ALFA-0701 CSR[24]
Other cardiac toxicity	[redacted]	[redacted]	ALFA-0701 CSR[24]
Central neurological toxicity	[redacted]	[redacted]	ALFA-0701 CSR[24]
Peripheral neurological toxicity	[redacted]	[redacted]	ALFA-0701 CSR[24]
Infections	[redacted]	[redacted]	ALFA-0701 CSR[24]
Hemorrhage	[redacted]	[redacted]	ALFA-0701 CSR[24]
Veno-occlusive disease	[redacted]	[redacted]	ALFA-0701 CSR[24]

CSR, clinical study report; DA, daunorubicin and cytarabine; GO, gemtuzumab ozogamicin.
[a] As-treated population.

GVHD is a clinically important complication of HSCT and therefore was captured in the model. The incidence of grade 3 or 4 acute GVHD for HSCT patients was estimated to be 15%,[132] with a mean duration of 2.5 months according to UK clinical opinion. The incidence of grade 3 or 4 chronic GVHD for HSCT patients was estimated to be 5.5%[133] based on the figures for 'extensive' chronic GVHD, which correspond to severe chronic GVHD. A mean duration of 9 months was used as confirmed with UK clinical experts. Acute GVHD occurs in the first 100 days post-HSCT, whereas chronic GVHD occurs after 100 days post-HSCT. Acute GVHD was assumed to occur in the first model cycle after leaving the HSCT procedure health state. Chronic GVHD was assumed to start 6 months after HSCT based on clinical feedback.

B.3.4 Measurement and validation of health effects

B.3.4.1. Mapping

No mapping exercises were performed for this submission.

B.3.4.2. Health-related quality-of-life studies

There were no studies identified by the HRQoL literature search that satisfied all the criteria for the NICE reference case. The literature review identified six publications that reported utility data and met the inclusion criteria, see Appendix H.[41,105,134-137] None of these studies reported utilities valued from the UK general population perspective. In addition, one further study was identified which reported QLQ-C30 data measured in a UK AML patient population but such data do not align with the NICE reference case.[138]

Two publications, Leunis A et al., 2014 and Cressman S et al., 2016,[41,137] report utility values that are the closest match to the NICE reference case. Both were evaluated for incorporation in the model as appropriate but only the study by Leunis A et al., 2014 was included as this focused on de novo AML cases using EQ-5D data (see appendix H).

Due to the known paucity of relevant HRQoL data, Pfizer commissioned a preference elicitation study to assign preference values to health states experienced by AML patients, as described in vignettes, from the perspective of the general UK

population (see Appendix H).[139] Time trade-off and visual analogue scale methodologies were applied in one-to-one, face-to-face interviews, following the guidance recommended by NICE.[110] A summary of the best available health-state utility values is shown in Appendix H.

In lack of published EQ-5D data to comply with the NICE reference case, the only relevant source of evidence to closely match our disease area and modelled heath states was NICE TA399.[66] Utility values were validated with a clinical expert in terms of applicability and relevance to the population of this submission. For more detail please read section B.3.4.4.

B.3.4.3. Adverse reactions

Due to the lack of published/clinical evidence the best source for adverse event disutilities was considered to be the published NICE TA399.[66] Therefore a mean utility decrement of 0.024 for all grade 3 and 4 AEs except for VOD is applied in the model. This decrement is applied in the model as a one-off for each adverse-event occurrence for each treatment group, taking account of the duration of events.

Disutility of VOD is applied as 0.208 this is referenced in defibrotide SMC submission[140] with the assumption that it is approximately the same as that associated with acute liver failure pre-transplant and is applied for a mean duration of 26.8 days.[141] This was considered a reasonable assumption by clinical experts.

B.3.4.4. Health-related quality-of-life data used in the cost-effectiveness analysis

Clinical experts (Table 50) believed that the health-state utility estimates from Pfizer's preference elicitation study were a reasonable representation of the HRQoL experienced by AML patients. However, EQ-5D health-state utility estimates from NICE TA399[66] were chosen for the base-case analysis to align with the NICE reference case. Utility values from the preference elicitation study were tested in scenario analyses as supporting evidence.

Clinical opinion was used to validate the assumptions required when using the health state utilities from NICE TA399:

- Utilities for AML patients who are not eligible to receive high-intensity chemotherapy were assumed to be the same for patients who are eligible for high-intensity chemotherapy.
- For the remission (off treatment) health state, it is assumed that the patient's quality of life resembles that of "remission" health state for the AML patient population in NICE TA399.
- For the relapse and refractory health states, it is assumed patients have the same quality of life as those in "post progression/relapse" health state of the NICE TA399 population.

The clinical expert suggested these are reasonable assumptions and agreed that the second is conservative since patients in the remission state (off treatment) will have a better quality of life compared to the NICE TA399 population that are in remission. The assumption that patients under azacitidine in remission remain on treatment and therefore suffer treatment side-effects of azacitidine and hospitalisations (admissions) was validated. Of the two mapping algorithms considered in NICE TA399,[66] results from the McKenzie and Van der pol[142] were selected as these were closer to values generated from the Pfizer's preference elicitation study (see Appendix H).[140]

For the functionally cured health state, age-adjusted EQ-5D values were used for the UK general population and calculated from the formula reported by Ara & Brazier.[143] A fixed age-adjusted value based on the mean baseline age was tested in a scenario analysis. Estimates selected for the model health states are presented in Table 41.

Table 41 Summary of utility values for the cost-effectiveness analysis

State	Utility Value: Mean (SE)	95% CI	Reference in Submission	Justification
Health state				
Chemotherapy treatment[a]	0.6574[c] (NA)	NA	B.3.4.4, B.3.4.2	Assumed utility for non-remission (partial remission or stable disease) from NICE TA399 represents HRQoL in this health state.[66]
Consolidation therapy	0.6574[b] (NA)	NA	B.3.4.4, B.3.4.2	Assumed utility for chemotherapy treatment represents HRQoL in this

State	Utility Value: Mean (SE)	95% CI	Reference in Submission	Justification
				health state.
HSCT procedure	0.6574[b] (NA)	NA	B.3.4.4, B.3.4.2	Assumed utility for chemotherapy treatment represents HRQoL in this health state.
GVHD (post HSCT)	0.67[e] (NA)	0.63-0.72	B.3.4.4, B.3.4.2	Used utility for AML patients with post-HSCT GVHD from Kurosawa (2014)[135]
CR or CRp	0.7400[c] (NA)	NA	B.3.4.4, B.3.4.2	Assumed utility for remission (CR or CRi) from NICE TA399 represents HRQoL in this health state.
Relapse	0.5680[c] (NA)	NA	B.3.4.4, B.3.4.2	Assumed utility for post-progression/relapse from NICE TA399 represents HRQoL in this health state.[66]
Refractory	0.5680[f] (NA)	NA	B.3.4.4, B.3.4.2	Assumed utility for relapse represents HRQoL in this health state.
Functionally cured	0.820[g] (NA)	NA	B.3.4.4, B.3.4.2	
Adverse reaction				
Skin toxicity	0.0207 (NA)	NA	B.3.4.3	Used mean utility decrement for grade 3 or 4 adverse events from NICE TA399.[66]
Mucosal toxicity	0.0207 (NA)	NA	B.3.4.3	
Pain	0.0207 (NA)	NA	B.3.4.3	
Nausea, vomiting, and diarrhea	0.0207 (NA)	NA	B.3.4.3	
Pulmonary toxicity	0.0207 (NA)	NA	B.3.4.3	
Cardiac rhythm disorder	0.0207 (NA)	NA	B.3.4.3	
Other cardiac toxicity	0.0207 (NA)	NA	B.3.4.3	
Central neurological toxicity	0.0207 (NA)	NA	B.3.4.3	
Peripheral neurological toxicity	0.0207 (NA)	NA	B.3.4.3	
Infections	0.0207 (NA)	NA	B.3.4.3	
Hemorrhage	0.0207 (NA)	NA	B.3.4.3	
Veno-occlusive disease	0.208 (NA)	NA	B.3.4.3	Used utility decrement for VOD Defibrotide SMC.[140]

AR, adverse reaction; CI, confidence interval; HS, health state; SE, standard error. CR, complete remission; CRp, complete remission with incomplete platelet recovery; GVHD, graft versus-host

disease; HSCT, hematopoietic stem-cell transplant; NICE, National Institute for Health and Care Excellence; SMC, Scottish Medicines Consortium; TTO, time trade-off; VAS, visual analog scale.

[a] Includes patients who are receiving induction or salvage chemotherapy.

[b] Assumed equal to chemotherapy treatment.

[c] Values from NICE TA 399,[66] using the mapping algorithm by McKenzie & Van der Pol, 2009.

[d] Assumed equal to CR or CRp.

[e] Value from Kurosawa et al., 2016.[136]

[f] Assumed equal to relapse.

[g] Age-adjusted EQ-5D value for UK general population calculated from the formula reported by Ara & Brazier, using mean patient age and gender. Value presented was calculated using baseline patient characteristics for all patients in the ALFA-0701 study.

B.3.5 Cost and Healthcare resource use identification measurement and validation

B.3.5.1. Resource identification, measurement and validation studies

A SLR was conducted to identify primary studies reporting costs and health care resource use. Electronic databases, annual proceedings of scientific meetings, and health technology assessments were searched (see Appendix H). Only 2 primary studies were identified related to the UK setting.[144,145] Both were abstracts and provided information on average lifetime costs for AML patients. Therefore, in addition to the SLR, the NICE website was searched to identify any relevant, recently published HTA submissions of de novo AML first-line treatments. Results of this structured search identified NICE TA399: Azacitidine for treating acute myeloid leukaemia with more than 30% bone marrow,[66] as the only relevant published submission. This population differs from the one of our trial as it also includes secondary AML cases, ineligible for transplantation with presence of at least 30% bone marrow blasts.

B.3.5.2. Intervention and comparator's costs and resource use

In the ALFA-0701 study, interventions could be administered for a maximum of two induction cycles and two consolidation cycles. GO was administered at the first induction cycle and at both consolidation cycles, with a maximum dose of 5 mg. GO was not administered at the second induction cycle; the same DA treatments were administered in both arms. All patients had discontinued their randomized treatment at the time of data cut-off. Treatment costs in the model account for delayed or missed chemotherapy cycles.

Drug costs for first-line treatments are calculated from the treatment regimens specified for each treatment course in the ALFA-0701 study. The mean drug cost per administered dose is calculated in two ways:

- Assuming no wastage and based on the minimum price per milligram (base case setting)
- Assuming wastage, the cost is calculated using the proportion of patients requiring each vial combination and the cost of each vial combination

Each dose of GO is 3 mg/m^2 with a dose cap of 5 mg (1 vial) (see Section B.3.2.4). Therefore, there is no wastage of GO unless a patient's BSA is below 1.66 m^2 (mean BSA in the ALFA-0701 study = [redacted]). Furthermore, clinical experts advised that little drug wastage is seen in clinical practice and so the base case analysis assumes no drug wastage. A scenario analysis is explored that assumes drug wastage. The drug acquisition costs used in cost calculations are presented in Table 42.

First-line treatment costs are calculated using treatment consumption data for the as-treated population from the ALFA-0701 study. This data are used in the model to calculate the proportion of patients who receive each treatment course for first and second induction, and first and second consolidation. Differences between arms using trial data are small and aligned with clinical expectations that proportion of patients receiving each treatment course in UK practice are expected to be similar between arms. Therefore pooled proportion data are used for the base-case analysis (see section B.3.7)

First-line treatments are administered in an inpatient setting. Drug-administration costs for first-line treatments are captured within the elective inpatient cost for AML patients.

Table 42 Drug acquisition costs: first-line therapies

Drug	Pack price	Source
GO (5-mg vial)	[redacted]	Pfizer
Daunorubicin (20-mg vial)	£65.00	BNF (2017)[146]
Cytarabine (2000 mg/5 mL solution, 5 vials)	£6.60	DoH (eMiT) (2017)[147]

BNF, British National Formulary; DoH, Department of Health; eMiT, electronic market information tool.

B.3.5.3. Subsequent Lines of Active Treatment and Noncurative Therapies

Relapsed and refractory patients who are deemed fit enough may receive high-intensity salvage therapy (see section B.3.3.7). The standard salvage therapy used in the UK is a combination of FLAG-Ida. All patients who receive salvage therapy in the model are assumed to receive FLAG-Ida, as validated by UK clinical experts (see Table 50). Patients either receive one or two cycles of FLAG-Ida, at their clinician's discretion. In the absence of available data, an assumption was made that patients receive a mean of 1.5 cycles. Clinical opinion validated this assumption as accurate. Salvage therapy is administered in an inpatient setting; drug-administration costs for salvage therapy are captured within the elective inpatient cost for AML patients.

Relapsed and refractory patients who are not deemed fit enough to receive salvage therapy instead receive non-curative therapies (including best supportive care) and palliative care. Clinical experts (see Table 50) advised that the three most commonly used therapies are hydroxycarbamide, low-dose cytarabine, and azacitidine; the experts estimated that these therapies are used in a 40:40:20 ratio, respectively. The model uses standard UK treatment protocols for these therapies. Patients who received salvage therapy and did not go on to receive an HSCT move to best supportive care only (hydroxycarbamide). Patients who received salvage therapy and did not go on to receive an HSCT move to best supportive care only (hydroxycarbamide). Non-curative therapies are continued until futility and are assumed to continue up to the start of the terminal care period. The model applies a terminal care cost of £6,659 (inflated from Addicott & Dewar)[148] for the last 8 weeks of a patient's life for those patients receiving non-curative therapy (including best supportive care). To calculate the duration of non-curative therapies prior to the terminal care cost (applied for 2 cycles), the model uses RMST estimates from the ALFA-0701 study for newly relapsed and refractory patients (see Appendix M). The base case used pooled RMST estimates of [redacted] for newly relapsed patients and [redacted] for refractory patients.

The drug acquisition costs used in cost calculations are presented in Table 43.

Table 43 Drug acquisition costs: second-line therapies

Drug	Pack price	Source
Salvage therapy		
Fludarabine (50 mg/2 mL concentrate, 1 vial)	£26.08	DoH (eMiT) (2017)[147]
Cytarabine (2000 mg/5 mL solution, 5 vials)	£6.60	DoH (eMiT) (2017)[147]
G-CSF (filgrastim) (30 million units/0.5 mL solution, 5 vials)	£49.30	BNF (2017)[146]
Idarubicin (5-mg powder for solution, 1 vial)	£87.36	BNF (2017)[146]
Non-curative therapies		
Low-dose cytarabine (100-mg vial)	£4.70	DoH (eMiT) (2017)[147]
Hydroxycarbamide (100 capsules)	£8.83	DoH (eMiT) (2017)[147]
Azacitidine (100-mg powder for suspension)	£321.00	BNF (2017)[146]

BNF, British National Formulary; DoH, Department of Health; eMiT, electronic market information tool.

B.3.5.4. Costs associated with HSCT

Patients can receive HSCT from the CR or CRp, relapse, or refractory health states (see section B.3.3.7). In the ALFA-0701 study, almost all transplants were allogenic and this is reflected in the model. Clinical experts (see Table 50) confirmed that these data were consistent with UK clinical practice. The cost considered for each HSCT was obtained from a NHS Blood and Transplant analysis (2014)[149] and inflated to 2016 costs. The cost was broken down into the cost of the transplantation procedure (which includes the post-transplantation recovery period) and the costs associated with the procedure (including complications and follow-ups) for time periods thereafter. The model applies a one-off cost for the transplantation procedure and recovery period whenever patients enter the HSCT health state (in which health state they remain for 1 cycle). The timing of the HSCT is tracked and monthly costs are applied over the 2-year period thereafter. The inflated costs used in the model for HSCT are as follows:

- HSCT procedure and recovery period: £60,892 (one-off)
- 0 to 6 months after the HSCT: £4,891 (per month)
- 6 to 12 months after the HSCT: £3,360 (per month)
- 12 to 24 months after the HSCT: £1,212 (per month)

No additional costs related to HSCT are applied after the 2-year period after an HSCT.

The model also takes into account transplant-related acute and chronic GvHD complications and because of the lack of UK data assigns an inflated cost of £26,889 from a French publication.[150]

B.3.5.5. Health state unit costs and resource use

Costs that are associated with managing patients with AML and that are not specifically related to systemic therapy are shown by health state in Table 44. This includes inpatient and outpatient attendances, disease-monitoring tests, and supportive therapies (antifungals and antibiotics). Clinical experts suggested that these resources should not differ between treatment arms (Table 50). For supportive therapies, the expert clinicians advised that posaconazole and gentamicin are the standard antifungals and antibiotics, respectively, used in UK clinical practice. Posaconazole and gentamicin treatment regimens were used for all supportive therapy cost calculations.

The cost of antifungals and antibiotics used to treat infections is already captured by the unit cost for managing infections (see section B.3.5.6) and the incidence of infections (see section B.3.3.10). Therefore, only prophylactic use of antifungals and antibiotics are added to the cost calculations. The mean duration of prophylaxis was estimated to be 21 days during each treatment phase by UK clinical experts. Drug administration costs for first-line and salvage therapies that are administered during inpatient stays are captured within the inpatient attendance costs.

In the case of non-curative therapies, clinical advice was required in order to best capture how these costs are applied in UK clinical practice. Expert clinicians validated that treatments for those patients are not administered during inpatient attendances. As per clinical advice, the cost for subsequent infusions of a chemotherapy cycle in an outpatient setting was applied for administering azacitidine which is commonly given on a day unit setting. No administration cost was applied for hydroxycarbamide because it is given as an outpatient in capsule form to take home. No administration cost was applied for low-dose cytarabine because it is given in the community by a relative, home nurse or by the patient themselves.

The expert clinicians (Table 50) advised that resource use, and therefore costs, were not expected to differ for either salvage therapy or non-curative therapy given to relapsed and refractory patients. Patients who are functionally cured were not expected to incur costs. Costs for patients in CR or CRp following HSCT are captured up to 2 years following the procedure; no additional costs are applied beyond this period (see section B.3.5.4).

The monthly resource use and costs that are applied in the model in each health state are presented in Table 44.

Resource-use data for blood products were taken from the ALFA-0701 study for the individual treatment arms (as-treated population). However, no resource use data for patients receiving non-curative therapy were available; estimates for azacitidine during progressive disease from NICE TA399 are implemented in the model.[66] Unit costs for red blood cells and platelets were obtained from NHS reference costs.[151] Costs for each treatment arm were calculated by multiplying the resource-use estimates with the unit costs. The resource-use inputs and resulting costs used in the model are presented in Table 45.

Table 44 Monthly resource use and costs, by health state

Resource Item	Frequency (per cycle)[c]	Measurement of Uncertainty (Distribution)[d]	Unit Costs	Source
Induction therapy: first induction				
Inpatient attendances			Total: £18,528	
Elective inpatient (days)	28	SE = 2.8 (normal)	£661.72	DoH (2016);[151] SA25G-M[e]
Outpatient attendances			Total: £233	
Consultant haematologist, first	1	SE = 0.1 (normal)	£196.64	DoH (2016);[151] WF01B
Specialist nurse (20 minutes)	1	SE = 0.1 (normal)	£36.00	PSSRU (2016);[152] band 6
Disease monitoring			Total: £1,225	
Bone marrow cytogenetics profile	1	SE = 5.3 (normal)	£16.88	DoH (2016);[151] DAPS01
Bone marrow aspirate & extraction	2	SE = 0.2 (normal)	£493.90 + £1.18	DoH (2016);[151] SA33Z & DAPS04
Full blood count & phlebotomy	28	SE = 2.8 (normal)	£3.37 + £3.10	DoH (2016);[151] DAPS08 & DAPS05
Biochemistry profile & phlebotomy	8	SE = 0.8 (normal)	£3.37 + £1.18	DoH (2016);[151] DAPS08 & DAPS04
Supportive therapies			Total: £1,971	
Antifungals (days)	21	SE = 2.8 (normal)	£491.20	BNF (2017)[146]
Antibiotics (days)	21	SE = 2.8 (normal)	£1.62-£4.61	BNF (2017)[146]
Induction therapy: second induction				
Inpatient attendances			Total: £12,573	
Elective inpatient (days)	19	SE = 1.9 (normal)	£661.72	DoH (2016);[151] SA25G-M[e]
Outpatient attendances			Total: £199	

Resource Item	Frequency (per cycle)[c]	Measurement of Uncertainty (Distribution)[d]	Unit Costs	Source
Consultant hematologist, follow-up	1	SE = 0.1 (normal)	£162.84	DoH (2016);[151] WF01A
Specialist nurse (20 minutes)	1	SE = 0.1 (normal)	£36.00	PSSRU (2016); band 6
Disease monitoring			Total: £1,208	
Bone marrow aspirate & extraction	2	SE = 0.2 (normal)	£493.90 + £1.18	DoH (2016);[151] SA33Z & DAPS04
Full blood count & phlebotomy	28	SE = 2.8 (normal)	£3.37 + £3.10	DoH (2016);[151] DAPS08 & DAPS05
Biochemistry profile & phlebotomy	8	SE = 0.8 (normal)	£3.37 + £1.18	DoH (2016);[151] DAPS08 & DAPS04
Supportive therapies			Total: £1,971	
Antifungals (days)	21	SE = 2.8 (normal)	£491.20	BNF (2017)[146]
Antibiotics (days)	21	SE = 2.8 (normal)	£1.62-£4.61	BNF (2017)[146]
CR or CRp: first consolidation				
Inpatient attendances			Total: £3,573	
Elective inpatient (days)	5.4	SE = 0.54 (normal)	£661.72	DoH (2016);[151] SA25G-M[e]
Outpatient attendances			Total: £199	
Consultant hematologist, follow-up	1	SE = 0.1 (normal)	£162.84	DoH (2016);[151] WF01A
Specialist nurse (20 minutes)	1	SE = 0.1 (normal)	£36.00	PSSRU (2016); band 6
Disease monitoring			Total: £495	
Bone marrow aspirate & extraction	1	SE = 0.1 (normal)	£493.90 + £1.18	DoH (2016);[151] SA33Z & DAPS04
Biochemistry profile & phlebotomy	1	SE = 0.1 (normal)	£3.37 + £1.18	DoH (2016);[151] DAPS08 & DAPS04

Resource Item	Frequency (per cycle)[c]	Measurement of Uncertainty (Distribution)[d]	Unit Costs	Source
Supportive therapies			Total: £1,971	
Antifungals (days)	21	SE = 2.8 (normal)	£491.20	BNF (2017)[146]
Antibiotics (days)	21	SE = 2.8 (normal)	£1.62-£4.61	BNF (2017)[146]
CR or CRp: second consolidation				
Inpatient attendances			Total: £3,573	
Elective inpatient (days)	5.4	SE = 0.54 (normal)	£661.72	DoH (2016); SA25G-M[e]
Outpatient attendances			Total: £199	
Consultant haematologist, follow-up	1	SE = 0.1 (normal)	£162.84	DoH (2016);[151] WF01A
Specialist nurse (20 minutes)	1	SE = 0.1 (normal)	£36.00	PSSRU (2016); band 6
Disease monitoring			Total: £500	
Bone marrow aspirate & extraction	1	SE = 0.1 (normal)	£493.90 + £1.18	DoH (2016);[151] SA33Z & DAPS04
Biochemistry profile & phlebotomy	1	SE = 0.1 (normal)	£3.37 + £1.18	DoH (2016);[151] DAPS08 & DAPS04
Supportive therapies			Total: £1,971	
Antifungals (days)	21	SE = 2.8 (normal)	£491.20	BNF (2017)[146]
Antibiotics (days)	21	SE = 2.8 (normal)	£1.62-£4.61	BNF (2017)[146]
CR or CRp: off treatment				
Outpatient attendances			Total: £99	
Consultant haematologist, follow-up	0.5	SE = 0.05 (normal)	£162.84	DoH (2016);[151] WF01A
Specialist nurse (20 minutes)	0.5	SE = 0.05 (normal)	£36.00	PSSRU (2016);[152] band 6

Resource Item	Frequency (per cycle)[c]	Measurement of Uncertainty (Distribution)[d]	Unit Costs	Source
Relapse or refractory: salvage therapy				
Inpatient attendances			Total: £16,543	
Elective inpatient (days)	25	SE = 2.5 (normal)	£661.72	DoH (2016);[151] SA25G-M[e]
Outpatient attendances			Total: £199	
Consultant haematologist, follow-up	1	SE = 0.1 (normal)	£162.84	DoH (2016);[151] WF01A
Specialist nurse (20 minutes)	1	SE = 0.1 (normal)	£36.00	PSSRU (2016); band 6
Disease monitoring			Total: £1,185	
Bone marrow aspirate & extraction	2	SE = 0.2 (normal)	£493.90 + £1.18	DoH (2016);[151] SA33Z & DAPS04
Full blood count & phlebotomy	28	SE = 2.8 (normal)	£3.37 + £3.10	DoH (2016);[151] DAPS08 & DAPS05
Biochemistry profile & phlebotomy	3	SE = 0.3 (normal)	£3.37 + £1.18	DoH (2016);[151] DAPS08 & DAPS04
Supportive therapies			Total: £1,971	
Antifungals (days)	21	SE = 2.8 (normal)	£491.20	BNF (2017)[146]
Antibiotics (days)	21	SE = 2.8 (normal)	£1.62-£4.61	BNF (2017)[146]
Relapse or refractory: non-curative therapy				
Drug administration[a]			Total: £1,484	
Subsequent infusion (outpatient)	7	Fixed	£211.99	DoH (2016);[151] SB15Z
Inpatient attendances			Total: £9,264	
Elective inpatient (days)	14[b]	SE = 1.0 (normal)	£661.72	DoH (2016);[151] SA25G-M[e]
Outpatient attendances			Total: £597	

Resource Item	Frequency (per cycle)[c]	Measurement of Uncertainty (Distribution)[d]	Unit Costs	Source
Consultant hematologist, follow-up	3	SE = 0.3 (normal)	£162.84	DoH (2016);[151] WF01A
Specialist nurse (20 minutes)	3	SE = 0.3 (normal)	£36.00	PSSRU (2016);[152] band 6
Disease monitoring			Total: £1,051	
Bone marrow aspirate & extraction	2	SE = 0.2 (normal)	£493.90 + £1.18	DoH (2016);[151] SA33Z & DAPS04
Full blood count & phlebotomy	8	SE = 0.8 (normal)	£3.37 + £3.10	DoH (2016);[151] DAPS08 & DAPS05
Biochemistry profile & phlebotomy	2	SE = 0.2 (normal)	£3.37 + £1.18	DoH (2016);[151] DAPS08 & DAPS04
Supportive therapies			Total: £1,971	
Antifungals (days)	21	SE = 2.8 (normal)	£491.20	BNF (2017)[146]
Antibiotics (days)	21	SE = 2.8 (normal)	£1.62-£4.61	BNF (2017)[146]

BNF, British National Formulary; CR, complete remission; CRp, complete remission with incomplete platelet recovery; DoH, Department of Health; PSSRU, Personal Social Services Research; SE, standard error

[a] Administration cost applied for azacitidine only.

[b] This estimate is based on clinical opinion that for these patients it is reasonable to expect two significant week-long inpatient admissions due to severe infection for the remainder of the patients' lifespan.

[c] Based on clinical opinion.

[d] Uncertainty information was not available; SE was estimated as 10% of the mean value.

[e] Weighted-average cost calculated.

Table 45 Mean Blood Products Usage and Costs per Patient

Treatment Cycle: Mean (SD)	Red Blood Cells		Platelets		Total Cost[b] (GO + DA)	Total Cost[b] (DA)	PSA Distribution	Source
	GO + DA (N = 131)	DA (N = 137)	GO + DA (N = 131)	DA (N = 137)				
First induction	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	Normal	ALFA-0701 CSR[24]
Second induction	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	Normal	ALFA-0701 CSR[24]
First consolidation	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	Normal	ALFA-0701 CSR[24]
Second consolidation	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	Normal	ALFA-0701 CSR[24]
Salvage therapy	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	Normal	ALFA-0701 CSR[24]
Noncurative therapy[a]	[redacted]	[redacted]	[redacted]	[redacted]	£1,647	£1,647	Normal	NICE TA399[66]

DA, daunorubicin and cytarabine; GO, gemtuzumab ozogamicin; NA, not available; NHS, National Health Service; PSA, probabilistic sensitivity analysis; SD, standard deviation.

[a] Resource use is applied per model cycle.

[b] Total costs for each treatment arm were calculate by multiplying the resource-use estimates with the unit costs for red blood cells and platelets. The unit costs were obtained from NHS reference costs.[151]

B.3.5.6. Adverse reaction unit costs and resource use

The cost of managing grade 3 or 4 TEAEs is calculated by combining the number of adverse events reported in the ALFA-0701 study with the estimates of the cost per event. AEs are assumed to occur within the first year of the model and as such are not discounted. Table 46 presents a summary of the data used for adverse-event costs in the model.

The costs were based on NHS Reference Costs, where available.[151] These costs are nationally representative of care in the UK and reflect observed length of stay for UK patients but are not specific to AML patients.

The cost of VOD was calculated based on the recommended diagnosis and treatment of VOD caused by GO in the AML 17 trial.[88] For diagnosis, a cost of £611.79 is included for an endoscopic ultrasound examination.[151] VOD associated with GO appears to be a milder form of the disease than VOD following HSCT, the AML 17 trial protocol recommended that a total of 10 mg/kg of defibrotide is administered daily for 7 days. Mean body weight data from the ALFA-0701 study are used to calculate the dose. An official list price for defibrotide is not available. In the AML 17 trial, defibrotide is supplied at a cost of £1,273 for 10, 200mg vials. However, NHS England Commissioning reports an expected list price of £365 per 200 mg vial;[153] the model uses this price.

Table 46 Cost of Adverse Events for First-Line AML Therapies

Adverse Event	Cost (£)	PSA Distribution	Source (Code)
Skin toxicity	1,586[a]	Normal	Department of Health (2016)[151] (JD07A-K)
Mucosal toxicity	1,493[a]	Normal	Department of Health (2016)[151] (FZ91A-M)
Pain	1,009[a]	Normal	Department of Health (2016)[151] (WH08A-B)
Nausea, vomiting, and diarrhea	1,493[a]	Normal	Department of Health (2016)[151] (FZ91A-M)
Pulmonary toxicity	1,527[a]	Normal	Department of Health (2016)[151] (DZ20D-F)
Cardiac rhythm disorder	997[a]	Normal	Department of Health (2016)[151] (EB07A-E)
Other cardiac toxicity	1,713[a]	Normal	Department of Health (2016)[151] (EB14A-E)
Central neurological toxicity	389	Normal	Department of Health (2016)[151] (VC12Z)
Peripheral neurological toxicity	389	Normal	Department of Health (2016)[151] (VC12Z)
Infections	1,938[a]	Normal	Department of Health (2016)[151] (WH07A-G)
Hemorrhage	1,251[a]	Normal	Department of Health (2016)[151] (SA02G-J)
Venu-occlusive disease	10,116.39[b]	Fixed	Calculation

AML, acute myeloid leukemia; PSA, probabilistic sensitivity analysis.

[a] Weighted-average cost calculated from the cost and total activity for each code.

[b] Includes cost for diagnosis using ultrasound and treatment with defibrotide per AML 17 trial protocol. The ultrasound cost is varied in the PSA but the total calculated cost is not.

B.3.6 Summary of base-case analysis inputs and assumptions

A comprehensive summary of variables that are used in the model are presented in Appendix N and Table 47 below for the subpopulation of favourable/intermediate cytogenetics profile known to benefit from GO as add-on therapy. All economic analyses were undertaken for the total population, summary of the values considered can be found in Appendix O.

Assumptions

All model assumptions were discussed in relevant subsections. For a comprehensive list of assumptions required in the model please see appendix N.

Table 47 Summary of Base-Case Adverse Events Applied in the Economic Model

Adverse Event	GO + DA: %	n/N	DA: %	n/N	Cost (£) Mean	Utility Decrement	
						Mean	SE
Skin toxicity	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	0.0207	0.00
Mucosal toxicity	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	0.0207	0.00
Pain	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	0.0207	0.00
Nausea, vomiting, and diarrhea	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	0.0207	0.00
Pulmonary toxicity	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	0.0207	0.00
Cardiac rhythm disorder	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	0.0207	0.00
Other cardiac toxicity	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	0.0207	0.00
Central neurological toxicity	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	0.0207	0.00
Peripheral neurological toxicity	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	0.0207	0.00
Infections	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	0.0207	0.00
Hemorrhage	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	0.0207	0.00
Veno-occlusive disease	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	0.0208	0.00
Source	ALFA-0701 CSR[24]		ALFA-0701 CSR[24]		Department of Health (2016)[151]	NICE TA399[66].	

DA, daunorubicin and cytarabine; GO, gemtuzumab ozogamicin; NICE, National Institute for Health and Care Excellence; SE, standard error.

B.3.7 Base-case incremental cost-effectiveness analysis results

Considering the evidence base and the modelling assumptions described in the sections above, results point out that adding GO in the clinical pathway is a cost-effective strategy to manage AML patients with a favourable and intermediate cytogenetics risk profile. GO + DA delivers approximately 1 additional QALY at the additional cost of approximately [redacted] per patient. This equates to an ICER of around £12,251 per average patient.

Table 48 Cost-effectiveness results: cytogenetic subpopulation

Technologies	Total costs (£)	Total LYG	Total QALYs	Incremental costs (£)	Incremental LYG	Incremental QALYs	ICER incremental (£/QALY)
GO + DA	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
DA	[redacted]	[redacted]	[redacted]				

ICER, incremental cost-effectiveness ratio; LYG, life years gained; QALYs, quality-adjusted life years.

Results for total population including cases with unfavourable cytogenetic profile can be found in Appendix O.

B.3.8 Sensitivity analyses

B.3.8.1. Probabilistic sensitivity analysis

Probabilistic sensitivity analysis was conducted with 1,000 iterations to capture stochastic uncertainty around model inputs. All parameters were varied simultaneously based on predefined probability distributions. Where uncertainty data was not available for an input, variability (i.e., SE) of 10% of the mean values was assumed.

Probabilistic sensitivity analysis results for the subpopulation population are presented on the cost-effectiveness plane in Figure 18 and the cost-effectiveness acceptability curve is presented in Figure 19. The mean probabilistic total costs were [redacted] (GO+DA) and [redacted] (DA). The mean probabilistic total QALYs were [redacted] (GO+DA) and [redacted] (DA); mean life years were [redacted] (GO+DA) and [redacted] (DA). The mean probabilistic ICER was £13,600 (95% credible interval [CrI]: [redacted] for costs; [redacted] for QALYs). Deterministic and probabilistic

cost effectiveness results are similar (11%). Increasing the numbers of simulations did not decrease this difference between ICERs. When survival input parameters were excluded from sampling during the PSA the difference was reduced to below 2%. A Cholesky decomposition is implemented in the model, which attempts to sample the random draws in a PSA while reflecting the correlation between survival parameters. However, this is not always enough to reduce the skew in the sampling distribution of mean ICERs produced by a PSA. At a willingness-to-pay threshold of £30,000, the probability of GO + DA being cost-effective was 80% (Figure 19).

PSA results for the all patients population are presented in Appendix O.4.1 for the total ALFA-0710 trial population.

Figure 18 Probabilistic Sensitivity Analysis Results Presented on the Cost-effectiveness Plane: Cytogenetic subpopulation

PSA, probabilistic sensitivity analysis; QALY, quality-adjusted life-year.

Figure 19 Cost-effectiveness acceptability curve: cytogenetic subpopulation

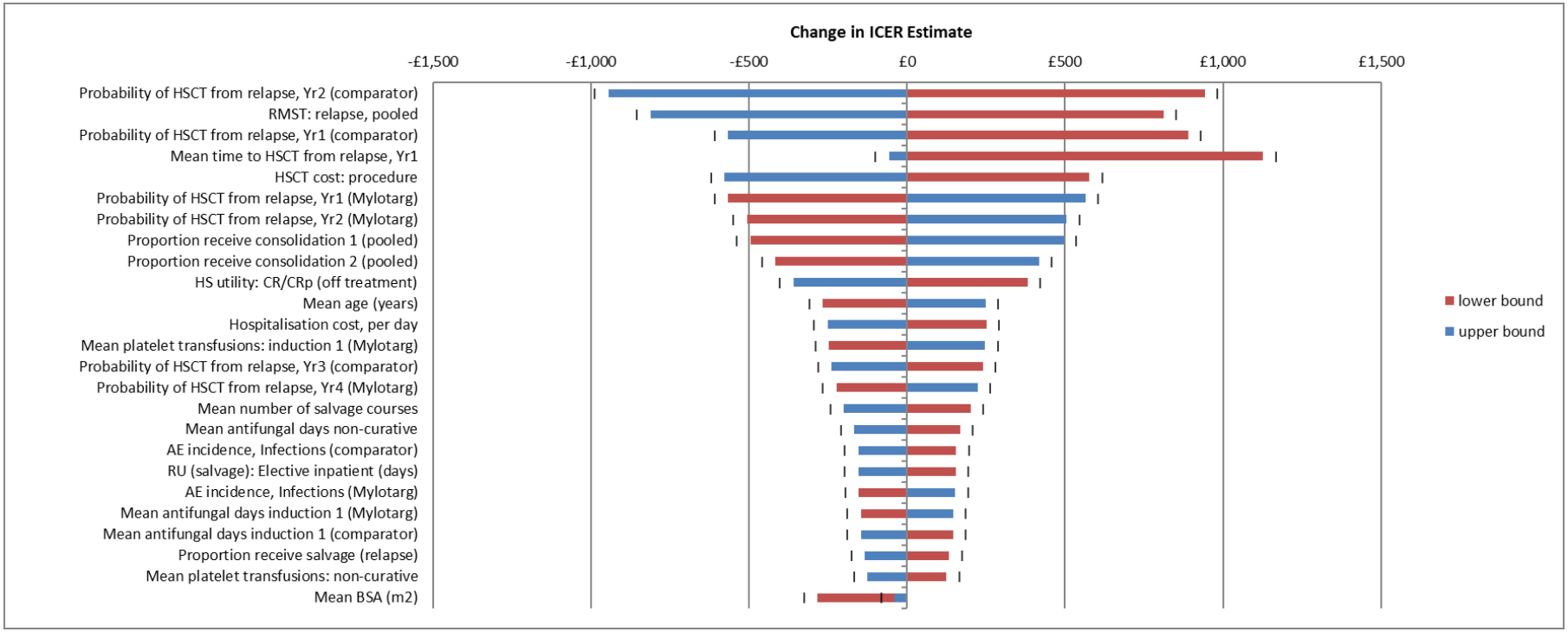
CEAC, cost-effectiveness acceptability curve; PSA, probabilistic sensitivity analysis; QALY, quality-adjusted life-year.

B.3.8.2. Deterministic sensitivity analysis

Automated univariate sensitivity analysis was performed in which parameters were varied by +/- 10%. Univariate sensitivity analysis results are presented in Figure 20. The tornado diagram demonstrated that the key areas of model sensitivity are HSCT probabilities from relapse in years 1 and 2 for the comparator and the RMST for relapse patients. The change in the ICER estimate for all parameters was less than £1,200.

Deterministic sensitivity analysis was also explored for the all patients group, results can be found in Appendix O.

Figure 20 Tornado Diagram: Cytogenetic subpopulation [£ per QALY]



AE, adverse event; BSA, body surface area; CR, complete remission; CRp, complete remission with incomplete platelet recovery; DA, daunorubicin and cytarabine; GO, gemtuzumab ozogamicin; GP, general population; HS, health state; HSCT, hematopoietic stem-cell transplant; ICER, incremental cost-effectiveness ratio; QALY, quality-adjusted life-year; RMST, restricted mean survival time; RU, resource use; SoC, standard of care.

Notes: The quadrant where the ICER falls is shown in the graph at the ends of each bar: I = quadrant 1 (GO + DA is more expensive and more effective than the comparator); II = quadrant 2 (GO + DA is dominated by the comparator); III = quadrant 3 (GO + DA is less expensive and less effective than the comparator); IV = quadrant 4 (GO + DA is dominant over the comparator). The age variable uses the 95% CI bounds from the ALFA-0701 trial.

B.3.8.3. Scenario analysis

Table 49 presents results from a range of scenario analyses. Univariate scenario analyses of key model drivers and areas of uncertainty were tested in the model to evaluate how these impact results. The average ICER ranged from £6,821 per patient, when the best fitting standard parametric survival functions (combined treatment-arm function for OS (CR or CRp) were used, assuming no cure rate, to £20,334, when the RMST for relapse patients was based on individual treatment arms. Other scenarios that resulted in variations of more than £5,000 per QALY gained from the base-case ICER were pooling HSCT probabilities from relapse and excluding all HSCTs. It is important to note that results were relatively stable to the survival method employed. Indeed, the MCM base case model gave the highest ICERs compared to the best fitting standard parametric curves and spline models (with and without cure point).

The model base case already reflects that some proportion of patients in each arm will receive HSCTs (section B.3.3.8) and this will have an effect on quality of life (section B.3.4.4) and costs (section B.3.5.4). The NICE scope (ID982)[154] requests a scenario that isolates the accrued costs and QALYs of those patients that receive a HSCT. A scenario is presented in Table 49 whereby no patients receive a HSCT in either arm of the model. This increases the ICER to £19,617, driven primarily by the significant decrease in relapse HSCT costs in the DA arm – i.e. more HSCTs occurred in relapse patients in the DA arm of ALFA-0701. An additional scenario simply adds together the accrued costs and QALYs associated with receiving a HSCT in the model as an average across all model patients. This includes costs and QALYs accrued in the transplant procedure, and post-HSCT CR or CRp with and without GVHD health states. It also includes the QALYs accrued in the functionally cured state for patients who received a HSCT. Costs and QALYs are lower in the GO + DA arm because less patients receive a HSCT: total costs are [redacted] (GO + DA) and [redacted] (DA); total QALYs [redacted] (GO + DA) and [redacted] (DA).

Cost-effectiveness scenario analyse for the total ALFA-0710 trial population can be found in Appendix O.

Table 49 Scenario Analyses: Cytogenetic subpopulation

Parameter	Base Case	Scenario Analysis	Δ Cost	Δ QALYs	ICER
Base-case results			[redacted]	[redacted]	£12,251
Response rates					
Pooling assumption	Pooled arms	Individual arms	[redacted]	[redacted]	£10,395
Survival functions: RFS and OS (CR or CRp)					
MCM functions	MCM log-normal	MCM Weibull[a]	[redacted]	[redacted]	£12,233
Flexible-spine functions	MCM log-normal	Best-fitting flexible-spine functions (cure applied)[b]	[redacted]	[redacted]	£10,621
Flexible-spine functions	MCM log-normal	Best-fitting flexible-spine functions (cure applied at 3 years)[c]	[redacted]	[redacted]	£10,724
Flexible-spine functions	MCM log-normal	Best-fitting flexible-spine functions (cure applied); OS (CR) combined function with treatment covariate[d]	[redacted]	[redacted]	£9,902
Flexible-spine functions	MCM log-normal	Best-fitting flexible-spine functions with (no cure applied)[e]	[redacted]	[redacted]	£10,895
Flexible-spine functions	MCM log-normal	Best-fitting flexible-spine functions (no cure applied); OS (CR) combined function with treatment covariate[f]	[redacted]	[redacted]	£9,889
Standard parametric functions	MCM log-normal	Best-fitting parametric functions (cure applied)[g]	[redacted]	[redacted]	£8,923
Standard parametric functions	MCM log-normal	Best-fitting parametric functions (cure applied at 3 years)[h]	[redacted]	[redacted]	£10,558
Standard parametric functions	MCM log-normal	Best-fitting parametric functions (cure applied); OS (CR) combined function with treatment covariate[i]	[redacted]	[redacted]	£8,911
Standard parametric functions	MCM log-normal	Best-fitting parametric functions (no cure applied)[j]	[redacted]	[redacted]	£7,669
Standard parametric functions	MCM log-normal	Best-fitting parametric functions (no cure applied); OS (CR) combined function with treatment covariate[k]	[redacted]	[redacted]	£6,821

Parameter	Base Case	Scenario Analysis	Δ Cost	Δ QALYs	ICER
Survival functions: OS (refractory)					
Assumption between treatment arms	Pooled arms	Individual arms[l]	[redacted]	[redacted]	£15,692
HSCT probabilities					
HSCT from relapse	Individual arms	Pooled arms	[redacted]	[redacted]	£19,787
NICE scope scenario: Include/exclude HSCTs	HSCTs included	No HSCTs occur in either arm	[redacted]	[redacted]	£19,617
NICE scope scenario: HSCT proportion only	All patients accrued costs and QALYs counted	Only patients who received HSCT accrued costs and QALYs counted	[redacted]	[redacted]	NA
Excess mortality HR for long-term AML survivors vs. general population					
HR	[redacted]	[redacted]	[redacted]	[redacted]	£11,795
HR	[redacted]	[redacted]	[redacted]	[redacted]	£12,531
HR	[redacted]	[redacted]	[redacted]	[redacted]	£13,325
HR	[redacted]	[redacted]	[redacted]	[redacted]	£14,022
HR	[redacted]	[redacted]	[redacted]	[redacted]	£14,650
HR	[redacted]	[redacted]	[redacted]	[redacted]	£15,225
HR	[redacted]	[redacted]	[redacted]	[redacted]	£15,758
Cost calculations					
RMST assumption (relapse and refractory patients)	Pooled arms	Individual arms	[redacted]	[redacted]	£20,334
Drug wastage	Excluded	Included	[redacted]	[redacted]	£11,965
First-line treatment courses assumption	Pooled arms	Individual arms	[redacted]	[redacted]	£11,489
Health-state utility weights					
Source	EQ-5D (default)[m]	EQ-5D (alternative)[n]	[redacted]	[redacted]	£12,132
Source	EQ-5D (default)[m]	Pfizer VAS	[redacted]	[redacted]	£15,180
Source	EQ-5D (default)[m]	Pfizer TTO	[redacted]	[redacted]	£14,094
Functionally cured health-state calculation	Age-adjusted	Fixed value	[redacted]	[redacted]	£11,967
Utility decrements for adverse events					
Source	EQ-5D	EQ-5D (alternative)[n]	[redacted]	[redacted]	£12,253

Parameter	Base Case	Scenario Analysis	Δ Cost	Δ QALYs	ICER
	(default)[m]				

CR, CRp, gemtuzumab ozogamicin; GO, HR, hazard ratio; HSCT, hematopoietic stem-cell transplant; ICER, incremental cost-effectiveness ratio; IRC, independent review committee; MCM, mixture cure model; OS, overall survival; QALY, quality-adjusted life-year; RMST, restricted mean survival time; TTO, time trade-off; VAS, visual analog scale.

[a] MCM Weibull function selected for RFS and OS (CR or CRp) for both treatment arms.

[b] Spline log-normal and spline Weibull functions selected as best fitting functions for RFS (GO and comparator arms, respectively); for OS (CR or CRp), spline Weibull and spline log-normal functions selected as best fitting for GO and comparator arms, respectively. Cure rate assumed at the end of trial follow-up; extrapolations for RFS and OS based on general populations mortality rates adjusted by a HR for excess mortality for long-term AML survivors.

[c] Spline log-normal and spline Weibull functions selected as best fitting functions for RFS (GO and comparator arms, respectively); for OS (CR or CRp), spline Weibull and spline log-normal functions selected as best fitting for GO and comparator arms, respectively. Cure rate assumed at 3 years; extrapolations for RFS and OS based on general populations mortality rates adjusted by a HR for excess mortality for long-term AML survivors.

[d] Spline log-normal and spline Weibull functions selected as best fitting functions for RFS (GO and comparator arms, respectively); for OS (CR or CRp), spline log-normal combined treatment-arm function selected as best fitting function. Cure rate assumed at the end of trial follow-up; extrapolations for RFS and OS based on general populations mortality rates adjusted by a HR for excess mortality for long-term AML survivors.

[e] Spline log-normal and spline Weibull functions selected as best fitting functions for RFS (GO and comparator arms, respectively); for OS (CR or CRp), spline Weibull and spline log-normal functions selected as best fitting for GO and comparator arms, respectively. No cure rate assumed at the end of trial follow-up; extrapolations for RFS and OS based on curve extrapolations (or general populations mortality rates adjusted by a HR for excess mortality for long-term AML survivors if higher than curve prediction).

[f] Spline log-normal and spline Weibull functions selected as best fitting functions for RFS (GO and comparator arms, respectively); for OS (CR or CRp), spline log-normal combined treatment-arm function selected as best fitting function. No cure rate assumed at the end of trial follow-up; extrapolations for RFS and OS based on curve extrapolations (or general populations mortality rates adjusted by a HR for excess mortality for long-term AML survivors if higher than curve prediction).

[g] Generalized gamma function selected as best fitting function for RFS and OS (CR or CRp) for both treatment arms. Cure rate assumed at the end of trial follow-up; extrapolations for RFS and OS based on general populations mortality rates adjusted by a HR for excess mortality for long-term AML survivors.

[h] Generalized gamma function selected as best fitting function for RFS and OS (CR or CRp) for both treatment arms. Cure rate assumed at 3 years; extrapolations for RFS and OS based on general populations mortality rates adjusted by a HR for excess mortality for long-term AML survivors.

[i] Generalized gamma function selected as best fitting function for RFS (both treatment arms); for OS (CR or CRp), generalized gamma combined treatment-arm function selected as best fitting function. Cure rate assumed at the end of trial follow-up; extrapolations for RFS and OS based on general populations mortality rates adjusted by a HR for excess mortality for long-term AML survivors.

[j] Generalized gamma function selected as best fitting function for RFS and OS (CR or CRp) for both treatment arms. No cure rate assumed at the end of trial follow-up; extrapolations for RFS and OS based on curve extrapolations (or general populations mortality rates adjusted by a HR for excess mortality for long-term AML survivors if higher than curve prediction).

[k] Generalized gamma function selected as best fitting function for RFS (both treatment arms); for OS (CR or CRp), generalized gamma combined treatment-arm function selected as best fitting function. No cure rate assumed at the end of trial follow-up; extrapolations for RFS and OS based on curve extrapolations (or general populations mortality rates adjusted by a HR for excess mortality for long-term AML survivors if higher than curve prediction).

[l] Gompertz function selected for both treatment arms.

[m] Values from NICE technology appraisal 399, using the mapping algorithm by McKenzie and Van der Pol (2009).

[n] Values from NICE technology appraisal 399, using the mapping algorithm by Proskorovsky et al. (2014).

B.3.9 Subgroup analysis

This economic evaluation has focused on the subpopulation defined in section B.3.2.1 (i.e. excluding patients with unfavourable cytogenetics profile) and there are no subgroup analyses to present here. Appendix O presents all analyses and results associated with the total population of the ALFA-0701 study.

B.3.10 Validation

Validation of de novo cost-effectiveness analysis

This economic evaluation has been informed by the relevant published literature, previously published NICE appraisals and feedback on model design from independent UK clinical and statistical experts. Model validation was performed in alignment with best practices[155] and survival model selection follows recommendations from NICE DSU TSD14[110] aligned with the NICE reference case. A brief description of validation tasks is presented in Table 50.

Table 51 compares base case model predictions with Kaplan-Meier data from the ALFA-0701 study. Model results are virtually identical to the ALFA-0701 KM data, which is expected because our chosen base case OS and RFS curves demonstrate a strong goodness of fit to the trial follow-up data. A comparison with the OS data provided by professor Robert Hills and used to calculate the excess mortality HR is warranted (section B.3.3.4). This pooled AML 10 to 16 trial data for patients with favourable and intermediate cytogenetics, conditional on survival to 5 years showed the following proportions alive: [redacted]. MCM model base case predictions of those alive, conditional on survival at 5 years, were similar to the pooled AML trial estimates: [redacted]. It is important to note that the model predictions in Table 50 are not comparable to the pooled AML trial data because these reflect all patients (i.e. including those that died before 5 years).

Table 50 Validation of the de novo cost-effectiveness analysis

Validation performed by	Nature of validation	Date(s)	Aspects covered
UK clinical expert Statistics expert (Professor Robert Hills)	Expert panel meeting	January 2016	UK treatment pathway, model structure, clinical assumptions, UK-specific input parameters
6 UK clinical experts	Expert panel meeting	July 2016	Model health states and descriptions for preference elicitation study, questionnaire for resource-use estimates
Health economics expert (Dr. Nicholas Latimer)	External reviewer	October 2016	Survival analysis plan
3 UK clinical experts 2 health economics experts	Advisory board meeting	February 2017	Clinical data; UK treatment pathway; model structure and assumptions; survival analyses and other transition probabilities; resource-use, costs, and utility estimates
UK clinical expert	Teleconference	August 2017	Clinical assumptions and data gaps
Health economics expert (Dr. Nicholas Latimer)	External reviewer	September 2017	Survival analysis plan for MCM functions
Statistics expert (Professor Robert Hills)	Teleconference	October 2017	Long-term survival projections and model data
UK clinical expert (Dr. Paul Cahalin)	Teleconference	October 2017	Validation of model's long-term extrapolated outcomes and clinical inputs
UK clinical expert	Meeting	October 2017	Validation of model's long-term extrapolated outcomes and clinical inputs
RTI Health Solutions	Quality control	October 2017	Checked input data against sources, reviewed model programming
Health economics expert	External reviewer	October 2017	Validated model assumptions, verified robustness of model clinical inputs, checked model survival outputs

Table 51 Comparison of model and trial outcomes across time for favourable/intermediate cytogenetic patients

Outcome	Source	3 months (%)	6 months (%)	12 months (%)	18 months (%)	60 months (%)	120 months (%)	240 months (%)
RFS	Trial	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]		
	Model	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
OS	Trial	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]		
	Model	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]

Trial OS and RFS are based on ALFA-0701 study KM curves. Model OS and RFS were derived from the MCM base case model.

Clinicians (see Table 50) validated the following as clinically plausible:

- cure rates predicted by the estimated base case MCM (Table 52)
- base case model extrapolations for OS and RFS for patients in complete remission
- base case model extrapolations for OS for refractory patients

Table 52 Statistical cure rates for MCM Log-normal Base Case

Outcome	GO	No GO	IRC trial CR +CRp (pooled)
All patients			
OS (CR)	[redacted]	[redacted]	[redacted]
RFS	[redacted]	[redacted]	
Cytogenetic subpopulation			
OS (CR)	[redacted]	[redacted]	[redacted]
RFS	[redacted]	[redacted]	

CR, complete response; RFS, relapse-free survival; OS, overall survival.

Clinicians also validated the view that patients are considered “functionally cured” in clinical practice after remaining in complete remission for 3 to 5 years.

External validation

The model structure was designed, refined and validated following a series of 3 advisory board meetings held by Pfizer (January 2016, July 2016, and lastly in

February 2017), Validation of the model was undertaken in terms of ensuring that the relevant population, comparator treatments, model structure, assumptions and input parameters were clinically plausible. Resource use data and assumptions were discussed and clinical opinion was incorporated in the model in cases where there was a lack of published data. The key modelling recommendation was to replace the simple partitioned survival model with a cohort state-transition model that could incorporate additional transitions. It was also advised that cure models should be explored as a way to capture the prominent plateaus observed in trial data.

The survival analysis plan was validated by an external health economic and statistical expert. An additional external health economics and modelling expert reviewed the model in order to ensure the validity of assumptions and inputs.

Quality control

Quality-control procedures in terms of model programming were performed by an independent modeller in accordance with a pre-specified test plan. These procedures were designed to ensure no coding errors occurred and to verify model output via sense checking.

B.3.11 Interpretation and conclusions of economic evidence

Comparison with previous economic evaluations

This is the first economic evaluation undertaken for de novo AML patients, thus it was not possible to compare results presented here with previous published analyses.

Relevance of economic evaluation to all groups of patients who could potentially use the technology

The results of the economic evaluation align with the decision problem statement and are presented for the population expected to gain the most clinical/economic benefit from first line GO add-on treatment – de novo AML patients with favourable and intermediate cytogenetic risk (i.e. excluding those with unfavourable cytogenetics profile).

Relevance (and generalisability) to clinical practice in England

The patients in the ALFA-0701 study (French patients aged between 50 and 70 years) were considered by clinical experts to be representative of the population expected to receive GO + DA in routine UK clinical practice in terms of their underlying risk for disease progression, risk for death, and response to treatment. The UK clinical experts did detect some differences in the treatment schema between the ALFA-0701 study and that in UK clinical practice:

- The fractionated induction GO dose was capped at 5 mg (the ongoing UK AML 18 and AML 19 trials use 2 doses of 3 mg/m^2; whereas in the previous UK AML 15 and AML 16 trials, a single dose of 3 mg/m^2 was used).
- GO was used in consolidation courses (GO was not used during consolidation in the UK trials).
- The DA 3+7 regimen was used as continuous infusion (the DA 3+10 regimen as a bolus dose was used in the UK trials).
- The cytarabine consolidation dose of 1 g/m^2 was used (a dose of 3 g/m^2 was used in the UK trials).
- The daunorubicin dose was split across two consolidation courses (daunorubicin was used only in the first consolidation course in the UK trials).

The UK clinical experts advised that it is difficult to define a standard practice because treatment is often guided by individual patient characteristics and thus expressed a preference for the fractioned dose as this would allow them more flexibility in clinical practice. Overall, clinical experts expressed that patients in UK clinical practice are not materially different from those of the ALFA-0701 study in terms of care and characteristics and so no generalisability issues were stressed.

Study Strengths

The structure of the model was informed by previous models,[66,104,142] and developed in accordance with extensive feedback from clinicians, and economic advisors (Section Table 50). Modelling techniques were used to extrapolate the data, representing the proportion of cured patients, so as to appropriately capture the benefit of add-on GO in extending relapse free survival compared to DA alone. The

model design is more advanced than a partitioned survival analysis and includes additional transitions to capture second-line treatments and HSCTs, based on analyses of patient-level data from the ALFA-0701 study.

Additional analyses were conducted to compensate for the lack of published data in this disease area. Pfizer commissioned and published a preference elicitation study[108,139] to estimate health state utilities for consideration in the model (Appendix H). No appropriate literature could be found for long-term excess mortality risk to match the population of this decision problem and so an estimate was calculated based on pooled UK AML trial data (section B.3.3.4). This approach is considered to provide the best estimate for long-term excess mortality as it is based on evidence form clinical studies conducted in the UK. Alternative values were considered in the analyses presented for completeness and to test the impact of different assumptions in the results.

Study Limitations

The AML treatment pathway is complex. Performing a sound cost-effectiveness analysis in this field requires advanced modelling techniques to capture outcomes over a patient's lifetime from diagnosis. The model has been built under the best available evidence and relied on clinical opinion where there were data gaps in order to address clinical uncertainty. Multiple UK Clinical experts were interviewed to ensure alignment.

Published estimates were considered and utilised in the absence of trial data that could not fully capture disease complexity. However, the selection of sources and validation of assumptions was based on external validation. For example, second-line CR or CRp for relapsed or refractory patients was not reported in the ALFA-0701 study; therefore, second-line CR or CRp was not explicitly modelled. However, since the main benefit of GO is prolonged RFS and EFS, it was important to capture second-line outcomes. To do so required the clinically validated assumption that relapsed or refractory patients who received a HSCT in the ALFA-0701 study, had first attained second line CR (section B.3.3.8).

In clinical practice, non-curative therapies are followed by palliative care until death. Using this assumption until the end of the model time horizon would have over-predicted the costs for non-curative therapies because some patients attain second-line CR or CRp and no longer require treatment. Since there were more relapses in the control arm, this assumption would have favoured GO + DA. Instead of using this approach, the model used the RMST from the ALFA-0701 study to calculate non-curative therapy costs.

Interpretation of economic evidence

The base-case analysis and sensitivity/scenario results suggest that the addition of GO to standard care is a cost-effective strategy to manage favourable/intermediate cytogenetic risk profile AML patients.

This is reflective of the survival data which shows that the addition of GO significantly improves time spent in relapse-free survival over the average patients lifetime (section B.3.3.5). Scenario analyses show that the cost-effectiveness gains from this are stable across a variety of survival methods (ICERs ranged from £6,821 to £12,233).

The lifetime incremental QALYs and life-years were [redacted] and [redacted], respectively, yielding an ICER of £12,251 per patient. The probability that GO is cost-effective at a willingness-to-pay threshold of £30,000 per QALY gained was 80%. In univariate sensitivity analyses, the change in the ICER for all parameters was less than £1,200. In scenario analyses, ICER values ranged from £6,821, when the best fitting standard parametric survival functions (combined treatment-arm function for OS (CR or CRp) were used, assuming no cure rate, to £20,334, when the RMST for relapse patients was based on individual treatment arms.

The probability that GO is cost-effective is 80% at a willingness-to-pay threshold of £30,000 /QALY. This reinforces the low risk in the decision of the NHS adopting GO. The addition of GO to standard chemotherapy in the NHS will significantly improve outcomes for patients with favourable/intermediate cytogenetics, essentially prolonging time in relapse-free survival for the large proportion of patients that obtain complete remission, whilst being considered a cost-effective use of NHS resources.

B.4 References

1. British Committee for Standards in H, Milligan DW, Grimwade D *et al.* Guidelines on the management of acute myeloid leukaemia in adults. *Br J Haematol* 2006;135:450-74.
2. Linenberger ML. CD33-directed therapy with gemtuzumab ozogamicin in acute myeloid leukemia: progress in understanding cytotoxicity and potential mechanisms of drug resistance. *Leukemia* 2005;19:176-82.
3. Ricart AD. Antibody-drug conjugates of calicheamicin derivative: gemtuzumab ozogamicin and inotuzumab ozogamicin. *Clin Cancer Res* 2011;17:6417-27.
4. De Kouchkovsky I, Abdul-Hay M. Acute myeloid leukemia: a comprehensive review and 2016 update. *Blood Cancer J* 2016;6:e441.
5. Laszlo GS, Estey EH, Walter RB. The past and future of CD33 as therapeutic target in acute myeloid leukemia. *Blood Rev* 2014;28:143-53.
6. Jilani I, Estey E, Huh Y *et al.* Differences in CD33 intensity between various myeloid neoplasms. *Am J Clin Pathol* 2002;118:560-6.
7. Ehninger A, Kramer M, Rollig C *et al.* Distribution and levels of cell surface expression of CD33 and CD123 in acute myeloid leukemia. *Blood Cancer J* 2014;4:e218.
8. Cowan AJ, Laszlo GS, Estey EH, Walter RB. Antibody-based therapy of acute myeloid leukemia with gemtuzumab ozogamicin. *Front Biosci (Landmark Ed)* 2013;18:1311-34.
9. Damle NK. Tumour-targeted chemotherapy with immunoconjugates of calicheamicin. *Expert Opin Biol Ther* 2004;4:1445-52.
10. Walter RB, Appelbaum FR, Estey EH, Bernstein ID. Acute myeloid leukemia stem cells and CD33-targeted immunotherapy. *Blood* 2012;119:6198-208.
11. Taksin AL, Legrand O, Raffoux E *et al.* High efficacy and safety profile of fractionated doses of Mylotarg as induction therapy in patients with relapsed acute myeloblastic leukemia: a prospective study of the alfa group. *Leukemia* 2007;21:66-71.
12. European Medicines Agency. Public summary of opinion on orphan designation: gemtuzumab ozogamicin for the treatment of acute myeloid leukaemia. Available from: http://www.ema.europa.eu/docs/en_GB/document_library/Orphan_designation/2009/10/WC500006263.pdf (accessed 26 July 2017).
13. Duncan N. Adult myeloid leukaemias pathogenesis, clinical features and classification. *Clinical Pharmacist* 2010;2:117–21.
14. Cancer Research UK. About acute myeloid leukaemia (AML). Available from http://www.cancerresearchuk.org/about-cancer/acute-myeloid-leukaemia-aml/about-acute-myeloid-leukaemia (accessed 28 June 2017).
15. Stone RM, O'Donnell MR, Sekeres MA. Acute myeloid leukemia. *Hematology Am Soc Hematol Educ Program* 2004:98-117.
16. Kumar CC. Genetic abnormalities and challenges in the treatment of acute myeloid leukemia. *Genes Cancer* 2011;2:95-107.
17. Hjelle SM, Forthun RB, Haaland I *et al.* Clinical proteomics of myeloid leukemia. *Genome Med* 2010;2:41.

18. Cancer Research UK. About acute lymphoblastic leukaemia (ALL). Accessed http://www.cancerresearchuk.org/about-cancer/acute-lymphoblastic-leukaemia-all/about (accessed on 22 August 2017).
19. Cancer Research UK. About chronic lymphocytic leukaemia. Accessed http://www.cancerresearchuk.org/about-cancer/chronic-lymphocytic-leukaemia-cll/about (accessed on 22 August 2017).
20. Hulegardh E, Nilsson C, Lazarevic V *et al.* Characterization and prognostic features of secondary acute myeloid leukemia in a population-based setting: a report from the Swedish Acute Leukemia Registry. *Am J Hematol* 2015;90:208-14.
21. Dohner H, Estey EH, Amadori S *et al.* Diagnosis and management of acute myeloid leukemia in adults: recommendations from an international expert panel, on behalf of the European LeukemiaNet. *Blood* 2010;115:453-74.
22. Dohner H, Estey E, Grimwade D *et al.* Diagnosis and management of AML in adults: 2017 ELN recommendations from an international expert panel. *Blood* 2017;129:424-47.
23. Acute myelogenous leukemia stages, Cancer Treatment Centers of America. Accessed at http://www.cancercenter.com/leukemia/stages/tab/acute-myelogenous-leukemia/ (accessed on 3 August 2017).
24. Pfizer, Full Clinical Study Report for Protocol WS936568 (ALFA-0701 [MyloFrance 3]), 15 March 2016.
25. European Medicines Agency, Orphan designation. Accessed http://www.ema.europa.eu/ema/index.jsp?curl=pages/regulation/general/general_content_000029.jsp&mid=WC0b01ac05800240ce (accessed 18 October 2017).
26. Li J, Smith A, Crouch S, Oliver S, Roman E. Estimating the prevalence of hematological malignancies and precursor conditions using data from Haematological Malignancy Research Network (HMRN). *Cancer Causes Control* 2016;27:1019-26.
27. Cancer Research UK, Acute myeloid leukaemia (AML) incidence statistics. Accessed http://www.cancerresearchuk.org/health-professional/cancer-statistics/statistics-by-cancer-type/leukaemia-aml/incidence#heading-Zero (accessed 18 October 2017).
28. Cancer Research UK, AML statistics. Accessed http://www.cancerresearchuk.org/health-professional/cancer-statistics/statistics-by-cancer-type/leukaemia-aml#heading-Zero (accessed on 18 October 2017).
29. Cancer Research UK. Acute Myeloid Leukemia (AML) Statistics. 2013; Accessed at http://www.cancerresearchuk.org/about-cancer/acute-myeloid-leukaemia-aml/survival (accessed on 7 September 2017)
30. Office for National Statistics. Cancer Registration Statistics, England, 2015. Accessed from: https://www.ons.gov.uk/peoplepopulationandcommunity/healthandsocialcare/conditionsanddiseases/datasets/cancerregistrationstatisticscancerregistrationstatisticsengland (accessed on 23 August 2017).
31. Burnett AK, Russell NH, Hills RK *et al.* Addition of gemtuzumab ozogamicin to induction chemotherapy improves survival in older patients with acute myeloid leukemia. *J Clin Oncol* 2012;30:3924-31.

32. Cancer Research UK, AML Mortality statistics. Accessed http://www.cancerresearchuk.org/health-professional/cancer-statistics/statistics-by-cancer-type/leukaemia-aml/mortality#heading-Zero (accessed 18 October 2017).
33. Bhayat F, Das-Gupta E, Smith C, McKeever T, Hubbard R. The incidence of and mortality from leukaemias in the UK: a general population-based study. *BMC Cancer* 2009;9:252.
34. National Cancer Intelligence Network Trends in incidence and outcome for haematological cancers in England: 2001–2010, Public Health England. Published November 2014.
35. Pfizer meta-analysis CSR: A patient-level meta-analysis of studies utilizing intensive induction chemotherapy with or without mylotarg (gemtuzumab ozogamicin) in patients with previously untreated acute myeloid leukemia report. PF-05208747.
36. Burnett AK, Hills RK, Milligan D *et al.* Identification of patients with acute myeloblastic leukemia who benefit from the addition of gemtuzumab ozogamicin: results of the MRC AML15 trial. *J Clin Oncol* 2011;29:369-77.
37. University College London Hospitals NHS Foundation Trust. What is acute myeloid leukaemia (AML)? Available from https://www.uclh.nhs.uk/OurServices/ServiceA-Z/Cancer/CBD/AML/Pages/WhatisAML.aspx (accessed 29 June 2017).
38. American Cancer Society. Acute myeloid leukemia (AML) signs and symptoms. Available from: https://www.cancer.org/cancer/acute-myeloid-leukemia/detection-diagnosis-staging/signs-symptoms.html (Accessed 16 May 2017).
39. Bloodwise. Signs and symptoms of acute myeloid leukemia. Accessed at: https://bloodwise.org.uk/acute-myeloid-leukaemia-adult-aml/signs-symptoms-and-diagnosis-acute-myeloid-leukaemia-aml (accessed on 29 June 2017).
40. Williams LA, Garcia-Gonzalez A, Ahaneku HO *et al.* A patient-reported outcome measure for symptoms and symptom burden of acute myeloid leukemia (AML) and myelodysplastic syndrome (MDS). Poster presented at the American Society of Hematology 57th Annual Meeting and Exposition, 5–8 December 2015, Orlando, Fl, USA. .
41. Leunis A, Redekop WK, Uyl-de Groot CA, Lowenberg B. Impaired health-related quality of life in acute myeloid leukemia survivors: a single-center study. *Eur J Haematol* 2014;93:198-206.
42. Sekeres MA, Stone RM, Zahrieh D *et al.* Decision-making and quality of life in older adults with acute myeloid leukemia or advanced myelodysplastic syndrome. *Leukemia* 2004;18:809-16.
43. Alibhai SM, Leach M, Gupta V *et al.* Quality of life beyond 6 months after diagnosis in older adults with acute myeloid leukemia. *Crit Rev Oncol Hematol* 2009;69:168-74.
44. Uyl-de Groot CA, Lowenberg B, Vellenga E *et al.* Cost-effectiveness and quality-of-life assessment of GM-CSF as an adjunct to intensive remission induction chemotherapy in elderly patients with acute myeloid leukemia. *Br J Haematol* 1998;100:629-36.
45. Schumacher A, Kessler T, Buchner T, Wewers D, van de Loo J. Quality of life in adult patients with acute myeloid leukemia receiving intensive and prolonged chemotherapy – a longitudinal study. *Leukemia* 1998;12:586-92.

46. Jia M, Li J, Chen C, Cao F. Post-traumatic stress disorder symptoms in family caregivers of adult patients with acute leukemia from a dyadic perspective. *Psychooncology* 2015;24:1754-60.
47. Zeidan AM, Mahmoud D, Kucmin-Bemelmans IT *et al.* Economic burden associated with acute myeloid leukemia treatment. *Expert Rev Hematol* 2016;9:79-89.
48. Uyl-de Groot CA, Gelderblom-den Hartog J, Huijgens PC, Willemze R, van Ineveld BM. Costs of diagnosis, treatment, and follow up of patients with acute myeloid leukemia in the netherlands. *J Hematother Stem Cell Res* 2001;10:187-92.
49. Irish W, Ryan M, Gache L *et al.* Acute myeloid leukemia (AML): a retrospective claims analysis of resource utilization and expenditures for de novo patients from first line induction to remission and relapse. Presented at: ISPOR; 2016; Washington, DC.
50. Leunis A, Blommestein HM, Huijgens PC *et al.* The costs of initial treatment for patients with acute myeloid leukemia in the Netherlands. *Leuk Res* 2013;37:245-50.
51. Bryant AL, Deal AM, Walton A *et al.* Use of ED and hospital services for patients with acute leukemia after induction therapy: one year follow-up. *Leuk Res* 2015;39:406-10.
52. Tennvall GR, Persson U, Nilsson B. The economic costs of acute myeloid leukemia in Sweden. *Int J Technol Assess Health Care* 1994;10:683-94.
53. Duncan N. Adult myeloid leukaemias: current and future treatments. *Clinical Pharmacist*2:122–27.
54. Breems DA, Van Putten WL, Huijgens PC *et al.* Prognostic index for adult patients with acute myeloid leukemia in first relapse. *J Clin Oncol* 2005;23:1969-78.
55. Dohner H, Weisdorf DJ, Bloomfield CD. Acute Myeloid Leukemia. *N Engl J Med* 2015;373:1136-52.
56. Yates JW, Wallace HJ, Jr., Ellison RR, Holland JF. Cytosine arabinoside (NSC-63878) and daunorubicin (NSC-83142) therapy in acute nonlymphocytic leukemia. *Cancer Chemother Rep* 1973;57:485-8.
57. Dombret H, Gardin C. An update of current treatments for adult acute myeloid leukemia. *Blood* 2016;127:53-61.
58. Popplewell LL, Forman SJ. Is there an upper age limit for bone marrow transplantation? *Bone Marrow Transplant* 2002;29:277-84.
59. Sorror ML, Estey E. Allogeneic hematopoietic cell transplantation for acute myeloid leukemia in older adults. *Hematology Am Soc Hematol Educ Program* 2014;2014:21-33.
60. The 2012 revised edition of the EBMT-ESH handbook on haematopoietic stem cell transplanation. Editors: J Apperley, E Carreras, E Gluckman, T Masszi. Chapter 19.1 Indictions for HSCT in adults: Acute Myeloid Leukaemia. Accessed at https://ebmtonline.forumservice.net/media/19_1/tex/content_alt/EBMT_Handbook2012_CHAP19_1.pdf (accessed on 14 August 2017).
61. Yanada M, Naoe T. Acute myeloid leukemia in older adults. *Int J Hematol* 2012;96:186-93.

62. Liersch R, Muller-Tidow C, Berdel WE, Krug U. Prognostic factors for acute myeloid leukaemia in adults – biological significance and clinical use. *Br J Haematol* 2014;165:17-38.
63. Redaelli A, Botteman MF, Stephens JM, Brandt S, Pashos CL. Economic burden of acute myeloid leukemia: a literature review. *Cancer Treat Rev* 2004;30:237-47.
64. The EBMT 2008 handbook – chapter 20: HSCT for acute myeloid leukaemia in adults. Available from: https://www.ebmt.org/Contents/Resources/Library/EBMTESHhandbook/Documents/EBMT2008_Cap20.pdf (accessed 14 August 2017).
65. Szer J. The prevalent predicament of relapsed acute myeloid leukemia. *Hematology Am Soc Hematol Educ Program* 2012;2012:43-8.
66. National Institute for Health and Care Excellence (NICE). Azacitidine for treating acute myeloid leukaemia with more than 30% bone marrow blasts (TA399), 27 July 2016. Available at: https://www.nice.org.uk/guidance/ta399 (Accessed 11 April 2017).
67. Myeloid leukemia – clinical diagnosis and treatment. Edited by Dr Steffen Koschmieder. Chapter 14 – Prognosis and survival in acute myelogenous leukemia, by Muath Dawod and Amr Hanbali.
68. Byrd JC, Mrozek K, Dodge RK *et al.* Pretreatment cytogenetic abnormalities are predictive of induction success, cumulative incidence of relapse, and overall survival in adult patients with de novo acute myeloid leukemia: results from Cancer and Leukemia Group B (CALGB 8461). *Blood* 2002;100:4325-36.
69. Medeiros BC, Othus M, Fang M, Roulston D, Appelbaum FR. Prognostic impact of monosomal karyotype in young adult and elderly acute myeloid leukemia: the Southwest Oncology Group (SWOG) experience. *Blood* 2010;116:2224-8.
70. Grimwade D, Hills RK, Moorman AV *et al.* Refinement of cytogenetic classification in acute myeloid leukemia: determination of prognostic significance of rare recurring chromosomal abnormalities among 5876 younger adult patients treated in the United Kingdom Medical Research Council trials. *Blood* 2010;116:354-65.
71. Cancer Genome Atlas Research N, Ley TJ, Miller C *et al.* Genomic and epigenomic landscapes of adult de novo acute myeloid leukemia. *N Engl J Med* 2013;368:2059-74.
72. Ostgard LS, Norgaard JM, Sengelov H *et al.* Comorbidity and performance status in acute myeloid leukemia patients: a nation-wide population-based cohort study. *Leukemia* 2015;29:548-55.
73. National Institute for Health and Care Excellence (NICE). NICE guideline NG47. Haematological cancers: improving outcomes. May 2016. Available at: https://www.nice.org.uk/guidance/ng47 (Accessed 11 April 2017).
74. Fey MF, Buske C, Group EGW. Acute myeloblastic leukaemias in adult patients: ESMO Clinical Practice Guidelines for diagnosis, treatment and follow-up. *Ann Oncol* 2013;24 Suppl 6:vi138-43.
75. National Institute for Health and Care Excellence (NICE). Azacitidine for the treatment of myelodysplastic syndromes, chronic myelomonocytic leukaemia and acute myeloid leukaemia (TA218). 23 March 2011. Available at: https://www.nice.org.uk/Guidance/TA218 (Accessed 11 April 2017).

76. Midostaurin for untreated acute myeloid leukaemia (NICE ID894). Accessed at: https://www.nice.org.uk/guidance/indevelopment/gid-ta10124 (accessed on 11 July 2017).
77. NICE Health Technology Appraisal Midostaurin for untreated acute myeloid leukaemia Final scope. Accessed at https://www.nice.org.uk/guidance/gid-ta10124/documents/final-scope (accessed on 11 July 2017).
78. EMA summary of product characteristics Dacogen (decitabine). Available from: http://www.ema.europa.eu/docs/en_GB/document_library/EPAR_-_Product_Information/human/002221/WC500133569.pdf (accessed 11 July 2017).
79. Chen Y, Yang T, Zheng X *et al.* The outcome and prognostic factors of 248 elderly patients with acute myeloid leukemia treated with standard-dose or low-intensity induction therapy. *Medicine (Baltimore)* 2016;95:e4182.
80. Zeidner JF, Foster MC, Blackford AL *et al.* Randomized multicenter phase II study of flavopiridol (alvocidib), cytarabine, and mitoxantrone (FLAM) versus cytarabine/daunorubicin (7+3) in newly diagnosed acute myeloid leukemia. *Haematologica* 2015;100:1172-9.
81. Ohtake S, Miyawaki S, Fujita H *et al.* Randomized study of induction therapy comparing standard-dose idarubicin with high-dose daunorubicin in adult patients with previously untreated acute myeloid leukemia: the JALSG AML201 Study. *Blood* 2011;117:2358-65.
82. Holowiecki J, Grosicki S, Giebel S *et al.* Cladribine, but not fludarabine, added to daunorubicin and cytarabine during induction prolongs survival of patients with acute myeloid leukemia: a multicenter, randomized phase III study. *J Clin Oncol* 2012;30:2441-8.
83. Mandelli F, Vignetti M, Suciu S *et al.* Daunorubicin versus mitoxantrone versus idarubicin as induction and consolidation chemotherapy for adults with acute myeloid leukemia: the EORTC and GIMEMA Groups Study AML-10. *J Clin Oncol* 2009;27:5397-403.
84. Saultz JN, Garzon R. Acute myeloid leukemia: a concise review. *J Clin Med* 2016;5.
85. Leopold LH, Willemze R. The treatment of acute myeloid leukemia in first relapse: a comprehensive review of the literature. *Leuk Lymphoma* 2002;43:1715-27.
86. Petersdorf SH, Kopecky KJ, Slovak M *et al.* A phase 3 study of gemtuzumab ozogamicin during induction and postconsolidation therapy in younger patients with acute myeloid leukemia. *Blood* 2013;121:4854-60.
87. Delaunay J, Recher C, Pigneux A. Addition of Gemtuzumab Ozogamycin to Chemotherapy Improves Event-Free Survival but Not Overall Survival of AML Patients with Intermediate Cytogenetics Not Eligible for Allogeneic Transplantation. Results of the GOELAMS AML 2006 IR Study. Blood (ASH Annual Meeting Abstracts). 2011;118:Abstract 79.
88. Burnett A, Cavenagh J, Russell N *et al.* Defining the dose of gemtuzumab ozogamicin in combination with induction chemotherapy in acute myeloid leukemia: a comparison of 3 mg/m2 with 6 mg/m2 in the NCRI AML17 Trial. *Haematologica* 2016;101:724-31.
89. Castaigne S, Pautas C, Terre C *et al.* Effect of gemtuzumab ozogamicin on survival of adult patients with de-novo acute myeloid leukaemia (ALFA-0701): a randomised, open-label, phase 3 study. *Lancet* 2012;379:1508-16.

90. Cardiff University Center for Clinical trials – AML19. Available from: http://www.cardiff.ac.uk/centre-for-trials-research/research/studies-and-trials/view/aml19 (accessed 12 August 2017).
91. Cancer Research UK, A trial looking at different treatments for acute myeloid leukaemia and high risk myelodysplastic syndrome (AML 19). Available from: http://www.cancerresearchuk.org/about-cancer/find-a-clinical-trial/a-trial-looking-at-different-treatments-for-acute-myeloid-leukaemia-and-high-risk-myelodysplastic#undefined (accessed 23 August 2017).
92. Cancer Research UK: A trial looking at treatment for older people with acute myeloid leukaemia and high risk myelodysplastic syndrome (AML 18). Available from: http://www.cancerresearchuk.org/about-cancer/find-a-clinical-trial/a-trial-looking-at-treatment-for-older-people-with-acute-myeloid-leukaemia-and-high-risk-myelodysplastic-syndrome-aml18 (accessed 5 June 2017).
93. Cardiff University Center for Clinical trials – AML18. Available from: http://www.cardiff.ac.uk/centre-for-trials-research/research/studies-and-trials/view/aml18 (accessed 12 August 2017).
94. Trial to test the effects of adding 1 of 2 new treatment agents to commonly used Chemotherapy Combinations (AML18). Available from: https://clinicaltrials.gov/ct2/show/NCT02272478 (accessed 12 August 2017).
95. Castaigne S, Pautas C, Terré C *et al.* Final Analysis of the ALFA 0701 Study [platform presentation]. Presented at the American Society of Hematology 56th Annual Meeting and Exposition, 6–9 December 2014, San Francisco, CA, USA
96. Cheson BD, Bennett JM, Kopecky KJ *et al.* Revised recommendations of the International Working Group for Diagnosis, Standardization of Response Criteria, Treatment Outcomes, and Reporting Standards for Therapeutic Trials in Acute Myeloid Leukemia. *J Clin Oncol* 2003;21:4642-9.
97. Castaigne S, Pautas C, Terre C *et al.* Fractionated Doses of Gemtuzumab Ozogamicin (GO) Combined to Standard Chemotherapy (CT) Improve Event-Free and Overall Survival in Newly-Diagnosed De Novo AML Patients Aged 50–70 Years Old: A Prospective Randomized Phase 3 Trial From the Acute Leukemia French Association (ALFA) [abstract]. Presented at the 53rd American Society of Hematology Annual Meeting, 10–13 December 2011, San Diego, Ca, USA.
98. Pfizer, Event-free survival and relapse-free survival, response rate data (30 April 2013 data cut-off) in the subpopulation of patients with favourable/intermediate cytogenetics (IRC analysis).
99. Pfizer, Treatment-emergent adverse events in patients with favourable/intermedated cytogenetics risk profile.
100. Ohtake S, Miyawaki S, Kiyoi H *et al.* Randomized trial of response-oriented individualized versus fixed-schedule induction chemotherapy with idarubicin and cytarabine in adult acute myeloid leukemia: the JALSG AML95 study. *Int J Hematol* 2010;91:276-83.
101. Farhat H, Reman O, Raffoux E *et al.* Fractionated doses of gemtuzumab ozogamicin with escalated doses of daunorubicin and cytarabine as first acute myeloid leukemia salvage in patients aged 50-70-year old: a phase 1/2 study of the acute leukemia French association. *Am J Hematol* 2012;87:62-5.

102. Obinutuzumab in combination with chlorambucil for untreated chronic lymphocytic leukaemia (NICE TA343). Available from: https://www.nice.org.uk/guidance/ta343 (accessed 20 September 2017).
103. Tikhonova IA, Hoyle MW, Snowsill TM *et al.* Azacitidine for Treating Acute Myeloid Leukaemia with More Than 30 % Bone Marrow Blasts: An Evidence Review Group Perspective of a National Institute for Health and Care Excellence Single Technology Appraisal. *Pharmacoeconomics* 2017;35:363-73.
104. Edlin R, et al. Azacitidine for the treatment of myelodysplastic syndromes, chronic myelomonocytic leukaemia and acute myeloid leukaemia. Health Technology Assessment Database. 2011(4). Available at, http://onlinelibrary.wiley.com/o/cochrane/clhta/articles/HTA-32011000499/frame.html. (Accessed, 10 October 2017).
105. Kurosawa S, Yamaguchi T, Miyawaki S *et al.* A Markov decision analysis of allogeneic hematopoietic cell transplantation versus chemotherapy in patients with acute myeloid leukemia in first remission. *Blood* 2011;117:2113-20.
106. Leunis A, Redekop WK, van Montfort KA, Lowenberg B, Uyl-de Groot CA. The development and validation of a decision-analytic model representing the full disease course of acute myeloid leukemia. *Pharmacoeconomics* 2013;31:605-21.
107. Wang HI, Aas E, Howell D *et al.* Long-term medical costs and life expectancy of acute myeloid leukemia: a probabilistic decision model. *Value Health* 2014;17:205-14.
108. Castejón N, Cappelleri JC, Cuervo J *et al.* Estimating health state utilities for patients with acute myeloid leukemia. Poster presented at the International Society for Pharmacoeconomics and outcomes research (ISPOR) 22nd Annual International Meeting. May 20–24, 2017, Boston, MA, USA.
109. Burnett AK, Goldstone A, Hills RK *et al.* Curability of patients with acute myeloid leukemia who did not undergo transplantation in first remission. *J Clin Oncol* 2013;31:1293-301.
110. National Institute for Health and Care Excellence. DSU TSD 14: Survival analysis for economic evaluations alongside clinical trials - extrapolation with patient level data (June 2011). Available at, http://scharr.dept.shef.ac.uk/nicedsu/wp-content/uploads/sites/7/2016/03/NICE-DSU-TSD-Survival-analysis.updated-March-2013.v2.pdf (Accessed 13 October 2017).
111. Achcar J, A. Coelho-Barros E, Mazucheli J. Cure fraction models using mixture and non-mixture models.2012.
112. Othus M, Barlogie B, Leblanc ML, Crowley JJ. Cure models as a useful statistical tool for analyzing survival. *Clin Cancer Res* 2012;18:3731-6.
113. Jia X, Sima CS, Brennan MF, Panageas KS. Cure models for the analysis of time-to-event data in cancer studies. *J Surg Oncol* 2013;108:342-7.
114. Martinez EZ, Achcar JA, Jacome AA, Santos JS. Mixture and non-mixture cure fraction models based on the generalized modified Weibull distribution with an application to gastric cancer data. *Comput Methods Programs Biomed* 2013;112:343-55.
115. Francisci S, Capocaccia R, Grande E *et al.* The cure of cancer: a European perspective. *Eur J Cancer* 2009;45:1067-79.

116. Peng Y, Dear KB, Denham JW. A generalized F mixture model for cure rate estimation. *Stat Med* 1998;17:813-30.
117. Gardin C, Chevret S, Pautas C *et al.* Superior long-term outcome with idarubicin compared with high-dose daunorubicin in patients with acute myeloid leukemia age 50 years and older. *J Clin Oncol* 2013;31:321-7.
118. Shah A, Andersson TM, Rachet B, Bjorkholm M, Lambert PC. Survival and cure of acute myeloid leukaemia in England, 1971-2006: a population-based study. *Br J Haematol* 2013;162:509-16.
119. Royston P, Parmar MK. Flexible parametric proportional-hazards and proportional-odds models for censored survival data, with application to prognostic modelling and estimation of treatment effects. *Stat Med* 2002;21:2175-97.
120. Lambert PC. Modeling of the cure fraction in survival studies. Available from, http://www.stata-journal.com/sjpdf.html?articlenum=st0131 (Accessed, 31 October 2017). *The Stata Journal* 2007;7:351-75.
121. Pfizer. Data on File. 2017.
122. Shimoni A, Labopin M, Savani B *et al.* Long-term survival and late events after allogeneic stem cell transplantation from HLA-matched siblings for acute myeloid leukemia with myeloablative compared to reduced-intensity conditioning: a report on behalf of the acute leukemia working party of European group for blood and marrow transplantation. *J Hematol Oncol* 2016;9:118.
123. Armenian SH, Sun CL, Vase T *et al.* Cardiovascular risk factors in hematopoietic cell transplantation survivors: role in development of subsequent cardiovascular disease. *Blood* 2012;120:4505-12.
124. Sun CL, Francisco L, Kawashima T *et al.* Prevalence and predictors of chronic health conditions after hematopoietic cell transplantation: a report from the Bone Marrow Transplant Survivor Study. *Blood* 2010;116:3129-39; quiz 377.
125. Cheng MJ, Hourigan CS, Smith TJ. Adult Acute Myeloid Leukemia Long-term Survivors. *J Leuk (Los Angel)* 2014;2.
126. Wingard JR, Majhail NS, Brazauskas R *et al.* Long-term survival and late deaths after allogeneic hematopoietic cell transplantation. *J Clin Oncol* 2011;29:2230-9.
127. Nivison-Smith I, Simpson JM, Dodds AJ *et al.* Relative survival of long-term hematopoietic cell transplant recipients approaches general population rates. *Biol Blood Marrow Transplant* 2009;15:1323-30.
128. Vanderwalde AM, Sun CL, Laddaran L *et al.* Conditional survival and cause-specific mortality after autologous hematopoietic cell transplantation for hematological malignancies. *Leukemia* 2013;27:1139-45.
129. Bhatia S, Robison LL, Francisco L *et al.* Late mortality in survivors of autologous hematopoietic-cell transplantation: report from the Bone Marrow Transplant Survivor Study. *Blood* 2005;105:4215-22.
130. Martin PJ, Counts GW, Jr., Appelbaum FR *et al.* Life expectancy in patients surviving more than 5 years after hematopoietic cell transplantation. *J Clin Oncol* 2010;28:1011-6.
131. Pond GR, Lipton JH, Messner HA. Long-term survival after blood and marrow transplantation: comparison with an age- and gender-matched normative population. *Biol Blood Marrow Transplant* 2006;12:422-9.

132. Battipaglia G, Labopin M, Candoni A *et al.* Risk of sinusoidal obstruction syndrome in allogeneic stem cell transplantation after prior gemtuzumab ozogamicin treatment: a retrospective study from the Acute Leukemia Working Party of the EBMT. *Bone Marrow Transplant* 2017;52:592-99.
133. NHS England. Clinical Commissioning Policy. Treatments for Graft versus Host Disease (GvHD) following Haematopoietic Stem Cell Transplantation. March 2017. Available at, https://www.england.nhs.uk/wp-content/uploads/2017/03/gvhd-heamatopoietic-stem-cell.pdf. (Accessed, 31 October 2017).
134. Kurosawa S, Yamaguchi T, Mori T *et al.* Patient-reported quality of life after allogeneic hematopoietic cell transplantation or chemotherapy for acute leukemia. *Bone Marrow Transplant* 2015;50:1241-9.
135. Kurosawa S, Yamaguchi H, Yamaguchi T *et al.* Decision analysis of allogeneic hematopoietic stem cell transplantation versus chemotherapy in cytogenetically standard-risk acute myeloid leukemia in first complete remission: The impact of FLT3-ITD profile. . *Blood* 2014;124.
136. Kurosawa S, Yamaguchi H, Yamaguchi T *et al.* Decision Analysis of Postremission Therapy in Cytogenetically Intermediate-Risk Acute Myeloid Leukemia: The Impact of FLT3 Internal Tandem Duplication, Nucleophosmin, and CCAAT/Enhancer Binding Protein Alpha. *Biol Blood Marrow Transplant* 2016;22:1125-32.
137. Cressman S, Savoie ML, Couban S *et al.* Health utility during the first two years of treatment of hematological malignancies. *Blood* 2016;128.
138. Watson M, Buck G, Wheatley K *et al.* Adverse impact of bone marrow transplantation on quality of life in acute myeloid leukaemia patients; analysis of the UK Medical Research Council AML 10 Trial. *Eur J Cancer* 2004;40:971-8.
139. Pfizer. Estimating the utility of health states associated with acute myeloid leukaemia (prepared by Laser Analytica). Data on file. 2016.
140. Scottish Medicines Consortium. defibrotide (Defitelio®), 9th June 2014. Available at, http://www.scottishmedicines.org.uk/SMC_Advice/Advice/967_14_defibrotide_Defitelio/defibrotide_Defitelio (Accessed, 5th Dec 2017).
141. National Institute for Health and Care Excellence. Inotuzumab ozogamicin for treating relapsed or refractory B-cell acute lymphoblastic leukaemia [ID893]. Committee Papers. Available at, https://www.nice.org.uk/guidance/gid-ta10091/documents/committee-papers-2 (Accessed 28 Nov 2017).
142. McKenzie L, van der Pol M. Mapping the EORTC QLQ C-30 onto the EQ-5D instrument: the potential to estimate QALYs without generic preference data. *Value Health* 2009;12:167-71.
143. Ara R, Brazier JE. Populating an economic model with health state utility values: moving toward better practice. *Value Health* 2010;13:509-18.
144. Wang H, Aas E, Roman E, Smith A. Comparison Of Different Costing Methods. *Value Health* 2015;18:A687.
145. Wang H, Aas E, Roman E, Smith A. Method Comparison Of Censoring Cost Analyses. *Value Health* 2015;18:A687.
146. British National Formulary. Available at, https://www.bnf.org/ (Accessed, 5th Nov 2017).

147. Department of Health. Drugs and pharmaceutical electronic market information (eMit). Updated 20 June 2017. Available at, https://www.gov.uk/government/publications/drugs-and-pharmaceutical-electronic-market-information-emit (Accessed, 17 Ocotber 20170.
148. Addicott R, Dewar S. Improving choice at end of life. King's Fund. 2008. Available at: https://www.kingsfund.org.uk/sites/default/files/improving-choice-end-of-life-descriptive-analysis-impact-costs-marie-curie-choice-programme-lincolnshire-rachael-addicot-steve-dewar-april-2008.pdf Accessed 17 October, 2017.
149. NHS Blood and Transplant. NHS Blood and Transplant Annual Report and Accounts 2014/15. HC162. Available at, https://www.gov.uk/government/uploads/system/uploads/attachment_data/file/443179/NHSBT_report_2014-15.pdf (Accessed 5th Nov 2017).
150. Esperou H, Brunot A, Roudot-Thoraval F *et al.* Predicting the costs of allogeneic sibling stem-cell transplantation: results from a prospective, multicenter, French study. *Transplantation* 2004;77:1854-8.
151. Department of Health. NHS reference costs (updated 15 December, 2016). Available at, https://www.gov.uk/government/collections/nhs-reference-costs (Accessed, 17 October, 2017).
152. Personal Social Services Research. Available at, http://www.pssru.ac.uk/. 2016.
153. NHS England. Clinical Commissioning Policy. Use of defibrotide in severe veno-occlusive disease following stem cell transplant. NHS England B04/P/c. January 2015. Available at: https://www.england.nhs.uk/commissioning/wp-content/uploads/sites/12/2015/01/b04-use-defibrotide.pdf. (Accessed August 23, 2017).
154. National Institute for Health and Care Excellence. Gemtuzumab ozogamicin for untreated acute myeloid leukaemia [ID982], Final scope (September 2017). Available at, https://www.nice.org.uk/guidance/gid-ta10142/documents/final-scope (Accessed, 5th Dec 2017).
155. Eddy DM, Hollingworth W, Caro JJ *et al.* Model transparency and validation: a report of the ISPOR-SMDM Modeling Good Research Practices Task Force--7. *Value Health* 2012;15:843-50.

B.5 Appendices

Appendix C: SmPC and EPAR

C1.1 SmPC

C1.2 EPAR

Appendix D: Identification, selection and synthesis of clinical evidence

D.1 Identification and selection of relevant studies (SLR)

D.1.1 Identification and selection of relevant RCT clinical evidence

D.1.2 Identification and selection of relevant non-RCT clinical evidence

D.2 Supporting materials for ALFA-0701 study 76

D.2.1 Participant flow in ALFA-0701

D.2.2 Patient characteristics in ALFA-0701

D.2.3 Efficacy endpoint data in ALFA-0701 at 1 August 2011 data cut-off

D.2.4 Efficacy endpoint data in ALFA-0701 by investigator assessment

D.3. Meta-analyses

D.3.1 Individual patient data meta-analysis

D.3.2 Meta-analyses identified in the SLR

D.3.3 Summary of efficacy results across all five meta-analyses

Appendix E: Subgroup analysis

E.1 Overall survival

E.2 Event-free survival

E.2.1 IRC analysis

NICE National Institute for Health and Care Excellence

Level 1A
City Tower
Manchester
M1 4BT
United Kingdom

+44 (0)300 323 0140

Single technology appraisal

Gemtuzumab ozogamicin for untreated acute myeloid leukaemia

Dear Company,

The Evidence Review Group, CRD/CHE University of York, and the technical team at NICE have looked at the submission received on 8 December 2017 from Pfizer. In general they felt that it is well presented and clear. However, the ERG and the NICE technical team would like further clarification on the clinical and cost effectiveness data (see questions listed at end of letter).

The ERG and the technical team at NICE will be addressing these issues in their reports.

Please provide your written response to the clarification questions by **5pm** on **26 February 2018**. Your response and any supporting documents should be uploaded to NICE Docs/Appraisals

Two versions of your written response should be submitted; one with academic/commercial-in-confidence information clearly marked and one with this information removed.

Please underline all confidential information, and separately highlight information that is submitted as commercial in confidence in turquoise, and all information submitted as academic in confidence in yellow.

If you present data that are not already referenced in the main body of your submission and that are academic/commercial in confidence, please complete the attached checklist for confidential information.

Please do not embed documents (PDFs or spreadsheets) in your response because this may result in them being lost or unreadable.

If you have any queries on the technical issues raised in this letter, please contact Kirsty Pitt, Technical Lead (kirsty.pitt@nice.org.uk). Any procedural questions should be addressed to Stephanie Callaghan, Project Manager (stephanie.callaghan@nice.org.uk).

Yours sincerely

Nicola Hay
Technical Adviser – Appraisals
Centre for Health Technology Evaluation
On behalf of:

NICE National Institute for Health and Care Excellence

Level 1A
City Tower
Manchester
M1 4BT
United Kingdom

+44 (0)300 323 0140

Frances Sutcliffe
Associate Director – Appraisals
Centre for Health Technology Evaluation

Encl. checklist for confidential information

Section A: Clarification on effectiveness data

Further reports required

A1. Please provide the full pack of references (including the appendix to the Castaigne 2012 article).

ALFA-0701 trial

A2. Please provide details of the criteria used for categorising patients into different cytogenetic profile groups (by Centre Hospitalier de Versailles, as presented in table 71, appendices). In addition, please explain the 'not available' category, i.e. were these patients not tested/were results inadequate?

A3. In figure 23 (page 76 of appendices) it states that [redacted] Please specify the reasons why patients did not receive GO.

A4. Table 11 (company submission) describes imputation for missing data. Please specify how many patients had (a) the day of the month and (b) the full date (i.e. day and month) imputed.

A5. Please present separate adverse event results for the 30 patients in the daunorubicin and cytarabine (DA) group who received GO as follow-up therapy, i.e. present tables 24 – 26 (company submission) separately for the GO + DA group, the DA group who did not subsequently receive GO and the DA group who subsequently received GO.

Subgroup analyses

A6. For event-free survival (EFS), most subgroup analyses are for the 01/08/11 data cut-off, rather than the 30/04/13 data cut-off (figure 31, appendices). On page 119 of the appendices, it states that results of the independent review committee (IRC) subgroup analysis for EFS at the 30/04/13 data cut-off were consistent with those of the investigator assessment at 01/08/11 data cut-off. Please provide the forest plot

NICE National Institute for Health and Care Excellence

Level 1A
City Tower
Manchester
M1 4BT
United Kingdom

+44 (0)300 323 0140

for the 30/04/13 data cut-off for EFS (as presented for overall survival (OS) and relapse-free survival (RFS)).

A7. White blood cell count is listed as a pre-planned subgroup on page 46 of the submission, but the results from this subgroup are not presented in the submission. Please present the OS, EFS, RFS and response rate results for the subgroup analysis based on white blood cell count (modified intention-to-treat population (mITT), 30/04/13 data cut-off).

Section B: Clarification on cost-effectiveness data

Survival analysis

B1. Please provide further justification for using the independent review committee rather than investigator assessments for the response rate, EFS and RFS analyses.

B2. **Priority question**: The cost-effectiveness model uses EFS, RFS and OS analyses based on a reference date of 30 April 2013. However, the clinical study report (CSR) (p134) reports an updated number of deaths at the retrospective data collection cut-off date of 01 Nov 2013.

(i) Please confirm which endpoints were assessed until 01 November 2013 and clarify why this additional period was not included in the analyses informing the model.

(ii) Please replicate table 25 and figure 7 in the CSR including deaths up to 01 November 2013.

(iii) Please report the revised statistical cure rates for OS based on the mixture cure model (MCM) lognormal, Weibull and generalised gamma functions.

(iv) Please provide an additional sensitivity analysis for the cost-effectiveness results which includes the updated number of deaths.

B3. **Priority question**: Please provide further justification for including patients with unknown cytogenetic results within the population considered in the cost-effectiveness analysis. Hills et al (2014) stated that “as an individual group we could not see them benefiting from GO” (HR for OS = 1.01 [0.84-1.22]). Please provide an additional sensitivity analysis for the cost-effectiveness results which excludes these patients.

B4. The ERG would like to further consider the potential impact of patient heterogeneity in the cost-effectiveness analysis. Hills et al (2014) also reported significant differences in the survival benefits between patients with favourable and intermediate risk disease. Please provide an additional sensitivity analysis for the cost-

NICE National Institute for Health and Care Excellence

Level 1A
City Tower
Manchester
M1 4BT
United Kingdom

+44 (0)300 323 0140

effectiveness analysis for the population at intermediate risk only (i.e. excluding favourable, unfavourable and unknown cytogenetic status).

B5. **Priority question**: Please replicate table 13 (p61) in the main submission for the following cytogenetic subgroups:

(i) Favourable/intermediate + unknown

(ii) Favourable only

(iii) Intermediate only

(iv) Unknown only

B6. **Priority question**: The ALFA-0701 trial specified that patients with favourable and intermediate-1 cytogenetic and molecular risk categories were not considered for stem cell transplant in first complete remission (CR), whereas patients with intermediate-2 or unfavourable cytogenetic and molecular risk categories who experienced a CR were considered for transplant. The submission also states that current European Society for Medical Oncology (ESMO)/European Leukemia Net (ELN) guidance recommends that haematopoietic stem cell transplant (HSCT) should be considered in all patients with intermediate risk disease.

The ERG considers that it is important to explore additional patient subgroups defined according to their potential eligibility to receive HSCT in the ALFA-0701 trial. To help assess the generalisability of the trial results in light of current guidelines for HSCT and UK clinical practice, please replicate table 13 (p61) in the main submission for the following subgroups:

(i) Favourable and intermediate-1

(ii) Intermediate-2 and unfavourable

(iii) Intermediate-1 only

(iv) Intermediate-2 only

B7. Please provide the Kaplan-Meier curves (with the number of patients at risk at each time point) for EFS, RFS and OS for the subgroups specified in questions B5 and B6.

B8. Please provide the separate technical report for the survival analysis methods (cited on p288 of the appendices).

NICE National Institute for Health and Care Excellence

Level 1A
City Tower
Manchester
M1 4BT
United Kingdom

+44 (0)300 323 0140

B9. **Priority question**: Please estimate and present the statistical cure rates (OS[CR] and RFS) for the MCM lognormal, Weibull and generalised gamma functions in the following cytogenetic subgroups:

(i) Favourable/intermediate (excluding unknown)

(ii) Favourable only

(iii) Intermediate only

(iv) Unknown only

B10. **Priority question**: Please provide an Excel file with the coefficients and variance/covariance matrices for the MCM lognormal, Weibull and generalised gamma functions in the subgroups specified in B9.

B11. **Priority question:** Please estimate and present the statistical cure rates (OS[CR] and RFS) for the MCM lognormal, Weibull and generalised gamma functions in the subgroups specified in B6.

B12. **Priority question:** Please provide an Excel file with the coefficient and variance/covariance matrices for the MCM lognormal, Weibull and generalised gamma functions in the subgroups specified in B6.

Haematopoietic stem cell transplantation

B13. Please clarify the following:

(i) Why the probability of HSCT in relapsed patients is modelled at specific time intervals from the point of randomisation as opposed to from the point of relapse.

(ii) What the number of patients at risk of HSCT (relapse patients) represents and why this differs from the total number 'N' reported in the annual breakdown (sheet pHSCT, economic model)

(iii) Whether data on HSCTs were collected after the original 2011 cut-off.

B14. **Priority question**: Please provide the Kaplan-Meier curves (with the number of patients at risk at each time point) for the time to HSCT from relapse (CR patients only) for each of the subgroups specified in B5.

B15. Please complete the following table on HSCT for the subgroups specified in B5 and B6.

NICE National Institute for Health and Care Excellence

Level 1A
City Tower
Manchester
M1 4BT
United Kingdom

+44 (0)300 323 0140

	GO + Daunorubicin + Cytarabine (N)	Daunorubicin + Cytarabine (N)
Patients receiving transplant	N (%)	N (%)
Status at transplant relative to EFS event *Before relapse* *After relapse* *After induction failure*	 N (%) N (%) N (%)	 N (%) N (%) N (%)

B16. Please provide the Kaplan-Meier curves (with the number of patients at risk at each time point) for the time to HSCT from relapse (as opposed to from randomisation) for each of the subgroups specified in B6.

B17. Please replicate figure 17 (p121 in the main submission) according to patients' status at transplant relative to EFS event: (i) before relapse (ii) after relapse; (iii) after induction failure.

Population

B18. Please provide the baseline characteristics (i.e. replicating the information in table 71 in the company submission appendices) of the following cytogenetic subgroups:

(i) Favourable only

(ii) Intermediate only

(iii) Unknown only

Modelling approach

B19. Please provide a clearer description of the assumptions underlying the economic model structure and how the patient distribution across health states is estimated, making reference to how survival data and rates of HSCT are used in within the model.

B20. Please explain the advantages of the proposed modelling approach compared with that of a more conventional partitioned survival model, making reference to the structural assumptions described in question B19.

Resource use

B21. Please provide additional information on the actual induction and consolidation treatment received by patients in the trial.

NICE National Institute for Health and Care Excellence

Level 1A
City Tower
Manchester
M1 4BT
United Kingdom

+44 (0)300 323 0140

(i) Please present the proportion of patients receiving induction course 1, induction course 2, consolidation course 1 and consolidation course 2 in the base-case population in the model.

(ii) Provide a description of how these figures were estimated in relation to the data provided in table 33 of the CSR.

(iii) In addition, please provide the number of each induction and consolidation course for the subgroups described in question B18.

B22. Please justify why assumptions were made in the model about the length of hospitalisation for patients during the initial treatment periods (induction and consolidation) rather than using resource use data collected in the trial itself (see table 43, CSR).

B23. Please replicate table 43 in the CSR for the subgroups specified in B5.

Section C: Textual clarifications and additional points

C1. Table 7 (company submission) states that outcomes were assessed using definitions provided in table 4. Should the table referred to be table 8?

C2. In the Castaigne 2012 article it states that 28 patients (20%) had unfavourable cytogenetics, rather than 27 patients (19.4%), as stated in table 71 (page 77 of appendices). Please confirm which figure is correct.

C3. Figure 22 in the appendices states that 24 publications were excluded with the reason 'Combinations of GO irrelevant to the UK setting'. However in the list below the figure (page 63 of the appendices, 'Complete reference lists for excluded studies: non-RCT evidence'), only 23 references are listed. Please clarify this discrepancy.

C4. On page 69 of the appendices, should the reason for exclusion of the study by Roboz et al. be 'not an intervention/comparator of interest (i.e. no GO at all)'?

C5. Please clarify whether note (a) or (b) under table 15 of the company submission presents the correct information.

NICE National Institute for Health and Care Excellence

Level 1A
City Tower
Manchester
M1 4BT
United Kingdom

+44 (0)300 323 0140

Single technology appraisal

Gemtuzumab ozogamicin for untreated acute myeloid leukaemia

Monday 26th of February 2018

Company response to ERG clarification questions (received 12th of February 2018)

Dear Dr Sutcliffe,

Thank you for the clarification questions and opportunity to provide further detail to Pfizer's evidence submission for gemtuzumab ozogamicin for untreated acute myeloid leukaemia.

Pfizer would like to highlight that it has been a challenge to address all the questions within the given time frame because the majority of the questions received asked for further subgroup analysis of the ALFA-0701 clinical data set which needed programming by our statistical team. Therefore, as agreed, responses to question B2 (ii), B7, B21 (iii); B23 are not included in this version but will be sent to NICE separately by 6th March 2018.

Please find the remainder of Pfizer's responses to the questions below. Supporting documents include:

- Three PDF documents relating to data requested in questions A1, A7 and B18.
- One Word document relating to data requested in B8
- Two Excel files relating to data requested in questions B10 and B12.

Sincerely,
[redacted]

NICE National Institute for Health and Care Excellence

Level 1A
City Tower
Manchester
M1 4BT
United Kingdom

+44 (0)300 323 0140

Section A: Clarification on effectiveness data

Further reports required

A1. Please provide the full pack of references (including the appendix to the Castaigne 2012 article).

ALFA-0701 trial

A2. Please provide details of the criteria used for categorising patients into different cytogenetic profile groups (by Centre Hospitalier de Versailles, as presented in table 71, appendices). In addition, please explain the 'not available' category, i.e. were these patients not tested/were results inadequate?

The criteria used by Centre Hospitalier de Versailles (CHV) for categorising patients into different cytogenetic profile groups is described on pg. 43 of the ALFA-0701 CSR and is as follows:

- Favourable cytogenetics include inv(16)/t(16;16), t(8;21), and t(15;17).
- Unfavourable cytogenetics include monosomy 5 or del(5q), monosomy 7 or del(7q), t(6;9), t(9;22), 3q26 abnormalities except t(3;5), 11q23 abnormalities except t(9;11), and complex karyotypes with 3 abnormalities or more.
- Intermediate cytogenetics include all other anomalies as well as normal karyotypes. Karyotype was classified as normal when at least 20 mitoses without chromosomal anomalies were observed in bone marrow samples.

It is important to note that subgroup analyses performed for the primary endpoint ,EFS, using NCCN risk classification and ELN risk classification were consistent with results derived as per CHV classification (see Figure 1 below).

With regard to the "not available" category, this included patients for whom the karyotype was not done, or done but inadequate.

A3. In figure 23 (page 76 of appendices) it states that [redacted]. Please specify the reasons why patients did not receive GO.

Reasons for not receiving GO during the induction course

For the induction course the reasons for patients not receiving GO are listed in the footnote for Figure 1 in the ALFA-0701 CSR (pg. 62); described as follows:
[redacted]

NICE National Institute for Health and Care Excellence

Level 1A
City Tower
Manchester
M1 4BT
United Kingdom

+44 (0)300 323 0140

Reasons for not receiving GO during consolidation courses

[redacted]

Of the 6 patients not receiving GO during consolidation 1 (C1) [redacted] (Pfizer data on file, 2017).

A4. Table 11 (company submission) describes imputation for missing data. Please specify how many patients had (a) the day of the month and (b) the full date (i.e. day and month) imputed.

As summarised in Table 11 of the company submission the Statistical Analysis Plan provided a plan for imputation of partially missing dates that were to be used in derivation of time-to-event endpoints. However, there were no partially missing dates for the date of induction failure, relapse date, or death so no imputation was required.

A5. Please present separate adverse event results for the 30 patients in the daunorubicin and cytarabine (DA) group who received GO as follow-up therapy, i.e. present tables 24 – 26 (company submission) separately for the GO + DA group, the DA group who did not subsequently receive GO and the DA group who subsequently received GO.

The requested data were not collected and so these analyses are not possible. This is because the safety reporting period in ALFA-0701 was from screening up to 28 days after the last dose of study drug in each treatment arm, except for data on veno-occlusive disease (VOD). Cross-over was not permitted in the ALFA-0701 study therefore patients in the DA arm who subsequently received GO as follow-up therapy did not receive GO as part of the ALFA-0701 study and therefore AEs were not collected for this group.

Data on VOD was collected until the patient's death or the retrospective data collection cut-off of 01 November 2013, whichever occurred first, to identify any late study drug toxicity associated with VOD (see pg.47 ALFA-0701 CSR). As reported in the company submission (pg. 83) the [redacted]

Subgroup analyses

A6. For event-free survival (EFS), most subgroup analyses are for the 01/08/11 data cut-off, rather than the 30/04/13 data cut-off (figure 31, appendices). On page 119 of the appendices, it states that results of the independent review committee (IRC) subgroup analysis for EFS at the 30/04/13 data cut-off were consistent with those of the investigator assessment at 01/08/11 data cut-off. Please provide the forest plot for the 30/04/13 data cut-off for EFS (as presented for overall survival (OS) and relapse-free survival (RFS)).

NICE National Institute for Health and Care Excellence

Level 1A
City Tower
Manchester
M1 4BT
United Kingdom

+44 (0)300 323 0140

Subgroup analyses of EFS based on events determined by the independent review using the 30 April 2013 censoring date are shown in Figure 1 below.

Figure 1 Forest Plot of Event-Free Survival per Independent Review – 30 April 2013 (mITT Population)

A7. White blood cell count is listed as a pre-planned subgroup on page 46 of the submission, but the results from this subgroup are not presented in the submission. Please present the OS, EFS, RFS and response rate results for the subgroup analysis based on white blood cell count (modified intention-to-treat population (mITT), 30/04/13 data cut-off).

In the statistical analysis plan, white blood cell count (<30 and >=30 per 109/L) was only pre-specified as a subgroup for response rate (pg. 32 ALFA-0701 statistical analysis plan). Please see PDF sent with responses for the data[redacted]

Section B: Clarification on cost-effectiveness data

Survival analysis

B1. Please provide further justification for using the independent review committee rather than investigator assessments for the response rate, EFS and RFS analyses.

In ALFA-0701 the primary study endpoint EFS was derived from investigator assessment (IA) with the retrospective efficacy data reviewed by a blinded independent review committee (BIRC). This was to address any possible bias introduced by the local investigators since ALFA-0701 was an open-label study. It is very common in open-label oncology trials to have a BIRC assessing outcomes such as response and relapse and for these to be the primary analyses for decision making. [redacted] The efficacy results from the BIRC review did support the IA results (summarised in Table 13 in the company submission). Further, the BIRC data were slightly more conservative.

NICE National Institute for Health and Care Excellence

Level 1A
City Tower
Manchester
M1 4BT
United Kingdom

+44 (0)300 323 0140

B2. **Priority question**: The cost-effectiveness model uses EFS, RFS and OS analyses based on a reference date of 30 April 2013. However, the clinical study report (CSR) (p134) reports an updated number of deaths at the retrospective data collection cut-off date of 01 Nov 2013.

(i) Please confirm which endpoints were assessed until 01 November 2013 and clarify why this additional period was not included in the analyses informing the model.

The reference date of 30 April 2013 is correct for EFS, RFS and OS as long term follow-up of patients beyond this date did not occur in the ALFA-0701 study. As documented on page 1 of the ALFA-0701 CSR the study completion date was 30 April 2013. As a result deaths occurring after this reference date were not included in the final OS analyses.

For the retrospective data collection, which uses the 1 Nov 2013 cut-off date, only deaths that were documented in the patient files were collected ([redacted]). However, as the study was already complete at that stage not all patients that had been included in the study were followed up to ensure that the records were up to date and so this dataset cannot be assumed to be complete. If there is any correlation between probability of being followed up and probability of death, any analyses conducted with this OS data would be affected by attrition bias.
It should be noted that the date of the final OS analysis was over 2.5 years after the last patient was enrolled. Pfizer considers that the OS data presented are mature and this has also been accepted by the regulatory agencies.

(ii) Please replicate table 25 and figure 7 in the CSR including deaths up to 01 November 2013.

As agreed during the call on the 15th of February, response to this question will be sent on 6 March 2018.

(iii) Please report the revised statistical cure rates for OS based on the mixture cure model (MCM) lognormal, Weibull and generalised gamma functions.

Pfizer do not consider it appropriate to conduct survival analyses on this dataset given the uncertainties. As stated in section B2 (i) above the five deaths recorded at the 1 Nov 2013 retrospective cut-off data do not represent long term OS follow-up in ALFA-0701 as patients in the study were only followed up until the end of the study which was 30 April 2013.

(iv) Please provide an additional sensitivity analysis for the cost-effectiveness results which includes the updated number of deaths.

Please see response in B2(ii) above.

NICE National Institute for Health and Care Excellence

Level 1A
City Tower
Manchester
M1 4BT
United Kingdom

+44 (0)300 323 0140

B3. **Priority question**: Please provide further justification for including patients with unknown cytogenetic results within the population considered in the cost-effectiveness analysis. Hills et al (2014) stated that "as an individual group we could not see them benefiting from GO" (HR for OS = 1.01 [0.84-1.22]). Please provide an additional sensitivity analysis for the cost-effectiveness results which excludes these patients.

Justification for including patients with unknown cytogenetics in the base-case population can be provided both from a UK clinical practice and a modeling perspective.

According to UK clinical expert opinion less than 10% of patients with de novo AML in the UK present with unknown cytogenetics (in line with the 9.2% included in ALFA-0701). An unknown classification may be a consequence of inadequate specimens or non-dividing cells making cytogenetic risk classification impossible. Depending on the severity of their symptoms these patients may need to be treated immediately rather than waiting for further confirmatory tests therefore it was considered to be appropriate to include these patients in our base-case population.
The mITT population in ALFA-0701 included 271 patients (GO + DA, n = 135; DA, n = 136). Including this "unknown" group in the base-case population was also done to increase the sample size for the analyses undertaken. The modelling approach undertaken broke down the base-case population into smaller subgroups e.g. response to treatment, relapse etc. (see B19) and the statistical modelling undertaken to estimate long term survival required a robust sample size for accurate predictions to be made. Adding this unknown group was also considered to be a conservative approach given the portion of patients with unfavourable cytogenetics within the unknown group would not benefit from GO.

The cost-effectiveness results excluding patients with unknown cytogenetics are reported in the Table 1 below. As per the company's base-case the following models were fit:

- MCM models for RFS and OS(CR) with the best fit statistics (MCM lognormal) and
- Standard parametric for pooled OS refractory with the best fit statistics (exponential).

Table 1 Cost-effectiveness results: favourable and intermediate patients (excluding unknown)

Technologies	Total costs (£)	Total LYG	Total QALYs	Incremental costs (£)	Incremental LYG	Incremental QALYs	ICER incremental (£/QALY)
GO + DA	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
DA	[redacted]	[redacted]	[redacted]				

NICE National Institute for Health and Care Excellence

Level 1A
City Tower
Manchester
M1 4BT
United Kingdom

+44 (0)300 323 0140

Please also note that in the Hills meta-analysis only the ALFA-0701 study included a dosing regimen for GO that is expected to be recommended by EMA. Additionally the patient population included does not directly match (i.e. not all studies included only de novo patients).

B4. The ERG would like to further consider the potential impact of patient heterogeneity in the cost-effectiveness analysis. Hills et al (2014) also reported significant differences in the survival benefits between patients with favourable and intermediate risk disease. Please provide an additional sensitivity analysis for the cost-effectiveness analysis for the population at intermediate risk only (i.e. excluding favourable, unfavourable and unknown cytogenetic status).

It's well established that cytogenetics profile is a powerful prognostic factor for predicting response to treatment and the durability of response in AML. Patients with favourable/intermediate cytogenetics profile have a better prognosis than those with unfavourable cytogenetics profile with regard to treatment response, risk of relapse and survival. Clear efficacy for GO has been demonstrated in both the favourable and intermediate subgroups including in the cited Hills (2014) meta-analysis.

According to UK clinical experts both patients with favourable and intermediate cytogenetics would be treated with GO in the UK given the available supportive evidence. Therefore we consider that the favourable/intermediate subgroup is the most clinically relevant subgroup in UK clinical practice and should be the basis of decision making.

The cost-effectiveness results for patients with intermediate cytogenetics are reported in Table 2 below. As per the company's base-case the following models were fit:

- MCM models for RFS and OS(CR) with the best fit statistics (MCM lognormal) and
- Standard parametric for pooled OS refractory with the best fit statistics (Exponential).

Table 2 Cost-effectiveness results: intermediate patients

Technologies	Total costs (£)	Total LYG	Total QALYs	Incremental costs (£)	Incremental LYG	Incremental QALYs	ICER incremental (£/QALY)
GO + DA	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
DA	[redacted]	[redacted]	[redacted]				

The higher ICER compared with the base-case of the submitted dossier [redacted] vs £12,151) reflects the compounded effect of removing patients with unknown (see B3) and favourable cytogenetics. Removing the latter reduced the estimated statistical cure rates for OS(CR) and RFS in the GO arm (see company's response to B9). Patients with favourable

NICE National Institute for Health and Care Excellence

Level 1A
City Tower
Manchester
M1 4BT
United Kingdom

+44 (0)300 323 0140

cytogenetics are expected to have better outcomes when treated with GO than those with intermediate cytogenetics and therefore this directional change in the ICER holds more clinical rationale but may still also be a consequence of a reduced sample size.

B5. **Priority question**: Please replicate table 13 (p61) in the main submission for the following cytogenetic subgroups:

Please note RFS only includes those that were in CR/CRp therefore CR+CRp; n%, can be derived from those values if required. The patient numbers in RFS for GO+DA and DA alone for each subgroup are reported in the tables below.
Our response to this question focuses on the BIRC data as per the justification given in the response to B1. If upon review of our response the IA data are still required we will send these.

(i) Favourable/intermediate + unknown

Table 3 Summary of efficacy endpoints: Favourable/intermediate + unknown

	GO + DA arm	***DA arm***	***Point***	***p value***
Randomized (n)	[redacted]	[redacted]		
[redacted]				
IRC assessment	[redacted]	[redacted]	[redacted]	[redacted]
[redacted]				
IRC assessment	[redacted]	[redacted]	[redacted]	[redacted]
[redacted]				
IRC assessment	[redacted]	[redacted]	[redacted]	[redacted]

(ii) Favourable only

Table 4 Summary of efficacy endpoints: Favourable only

	GO + DA arm	***DA arm***	***Point***	***p value***
[redacted]	[redacted]	[redacted]		
[redacted]				
[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
[redacted]				
[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
[redacted]				
[redacted]	[redacted]	[redacted]	[redacted]	[redacted]

NICE National Institute for Health and Care Excellence

Level 1A
City Tower
Manchester
M1 4BT
United Kingdom

+44 (0)300 323 0140

(iii) Intermediate only

Table 5 Summary of efficacy endpoints: intermediate only

	GO + DA arm	*DA arm*	*Point*	*p value*
[redacted]				
[redacted]				
[redacted]				
[redacted]				
[redacted]				
[redacted]				
[redacted]				
[redacted]				

(iv) Unknown only

Table 6 Summary of efficacy endpoints: unknown only

	GO + DA arm	*DA arm*	*Point*	*p value*
[redacted]				
[redacted]				
[redacted]				
[redacted]				
[redacted]				
[redacted]				
[redacted]				
[redacted]				

B6. **Priority question**: The ALFA-0701 trial specified that patients with favourable and intermediate-1 cytogenetic and molecular risk categories were not considered for stem cell transplant in first complete remission (CR), whereas patients with intermediate-2 or unfavourable cytogenetic and molecular risk categories who experienced a CR were considered for transplant. The submission also states that current European Society for Medical Oncology (ESMO)/European Leukemia Net (ELN) guidance recommends that haematopoietic stem cell transplant (HSCT) should be considered in all patients with intermediate risk disease.

NICE National Institute for Health and Care Excellence

Level 1A
City Tower
Manchester
M1 4BT
United Kingdom

+44 (0)300 323 0140

The ERG considers that it is important to explore additional patient subgroups defined according to their potential eligibility to receive HSCT in the ALFA-0701 trial. To help assess the generalisability of the trial results in light of current guidelines for HSCT and UK clinical practice, please replicate table 13 (p61) in the main submission for the following subgroups:

According to UK clinical experts HSCT is recommended in patients with unknown cytogenetics if it is feasible. In general, there is no indication for HSCT in those with favourable risk AML. With regards to those with intermediate-risk AML, this is a very heterogeneous group and decisions regarding HSCT are made on an individualised basis and consideration may depend on response to induction therapy therefore in the UK not all of these patients will be considered for HSCT as is stated in EU guidelines.

(i) Favourable and intermediate-1

Table 7 Summary of efficacy endpoints: Favourable and intermediate-1

	GO + DA arm	***DA arm***	***Point***	***p value***
[redacted]				
[redacted]				
[redacted]				
[redacted]				
[redacted]				
[redacted]				
[redacted]				

(ii) Intermediate-2 and unfavourable

Table 8 Summary of efficacy endpoints: Intermediate-2 and unfavourable

	GO + DA arm	***DA arm***	***Point***	***p value***
[redacted]				
[redacted]				
[redacted]				
[redacted]				
[redacted]				
[redacted]				
[redacted]				

(iii) Intermediate-1 only

NICE National Institute for Health and Care Excellence

Level 1A
City Tower
Manchester
M1 4BT
United Kingdom

+44 (0)300 323 0140

Table 9 Summary of efficacy endpoints: intermediate-1 only

	GO + DA arm	DA arm	Point	p value
[redacted]				
[redacted]				
[redacted]				
[redacted]				
[redacted]				
[redacted]				
[redacted]				

(iv) Intermediate-2 only

Table 10 Summary of efficacy endpoints: intermediate-2 only

[redacted]				
[redacted]				
[redacted]				
[redacted]				
[redacted]				
[redacted]				
[redacted]				

B7. Please provide the Kaplan-Meier curves (with the number of patients at risk at each time point) for EFS, RFS and OS for the subgroups specified in questions B5 and B6.

As agreed, Pfizer will not provide a response to this question due to not being able to complete the analyses by the set deadline.

B8. Please provide the separate technical report for the survival analysis methods (cited on p288 of the appendices).

Please see attachment sent with response.

B9. **Priority question**: Please estimate and present the statistical cure rates (OS[CR] and RFS) for the MCM lognormal, Weibull and generalised gamma functions in the following cytogenetic subgroups:

(i) Favourable/intermediate (excluding unknown)

NICE National Institute for Health and Care Excellence

Level 1A
City Tower
Manchester
M1 4BT
United Kingdom

+44 (0)300 323 0140

(ii) Favourable only

(iii) Intermediate only

(iv) Unknown only

[redacted]

The estimated MCM statistical cure rates for both OS [CR] and RFS for the "favourable and intermediate" and "intermediate" subgroups are presented below. The percentage point difference in cure rates between arms is also presented.
For comparison, the estimated statistical cure rates in the base-case MCM lognormal model of the company submission (favourable, intermediate and unknown) were as follows

[redacted]

[redacted]

Favourable and intermediate subgroup results

Please see the estimated cure rates for the favourable and intermediate subgroup in Table 11 below. Fit statistics (AIC/BIC) favoured the lognormal MCM for both RFS and OS in the intermediate and favourable subgroup analysis.

Table 11 Estimated statistical cure rates for the "favourable and intermediate" subgroup (CR only)

Intermediate only subgroup (n=xxx)		GO (n=xxx)		No GO (n=xxx)		Difference in cure rates (GO vs no GO)
		Cure rate	SE	Cure rate	SE	
RFS	Weibull MCM	[redacted]				
	Lognormal MCM					
	Gen. gamma MCM					
OS[CR]	Weibull MCM					
	Lognormal MCM					
	Gen. gamma MCM					

Intermediate subgroup results

Fit statistics (AIC/BIC) favoured the lognormal MCM for both RFS and OS. STATA was unable to fit the generalised gamma functions for RFS because during estimation (MLE) the likelihood function could not converge to produce estimates. This also occurred in some of the analyses for the full mITT population (see Appendix section M.3.2 of the submitted dossier).

NICE National Institute for Health and Care Excellence

Level 1A
City Tower
Manchester
M1 4BT
United Kingdom

+44 (0)300 323 0140

Table 12 Estimated statistical cure rates for the "intermediate" subgroup (CR only)

Intermediate only subgroup (n=xxx)		GO (n=xxx)		No GO (n=xxx)		Difference in cure rates (GO vs no GO)
		Cure rate	SE	Cure rate	SE	
RFS	Weibull MCM	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
	Lognormal MCM	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
	Gen. gamma MCM	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
OS[CR]	Weibull MCM	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
	Lognormal MCM	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
	Gen. gamma MCM	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]

B10. **Priority question**: Please provide an Excel file with the coefficients and variance/covariance matrices for the MCM lognormal, Weibull and generalised gamma functions in the subgroups specified in B9.

Please see Excel sheet sent with the response.

B11. **Priority question:** Please estimate and present the statistical cure rates (OS[CR] and RFS) for the MCM lognormal, Weibull and generalised gamma functions in the subgroups specified in B6.

Given the small number of patients in the intermediate-2 only subgroup (n=24 in total; GO+DA arm n=13 and DA arm n=11), it was not possible to fit MCM models. For the remaining subgroups the results are presented below.

Favourable and intermediate-1 subgroup results

Statistical cure rates for the favourable and intermediate 1 subgroup are presented in the table below. The functions with the best statistical fit (AIC/BIC) are in bold. Fit statistics (AIC/BIC) favoured the lognormal model and these are provided in the appendix. STATA was again unable to fit the generalised gamma functions for RFS.

Table 13 Estimated statistical cure rates for the "favourable and intermediate-1" subgroup (CR only)

Intermediate only subgroup (n=xxx)		GO (n=xxx)		No GO (n=xxx)		Difference in cure rates (GO vs no GO)
		Cure rate	SE	Cure rate	SE	
RFS	Weibull MCM	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
	Lognormal MCM	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
	Gen. gamma MCM	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
OS[CR]	Weibull MCM	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
	Lognormal MCM	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
	Gen. gamma MCM	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]

NICE National Institute for Health and Care Excellence

Level 1A
City Tower
Manchester
M1 4BT
United Kingdom

+44 (0)300 323 0140

Intermediate-2 and unfavourable subgroup results

Statistical cure rates for the intermediate-2 and unfavourable subgroup are presented in the table below. The functions with the best statistical fit (AIC/BIC) are in bold. STATA was again unable to fit the generalised gamma functions for OS(CR).

Table 14 Estimated statistical cure rates for the "intermediate 2 and unfavourable" subgroup (CR only)

Intermediate only subgroup (n=xxx)		GO (n=xxx)		No GO (n=xxx)		Difference in cure rates (GO vs no GO)
		Cure rate	SE	Cure rate	SE	
RFS	Weibull MCM	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
	Lognormal MCM	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
	Gen. gamma MCM	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
OS[CR]	Weibull MCM	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
	Lognormal MCM	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
	Gen. gamma MCM	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]

[redacted]

Intermediate 1 only subgroup results

Statistical cure rates for the Intermediate 1 only subgroup are presented in the table below. The functions with the best statistical fit (AIC/BIC) are again in bold. STATA was unable to fit the generalised gamma functions for OS(CR). The sample sizes are again low and so these results should be interpreted with caution.

Table 15 Estimated statistical cure rates for the intermediate-1 subgroup (CR only)

Intermediate only subgroup (n=xxx)		GO (n=xxx)		No GO (n=xxx)		Difference in cure rates (GO vs no GO)
		Cure rate	SE	Cure rate	SE	
RFS	Weibull MCM	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
	Lognormal MCM	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
	Gen. gamma MCM	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
OS[CR]	Weibull MCM	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
	Lognormal MCM	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
	Gen. gamma MCM	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]

B12. **Priority question:** Please provide an Excel file with the coefficient and variance/covariance matrices for the MCM lognormal, Weibull and generalised gamma functions in the subgroups specified in B6.

NICE National Institute for Health and Care Excellence

Level 1A
City Tower
Manchester
M1 4BT
United Kingdom

+44 (0)300 323 0140

See attached Excel document

Haematopoietic stem cell transplantation

B13. Please clarify the following:

(i) Why the probability of HSCT in relapsed patients is modelled at specific time intervals from the point of randomisation as opposed to from the point of relapse.

Time-to-HSCT curves were derived from the ALFA-0701 trial data. Utilisation in a Markov model was discussed during an advisory board meeting held on 21 February 2017 with clinical and health economic advisors. Fitting time-to-HSCT curves from the point of relapse was considered. However, HSCT probabilities would need to be applied for the proportion of relapses that occurred at each cycle in the model over 4 years (i.e. 60 cycles), meaning that each of these (up to) 60 sets of patients would need to be tracked separately. It was decided that having two time-dependent probabilities (time of relapse and time-to-HSCT) would add to the complexity of the model for little additional explanatory power. For example, there may be patients who relapse later in the model who are older/less fit and so may be less likely to receive a HSCT and so applying time-to-HSCT curves (from the time of relapse) at each cycle would also require simplifying assumptions.
Time-to-HSCT curves were fitted from the time of randomisation and annual HSCT probabilities applied for relapse patients. Clinical advisors agreed that using annual HSCT probabilities from the relapse health state was a reasonable approach to capture the differences between treatment arms.

Annual HSCT probabilities are calculated using the number of patients who attained induction success (i.e. the number of patients eligible for relapse) as the denominator in the calculation. For example, the year 1 HSCT probability for the GO arm is calculated as the total HSCTs occurring in year 1 in the GO arm divided by the number of patients with induction success in the GO arm. This ensures that the total number of HSCTs always match the trial in each arm.

The annual HSCT probabilities were originally calculated using the proportion of relapsed patients alive at each HSCT time point (the number at risk) in the time-to-HSCT analyses. The probabilities were in turn applied to the proportion of patients in the relapse health state at each annual time point specified in the model. However, the numbers of HSCTs predicted by the model were not matching the actual number of HSCTs seen in the trial (nor reflecting the difference between treatment arms). This was because of the following reasons:

- aggregating the HSCTs into annual probabilities and multiplying by the proportion of patients in the relapse health state at certain time points was an approximation;

NICE National Institute for Health and Care Excellence

Level 1A
City Tower
Manchester
M1 4BT
United Kingdom

+44 (0)300 323 0140

- different proportions of patients were predicted to be in the relapse health state over time depending on the survival functions selected in the model.

(ii) What the number of patients at risk of HSCT (relapse patients) represents and why this differs from the total number 'N' reported in the annual breakdown (sheet pHSCT, economic model)

The number at risk refers to the number of living relapsed patients eligible for HSCT (i.e., who have not already received HSCT) at each time that a HSCT occurs in the model. The number at risk was originally used to calculate the HSCT probabilities (as described in previous response) but is no longer used in the model. The N refers to the total number of patients who achieved CR or CRp (induction success). This is now used so that the numbers of HSCTs match the trial, as described in the previous response.

(iii) Whether data on HSCTs were collected after the original 2011 cut-off.

Yes, data on HSCT were collected after the original 2011 cut-off as part of the retrospective data collection.

B14. **Priority question**: Please provide the Kaplan-Meier curves (with the number of patients at risk at each time point) for the time to HSCT from relapse (CR patients only) for each of the subgroups specified in B5.

Time to HSCT (TTHSCT) has been calculated and plotted for:

- All remission patients who relapsed, regardless of whether they had a HSCT or not (n=118, GO arm n=58; DA arm n=60).
- Remission patients who relapsed and had a HSCT (n=35, GO arm n=14; DA arm n=21). In this case, the TTHSCT is the actual time-to-SCT (from relapse date until SCT date).

Favourable + Intermediate + Unknown

Figure 2 Time to HSCT from relapse: Favourable + Intermediate + Unknown (HSCT + no HSCT patients)

NICE National Institute for Health and Care Excellence

Level 1A
City Tower
Manchester
M1 4BT
United Kingdom

+44 (0)300 323 0140

Figure 3 Time to HSCT from relapse: Favourable + Intermediate + Unknown (HSCT patients only)

Favourable only

[redacted]

Figure 4 Time to HSCT from relapse: Favourable only (HSCT + no HSCT patients)

Intermediate only

NICE National Institute for Health and Care Excellence

Level 1A
City Tower
Manchester
M1 4BT
United Kingdom

+44 (0)300 323 0140

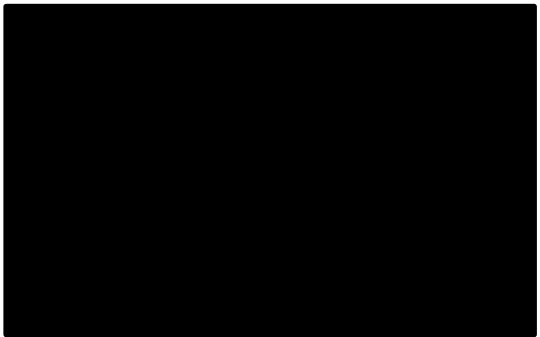
Figure 5 Time to HSCT from relapse: intermediate only (HSCT + no HSCT patients)

Figure 6 Time to HSCT from relapse: intermediate only (HSCT only patients)

Unknown only

Figure 7 Time to HSCT from relapse: intermediate only (HSCT + no HSCT patients)

Figure 8 Time to HSCT from relapse: unknown only (HSCT patients only)



NICE National Institute for Health and Care Excellence

Level 1A
City Tower
Manchester
M1 4BT
United Kingdom

+44 (0)300 323 0140

B15. Please complete the following table on HSCT for the subgroups specified in B5 and B6.

Table 16 Patients undergoing HSCT relative to their EFS event: favourable/Intermediate + unknown

Favourable only	GO	DA
Patients receiving transplant	[redacted]	[redacted]
Status at transplant relative to EFS event *Before relapse* *After relapse* *After induction failure*	[redacted]	[redacted]

Table 17 Patients undergoing HSCT relative to their EFS event: favourable only

Favourable only	GO	DA
Patients receiving transplant	[redacted]	[redacted]
Status at transplant relative to EFS event *Before relapse* *After relapse* *After induction failure*	[redacted]	[redacted]

Table 18 Patients undergoing HSCT relative to their EFS event: intermediate only

Favourable only	GO	DA
Patients receiving transplant	[redacted]	[redacted]
Status at transplant relative to EFS event *Before relapse* *After relapse* *After induction failure*	[redacted]	[redacted]

NICE National Institute for Health and Care Excellence

Level 1A
City Tower
Manchester
M1 4BT
United Kingdom

+44 (0)300 323 0140

Table 19 Patients undergoing HSCT relative to their EFS event: Unknown only

Favourable only	GO	DA
Patients receiving transplant	[redacted]	[redacted]
Status at transplant relative to EFS event *Before relapse* *After relapse* *After induction failure*	[redacted]	[redacted]

Table 20 Patients undergoing HSCT relative to their EFS event: favourable and intermediate I

Favourable only	GO	DA
Patients receiving transplant	[redacted]	[redacted]
Status at transplant relative to EFS event *Before relapse* *After relapse* *After induction failure*	[redacted]	[redacted]

Table 21 Patients undergoing HSCT relative to their EFS event: Intermediate II and Unfavourable

Favourable only	GO	DA
Patients receiving transplant	[redacted]	[redacted]
Status at transplant relative to EFS event *Before relapse* *After relapse* *After induction failure*	[redacted]	[redacted]

Table 22 Patients undergoing HSCT relative to their EFS event: Intermediate I only

Favourable only	GO	DA
Patients receiving transplant	[redacted]	[redacted]
Status at transplant relative to EFS event *Before relapse* *After relapse* *After induction failure*	[redacted]	[redacted]

NICE National Institute for Health and Care Excellence

Level 1A
City Tower
Manchester
M1 4BT
United Kingdom

+44 (0)300 323 0140

Table 23 Patients undergoing HSCT relative to their EFS event: Intermediate II only

Favourable only	GO	DA
Patients receiving transplant	[redacted]	[redacted]
Status at transplant relative to EFS event *Before relapse* *After relapse* *After induction failure*	[redacted]	[redacted]

B16. Please provide the Kaplan-Meier curves (with the number of patients at risk at each time point) for the time to HSCT from relapse (as opposed to from randomisation) for each of the subgroups specified in B6.

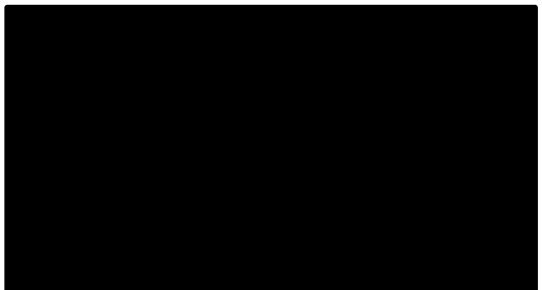
Time to HSCT (TTHSCT) has been calculated and plotted for:

- All remission patients who relapsed, regardless of whether they had a HSCT or not (n=118, GO arm n=58; DA arm n=60).
- Remission patients who relapsed and had a HSCT (n=35, GO arm n=14; DA arm n=21). In this case, the TTHSCT is the actual time-to-SCT (from relapse date until SCT date).

Favourable + Intermediate I

Figure 9 Time to HSCT from relapse: Favourable + Intermediate-1 only (HSCT + no HSCT patients)

Figure 10 Time to HSCT from relapse: Favourable + Intermediate-1 only (HSCT only patients)



Level 1A
City Tower
Manchester
M1 4BT
United Kingdom

+44 (0)300 323 0140

Intermediate-II + Unfavourable

Figure 11 Time to HSCT from relapse: Intermediate-II + Unfavourable (HSCT + no HSCT patients)

Figure 12 Time to HSCT from relapse: Intermediate-II + Unfavourable (HSCT only patients)

Intermediate-I only

Figure 13 Time to HSCT from relapse: Intermediate-I only (HSCT + no HSCT patients)

NICE National Institute for Health and Care Excellence

Level 1A
City Tower
Manchester
M1 4BT
United Kingdom

+44 (0)300 323 0140

Figure 14 Time to HSCT from relapse: Intermediate-I only (HSCT only patients)

Intermediate II only

Figure 15 Time to HSCT from relapse: Intermediate II only (HSCT + no HSCT patients)

Figure 16 Time to HSCT from relapse: Intermediate II only (HSCT only patients)

NICE National Institute for Health and Care Excellence

Level 1A
City Tower
Manchester
M1 4BT
United Kingdom

+44 (0)300 323 0140

B17. Please replicate figure 17 (p121 in the main submission) according to patients' status at transplant relative to EFS event: (i) before relapse (ii) after relapse; (iii) after induction failure.

(i) Before relapse

Figure 17 OS, Kaplan-Meier Curve, HSCT Patients from the Time of HSCT (HSCT before relapse)

(ii) After relapse

Figure 18 OS, Kaplan-Meier Curve, HSCT Patients from the Time of HSCT (HSCT after relapse)

(iii) After induction failure

Figure 19 OS, Kaplan-Meier Curve, HSCT Patients from the Time of HSCT (HSCT after induction failure)

NICE National Institute for Health and Care Excellence

Level 1A
City Tower
Manchester
M1 4BT
United Kingdom

+44 (0)300 323 0140

Population

B18. Please provide the baseline characteristics (i.e. replicating the information in table 71 in the company submission appendices) of the following cytogenetic subgroups:

(i) Favourable only

(ii) Intermediate only

(iii) Unknown only

Please see attachment [redacted]

Modelling approach

B19. Please provide a clearer description of the assumptions underlying the economic model structure and how the patient distribution across health states is estimated, making reference to how survival data and rates of HSCT are used in within the model.

Patients enter the model and begin induction with DA + GO or DA alone. After two cycles, patients are grouped based on response to induction therapy from the ALFA-0701 trial: CR or CRp (induction success) or refractory (induction failure). The trial population was split into 2 groups based on induction success/failure and separate survival analyses and HSCT calculations were performed for each group.

Transition probabilities for relapses and deaths occurring after the induction treatment phase were calculated from the RFS and OS curves. Separate OS analyses were performed for patients with induction success/failure. RFS was used for induction success patients but not for refractory patients because relapses following second-line treatment resulting in a CR or CRp were not captured in the trial. Therefore, after induction success, relapse and death rates are governed by RFS and OS curves for patients in the CR or CRp (on- or off-treatment) health state. After induction failure, death for patients in the refractory health state is governed by the separate OS curve.

RFS and OS curves determine the transitions between the CR or CRp, relapse, refractory, and dead health states. Patients can also undergo HSCT and HSCT probabilities were calculated for patients in the CR or CRp, relapse, and refractory health states. Patients

NICE National Institute for Health and Care Excellence

Level 1A
City Tower
Manchester
M1 4BT
United Kingdom

+44 (0)300 323 0140

transition to a HSCT health state that is associated with separate costs and quality of life. However, separate survival outcomes are not modelled for HSCT patients. HSCTs are tracked as sub-states within the CR or CRp, relapse, and refractory health states and aggregated for cost and QALY accrual. HSCT patients become functionally cured after 60 months in the base-case model – these patients are again tracked as a sub-state of the core health states and aggregated for cost and QALY accrual.

Including a separate survival curve for HSCT patients (calculated for all HSCTS, from the time of HSCT) was considered. This would have added considerable complexity to the model because matrices would need to have been included to calculate the survival of each HSCT patient from the time of HSCT. It was determined that this would not have added to the accuracy of the model because the OS curves would not add up to the OS curve for all patients in the trial.

Patients who receive HSCT are expected to have better survival outcomes than patients who don't (particularly versus relapse and refractory patients receiving non-curative therapy). Since survival for HSCT and non-HSCT patients is governed by the same underlying survival curves, adjustments were made to improve survival for HSCT versus non-HSCT patients. These are detailed in Section B.3.3.9 (pg. 120) of the submission.

B20. Please explain the advantages of the proposed modelling approach compared with that of a more conventional partitioned survival model, making reference to the structural assumptions described in question B19.

The model structure was designed to accurately capture the patient pathway for previously untreated AML patients in the UK and the benefits of GO. The model health states were informed and validated by UK clinical experts at an advisory board meeting held by Pfizer in July 2016. Health states that were deemed to have important differences in costs and/or HRQoL were included in the model.

A statistically significant improvement in induction success rates was not observed for GO. The key benefit was improved RFS which meant that patients receiving GO remained in CR for longer and delayed or avoided relapses. Moreover, patients with induction failure (refractory) had no benefit from receiving GO. Therefore, improvements in OS were attributable to patients with induction success. Performing separate survival analyses for patients with induction success/failure allowed us to isolate the refractory patients (for whom no difference was expected between treatment arms) and accurately model patient movements between the CR/CRp and relapse health states. Using a traditional 3-state partitioned survival model would not have allowed us to differentiate between relapse and refractory patients and accurately calculate differences in costs and QALYs between treatment arms for these patients.

NICE National Institute for Health and Care Excellence

Level 1A
City Tower
Manchester
M1 4BT
United Kingdom

+44 (0)300 323 0140

Salvage therapy and subsequent HSCT for relapse patients were considered important cost drivers because GO delays and avoids relapses. As stated, separate survival analyses for induction success/failure patients allowed us to model the accrual of costs and QALYs in separate relapse and refractory health states. It was also important to capture the different costs and QALYs for HSCT patients and patients classified as being functionally cured. Transitions to these health states from CR/CRp, relapse, refractory were tracked as sub-states.

The model was programmed as a cohort state-transition model rather than as a dual-partitioned survival model (i.e. a partitioned survival model with separate curves for induction success/failure) to allow the flexibility to incorporate the transition probabilities for the HSCT and functionally cured health states, and to adjust survival for HSCT and non-HSCT patients. The decision to program the model as a cohort state-transition model was made following discussions with and recommendations from health economic advisors during an advisory board meeting held on 21 February, 2017.

Resource use

B21. Please provide additional information on the actual induction and consolidation treatment received by patients in the trial.

(i) Please present the proportion of patients receiving induction course 1, induction course 2, consolidation course 1 and consolidation course 2 in the base-case population in the model.

The requested data are presented in the resource use worksheet in the company's economic model. In summary, the base case analysis uses pooled estimates across the treatment arms because clinical experts did not expect GO to affect the proportion of patients who receive each first-line treatment course. This is a conservative assumption because more patients received a second induction course and more consolidation courses in the DA arm than in the ALFA-0701 trial. The pooled base case data are as follows:

- Induction 1: 100%
- Induction 2: 19.8%
- Consolidation 1: 67.6%
- Consolidation 2: 59.7%

(ii) Provide a description of how these figures were estimated in relation to the data provided in table 33 of the CSR.

NICE National Institute for Health and Care Excellence

Level 1A
City Tower
Manchester
M1 4BT
United Kingdom

+44 (0)300 323 0140

Data on the proportion of patients who received treatment (as reported in Table 33 in the ALFA-0701 CSR) were used in the model. Separate values for daunorubicin and cytarabine were not used in the model and when the values differed the higher value was taken as the pooled base-case value. For example, 98.5% daunorubicin for the DA arm in induction 1 and 100% for cytarabine (i.e. the higher value).

The values for induction 2 are found in the text below Table 33 of the CSR: [redacted](Table 14.4.1.2.3)".

[redacted]

Consequently, the proportion of patients receiving each consolidation course was down-weighted as follows:

[redacted]

(iii) In addition, please provide the number of each induction and consolidation course for the subgroups described in question B18.

As agreed, Pfizer will not provide a response to this question due to not being able to complete by the deadline.

B22. Please justify why assumptions were made in the model about the length of hospitalisation for patients during the initial treatment periods (induction and consolidation) rather than using resource use data collected in the trial itself (see table 43, CSR).

Table 43 in the ALFA-0701 CSR reports aggregated median time in hospital data that includes hospitalisations for planned study treatments, hospitalisations for adverse events, and hospitalizations due to worsening of AML. These events were captured if they occurred up to 28 days after the last study dose (for induction, consolidation and salvage treatment phases). Therefore, the data in Table 43 do not fully align with what is required for the model health states and timings. The model also required separate hospitalisation data for salvage treatment for the relapse and refractory health states (these are not reported in Table 43). Furthermore, patients can relapse throughout the model and so the timing related information in Table 43 (i.e. 28 days after the last study dose) did not align with the requirements of the model.

It was important to accurately capture the cost of salvage therapy for relapse patients, particularly because improved RFS is the key benefit of GO. Using data from Table 43 would not have given the granularity to measure the cost differences for patients in the relapse health state. Furthermore, patients relapse throughout the trial and using the aggregated data (with a time cut-off) would not have captured all hospitalisations due to salvage therapy and these could not have been distinguished for relapse / refractory patients. Including

NICE National Institute for Health and Care Excellence

Level 1A
City Tower
Manchester
M1 4BT
United Kingdom

+44 (0)300 323 0140

separate estimates for salvage therapy - in addition to using data from Table 43 (which includes salvage) for induction and consolidation - would have resulted in double-counting costs.

The NHS Reference Costs used in the model for adverse events include hospitalisation costs, and the data in Table 43 included hospitalisations for adverse events. Therefore, including both hospitalisations and adverse events from the trial would result in double-counting of costs for adverse events. Hospitalisation estimates from clinical experts were for treatment administration only.

B23. Please replicate table 43 in the CSR for the subgroups specified in B5.

As agreed, Pfizer will not provide a response to this question due to not being able to complete by the deadline.

Section C: Textual clarifications and additional points

C1. Table 7 (company submission) states that outcomes were assessed using definitions provided in table 4. Should the table referred to be table 8?

Yes, the table referred to in the company submission is Table 8.

C2. In the Castaigne 2012 article it states that 28 patients (20%) had unfavourable cytogenetics, rather than 27 patients (19.4%), as stated in table 71 (page 77 of appendices). Please confirm which figure is correct.

The correct figure is 28 patients (20%).

C3. Figure 22 in the appendices states that 24 publications were excluded with the reason 'Combinations of GO irrelevant to the UK setting'. However in the list below the figure (page 63 of the appendices, 'Complete reference lists for excluded studies: non-RCT evidence'), only 23 references are listed. Please clarify this discrepancy.

While the study listing is correct, the PRISMA diagram inaccurately included a duplicate citation. Please see Figure 2 below for the corrected PRISMA diagram below.

Figure 20 PRISMA Diagram for non-RCT evidence

NICE National Institute for Health and Care Excellence

Level 1A
City Tower
Manchester
M1 4BT
United Kingdom

+44 (0)300 323 0140

C4. On page 69 of the appendices, should the reason for exclusion of the study by Roboz et al. be 'not an intervention/comparator of interest (i.e. no GO at all)'?

Roboz et al. reports results for **relapsed patients** receiving a variety of treatments categorized as either low-intensity or standard-intensity salvage therapies. Although outcomes are not presented by a specific treatment regimen, it is noted that GO is among the treatments included. Therefore, it is appropriate to exclude Roboz et al. 2012 as a study that evaluates 'relapsed/refractory/second-line' patients as stated in the company submission.

C5. Please clarify whether note (a) or (b) under table 15 of the company submission presents the correct information.

Note (b) is correct. i.e. the main cause of censoring was patients being event free at the 30 April 2013 reference date (GO + DA: 21.5%; DA arm: 13.2%).

NICE National Institute for Health and Care Excellence

Level 1A
City Tower
Manchester
M1 4BT
United Kingdom

+44 (0)300 323 0140

Appendices

1. **Appendix for B9**

Statistical tests for including/excluding patients with unknown cytogenetics

[redacted]

Figure 21 RFS KM curves when including and excluding patients with unknown cytogenetics

Table 24 Cox proportional hazards model (RFS including vs excluding unknowns)

Cox model	coef	exp(coef)	Logrank test (p-value)
	[redacted]	[redacted]	[redacted]

Figure 22 OS(CR) KM curves when including and excluding patients with unknown cytogenetics

Table 25 Cox proportional hazards model (OS(CR) including vs excluding unknowns)

Cox model	coef	exp(coef)	Logrank test (p-value)
	[redacted]	[redacted]	[redacted]

Intermediate and favourable RFS

NICE National Institute for Health and Care Excellence

Level 1A
City Tower
Manchester
M1 4BT
United Kingdom

+44 (0)300 323 0140

Figure 23 MCM Weibull; RFS; intermediate and favourable

Figure 24 MCM Lognormal; RFS; intermediate and favourable

Figure 25 MCM generalised gamma; RFS; intermediate and favourable

Table 26 Fit statistics RFS (intermediate and favourable)

	GO				No GO			
	AIC	Rank	BIC	Rank	AIC	Rank	BIC	Rank
Weibull MCM	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]

Level 1A
City Tower
Manchester
M1 4BT
United Kingdom

+44 (0)300 323 0140

Lognormal MCM	[redacted]							
Gen. gamma MCM	[redacted]							

Intermediate and favourable OS(CR)

Figure 26 MCM Weibull; OS(CR); intermediate and favourable

Figure 27 MCM Lognormal; OS(CR); intermediate and favourable

Figure 28 MCM generalised gamma; OS(CR); intermediate and favourable

Table 27 Fit statistics OS(CR) (intermediate and favourable)

	GO	No GO

NICE National Institute for Health and Care Excellence

Level 1A
City Tower
Manchester
M1 4BT
United Kingdom

+44 (0)300 323 0140

	AIC	Rank	BIC	Rank	AIC	Rank	BIC	Rank
Weibull MCM	[redacted]							
Lognormal MCM								
Gen. gamma MCM								

Intermediate only RFS

Figure 29 MCM Weibull; RFS; intermediate only

[redacted]

Figure 30 MCM Lognormal; RFS; intermediate only

[redacted]

Table 28 Fit statistics RFS ; intermediate only

	GO				No GO			
	AIC	Rank	BIC	Rank	AIC	Rank	BIC	Rank
Weibull MCM	[redacted]							
Lognormal MCM								
Gen. gamma MCM								

Intermediate only OS(CR)

NICE National Institute for Health and Care Excellence

Level 1A
City Tower
Manchester
M1 4BT
United Kingdom

+44 (0)300 323 0140

Figure 31 MCM Weibull; OS(CR); intermediate only

Figure 32 MCM Lognormal; OS(CR); intermediate only

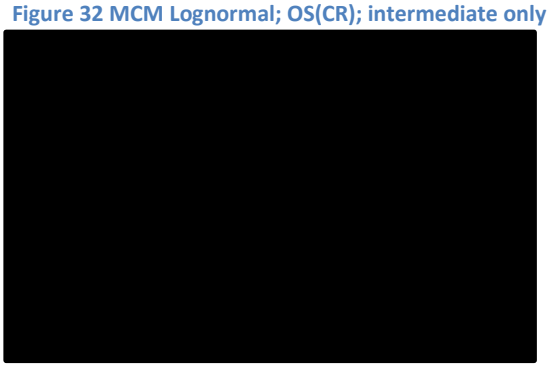


Figure 33 MCM generalised gamma; OS(CR); intermediate only

Table 29 Fit statistics OS(CR) ; intermediate only

	GO				No GO			
	AIC	Rank	BIC	Rank	AIC	Rank	BIC	Rank
Weibull MCM	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
Lognormal MCM	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
Gen. gamma MCM	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]

NICE National Institute for Health and Care Excellence

Level 1A
City Tower
Manchester
M1 4BT
United Kingdom

+44 (0)300 323 0140

2. **Appendix for B11**

Favourable and intermediate 1RFS

Figure 34 MCM Weibull; RFS; favourable and intermediate 1

Figure 35 MCM Lognormal; RFS; favourable and intermediate 1

Table 30 Fit statistics RFS; Favourable and intermediate 1

	GO				No GO			
	AIC	Rank	BIC	Rank	AIC	Rank	BIC	Rank
Weibull MCM	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
Lognormal MCM	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
Gen. gamma MCM	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]

Favourable and intermediate 1 OS(CR)

Figure 36 MCM Weibull; OS(CR); favourable and intermediate 1

NICE National Institute for Health and Care Excellence

Level 1A
City Tower
Manchester
M1 4BT
United Kingdom

+44 (0)300 323 0140

Figure 37 MCM Lognormal; OS(CR); favourable and intermediate 1

Figure 38 MCM generalised gamma; OS(CR); favourable and intermediate 1

Table 31 Fit statistics OS(CR); Favourable and intermediate 1

	GO				No GO			
	AIC	Rank	BIC	Rank	AIC	Rank	BIC	Rank
Weibull MCM	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
Lognormal MCM	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
Gen. gamma MCM	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]

Intermediate 2 and unfavourable RFS

Figure 39 MCM Weibull; RFS; intermediate 2 and unfavourable

NICE National Institute for Health and Care Excellence

Level 1A
City Tower
Manchester
M1 4BT
United Kingdom

+44 (0)300 323 0140

Figure 40 MCM Lognormal; RFS; intermediate 2 and unfavourable

Figure 41 MCM generalised gamma; RFS; intermediate 2 and unfavourable

Table 32 Fit statistics RFS: intermediate 2 and unfavourable

	GO				No GO			
	AIC	Rank	BIC	Rank	AIC	Rank	BIC	Rank
Weibull MCM	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
Lognormal MCM	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
Gen. gamma MCM	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]

NICE National Institute for Health and Care Excellence

Level 1A
City Tower
Manchester
M1 4BT
United Kingdom

+44 (0)300 323 0140

Intermediate 2 and unfavourable OS(CR)

Figure 42 MCM Weibull; OS(CR); intermediate 2 and unfavourable

Figure 43 MCM Lognormal; OS(CR); intermediate 2 and unfavourable

Table 33 Fit statistics OS(CR): intermediate 2 and unfavourable

	GO				No GO			
	AIC	Rank	BIC	Rank	AIC	Rank	BIC	Rank
Weibull MCM	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
Lognormal MCM	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
Gen. gamma MCM	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]

Intermediate 1 only RFS

Figure 44 MCM Weibull; RFS; intermediate 1 only

NICE National Institute for Health and Care Excellence

Level 1A
City Tower
Manchester
M1 4BT
United Kingdom

+44 (0)300 323 0140

Figure 45 MCM Lognormal; RFS; intermediate 1 only

Figure 46 MCM generalised gamma; RFS; intermediate 1 only

Table 34 Fit statistics RFS; intermediate 1 only

	GO				No GO			
	AIC	Rank	BIC	Rank	AIC	Rank	BIC	Rank
Weibull MCM	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
Lognormal MCM	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
Gen. gamma MCM	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]

Intermediate 1 only OS(CR)

Figure 47 MCM Weibull; OS(CR); intermediate 1 only

NICE National Institute for Health and Care Excellence

Level 1A
City Tower
Manchester
M1 4BT
United Kingdom

+44 (0)300 323 0140

Figure 48 MCM Lognormal; OS(CR); intermediate 1 only

Table 35 Fit statistics OS(CR); intermediate 1 only

	GO				No GO			
	AIC	Rank	BIC	Rank	AIC	Rank	BIC	Rank
Weibull MCM	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
Lognormal MCM	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
Gen. gamma MCM	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]

NICE National Institute for Health and Care Excellence

Patient organisation submission

Gemtuzumab ozogamicin for untreated acute myeloid leukaemia [ID982]

Thank you for agreeing to give us your organisation's views on this technology and its possible use in the NHS.

You can provide a unique perspective on conditions and their treatment that is not typically available from other sources.

To help you give your views, please use this questionnaire with our guide for patient submissions.

You do not have to answer every question – they are prompts to guide you. The text boxes will expand as you type.

Information on completing this submission

- Please do not embed documents (such as a PDF) in a submission because this may lead to the information being mislaid or make the submission unreadable
- We are committed to meeting the requirements of copyright legislation. If you intend to include **journal articles** in your submission you must have copyright clearance for these articles. We can accept journal articles in NICE Docs.
- Your response should not be longer than 10 pages.

About you	
1.Your name	[redacted]

2. Name of organisation	Leukaemia Care
3. Job title or position	[redacted]
4a. Brief description of the organisation (including who funds it). How many members does it have?	Leukaemia Care is a national blood cancer charity, first registered with the Charity Commission in 1969. We work to ensure that everybody has the right information, advice and support. Our key services include: Freephone helpline, Nurse Advisor, LiveChat, Nationwide Support Groups, Conferences, Campaigning and Advocacy, Buddy Support and Patient Booklets. Over 85% of our funding comes from our own fundraising activities and those of our volunteers. This includes a wide range of activities – such as legacies, community events, marathons, recycling campaigns etc. Leukaemia CARE also receives funding from a wide range of pharmaceutical companies, which in total represent approximately 15% of our annual income. Any funds received from the pharmaceutical industry are in accordance with the ABPI Code of Practice and the Leukaemia Care Code of Practice, our voluntary commitment that governs how we work with, and accept funding from, the pharmaceutical industry: www.leukaemiacare.org.uk/resources/code-of-practice
4b. Do you have any direct or indirect links with, or funding from, the tobacco industry?	N/A

5. How did you gather information about the experiences of patients and carers to include in your submission?	Information primarily gathered through Leukaemia Care patient experience survey – 'Living with Leukaemia' (www.leukaemiacare.org.uk/living-with-leukaemia). The survey was run from September to December 2016, as a follow up to NHS England's annual Cancer Patient Experience Survey (NCPES). The Leukaemia Care survey involved 85 questions and had responses from 2519 blood cancer patients including 373 AML patients. The results of this survey have been used to inform our submission. Additionally, we have gathered information through our helpline, support groups, communication with our membership and one to one discussion with patients. We also work closely with other patient groups and share expertise.
Living with the condition	
6. What is it like to live with the condition? What do carers experience when caring for someone with the condition?	Acute myeloid leukaemia (AML) accounts for around a third of cases of leukaemia in adults, with approximately 3000 people diagnosed in the UK each year. Approximately two thirds of patients are diagnosed aged 65 and over. Due to the rapidly progressing nature of the condition, 54% of patients had experienced symptoms for less than a month before visiting their GP. Common symptoms experienced prior to diagnosis include fatigue (70%), weakness/breathlessness (56%) and bruising or bleeding easily (31%). "I was experiencing all the common symptoms of AML; fatigue, shortness of breath, bruising easily. I was also at an increased risk of infection."

	The NCIN 'Routes to Diagnosis' report shows that 53% of AML patients are diagnosed via emergency presentation, compared to a cancer average of 22%. Being told you have cancer can be very upsetting. It can be especially difficult with acute leukaemia as you often get ill suddenly, and have to start treatment quickly (79% of AML patients start treatment within a week of diagnosis). There is usually very little time to take in information and start to cope with it. As a result, 51% of AML patients report being depressed or anxious more often since diagnosis. The emotional impact does not affect the patient in isolation. A diagnosis can place huge emotional strain on families, many of whom may also be affected. The most common symptoms encountered by patients since their diagnosis include fatigue (73%), weakness or breathlessness (51%), bruising or bleeding (37%) and infections (32%). AML also has a practical impact, with 52% of patients experiencing pain as a direct result of their condition. Additionally, 51% of AML patients have difficulty moving around and 59% of patients have difficulty performing some of their daily routines, such as cooking or cleaning. Of those in work or education before their diagnosis, 78% have been impacted (32% reduced hours, 45% no longer able to work or continue education). Consequently, 53% of patients reported a negative financial impact as a result of having AML (increased costs or reduced income).

7. What do patients or carers think of current treatments and care available on the NHS?	There has been limited progress in the treatment of AML for decades. In 2014 there were 3072 new cases of AML diagnosed in the UK. In 2014, there were around 2,516 deaths from AML in the UK. As such, there is an urgent need for improvements.
8. Is there an unmet need for patients with this condition?	Yes.
9. What do patients or carers think are the advantages of the technology?	Gemtuzumab ozogamicin (in combination with chemotherapy) offers patients a survival benefit (EFS), (17.3 months v 9.5 months). AML has extremely poor outcomes and high unmet need, with little progress in decades.
10. What do patients or carers think are the disadvantages of the technology?	Early usage of gemtuzumab ozogamicin (monotherapy) resulted in a withdrawal from the market. However, since then, gemtuzumab ozogamicin is being used at a lower dose (in combination with chemotherapy) and in a different patient population.

	Gemtuzumab ozogamicin has a series of side effects – infections, liver damage, veno-occlusive disease, bleeding. However, 80% of AML patients reported that they would be willing to experience additional side effects for a more effective treatment.
11. Are there any groups of patients who might benefit more or less from the technology than others? If so, please describe them and explain why.	
12. Are there any potential equality issues that should be taken into account when considering this condition and the technology?	

13. Are there any other issues that you would like the committee to consider?	

Key messages

15. In up to 5 bullet points, please summarise the key messages of your submission:

- Approximately 3000 people are diagnosed with AML in the UK each year. Two thirds of patients are diagnosed aged 65 and over. Due to the rapidly progressing nature of the condition, 54% of patients had experienced symptoms for less than a month before visiting their GP.
- AML patients experience a range of symptoms, as well as both a practical (pain, mobility) and financial impact. Common symptoms include fatigue (73%), weakness or breathlessness (51%), bruising or bleeding (37%) and infections (32%). Being diagnosed with AML also has an emotional impact, with 51% of AML patients report being depressed or anxious more often since diagnosis.
- Gemtuzumab ozogamicin (in combination with chemotherapy) offers patients a survival benefit (EFS). AML has extremely poor outcomes and high unmet need, with little progress in decades.

Thank you for your time. Please log in to your NICE Docs account to upload your completed submission.

NICE National Institute for Health and Care Excellence

Professional organisation submission

Leukaemia (acute myeloid, untreated) – gemtuzumab ozogamicin (ID982)

Thank you for agreeing to give us your organisation's views on this technology and its possible use in the NHS.

You can provide a unique perspective on the technology in the context of current clinical practice that is not typically available from the published literature.

To help you give your views, please use this questionnaire. You do not have to answer every question – they are prompts to guide you. The text boxes will expand as you type.

Information on completing this submission

- Please do not embed documents (such as a PDF) in a submission because this may lead to the information being mislaid or make the submission unreadable
- We are committed to meeting the requirements of copyright legislation. If you intend to include **journal articles** in your submission you must have copyright clearance for these articles. We can accept journal articles in NICE Docs.
- Your response should not be longer than 13 pages.

About you	
1. Your name	[redacted], submitting on behalf of:
2. Name of organisation	NCRI-ACP-RCP

3. Job title or position	[redacted]
4. Are you (please tick all that apply):	☐ an employee or representative of a healthcare professional organisation that represents clinicians? **x** a specialist in the treatment of people with this condition? ☐ a specialist in the clinical evidence base for this condition or technology? ☐ other (please specify):
5a. Brief description of the organisation (including who funds it).	The AML Working Group (Clinical Studies Group) is funded by the National Cancer Research Institute. It comprises clinical haematologists with a specialist interest in acute myeloid leukaemia (representing the majority of large AML-treating centres in the UK including all 3 devolved nations), specialist AML laboratory scientists and trials designers / statisticians. Over the last 30-40 years the AML Working Group has designed and overseen the MRC/NCRI national AML trials including randomised assessments of gemtuzumab ozogamicin (GO) in newly-diagnosed patients aged <60 (AML15, 17 studies) and >60 years (AML16, 18).
5b. Do you have any direct or indirect links with, or funding from, the tobacco industry?	No
The aim of treatment for this condition	
6. What is the main aim of treatment? (For example, to stop progression, to improve mobility, to cure the condition,	The aim of intensive AML therapy is curative. Firstly by achieving remission and then by giving further chemotherapy (with or without the addition of allogeneic stem cell transplant) to prevent relapse. The largest trials of GO as part of intensive chemotherapy for newly-diagnosed AML patients were carried out by UK haematologists under the auspices of the NCRI funded trials in both young (n=1113) and older patients (n=1115).

or prevent progression or disability.)	In a patient based meta-analysis for all trials, which was led from the UK (Hills et al Lancet Oncology 2014 15(9), 986), the addition of GO significantly reduced the risk of relapse (OR 0.81, p=0.0001), and hence improved relapse-free (OR 0.87, p=0.005) and overall survival at 5 years (OR 0.90, p=0.01). The survival benefit was most apparent in patients with favourable risk cytogenetics (OR 0.47, p=0.0006), and also seen in those with intermediate risk disease (OR 0.84, p=0.005), but not in those with adverse risk AML (OR 0.99, p=0.9). The NCRI trials also established that a higher dose of GO (6mg/m2) was not superior to the standard UK dose of 3mg/m2. Currently all UK NCRI trial patients (AML18 and AML19 trials) receive GO with chemotherapy as part of a randomised assessment aimed at establishing whether dosing on 2 days of induction chemotherapy (days 1 and 4) is better than for the single dose (day 1) used in the previous UK trials.
7. What do you consider a clinically significant treatment response? (For example, a reduction in tumour size by x cm, or a reduction in disease activity by a certain amount.)	Achievement of complete remission. Improving and event free survival and overall survival.
8. In your view, is there an unmet need for patients and healthcare professionals in this condition?	Yes. Intensive treatment for AML with combination chemotherapy is essentially unchanged in 40 years and although survival has gradually improved this has been largely due to improvements in supportive care and the judicious application of allogeneic stem cell transplantation. Although most patients will enter complete remission with currently available chemotherapy, 50% of younger patients (<60years) and 80% of older patients will relapse. Mylotarg has been seen to reduce the risk of relapse in important subgroups of patients and hence reduces the requirement for salvage chemotherapy (and subsequent BMT in second remission).
What is the expected place of the technology in current practice?	

9. How is the condition currently treated in the NHS?	The majority of UK patients are entered into NCRI AML trials. Currently these comprise AML 19 for younger patients and AML 18 for patients >60 years. Both trials currently include Mylotarg in induction for patients without a known adverse karyotype at diagnosis. Also a paediatric AML study (MyeChild) currently uses GO in induction. Outside the setting of clinical trials, the majority of patients who are considered suitable for intensive chemotherapy currently receive 'standard' combination chemotherapy (daunorubicin + cytarabine) without the addition of Mylotarg.
• Are any clinical guidelines used in the treatment of the condition, and if so, which?	The AML18 and 19 study protocols are used as 'guidelines' by most UK haematologists. The ELN AML guidelines published in 2017 were written before the approval of GO in the US
• Is the pathway of care well defined? Does it vary or are there differences of opinion between professionals across the NHS? (Please state if your experience is from outside England.)	Yes – the care pathway is generally well-defined. Patients are treated in larger centres with experience of treating AML; usually these are centres that participate in the NCRI AML trials. All trials of GO to date have shown that the 20% of patients with adverse risk disease derive no benefit from the addition of GO, although it does not harm them. This raises the issue of the desirability of being able to access cytogenetic results very promptly at diagnosis before starting treatment, which is not standard practice, in order to be able to determine which patients should receive Mylotarg with their induction chemotherapy. This may cause some difficulty for cytogenetic lab services to produce the result within the required 48 hours, although in many AML cases with 'less proliferative' disease the start of chemotherapy may safely be delayed by a few days without compromising patient outcome while cytogenetic analysis is performed.
• What impact would the technology have on the current pathway of care?	Very little direct impact. As discussed above, there may be greater onus placed on rapid turnaround of cytogenetic analysis prior to starting chemotherapy.

10. Will the technology be used (or is it already used) in the same way as current care in NHS clinical practice?	Yes. The expectation would be that GO would be administered as part of induction chemotherapy, as has been the case in the majority of NCRI AML trials over the last 10-15 years. Most haematologists / haematology units are very used to administering the treatment in this way.
• How does healthcare resource use differ between the technology and current care?	Mylotarg would be added to standard AML induction therapies. This would primarily be with the induction combination of daunorubicin/ cytarabine (DA 3+10), or alternatively as part of 'FLAG-Ida' (fludarabine/cytarabine/G-CSF/idarubicin. Based on data from AML16 and 17 trials, adding Mylotarg would not be anticipated to extend inpatient stays or significantly alter the use of supportive care (eg. blood products).
• In what clinical setting should the technology be used? (For example, primary or secondary care, specialist clinics.)	Secondary/tertiary care – inpatient.
• What investment is needed to introduce the technology? (For example, for facilities, equipment, or training.)	Very little is required. Haematologists and pharmacy departments already have significant experience of using GO from NCRI AML clinical trials over the last 10-15 years.
11. Do you expect the technology to provide clinically meaningful benefits compared	The NCRI AML15 trial demonstrated a significant benefit of giving Mylotarg when combined with DA, ADE or FLAG-Ida chemotherapy in at least 70% of younger patients with AML; this has been reinforced in older patients by the results of the NCRI AML16 trial, and a recent meta-analysis of all similar trials in adults published by Hills et al (referenced above). Included in this experience is that of the French ALFA group who demonstrated significantly improved survival in patients aged 50-70 given fractionated doses of Mylotarg. In the latter trial although CR with or without platelet recovery

with current care?	and early deaths were similar, patients in the Mylotarg arm had significantly improved median event-free survival (19.6 vs 11.9 months; p=0.00018) and overall survival (34 vs 19.2 months; p=0.046), with a sub-analysis revealing benefit limited to patients with favourable and intermediate-risk karyotype. This fractionated dose of GO is the one that is likely to be licenced as it gave the clearest benefit
• Do you expect the technology to increase length of life more than current care?	Yes. A meta-analysis of 3,325 patients from five randomised studies in untreated AML (aged 18–84) concluded that Mylotarg improved overall survival in patients with favourable and intermediate-risk karyotype when combined with standard induction chemotherapy
• Do you expect the technology to increase health-related quality of life more than current care?	Yes – primarily by reducing relapse which is a devastating event for patients.
12. Are there any groups of people for whom the technology would be more or less effective (or appropriate) than the general population?	Yes. The survival benefit was most apparent in patients with favourable risk cytogenetics (OR 0.47, p=0.0006), and also seen in those with intermediate risk cytogenetics (OR 0.84, p=0.005), but not in those with adverse risk AML (OR 0.99, p=0.9). In both the current AML18 and 19 trials, centres are encouraged to await cytogenetic results before treating with Mylotarg, although this is sometimes not clinically feasible in the minority of patients who have a more proliferative / aggressive presentation of AML where immediate treatment is required.
The use of the technology	
13. Will the technology be easier or more difficult to use for patients or healthcare	The acute infusion-related toxicities seen with GO are transient and usually respond to standard interventions. The toxicity of greatest concern is the development of hepatic veno-occlusive disease (VOD). The risk of VOD appears to be relatively low when individual doses of no greater than 3 mg/m2 are used in

professionals than current care? Are there any practical implications for its use (for example, any concomitant treatments needed, additional clinical requirements, factors affecting patient acceptability or ease of use or additional tests or monitoring needed.)	combination with conventional therapy as part of initial therapy of AML (as would be the case under the terms of the license).
14. Will any rules (informal or formal) be used to start or stop treatment with the technology? Do these include any additional testing?	Karyotypic (cytogenetic) analysis is a standard part of AML work up at the time of initial diagnosis so no additional testing is required. As stated above though, if Mylotarg is to be restricted (in line with clinical trial results) to patients with favourable or intermediate risk cytogenetics, the more rapid turnaround of these tests may create an administrative/workload challenge for genetics laboratories.
15. Do you consider that the use of the technology will result in any substantial health-related benefits that are unlikely to be included in the quality-adjusted life year	No

(QALY) calculation?	
16. Do you consider the technology to be innovative in its potential to make a significant and substantial impact on health-related benefits and how might it improve the way that current need is met?	Yes. This is highly innovative in that it will be the first routine application of antibody-directed chemotherapy in the treatment of AML (or indeed for any cancer).
• Is the technology a 'step-change' in the management of the condition?	Yes. GO is the first drug (in 30+ years of clinical trials) that has consistently shown survival benefit when added to standard induction chemotherapy in newly-diagnosed AML.
• Does the use of the technology address any particular unmet need of the patient population?	As above
17. How do any side effects or adverse effects of the technology affect the	The acute infusion-related toxicities seen with GO are transient and usually respond to standard interventions. The toxicity of greatest concern is the development of liver veno-occlusive disease (VOD). The risk of VOD appears to be low when individual doses of no greater than 3 mg/m2 are used in

management of the condition and the patient's quality of life?	combination with conventional therapy as part of initial therapy of AML. This has been a rare event in the AML 18 and 19 trials
Sources of evidence	
18. Do the clinical trials on the technology reflect current UK clinical practice?	The UK has led clinical trials involving the application of this technology. Given that the majority of newly-diagnosed AML patients in the UK continue to be treated according to NCRI AML trials protocols, this clearly reflects UK clinical practice.
• If not, how could the results be extrapolated to the UK setting?	Not applicable.
• What, in your view, are the most important outcomes, and were they measured in the trials?	Relapse free survival and Overall survival - both were measured (and improved by Mylotarg) in the relevant NCRI (and ALFA group) trials.
• If surrogate outcome measures were used, do they adequately predict long-term clinical outcomes?	Not applicable.
• Are there any adverse effects that were not apparent in clinical trials but have come to light	No.

subsequently?	
19. Are you aware of any relevant evidence that might not be found by a systematic review of the trial evidence?	No.
20. Are you aware of any new evidence for the comparator treatment(s) since the publication of NICE technology appraisal guidance [TAXXX]?	No.
21. How do data on real-world experience compare with the trial data?	Given that we think that 85% of younger patients with AML have been entered into the national trials over the years, we feel that a particular strength of the NCRI clinical trials is that they capture 'UK real world data' rather than that of a highly-selected clinical trial population.
Equality	
22a. Are there any potential equality issues that should be taken into account when considering this treatment?	No.

22b. Consider whether these issues are different from issues with current care and why.	Not applicable.

Key messages

24. In up to 5 bullet points, please summarise the key messages of your submission.

- Our UK trials, and the international meta-analysis, show that gemtuzumab ozogamicin reduces the incidence of relapse and improves overall survival when added to induction chemotherapy for patients with favourable and intermediate risk disease karyotype.
- A dose level of 3mg/m2 given once or twice (to be determined by ongoing UK AML18 and 19 trials) is well tolerated.
- By reducing relapse, the significant expense of relapse (salvage) treatment, which will include further chemotherapy and allogeneic stem cell transplantation is saved
- UK haematologists and pharmacists are familiar with the drug.
-

Thank you for your time.

Please log in to your NICE Docs account to upload your completed submission.

NICE National Institute for Health and Care Excellence

Professional organisation submission

Gemtuzumab ozogamicin for untreated acute myeloid leukaemia [ID982]

Thank you for agreeing to give us your organisation's views on this technology and its possible use in the NHS.

You can provide a unique perspective on the technology in the context of current clinical practice that is not typically available from the published literature.

To help you give your views, please use this questionnaire. You do not have to answer every question – they are prompts to guide you. The text boxes will expand as you type.

Information on completing this submission

- Please do not embed documents (such as a PDF) in a submission because this may lead to the information being mislaid or make the submission unreadable
- We are committed to meeting the requirements of copyright legislation. If you intend to include **journal articles** in your submission you must have copyright clearance for these articles. We can accept journal articles in NICE Docs.
- Your response should not be longer than 13 pages.

About you	
1. Your name	**Nigel Russell**
2. Name of organisation	

3. Job title or position	
4. Are you (please tick all that apply):	☐ an employee or representative of a healthcare professional organisation that represents clinicians? x☐ a specialist in the treatment of people with this condition? ☐ a specialist in the clinical evidence base for this condition or technology? ☐ other (please specify):
5a. Brief description of the organisation (including who funds it).	
5b. Do you have any direct or indirect links with, or funding from, the tobacco industry?	**no**
The aim of treatment for this condition	
6. What is the main aim of treatment? (For example, to stop progression, to improve mobility, to cure the condition, or prevent progression or	The aim of therapy is curative. Firstly by achieving remission and then by giving further chemotherapy to prevent relapse. The largest trials of GO were carried out by UK haematologists under the auspices of the NCRI funded trials in both young (n=1113) and older patients (n=1115) In a patient based meta-analysis for all trials, which was lead from the UK (Hills et al Lancet Oncology 2014 15(9), 986), the addition of GO significantly reduced the risk of relapse (OR 0.81, p=0.0001), and hence improved relapse-free (OR 0.87, p=0.005) and overall survival at 5 years (OR 0.90, p=0.01). The survival benefit was most

disability.)	apparent in patients with favourable risk cytogenetics (OR 0.47, p=0.0006), and also seen in those with intermediate risk disease (OR 0.84, p=0.005), but not in those with adverse risk AML (OR 0.99, p=0.9). The UK haematologists established that the higher dose (6mg/msq) was not superior than the UK dose of 3mg/msq. Currently all UK trial patients receive GO where the aim is to establish whether dosing on 2 days if induction chemotherapy is better than for one day as in the previous UK trials.
7. What do you consider a clinically significant treatment response? (For example, a reduction in tumour size by x cm, or a reduction in disease activity by a certain amount.)	Complete remission and event free survival
8. In your view, is there an unmet need for patients and healthcare professionals in this condition?	Yes. The therapy for AML is essentially unchanged in 40 years and although survival has improved this is mainly due to better supportive care. Although most patients will enter complete remission with currently available chemotherapy, 50% of younger patients (<60years) and 80% of older patients will relapse. Mylotarg reduces the risk of relapse in important subgroups of patients and hence reduces the requirement for salvage chemotherapy and BMT
What is the expected place of the technology in current practice?	
9. How is the condition currently treated in the NHS?	Many patients are entered into NCRI AML trials. Currently these are AML 19 for younger patients and AML 18 for patients >60 years. Both trials include Mylotarg in induction for patients without a known adverse karyotype at diagnosis. Also a paediatric AML study (MyChild) uses GO in induction.
• Are any clinical	AML18 and 19 are used as guidelines. The ELN AML guidelines published in 2017 were written before the

guidelines used in the treatment of the condition, and if so, which?	approval of GO in the US
• Is the pathway of care well defined? Does it vary or are there differences of opinion between professionals across the NHS? (Please state if your experience is from outside England.)	It is. Patients are treated in centres with experience of treating AML usually these are centres that participate in NCRI AML trials. All trials have shown that the 20% of patients with adverse risk disease derive no benefit from the addition of GO, it does not harm them. This raises the issue of getting the cytogenetic data very promptly before starting treatment, which is not standard practice.. It may cause difficulty for cytogenetic lab services to produce the result in the required 48 hours. I
• What impact would the technology have on the current pathway of care?	None
10. Will the technology be used (or is it already used) in the same way as current care in NHS clinical practice?	Yes
• How does healthcare resource use differ between the technology and current care?	Mylotarg would be added to standard AML induction therapies. This would be the combination of daunorubicin/ Ara-C (DA 3+10) or fludarabine/Ara-C/G-CSF/idarubicin (Flag-Ida)

• In what clinical setting should the technology be used? (For example, primary or secondary care, specialist clinics.)	Secondary/tertiary care
• What investment is needed to introduce the technology? (For example, for facilities, equipment, or training.)	I think very little is required. Haematologists already have significant experience of using GO from NCRI AML clinical trials
11. Do you expect the technology to provide clinically meaningful benefits compared with current care?	The NCRI AML15 trial demonstrated a significant benefit of giving Mylotarg when combined with DA, ADE or FLAG-Ida chemotherapy in at least 70% of patients with AML; this has been reinforced by the results of the NCRI AML16 trial, and a recent meta-analysis of all similar trials in adults published by Hills et al (referenced above). Included in this experience is that of the French ALFA group who demonstrated significantly improved survival in patients aged 50-70 given fractionated doses of Mylotarg. In the latter trial although CR with or without platelet recovery and early deaths were similar, patients in the Mylotarg arm had significantly improved median event-free survival (19.6 vs 11.9 months; p=0.00018) and overall survival (34 vs 19.2 months; p=0.046), with a sub-analysis revealing benefit limited to patients with favorable and intermediate-risk karyotype. This fractionated dose of GO is the one that is likely to be licenced as it gave the clearest benefit
• Do you expect the technology to increase length of life more than current care?	Yes. A meta-analysis of 3,325 patients from five randomised studies in untreated AML (aged 18–84) concluded that Mylotarg improved overall survival in patients with favourable and intermediate-risk karyotype when combined with standard induction chemotherapy
• Do you expect the technology to increase	Yes by reducing relapse which is a devastating event for patients

health-related quality of life more than current care?	
12. Are there any groups of people for whom the technology would be more or less effective (or appropriate) than the general population?	Yes. The survival benefit was most apparent in patients with favorable risk cytogenetics (OR 0.47, p=0.0006), and also seen in those with intermediate risk cytogenetics (OR 0.84, p=0.005), but not in those with adverse risk AML (OR 0.99, p=0.9). In both the AML18 and 19 trials currently running in the UK centres are encouraged to await for cytogenetic results before treating with Mylotarg although this is not always clinically feasible
The use of the technology	
13. Will the technology be easier or more difficult to use for patients or healthcare professionals than current care? Are there any practical implications for its use (for example, any concomitant treatments needed, additional clinical requirements, factors affecting patient acceptability or ease of use or additional	The acute infusion-related toxicities seen with GO are transient and usually respond to standard interventions. The toxicity of greatest concern is the development of liver veno-occlusive disease (VOD). The risk of VOD appears to be relatively low when individual doses of no greater than 3 mg/m2 are used in combination with conventional therapy as part of initial therapy of AML.

tests or monitoring needed.)	
14. Will any rules (informal or formal) be used to start or stop treatment with the technology? Do these include any additional testing?	Karyotypic analysis is a standard part of AML work up so no additional testing is required
15. Do you consider that the use of the technology will result in any substantial health-related benefits that are unlikely to be included in the quality-adjusted life year (QALY) calculation?	No
16. Do you consider the technology to be innovative in its potential to make a significant and substantial impact on health-related benefits and how might it	Yes this is the first application of antibody directed chemotharapy in the treatment of AML(or indeed for any cancer)

improve the way that current need is met?	
• Is the technology a 'step-change' in the management of the condition?	Yes it is the first drug that has been consistently shown benefit in AML when added to standard induction chemotherapy
• Does the use of the technology address any particular unmet need of the patient population?	As above
17. How do any side effects or adverse effects of the technology affect the management of the condition and the patient's quality of life?	The acute infusion-related toxicities seen with GO are transient and usually respond to standard interventions. The toxicity of greatest concern is the development of liver veno-occlusive disease (VOD). The risk of VOD appears to be low when individual doses of no greater than 3 mg/m2 are used in combination with conventional therapy as part of initial therapy of AML. This has been a rare event in the AML 18 and 19 trials
Sources of evidence	
18. Do the clinical trials on the technology reflect current UK clinical practice?	The UK has led clinical trials involving the application of this technology

• If not, how could the results be extrapolated to the UK setting?	The UK has widespread experience in the use of this drug in AML
• What, in your view, are the most important outcomes, and were they measured in the trials?	Relapse free survival and Overall survival
• If surrogate outcome measures were used, do they adequately predict long-term clinical outcomes?	N/A
• Are there any adverse effects that were not apparent in clinical trials but have come to light subsequently?	No
19. Are you aware of any relevant evidence that might not be found by a systematic review of the trial evidence?	No
20. Are you aware of any new evidence for the comparator	No

treatment(s) since the publication of NICE technology appraisal guidance [TAXXX]?	
21. How do data on real-world experience compare with the trial data?	Given that we think that 85% of younger patients with AML are entered into the national trial, I think the clinical trial results from the UK capture real world data
Equality	
22a. Are there any potential equality issues that should be taken into account when considering this treatment?	No
22b. Consider whether these issues are different from issues with current care and why.	
Key messages	

24. In up to 5 bullet points, please summarise the key messages of your submission.

- Our UK trails and the international meta-analysis show that GO reduces relapse and improves survival when added to induction chemotherapy
- A dose level of 3mg/sqm given once or twice (Which will be determined by ongoing UK trial) is well tolerated.
- By reducing relapse, treatment expense of relapse, (which will include further chemotherapy and transplantation) is saved
- UK haematologists and pharmacists are familiar with the drug.
-

Thank you for your time.

Please log in to your NICE Docs account to upload your completed submission.

CONFIDENTIAL UNTIL PUBLISHED

Evidence Review Group's Report
Gemtuzumab ozogamicin for untreated acute myeloid leukaemia

Produced by	CRD and CHE Technology Assessment Group, University of York, Heslington, York YO10 5DD
Authors	Ros Wade, Research Fellow, CRD Lindsay Claxton, Research Fellow in Health Economics, CRD Ruth Walker, Research Fellow, CRD James Altunkaya, NIHR Research Methods Fellow, CHE Marta Soares, Senior Research Fellow, CHE Kath Wright, Information Specialist, CRD Robert Hodgson, Lead Health Economist, CRD Mark Simmonds, Senior Research Fellow, CRD Stephen Palmer, Professor of Health Economics, CHE
Correspondence to	Ros Wade, Research Fellow, CRD, University of York, York YO10 5DD
Date completed	22/03/2018

Source of funding
This report was commissioned by the NIHR HTA Programme as project number 16/134/13.

Declared competing interests of the authors
None.

Acknowledgements
We would like to thank Dr Eleni Tholouli, Consultant Haematologist at Manchester Royal Infirmary for clinical advice throughout the 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:

Wade R, Claxton L, Walker R, Altunkaya J, Soares M, Wright K, Hodgson R, Simmonds M and Palmer S. Gemtuzumab ozogamicin for untreated acute myeloid leukaemia: A Single Technology Appraisal. CRD and CHE Technology Assessment Group, 2018.

Contributions of authors

Ros Wade and Ruth Walker wrote the clinical effectiveness sections of the report. Lindsay Claxton wrote the cost effectiveness sections and conducted the economic analyses. James Altunkaya, Marta Soares and Robert Hodgson provided methodological and technical support for the cost-effectiveness sections and model validation. Kath Wright wrote the sections on the search strategies. Mark Simmonds provided advice, commented on drafts of the report and took overall responsibility for the clinical effectiveness sections. Stephen Palmer provided advice, contributed to drafts of the report and took overall responsibility for the cost effectiveness sections.

Note on the text

All commercial-in-confidence (CIC) data have been highlighted in [redacted] all academic-in-confidence (AIC) data are highlighted in [redacted]

Copyright statement

Table of Contents

Table of Tables

Table of Figures

List of abbreviations

ADE	Daunorubicin, cytarabine and etoposide
AE	Adverse event
AIC	Akaike information criterion
AML	Acute myeloid leukaemia
ANC	Absolute neutrophil count
AraC	Cytarabine
BCSH	British Committee for Standards in Haematology
BIC	Bayesian information criterion
BMA	Bone marrow aspirate
BNF	British National Formulary
BSA	Body surface area
BSC	Best supportive care
CEA	Cost-effectiveness analysis
CEAC	Cost-effectiveness acceptability curve
CEBPA	CCAAT/enhancer-binding protein gene
CHMP	Committee for Medicinal Products for Human Use
CHV	Centre Hospitalier de Versailles
CI	Confidence interval
CR	Complete remission
CRp	Complete remission with incomplete platelet recovery
CS	Company submission
CSR	Clinical study report
DA	Daunorubicin and cytarabine
DClo	Daunorubicin and clofarabine
DNR	Daunorubicin
EFS	Event-free survival
ELN	European LeukaemiaNet
EMA	European Medicines Agency
eMIT	Electronic market information tool
EORTC	European Organisation for Research and Treatment of Cancer
EPAR	European public assessment report
EQ-5D	EuroQol 5-dimension quality of life questionnaire
ERG	Evidence Review Group
FLAG-Ida	Fludarabine, cytarabine, G-CSF and idarubicin

FDA	Food and Drug Administration
FLAG-Ida	Fludarabine, cytarabine, G-CSF and idarubicin
FLT3-ITD	Internal tandem duplication of the FMS-like tyrosine kinase 3 gene
G-CSF	Granulocyte colony-stimulating factor
GO	Gemtuzumab ozogamicin
GVHD	Graft versus host disease
HR	Hazard ratio
HRQoL	Health-related quality of life
HSCT	Haematopoietic stem cell transplantation
IA	Investigator assessment
ICER	Incremental cost-effectiveness ratio
IDAC	Intermediate dose cytarabine
IPD	Individual patient data
IRC	Independent review committee
ITT	Intention-to-treat
MCM	Mixture cure models
mg	milligram
MFI	Mean fluorescence intensity
mITT	Modified intention-to-treat
MLL	Myeloid/lymphoid leukaemia gene
MRC	Medical Research Council
MRD	Minimal residual disease
NCCN	National Comprehensive Cancer Network
NHS	National Health Service
NICE	National Institute for Health and Care Excellence
NPM1	Nucleophosmin-1 gene
NR	Not reported
OR	Odds ratio
OS	Overall survival
PartSA	Partitioned survival analysis
PFC	Points for clarification
PRISMA	Preferred Reporting Items for Systematic Review and Meta-analyses
PSA	Probabilistic sensitivity analysis
PSS	Personal Social Services
PSSRU	Personal Social Services Research Unit

QALY	Quality-adjusted life year
RCT	Randomised controlled trial
RFS	Relapse-free survival
RMST	Restricted mean survival time
SAE	Serious adverse event
SD	Standard deviation
SE	Standard error
SLR	Systematic literature review
SmPC	Summary of product characteristics
SOC	System organ class
STA	Single Technology Appraisal
TEAE	Treatment-emergent adverse event
ULN	Upper limit of normal
VOD	Veno-occlusive disease
WBC	White blood cell
WT1	Wilms' tumour suppressor gene

1 Summary

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

The patient population addressed in the company submission (CS) is adult patients not known to have unfavourable cytogenetics, with previously untreated, de novo acute myeloid leukaemia (AML). The CS notes that this is a subpopulation of the anticipated marketing authorisation for gemtuzumab ozogamicin (GO) and the population described in the final scope issued by NICE, which was adults with untreated AML. The CS claims that patients known to have unfavourable cytogenetics would not be treated with GO plus intensive chemotherapy in NHS clinical practice, which was supported by the clinical advisor to the Evidence Review Group (ERG). However, in view of the very short timeframe between diagnosis and treatment in patients with AML, the requirement for cytogenetic test results prior to commencing treatment with GO could potentially delay the start of treatment.

The expected date for the European Medicine Agency (EMA) marketing authorisation for GO is Q2 2018. On 22 February 2018 the Committee for Medicinal Products for Human Use (CHMP) adopted a positive opinion, recommending the granting of a marketing authorisation for GO in 'combination therapy with daunorubicin (DNR) and cytarabine (AraC) for the treatment of patients age 15 years and above with previously untreated, de novo CD33-positive acute myeloid leukaemia (AML), except acute promyelocytic leukaemia (APL)'. The restriction to CD33-positive AML is a narrower population than that addressed in the CS.

The intervention specified in the NICE scope is GO in combination with chemotherapy. The intervention addressed in the CS is GO in combination with daunorubicin plus cytarabine (DA; intensive chemotherapy), which matches the anticipated marketing authorisation for GO.

The comparator addressed in the CS is DA, which is narrower than that specified in the NICE scope. However, the clinical advisor to the ERG confirmed that DA is the standard chemotherapy regimen used for patients with de novo AML who can receive intensive chemotherapy, therefore, the comparator addressed in the CS was appropriate.

The outcome measures specified in the NICE scope were reported in the CS, although health related quality of life (HRQoL) outcomes were not presented in the CS as the pivotal study presented in the submission did not collect HRQoL or utility data.

1.2 Summary of clinical effectiveness evidence submitted by the company

The company conducted a systematic review to identify evidence on the comparative efficacy and safety of GO + DA versus DA for the treatment of patients with previously untreated de novo AML. Eight randomised controlled trials (RCTs) were included in the systematic review. The ALFA-0701

trial is the pivotal study used to support the EMA marketing authorisation. The remaining seven RCTs did not use a dose or dosing schedule of GO that is expected to be approved by the EMA; therefore, only the ALFA-0701 trial is the primary source of clinical effectiveness evidence for the appraisal.

ALFA-0701 was a phase 3 multicentre open-label RCT undertaken at 26 haematology centres in France. Patients aged 50-70 years of age were randomised to GO + DA or DA alone.

The ALFA-0701 trial demonstrated that GO + DA was effective at improving event-free survival (EFS) by approximately [redacted] and relapse-free survival (RFS) by approximately [redacted], compared with DA alone, in the overall patient population. Whilst overall survival (OS) and response rate appeared better in the GO + DA arm, these results did not reach statistical significance [redacted].

EFS and RFS results were statistically significantly improved in the GO + DA arm for the subgroup of patients with favourable/intermediate cytogenetic risk, to a similar extent as the overall population, with OS and response rate also improved in the GO + DA arm, but not reaching statistical significance, consistent with the overall results. However, for patients with an unfavourable cytogenetic profile, OS [redacted] and RFS [redacted] outcomes appeared to be worse in the GO + DA arm, compared with the DA arm, whilst EFS results were similar [redacted].

Additional analyses provided to the ERG on request showed that the benefit seen in patients with an intermediate-1 cytogenetic and molecular risk profile was not found in patients with an intermediate-2 cytogenetic and molecular risk profile, suggesting potentially important heterogeneity in the broader 'intermediate' cytogenetic subgroup.

Whilst the proportion of patients experiencing an adverse event was similar between treatment groups, the proportion of patients experiencing a serious adverse event was higher in the GO + DA arm than the DA arm [redacted]. The most common serious adverse event in the GO + DA arm was thrombocytopenia [redacted]. All [redacted] patients who experienced veno-occlusive disease (VOD) had received GO. A higher proportion of patients in the GO + DA arm permanently discontinued treatment because of an adverse event than in the DA arm [redacted].

The CS also presented the results of an individual patient data (IPD) meta-analysis, conducted by Pfizer for use in regulatory submissions. Overall survival was statistically significantly improved in the GO arm for the overall population in the IPD meta-analysis [redacted] and for the subpopulation of patients with a favourable or intermediate cytogenetic profile [redacted], but not for patients with an unfavourable cytogenetic profile [redacted], supporting the evidence from the ALFA-0701 trial.

The ALFA-0701 trial and the IPD meta-analysis did not report quality of life outcomes.

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

The evidence for the clinical effectiveness of GO + DA is based on one reasonably good quality open-label RCT and the results are likely to be reliable. The ALFA-0701 trial included GO + DA at the recommended dose, which the company states is in line with the anticipated marketing authorisation.

The ALFA-0701 trial included patients aged 50-70 years with previously untreated de novo AML. This is a sub-population of the patient population described in the final NICE scope, and the anticipated marketing authorisation, which specifies patients age 15 years and above. The clinical advisor to the ERG stated that GO is unlikely to work differently in patients under 50 years old, and that GO is unlikely to be used extensively in patients over 70 years old, due to the poorer prognosis of older patients. The majority of patients diagnosed with AML are over 50 years of age, therefore, the ALFA-0701 trial is likely to be reflective of the majority of patients who would be eligible for treatment with GO + DA in clinical practice.

The ALFA-0701 trial included patients with AML, regardless of CD33-positivity, so patients in the trial who were not CD33-positive would not be eligible for GO + DA in clinical practice under the anticipated licence. [redacted]

The IPD meta-analysis, presented as supporting evidence, included patients aged 15 years or older with newly diagnosed AML (either de novo or secondary), or high-risk myelodysplastic syndrome (MDS), which is a broader population than that defined in the decision problem or the anticipated licence. Therefore, the results may not be entirely generalisable to patients eligible for GO + DA in clinical practice.

1.4 Summary of cost effectiveness submitted evidence by the company

The company's economic submission included a systematic review of published evidence on the cost-effectiveness, health-related quality of life, resource use and costs associated with GO in the treatment of AML. The review identified a number of economic evaluations of other therapies for AML, including UK based economic evaluations which were used to inform model parameters in the analysis, but did not identify any relevant economic assessments of GO.

The cost effectiveness of GO+DA compared with DA alone was informed by an economic evaluation conducted by the company. The population included in the company's decision problem and economic model comprises patients not known to have unfavourable cytogenetics. The company's model used a semi-Markov cohort state-transition model. The model used the time-to-event data from the ALFA-0701 clinical trial to determine the movement of patients between the health states. The model structure consisted of the following health states: (i) induction therapy; (ii) complete remission (CR/CRp): consolidation therapy; (iii) CR/CRp: off-treatment; (iv) refractory (receiving salvage therapy); (v) refractory (receiving non-curative therapy); (vi) hematopoietic stem cell transplant (HSCT); (vii) post-HSCT CR/CRp (without graft versus host disease (GVHD)); (viii) post-HSCT CR/CRp (with GVHD); (ix) relapse (receiving salvage therapy), (x) relapse (receiving non-curative therapy); (xi) functionally cured; and (xii) death. Patients could receive (hematopoietic stem cell transplant) HSCT after achieving complete remission (CR/CRp) after induction therapy, after receiving salvage therapy if they were refractory to first-line therapy, and after receiving salvage therapy if they experienced a relapse.

The efficacy data, treatment and comparator dosage, proportions of patients receiving induction and consolidation therapy, rates of HSCT, adverse event rates and patient characteristics (age, gender, weight, body surface area) used in the economic model were sourced from the ALFA-0701 clinical trial, with the remaining inputs informed by studies identified in the cost-effectiveness review and other sources. Overall survival and relapse-free survival was estimated using mixture cure models fit to Kaplan-Meier data from the ALFA-0701 trial. Patients remaining alive or relapse-free at 60 months after induction therapy or after HSCT were assumed to be cured and experienced general population mortality adjusted to reflect excess mortality in AML survivors. Probabilities of HSCT were estimated from ALFA-0701 trial data and were applied at distinct time points during the first 60 months in the model.

In the base-case analysis of patients excluding those with known, unfavourable cytogenetics, the company found GO+DA to be more costly (cost difference of £[redacted]) and more effective ([redacted] QALY gain) compared with DA alone. The deterministic base-case incremental cost-effectiveness ratio (ICER) was £12,251 per QALY, and the mean probabilistic ICER was £13,600 per quality-

adjusted life year (QALY). The predicted probability that GO+DA was cost-effective compared with DA alone was 65.7% at a cost-effectiveness threshold of £20,000 per QALY and 80.3% at a cost-effectiveness threshold of £30,000 per QALY. The company reported that the most influential parameters in the one way sensitivity analysis included the probability of HSCT from relapse in years 1 and 2 for the DA group, and the restricted mean survival time (RMST) for relapsed patients. In a series of scenario analyses, the ICER were demonstrated to vary between £6,821 (best fitting standard parametric functions for RFS and OS in CR/CRp patients) and £20,334 per QALY (when the RMST for relapsed and refractory patients was based on individual treatment arms). The company also presented the results of the analysis for all patients, which had an associated ICER of £20,457. The ICER was higher in the broader patient population due to higher incremental costs (cost difference of £[redacted]) and fewer incremental QALYs ([redacted] QALY gain), due to the lower effect of GO+DA in patients with unfavourable cytogenetics.

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

The economic analysis presented by the company was considered to meet the decision problem specified in NICE's scope. However, the ERG identified a number of key uncertainties.

The ERG considers that the state-transition modelling approach introduces unnecessary complexity compared to a simpler partitioned survival analysis (PartSA) approach. A series of structural assumptions (e.g. sub-states and HSCT) were imposed which ultimately resulted in the same independence assumption between clinical events (EFS and OS) that underpin the PartSA approach. Additionally, the model did not include an explicit structural link between relapse and HSCT. The important cost-offsets assumed for HSCT are predicated on the functional cure assumption (i.e. that those patients who have not relapsed by five years will never relapse and hence will not require HSCT at some point in the future).

While the ERG agrees with the company's decision to exclude patients with known unfavourable cytogenetics, the ERG does not believe that the company has sufficiently addressed the heterogeneity in the subgroup of patients with unknown cytogenetics and within the intermediate population. The ERG considers that there remains significant heterogeneity within the base-case population which may have important implications concerning the difference in the cure fraction for further subgroups within the overall population.

The ERG considers that there remains significant uncertainty surrounding the long-term morbidity and survival of functionally cured patients. This was addressed by the company by applying an adjustment for excess mortality for functionally cured patients for OS; however, the use of general population quality of life was not internally consistent with the excess mortality. Given that

functionally cured patients are assumed to be at higher mortality risk than the general population, the ERG concluded that it would appear reasonable to assume that functionally cured patients would have lower quality of life than that of the general population.

The ERG identified several areas of uncertainty in some of the parameter values used in the model. The first of these related to the costing assumptions for veno-occlusive disease (VOD) (specifically the inclusion of VOD for DA alone) and for HSCT used in the model. Secondly, for the response data and for induction and consolidation therapies, the ERG did not consider that it was necessary or appropriate to pool response data. Even if the differences in rates are not considered clinically important, any differences should still be considered and the lack of statistical significance should be reflected in the distributions assigned to these parameters. The ERG also considers that that the initial treatment costs of the induction and consolidation therapies should be based directly on the investigator-assessed (IA) response outcomes, rather than attempting to adjust the independent review committee (IRC)-assessed response outcomes in the manner proposed by the company. Lastly, the ERG identified additional uncertainties in the estimate of the HR associated with excess mortality in functionally cured patients due to the small number of patients on which this was based.

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

1.6.1 Strengths

The clinical effectiveness evidence is derived from a reasonably good quality RCT which included GO + DA at the recommended dose, in line with the anticipated marketing authorisation.

The company's economic submission met the requirements of the NICE reference case. The company submission was informed by data from a reasonably good quality RCT comparing the two interventions head to head. Where the pivotal trial for GO did not provide parameter estimates, these were developed in accordance with clinical and economist involvement. The economic model accommodated a number of key clinical elements of the treatment of AML and incorporated a range of scenario analyses that allowed the impact of alternative assumptions to be explored. The predictions for the company base-case demonstrated internal validity and consistency with the ALFA-0701 trial results. The company leveraged the use of more complex survival models, including mixture cure models and spline-based analyses, allowing the complex instantaneous risk of events associated with this disease area to be captured. The company provided extensive additional evidence and analyses in response to the ERG's points for clarification.

1.6.2 Weaknesses and areas of uncertainty

The ALFA-0701 trial included only patients aged 50-70 years, whilst the anticipated marketing authorisation includes patients age 15 years and above; however, the majority of patients diagnosed

with AML are over 50 years of age, therefore the population included in the trial is likely to be reflective of the majority of patients eligible for GO in clinical practice. The main clinical evidence is based on a single trial, although its results were supported by those from the IPD meta-analysis. The IPD meta-analysis included patients with de novo or secondary AML or high-risk myelodysplastic syndrome, which is a broader population than the anticipated licence, so results may not be entirely generalisable to patients eligible for GO in clinical practice.

The ALFA-0701 trial and the IPD meta-analysis did not report quality of life outcomes.

The principle weakness of the economic evidence submitted by the company relates to the model structure. In particular, there was a lack of an explicit structural link between a number of key model parameters, most importantly between relapse and HSCT. While the company base-case analysis provided predictions that demonstrated internal validity and consistency with the ALFA-0701 trial results, the absence of a structural link restricted the ability of the model to explore alternative scenarios in an appropriate manner, and, therefore, to fully capture the uncertainty in the modelled results.

There are significant areas of uncertainty in the cost-effectiveness analysis. Firstly, there was heterogeneity within the base-case population which may have important implications concerning the difference in the cure fraction for further subgroups within the overall population and the cost-effectiveness of GO+DA in these groups. A second area relates to uncertainty regarding the long-term morbidity and survival of functionally cured patients. There were also uncertainties surrounding the costing assumptions for VOD and for HSCT.

1.7 Summary of exploratory and sensitivity analyses undertaken by the ERG

The ERG did not conduct any further sensitivity analyses relating to clinical effectiveness.

The ERG conducted a series of exploratory analyses exploring the robustness of the cost-effectiveness results to specific assumptions and additional uncertainties identified by the ERG. The ERG was unable to explore all the uncertainties identified in the CS, in particular with regard to the rate of HSCT and structural links between key model parameters, due to constraints relating to the available data and the model structure. The scenarios were not associated with substantial differences to the ICER; however, the ERG notes that due to the lack of structural links between key model parameters, the model has limited sensitivity to certain changes. The scenarios associated with the greatest impact on cost-effectiveness outcomes related to changes made to the HSCT costs, the quality of life in functionality cured patients and to the use of individual rates of response. The ERG also presented an

alternative base-case based on a combination of a number of these scenario analyses. The ERG's base-case makes the following amendments to the company's revised base-case:

1. Corrections for calculation errors;
2. Number of induction and consolidation therapy courses modelled for each arm individually, using the trial-observed rates;
3. Rates of response to treatment modelled for each arm individually;
4. The initial cost of HSCT estimated from NHS Reference Costs;
5. Removal of VOD events in the DA treatment group;
6. Exclusion of GVHD-specific costs;
7. Inclusion of hospital costs for the treatment of VOD;
8. Quality of life in functionally cured patients based on the utility value for off-treatment CR/CRp patients, and further adjusted for age;
9. Long-term mortality in functionally cured patients adjusted for excess mortality using the ERG-calculated hazard ratio.

The results of these scenario analyses including the ERG's base-case are summarised in Table 1. Due to time constraints, deterministic ICERs are presented throughout, with the exception of the ERG alternative base-case, which is based on the probabilistic analysis.

The ERG alternative base-case analysis estimated GO+DA to be more costly (cost difference [redacted]) and more effective ([redacted] QALY gain) compared with DA alone, and suggests that the ICER for GO+DA compared with DA is £17,956 per QALY.

The ERG also carried out a further series of exploratory subgroup analyses to explore the impact of heterogeneity between the subgroups included in the base-case population and possible variability within the intermediate group. The ICER of GO+DA versus DA alone varied between £16,343 (intermediate-1 only, defined by cytogenetic and molecular test) and £31,709 (intermediate only, defined by cytogenetic results only), allowing the ERG to conclude that further risk stratification using genetic and molecular testing may provide a clearer separation between subgroups. However, these findings can only be considered indicative due to data limitations, and uncertainties also remain concerning the practicality and feasibility of introducing additional risk stratification within routine clinical practice.

Table 1 Summary of ERG exploratory analyses

Scenarios	Treatments	Costs	QALYs	Inc. cost	Inc. QALY	ICER	Change in ICER
CS base case[$] (corrected)	***GO+DA***	[redacted]	[redacted]	[redacted]	[redacted]	£13,561	*n/a*
	DA	[redacted]	[redacted]	-	-	-	-
Individual rates for courses of treatment	GO+DA	[redacted]	[redacted]	[redacted]	[redacted]	£14,249	£688
	DA	[redacted]	[redacted]	-	-	-	-
Rate of response to treatment for individual arms	GO+DA	[redacted]	[redacted]	[redacted]	[redacted]	£10,526	-£3,035
	DA	[redacted]	[redacted]	-	-	-	-
Alternative HSCT costs	GO+DA	[redacted]	[redacted]	[redacted]	[redacted]	£16,003	£2,442
	DA	[redacted]	[redacted]	-	-	-	-
Without GVHD costs	GO+DA	[redacted]	[redacted]	[redacted]	[redacted]	£14,020	£459
	DA	[redacted]	[redacted]	-	-	-	-
Excluding VOD events in the DA alone group	GO+DA	[redacted]	[redacted]	[redacted]	[redacted]	£13,704	£143
	DA	[redacted]	[redacted]	-	-	-	-
Include hospital costs for the treatment of VOD	GO+DA	[redacted]	[redacted]	[redacted]	[redacted]	£13,733	£172
	DA	[redacted]	[redacted]	-	-	-	-
Alternative utility values for functionally cured patients	GO+DA	[redacted]	[redacted]	[redacted]	[redacted]	£13,878	£317
	DA	[redacted]	[redacted]	-	-	-	-
Alternative utility values for functionally cured patients, age-adjusted	GO+DA	[redacted]	[redacted]	[redacted]	[redacted]	£15,279	£1,718
	DA	[redacted]	[redacted]	-	-	-	-
ERG-estimated hazard ratio for long-term survival	GO+DA	[redacted]	[redacted]	[redacted]	[redacted]	£14,337	£776
	DA	[redacted]	[redacted]	-	-	-	-
ERG alternative base-case*	GO+DA	[redacted]	[redacted]	[redacted]	[redacted]	£17,956	£4,395
	DA	[redacted]	[redacted]	-	-	-	-
* based on probabilistic analysis							

2 Background

2.1 Critique of company's description of underlying health problem

The company's description of the underlying health problem is appropriate and relevant to the decision problem under consideration.

Leukaemia is a type of blood cancer that originates in the bone marrow. It is characterised by abnormal differentiation of haematopoietic stem cells and subsequent clonal over-proliferation of blood cells that cannot properly mature.[1, 2] Acute myeloid leukaemia (AML) involves overproduction of immature granulocytic and monocytic white blood cells, known as blasts.[1-3] AML is primarily a disease of the elderly, with incidence rising gradually from 40-44 years of age and then more steeply from 55-69 years of age.[4] In 2014 the incidence of AML in England was 5.2 per 100,000 population.[5] In 2015 there were 2471 new cases of AML in England;[4, 6] 73.6% patients are likely to have de novo AML (based on a large population-based study from Sweden),[7] therefore, 1819 patients would have been expected to have de novo AML.

The symptoms of AML include fever, fatigue, difficulty breathing, weight loss, bruising, bleeding, and aches and pains in the bones and joints.[2, 8-11] Prognosis is poor, particularly in older patients; 5-year overall survival is around 41% in patients aged 25-64 years and around 6% in patients aged 65 years or older.[12]

2.2 Critique of company's overview of current service provision

The company's overview of current service provision is appropriate and relevant to the decision problem under consideration.

The aim of treatment in AML is to achieve and maintain complete remission (CR). Treatment comprises two phases: induction (which aims to clear the bone marrow of all haematopoietic cells) followed by consolidation in patients who achieve CR (which aims to increase the durability of remission by eliminating all remaining disease).[11, 13] The duration of first CR is positively correlated with survival; patients who achieve CR for 3 consecutive years are unlikely to relapse.[11]

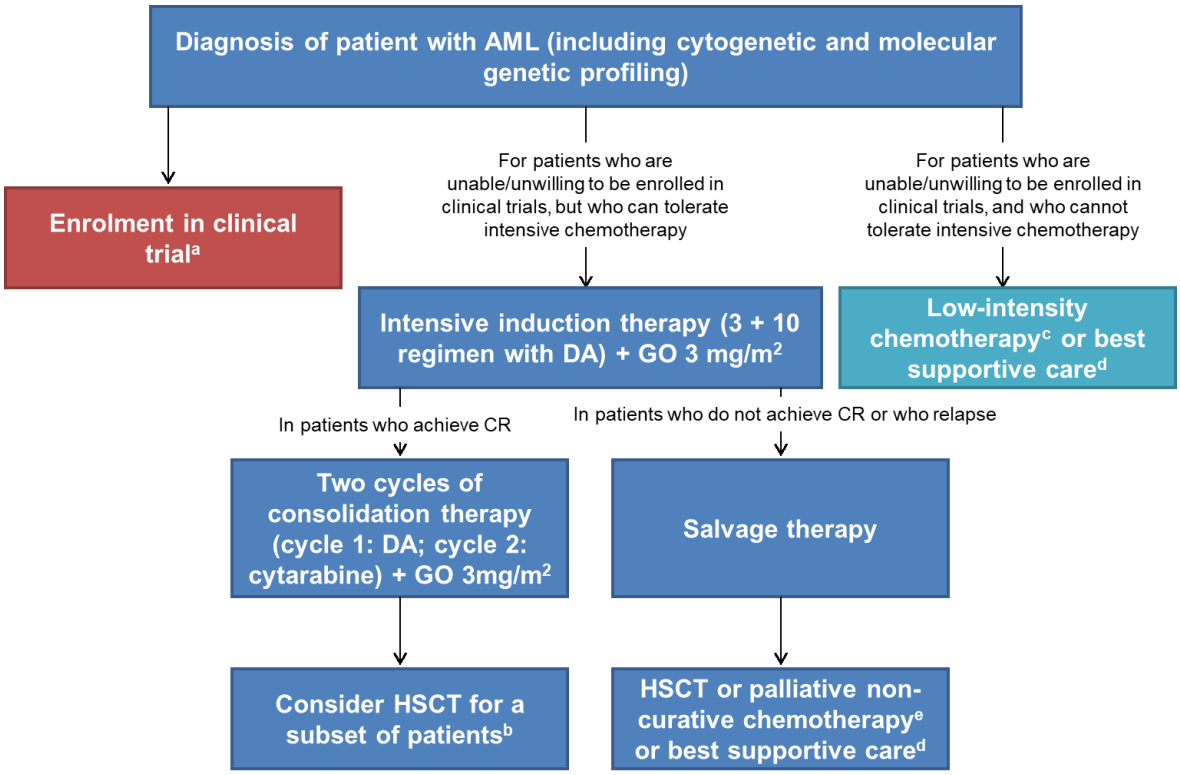
The company submission (CS) states that approximately 80% of UK AML patients are treated in the context of clinical trials. For patients who are not enrolled in clinical trials, who are able to tolerate intensive chemotherapy, the British Committee for Standards in Haematology (BCSH) guidelines recommend induction chemotherapy with daunorubicin (or another anthracycline) plus cytarabine (DA), administered over 3 and 10 days, respectively (known as the 3 + 10 regimen) or 3 and 7 days, respectively (known as the 3 + 7 regimen).[14] For patients unable to receive intensive chemotherapy, treatment with palliative, low intensity chemotherapy can be considered. Haematopoietic stem cell

transplant (HSCT) is reserved for patients who are at high risk of relapse (such as those with unfavourable cytogenetic profile) once they have achieved CR, or can be used in patients who relapse or who do not achieve CR following salvage therapy. The clinical advisor to the Evidence Review Group (ERG) confirmed that this is consistent with the treatment pathway for patients treated in NHS practice.

The CS states that gemtuzumab ozogamicin (GO) is an antibody-drug conjugate that combines a humanized, anti-CD33, monoclonal antibody with calicheamicin, a potent cytotoxic agent that causes DNA damage. CD33 is a sialic acid-dependent adhesion protein that is highly expressed on the surface of AML blast cells and on some leukaemic stem cells.[15, 16] The CD33 component of GO enables targeted delivery of calicheamicin to CD33-positive blast cells. *In vitro* studies showed that after a 3 mg/m^2 dose of GO, re-expression of CD33 to nearly pre-treatment levels occurred after 72 hours, which led to the hypothesis that repeated administration of lower, near-saturating, fractionated doses of GO may enable increased drug internalisation, while improving safety versus a higher, unfractionated dosing regimen. This fractionated 3 mg/m^2 dose is used in the ALFA-0701 trial, which is the pivotal trial on which regulatory submissions have been based and the dosing schedule is consistent with the anticipated European Medicines Agency (EMA) marketing authorisation.

The proposed position of GO in the treatment pathway is alongside DA. The current treatment pathway with the proposed placement of GO + DA is presented in Figure 1 (Figure 4 of the CS). This treatment pathway appears generally appropriate for patients with de novo AML.

Figure 1: Clinical pathway for the treatment of patients with de novo AML and the positioning of GO + DA in the clinical pathway (Figure 4 of CS)



[a]Guidelines recommend that patients are first enrolled in a clinical trial; [b]HSCT is offered in patients with unfavourable cytogenetics after first CR and considered in patients with intermediate cytogenetics; [c]Low-intensity chemotherapy options include low-dose cytarabine, azacitidine or hydroxycarbamide; [d]Best supportive care options include transfusion support and hydroxycarbamide; [e]Palliative non-curative chemotherapy can include azacitidine.

2.2.1 Cytogenetic and molecular testing

Cytogenetic testing is used to predict response to treatment; patients with a favourable or intermediate cytogenetic profile have a better prognosis than those with an unfavourable cytogenetic profile with regard to treatment response, risk of relapse and survival.[3, 17, 18] The incidence of unfavourable cytogenetic abnormalities increases with increasing age.[11] However, not all patients will receive a cytogenetic profile classification and will be classified as having an unknown cytogenetic risk.

The particular cytogenetic profile classification that a patient receives is based on the system that is used to define their karyotype; the revised Medical Research Council (MRC) classification is based on cytogenetic abnormalities and stratifies patients into favourable, intermediate or adverse (referred to as 'unfavourable' in the CS) cytogenetics profiles.[19] The European LeukemiaNet (ELN) classification recommends that both cytogenetic and genetic abnormalities are taken into account (using molecular testing) when assigning patients to favourable, intermediate-1, intermediate-2 or adverse cytogenetics profiles.[20, 21]

3 Critique of company's definition of decision problem

3.1 Population

The patient population addressed in the CS is adult patients not known to have unfavourable cytogenetics, with previously untreated, de novo AML. The CS notes that this is a subpopulation of the anticipated marketing authorisation for GO + DA and the population described in the final scope issued by NICE, which was adults with untreated AML. The CS claims that patients known to have unfavourable cytogenetics would not be treated with GO plus intensive chemotherapy in NHS clinical practice (as confirmed to the company by clinicians treating patients with AML) and that the subpopulation reflects where GO + DA provides clinical benefit and, consequently, optimises the cost effectiveness of GO + DA versus DA alone.

The clinical advisor to the ERG advised that cytogenetic test results should usually be available within five days. However, in view of the very short time between diagnosis and treatment in patients with AML, the requirement for cytogenetic test results prior to commencing treatment with GO could potentially delay the start of treatment. The way in which cytogenetic testing is used in the treatment of AML will change from informing prognostics, to guiding initial therapy. There is some evidence to suggest that delaying treatment initiation for up to a week does not significantly affect prognosis.[22]

The clinical effectiveness evidence presented is primarily from a single randomised controlled trial (RCT); the ALFA-0701 trial. The ERG notes that in the ALFA-0701 trial cytogenetic results were not available for 25/271 (9%) patients in the modified intention to treat (mITT) population; this included patients for whom the karyotype was not done, or done but inadequate. The clinical advisor to the ERG commented that in clinical practice around 5% of patients are likely to fail cytogenetic testing (have an 'unknown' risk); these patients may be initially treated as intermediate risk patients, until additional molecular test results are available.

In the clinical effectiveness section of the CS, efficacy results for the subgroup of patients with favourable and intermediate cytogenetics are presented separately, excluding those with unknown cytogenetics. In the economic model, patients with favourable, intermediate and unknown cytogenetics are used. The subgroup used in the model reflects the population defined in the decision problem, i.e. patients not known to have unfavourable cytogenetics.

On Day 180 of the EMA regulatory process, after the company had submitted their evidence to NICE, the population was extended to also include patients aged 15-17 years old but restricted in terms of disease characteristics to CD33-positive AML. On 22 February 2018 the Committee for Medicinal Products for Human Use (CHMP) adopted a positive opinion, recommending the granting of a

marketing authorisation for GO in 'combination therapy with daunorubicin (DNR) and cytarabine (AraC) for the treatment of patients age 15 years and above with previously untreated, de novo CD33-positive acute myeloid leukaemia (AML), except acute promyelocytic leukaemia (APL)'.[23]

The ALFA-0701 trial included patients aged 50-70 years with previously untreated de novo AML. Therefore, this is a sub-population of the patient population described in the final NICE scope, in which no age restriction was applied, and the anticipated marketing authorisation. The clinical advisor to the ERG stated that GO is unlikely to work differently in patients under 50 years old, and that GO is unlikely to be used extensively in patients over 70 years old, due to the poorer prognosis of older patients, who are more likely to have comorbidities and be less able to tolerate intensive chemotherapy. The majority of patients diagnosed with AML are over 50 years of age, therefore, the population included in the ALFA-0701 trial is likely to be reflective of the majority of patients who would be eligible for treatment with GO + DA in clinical practice.

The ALFA-0701 trial included patients with AML, regardless of CD33-positivity, therefore, patients in the trial who were not CD33-positive would not be eligible for GO + DA in clinical practice under the anticipated licence.

The CS also included an individual patient data (IPD) meta-analysis as supporting evidence in Appendix D.3.1 of the CS. Patients included in the IPD meta-analysis were aged 15 years or older with newly diagnosed AML (either de novo or secondary), or high-risk myelodysplastic syndrome (MDS), which is a broader population than that defined in the decision problem or the anticipated licence.

3.2 Intervention

The intervention specified in the NICE scope is GO in combination with chemotherapy. The intervention addressed in the CS is GO in combination with daunorubicin plus cytarabine (DA; intensive chemotherapy). GO does not currently have EMA marketing authorisation for the indication in this submission. On 22 February 2018 the CHMP adopted a positive opinion, recommending the granting of a marketing authorisation for GO in combination with DA.[23] The anticipated date of EMA approval is in Q2 2018.

The CS states that GO is administered as an intravenous infusion in combination with DA. The recommended dose is 3 mg/m^2/dose (up to a maximum of 5 mg/dose) infused over 2 hours on days 1, 4 and 7 of the induction therapy course and as a 3 mg/m^2/dose (up to a maximum of 5 mg/dose) infused over 2 hours on day 1 of consolidation therapy, for up to two cycles. The company states that this is in line with the anticipated marketing authorisation. This dose was used in the ALFA-0701

trial. Two ongoing trials which include UK treatment centres, AML18 and AML19, use different dosing regimens; both trials include a comparison of single dose GO (3 mg/m2, up to a maximum of 5 mg/m2) with a fractionated dose (3 mg/m2 on 2 days, up to a maximum of 5 mg/m2 per day) as part of standard induction chemotherapy with DA. Therefore, there may be implications for dosing in practice, if GO is approved, as UK clinicians are currently using the AML18 and AML19 doses, rather than the dose used in the ALFA-0701 trial, which is the dose in the anticipated marketing authorisation.

The IPD meta-analysis presented as supporting evidence included different doses of GO; 3 mg/m^2 single dose, 3 mg/m^2 fractionated dose and 6 mg/m^2 single dose. The chemotherapy regimen used alongside GO also differed in the trials included in the IPD meta-analysis, as detailed below in Section 3.3.

3.3 Comparators

The comparator specified in the NICE scope is "established clinical management without GO, including but not limited to amsacrine, cytarabine, daunorubicin, etoposide, fludarabine, idarubicin, mitoxantrone, thioguanine and midostaurin (subject to ongoing NICE appraisal)". The comparator addressed in the CS is that used in the ALFA-0701 trial: 60 mg/m^2/day of daunorubicin on days 1–3 and cytarabine 200 mg/m^2/day on days 1–7 as induction therapy and two consolidation courses (1) 60 mg/m^2 of daunorubicin on day 1 and cytarabine 1000 mg/m^2/every 12 hours on days 1–4, (2) 60 mg/m^2/day of daunorubicin on days 1–2 and cytarabine 1000 mg/m^2/every 12 hours on days 1–4.

The CS states that the rationale for the difference is that NHS England clinical expert opinion is that outside of clinical trials the standard of care for intensive chemotherapy is DA. The clinical advisor to the ERG confirmed that DA is the standard chemotherapy regimen used for patients with de novo AML who can receive intensive chemotherapy. For induction therapy, the 2017 ENL guidelines recommend 3 days of an anthracycline, such as 60 mg/m^2 daunorubicin, and 100-200 mg/m^2 cytarabine on a 3+7 regimen.[21] The 2006 BCSH guidelines recommend 200 mg/m^2 cytarabine with an anthracycline/anthracycline-like drug, usually 45-60mg/m^2 of daunorubicin given in a 3+7 or 3+10 regimen. Evidence suggests that a 3+10 regimen is not superior to a 3+7 regimen.[14] As the company's comparator for induction therapy appears to reflect clinical practice in the UK, the ERG considers this to be appropriate.

For consolidation therapy regimens, the 2017 ELN guidelines recommend 2 – 4 cycles of intermediate dose (1000-1500 mg/m^2 on days 1-3 or days 1-5 or 6) cytarabine (IDAC).[21] The company's comparator for consolidation therapy included daunorubicin alongside IDAC for 2 courses. Clinical advice to the ERG was that standard consolidation therapy in the UK would be 3 courses of

consolidation therapy; 1 course of daunorubicin alongside cytarabine and 2 courses of high dose cytarabine. The ELN guidelines state that consolidation regimens include single-agent cytarabine at high doses and multiagent chemotherapy, which lead to similar outcomes. Therefore, as the consolidation therapy used in the trial is similar to that used in clinical practice, and recommended in the ELN guidelines, it is likely to be appropriate.

The IPD meta-analysis presented as supporting evidence included different chemotherapy regimens as comparators and alongside GO; DA, daunorubicin + cytarabine + etoposide (ADE), fludarabine + cytarabine + G-CSF + idarubicin (FLAG-Ida), daunorubicin + clofarabine (DClo) or DA + growth factor.

3.4 Outcomes

The outcome measures specified in the NICE scope were reported in the CS; event-free survival (EFS), overall survival (OS), disease-free survival, adverse effects of treatment and health-related quality of life (HRQoL). In addition, the CS included relapse-free survival (RFS) and response rate. The primary outcome in the ALFA-0701 trial was EFS, secondary outcomes were OS, haematological remission (assessed as the proportion of patients achieving complete remission/complete remission with incomplete platelet recovery (CR/CRp)), RFS and adverse events. The ALFA-0701 trial did not collect HRQoL or utility data.

The CS states that patients who have been in CR for three consecutive years have little risk of relapsing and UK clinical experts consider patients to be 'functionally cured' when they have been in CR for three to five years, meaning that these patients' mortality risk can be considered equivalent to that of the general population. However, the CS acknowledges that there may be some excess long-term mortality risk associated with cured AML patients, because such patients may be more susceptible to cancer occurring or re-occurring or may experience a higher risk of mortality associated with their cancer treatments. In the economic model, patients are considered to transition to a functionally cured health state if they have remained in CR or CRp for between three and five years. The base-case assumed five years and applied a hazard ratio of [redacted] corresponding to a calculated excess mortality risk of [redacted] compared to that of the general population.

The independent review committee (IRC) analyses presented in the CS and used in the economic model at the 30 April 2013 data cut-off had a minimum follow-up of three years. There is evidence to suggest that, whilst uncommon, some late relapse in AML does occur in patients who have been in remission for more than five years, for example in Gardin et al. (2013), estimated relative survival appears to plateau at seven to eight years rather than three to five.[24] Other evidence suggests the risk of late relapse (after five years) to be between 1.16 and 3%.[25] Whilst not substantially likely, it is

therefore possible that the clinical evidence from the ALFA-0701 trial did not capture all relevant long-term events and some patients that were deemed functionally cured at last follow-up could have relapsed at a later time point, requiring further treatment.

3.5 Other relevant factors

The CS stated that there are no equality considerations expected for GO.

3.6 Summary

The clinical effectiveness evidence was based on a single trial, with a population aged 50-70 years, which is a sub-population of the patient population described in the NICE scope, in which no age restriction was applied. The clinical advisor to the ERG advised that GO is unlikely to be less effective in patients under 50 years old and would be unlikely to be used extensively in patients over 70 years old.

The ALFA-0701 trial included patients with AML, regardless of CD33-positivity. This is contrary to the anticipated marketing authorisation, which restricts the use to patients with CD33-positive AML.

The decision problem restricts the use of GO to patients not known to have unfavourable cytogenetics, which differs from the anticipated marketing authorisation, which does not restrict patients based on their cytogenetic profile.

HRQoL was not assessed in the ALFA-0701 trial. Patients who have been in remission for 3-5 years are considered to be ‘functionally cured’. However, whilst uncommon, some patients may relapse later than 5 years; these longer term events may not have been captured in the ALFA-0701 trial data.

4 Clinical Effectiveness

This section contains a critique of the methods of the systematic review of clinical effectiveness data, followed by a description and critique of the trials included in the review, including a summary of their quality and results and the results of any synthesis of studies. The ERG's conclusions on the clinical effectiveness of GO for treating de novo untreated AML are presented at the end of this section.

4.1 Critique of the methods of review(s)

The company conducted two systematic literature reviews, one designed to identify all published RCTs relating to the efficacy and safety of treatments for adult patients with previously untreated, de novo AML and one designed to identify all published non-RCT evidence relating to the efficacy and safety of treatments for adult patients with previously untreated, de novo AML.

4.1.1 Search strategy

The literature searching for each systematic review was carried out and reported separately. Section D.1.1.1 of the CS appendices reports on the search strategy for RCT clinical evidence and Section D.1.2.1 reports on the search strategy for non-RCT clinical evidence. The search strategies used and the number of records identified for each database are reported in Tables 53 to 55 and Tables 66 to 68 of the CS, respectively.

For the systematic review of RCT evidence an appropriate range of databases, trial registries and conference proceedings were searched in May 2017. The overall structure of the database search strategies was appropriate: terms for AML were combined with terms for GO and other relevant drug interventions (daunorubicin, cytarabine, FLAG protocol). A search filter was included in the strategy to restrict the results to RCTs. The strategies contained relevant subject headings, text word searches and synonyms. There appears to be no errors in how the search sets were combined or typographical errors within the search terms. The number of records identified matches the number reported in the Preferred Reporting Items for Systematic Review and Meta-analyses (PRISMA) diagram (Figure 21 of the CS).

While the CS states that systematic reviews identified by the literature searches were reviewed for additional publications, the search strategies of MEDLINE and Embase were restricted to identifying RCTs so it is unlikely that all systematic reviews would have been identified.

For the systematic review of non-RCT evidence, an appropriate range of databases and conference proceedings were searched in May 2017. The overall structure of the database search strategies was appropriate: terms for AML were combined with terms for GO and other relevant drug interventions

(daunorubicin, cytarabine, FLAG protocol). Various limits were included in the search strategies to exclude reports of RCTs, case reports, and animal studies from the results. The strategies contained relevant subject headings, text word searches and synonyms. There appears to be no errors in how the search sets are combined or typographical errors within the search terms themselves. The number of records identified matches the number reported in the PRISMA diagram (Figure 22 of the CS).

4.1.2 Inclusion criteria

The inclusion criteria for the systematic review of RCT evidence were presented in Table 57, Section D.1.1.2 of the CS appendices. Eligible studies were phase II and III RCTs of GO (monotherapy or combination therapy) + DA using the comparators DA or FLAG-Ida, for adult patients ≥ 18 years old with newly diagnosed AML receiving induction therapy. The inclusion criteria for the systematic review of non-RCT evidence were presented in Table 70, Section D.1.2.2 of the CS appendices. Eligible studies were uncontrolled, or non-randomised interventional studies or single-arm trials of GO (monotherapy or combination therapy) + DA using the comparators DA or FLAG-Ida, for adult patients ≥ 18 years old with newly diagnosed AML receiving induction therapy. However, studies investigating FLAG-Ida were subsequently excluded from both systematic reviews, post hoc. To be eligible for inclusion, studies had to report at least one relevant efficacy or safety outcome. Only studies reported in English were eligible for inclusion.

The methods used to screen and select relevant studies was to a good standard, with two blinded independent reviewers screening titles and abstracts and full texts using pre-defined inclusion criteria, with disagreements resolved by a third reviewer. A complete list of studies excluded at the full paper stage, with the reason for exclusion, was provided for both systematic reviews in Sections D.1.1.3 and D.1.2.3.

4.1.3 Data extraction

Data were extracted by a single reviewer and verified by a second, reducing the risk of error and bias in data extraction. Adequate data for the ALFA-0701 trial were presented in the CS with a detailed summary of trial methods presented in Table 7.

4.1.4 Quality assessment

The ALFA-0701 trial was assessed for quality using appropriate criteria for RCTs; the trial was reasonably good quality (see Section 4.2.3 for further details). The quality assessment results are presented in Table 12 of the CS, which was checked by the ERG.

4.1.5 Evidence synthesis

Eight RCTs were included in the systematic review of RCTs, however, seven RCTs did not use a dose or dosing schedule that is expected to be approved by the EMA. Therefore only the ALFA-0701 trial was used to address the decision problem in the CS. Four of the excluded RCTs, along with the ALFA-0701 trial, were included in an IPD meta-analysis, presented in Section D3 of the CS appendices and described in Section 4.3.1 of this report.

4.1.6 Conclusions from critique of systematic review methods

The company conducted two systematic literature reviews, one for RCT evidence and one for non-RCT evidence, to a good standard. The search strategy was generally appropriate, it is unlikely that any relevant RCTs of GO have been missed. Inclusion and exclusion criteria for the reviews were clearly stated and were generally appropriate, although studies investigating FLAG-Ida were initially included in the reviews and subsequently excluded, post hoc. Data extraction was undertaken by one reviewer and checked by a second, reducing the risk of error and bias, although it is unclear whether the same process was used for quality assessment. Adequate details of the methods of the ALFA-0701 trial were presented, along with a table of quality assessment results; the trial was reasonably good quality. Seven RCTs of GO that met inclusion criteria for the review were deemed to be unable to address the decision problem in the CS and were not included in the economic model, because they did not use a dose or dosing schedule of GO that is expected to be approved by the EMA. However, four of the trials, along with the ALFA-0701 trial, were included in an IPD meta-analysis, presented as supporting evidence.

4.1.7 Ongoing studies

The systematic reviews identified two ongoing trials of GO: AML18 and AML19. These are described in Section D.1.1.4 of the CS appendices.

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

4.2.1 Trials included in the review

Eight RCTs were included in the systematic review of RCT evidence and no non-RCTs were identified for inclusion in the review of non-RCT evidence. The ALFA-0701 trial is the pivotal study used to support the EMA marketing authorisation. The remaining seven RCTs: MRC AML15, NCRI AML16, MRC AML17, SWOG S0106, GOELAMS AML 2006 IR, AML18 and AML19 did not use a dose or dosing schedule of GO that is expected to be approved by the EMA. Therefore, only the ALFA-0701 trial is the primary source of clinical effectiveness evidence for the appraisal. The other seven RCTs were described in Section D.1.1.4 of the CS appendices.

4.2.2 The ALFA-0701 trial

ALFA-0701 was a phase 3 multicentre open label RCT undertaken at 26 haematology centres in France. France. Patients were enrolled between January 2008 and November 2010. The trial methodology is summarised in Figure 2: ALFA-0701 study design (Figure 5 of CS)

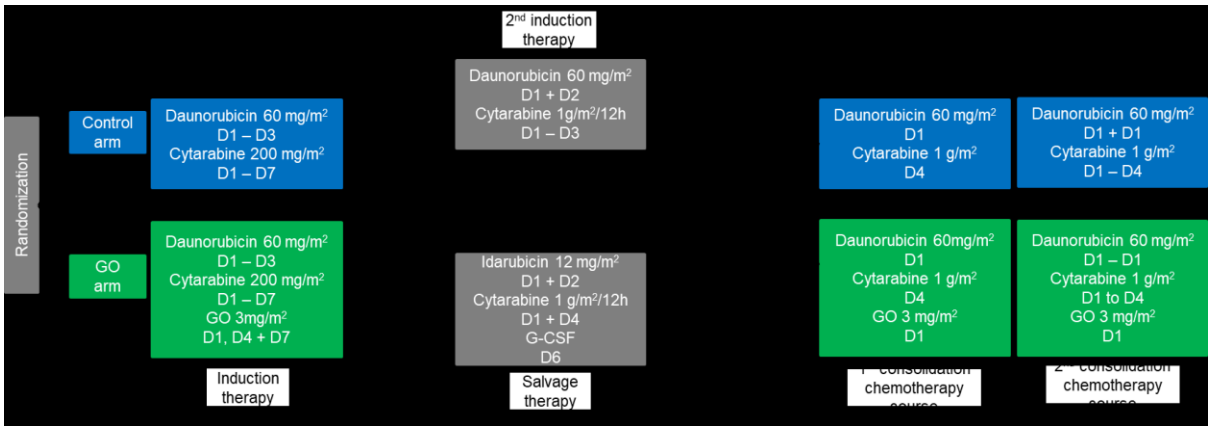


Table 2 (adapted from Table 7 of the CS) and the trial design is presented as Figure 2 (Figure 5 of the CS).

The ERG notes that only patients with intermediate-2 or unfavourable cytogenetic and molecular risk categories were eligible for HSCT, after achieving CR after induction therapy or salvage therapy (Figure 2: ALFA-0701 study design (Figure 5 of CS)

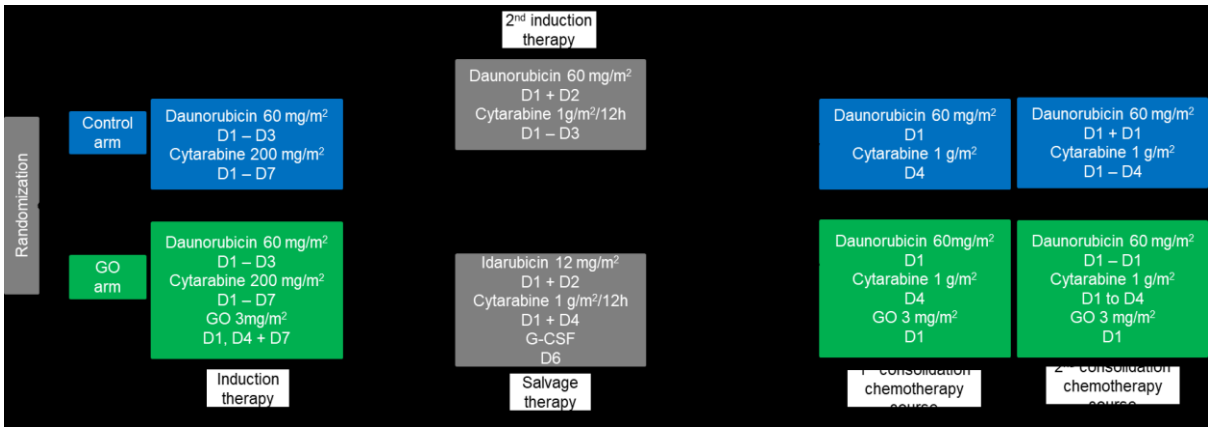


Table 2). The clinical advisor to the ERG stated that this is reflective of clinical practice, where only poor risk and some intermediate risk patients receive HSCT.

Figure 2: ALFA-0701 study design (Figure 5 of CS)

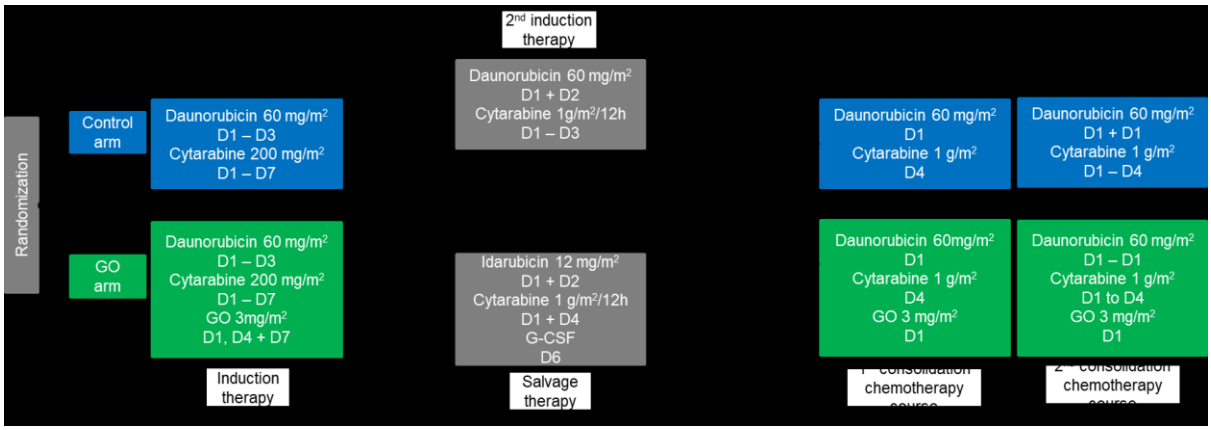


Table 2: Summary of ALFA-0701 trial methodology (adapted from Table 7 of CS)

Trial design	Investigator-sponsored, randomized (1:1 randomization), open-label, phase 3 trial comparing GO + DA versus DA alone in patients aged 50–70 years of age with de novo, untreated AML
Trial drugs **Intervention(s) (n = 135)** **Comparator(s) (n = 136)**	**Induction therapy:** (DA arm) Intravenous daunorubicin (60 mg/m² on days 1–3) + cytarabine (200 mg/m² as continuous infusion for 7 days (GO + DA arm) Intravenous fractionally dosed GO (3 mg/m² [maximum dose 5 mg] infused over 2 h on days 1, 4 and 7; **Salvage therapy:** Patients who did not achieve a CR after the first course of induction therapy and who did not receive the second course of induction therapy could receive a salvage course: idarubicin (12 mg/m^2; days 1 and 2), cytarabine (1 g/m^2 twice a day; days 1 and 4) and granulocyte colony-stimulating factor (G-CSF) (day 6). To be eligible for salvage therapy, patients needed an ECOG PS score of < 3 and creatinine clearance > 30 mL/min. Patients who did not respond to induction therapy (including salvage course) discontinued study treatment. Patients who achieved CR/CRp after induction or salvage therapy went on to receive two courses of consolidation therapy according to their initial randomization. **Consolidation therapy (two courses):** (DA arm) Intravenous daunorubicin (60 mg/m² for 1 day [first course] or 2 days [second course]) in combination with intravenous cytarabine (1000 mg/m² per 12 h, infused over 2 h on days 1–4), without (GO + DA arm)Intravenous GO (3 mg/m² on day 1 of each course of consolidation therapy;
Permitted and disallowed concomitant medication	**HSCT:** Patients who experienced CR could be considered for allogeneic transplant according to PS, age, the existence or not of a related donor, and cytogenetic and molecular risk categories. Patients with favourable and intermediate-1 cytogenetic and molecular risk categories were not to be sent for transplant in first CR; patients with intermediate-2 or unfavourable cytogenetic and molecular risk categories who experienced a CR were considered for transplant if qualified by other criteria.

	Permitted and disallowed concomitant medications were not specified
Duration of study	Minimum duration of follow-up: 3 years
Blinding	Open-label: the study treatment was not blinded to patients or investigators A retrospective analysis to confirm investigator-collected data and to collect additional data as per regulatory requirements was performed and investigators of the IRC were blinded to the treatment arms
Primary outcomes	EFS: time from date of randomisation to date of induction failure, relapse or death due to any cause, whichever came first.
Secondary/tertiary outcomes	OS: time from randomization to the date of death due to any cause. Haematological remission: assessed as the proportion of patients achieving CR/CRp; RFS: defined for patients experiencing remission as time from the date of remission to the date of relapse or death from any cause, whichever came first. Safety: A retrospective safety data assessment was conducted to collect treatment-emergent AEs (TEAEs) of special interest, based on review of patient medical files. For veno-occlusive disease (VOD), data were collected until the patient's death or the retrospective data cut-off of 1 November 2013, whichever occurred first.
Pre-planned subgroups	Age, WBC count, ECOG PS, CD33 expression, CD33 MFI ratio, NCCN risk classification, ELN risk classification, FLT3-ITD status, NPM1 status, CEBPA status, MLL status, WT1 status, cytogenetics profile as classified by CHV, genotype Subgroup analysis according to IRC was evaluated in all subgroups with the exception of CD33 MFI ratio

4.2.2.1 Study endpoints

After data were transferred to Pfizer, additional analyses and data collection were conducted. Two independent reviewers from the IRC undertook blinded analysis of ALFA-0701 data, with discrepancies resolved by a third.

The CS presents data from the IRC analyses at the 30 April 2013 data cut-off (minimum length of follow-up 3 years), the economic model also uses these analyses. In addition, data from investigator assessment (IA) analyses at a cut-off date of 1 August 2011 have also been presented for some of the outcomes; these are the analyses presented in the publication for the ALFA-0701 trial.[26] Details of the efficacy outcomes assessed by the IRC and IA are presented in Table 8 of the CS.

4.2.2.2 Trial populations

As discussed in Section 3.1, the population in the ALFA-0701 trial comprised only a subset of the anticipated licenced population, as only patients aged 50-70 years were included in the trial. The clinical advisor to the ERG advised that older patients (over 70 years) more often have unfavourable cytogenetic profiles and are more likely to have co-morbidities, therefore are less able to tolerate intensive chemotherapy. The majority of patients diagnosed with AML are over 50 years of age, therefore, the population in the trial is likely to be reflective of the majority of patients who would be eligible for treatment with GO + DA in clinical practice. The anticipated marketing authorisation for GO specifies patients with CD33-positive AML, whilst the trial also included AML patients who were not CD33-positive, therefore, a small proportion of patients in the trial would not be eligible for GO + DA in clinical practice.

Originally the trial included 280 patients, however informed consent forms were not transferred for nine patients when data were transferred to Pfizer. The data were therefore analysed for 271 patients in a modified intention-to-treat (mITT) population. Participant baseline characteristics are summarised in Table 3 for the mITT population (adapted from Table 71 of CS appendices).

Participant characteristics were generally balanced at baseline, although a higher proportion of patients in the GO + DA arm were male (54.8% versus 44.1%) and aged ≥ 60 years (71.9% versus 61.8%). Median white blood cell count was also higher in the GO + DA arm compared to the DA arm (median 5.8 x 10^9/L versus 4.1 x 10^9/L). Older patients generally have less favourable outcomes, so this baseline imbalance in age could have resulted in potentially poorer outcomes in the GO + DA arm. A similar proportion of patients had favourable/intermediate cytogenetic profiles in each arm (69.6% and 69.9% in the GO + DA and DA arms, respectively).

More patients in clinical practice have favourable risk cytogenetics (approximately 20-25%) than in the trial (2.2-4.4%), whilst in the trial a much higher proportion of patients had intermediate risk cytogenetics (65-67%) than in practice (approximately 45-55%).

As discussed in Section 3.1, the clinical effectiveness section of the CS presents subgroup efficacy results for the patients with favourable and intermediate cytogenetics, excluding those with unknown cytogenetics, whilst in the economic model patients with favourable, intermediate and unknown cytogenetics are used. The subgroup used in the model reflects the population defined in the decision problem, i.e. patients not known to have unfavourable cytogenetics.

Table 3: Characteristics of participants across treatment groups (mITT population) (adapted from Table 71 of CS)

Baseline characteristic	GO + DA (n = 135)	DA (n = 136)
Age, years, median (range)	62.0 (50–70)	61.0 (50–70)
< 60, n (%)	38 (28.1)	52 (38.2)
≥ 60, n (%)	97 (71.9)	84 (61.8)
Male, n (%)	74 (54.8)	60 (44.1)
ECOG PS, n (%)		
0–1	121 (89.6)	117 (86.0)
≥ 2	14 (10.4)	18 (13.2)
Missing	0 (0.0)	1 (0.7)
WBC count, × 10^9/L, median (IQR)	5.8 (0.5–151.0)	4.1 (0.1–180.5)
Cytogenetics n (%)		
Favourable	3 (2.2)	6 (4.4)
Intermediate	91 (67.4)	89 (65.4)
Unfavourable	27 (20.0)	30 (22.1)
Not available	14 (10.4)	11 (8.1)
CD33 expression, positivity		
< 30%	17 (12.6)	20 (14.7)
≥ 30%	83 (61.5)	74 (54.4)
< 70%	37 (27.4)	31 (22.8)
≥ 70%	63 (46.7)	63 (46.3)

4.2.3 Summary of the quality of the ALFA-0701 trial

The CS summarises the quality assessment of ALFA-0701 in Table 12, Section B.2.5. The trial was a large open-label RCT. The randomisation process seems adequate, as it was performed using a computer generated sequence. Baseline characteristics were generally balanced, however, a higher proportion of patients in the GO + DA arm were male, aged ≥ 60 years and had a higher median WBC count. Prognosis worsens with increasing age so any bias introduced by this baseline imbalance would be expected to be in favour of the DA arm.

Participants and care providers were not blinded to the treatment allocation. The open-label design has the potential to introduce bias because clinicians being aware of the treatment a patient is receiving could influence the care the patient receives. Equally the behaviour of the patient may be altered by the knowledge of the treatment received. A retrospective analysis was conducted by an independent review committee to verify the collected data (IRC analysis). At this point investigators were blinded to the treatment arms to ensure unbiased assessment of data. The ERG deems this to be sufficient in

terms of data assessment, although further treatment decisions made in practice were based on the unblinded investigator analysis. However, results were broadly similar between IRC and IA, as shown in Table 13 of the CS. Attrition was similar between treatment arms and there is no evidence that additional outcomes were measured and not reported.

Overall the ERG believes the trial was well conducted and has a low risk of bias, up to the limits of its open-label design.

4.2.4 Summary of the results of the ALFA-0701 trial

Throughout this section, we focus on the results from the mITT population using the data cut-off of 30 April 2013. The April 2013 data-cut is preferred as additional outcomes are captured and therefore the estimate produced is more conservative than that produced using the earlier 1 August 2011 cut-off. Data using the 1 August 2011 cut-off are summarised in Appendix D.2.3 of the CS.

The CS reported both results according to the IRC assessment, and the assessment made by the original investigators (IA). In this section we focus on the IRC assessment, as these results are less likely to be influenced by lack of blinding, and were similar, but generally slightly more conservative than those of the IA.

4.2.4.1 Efficacy results

Key clinical effectiveness results are presented in Section B.2.6 of the CS. The overall summary of efficacy end-points in ALFA-0701 are presented in *** [redacted] 4.

[redacted] The median EFS in the GO + DA arm was [redacted] compared to [redacted] in the DA arm. Relapse free survival (RFS) also showed a statistically significant improvement, with a median RFS in the GO + DA arm of [redacted] compared with [redacted] in the DA arm.

Results for OS and overall response rate (CR + CRp) favoured the GO + DA arm, but the difference between treatment groups was not statistically significant. Patient reported outcomes and health related quality of life (HRQoL) were not assessed in the ALFA-0701 trial. [redacted] patients in the DA arm received GO as follow-up therapy through a compassionate use programme, which may have confounded overall survival results. Similarly, most patients ([redacted] in GO + DA arm and [redacted] in DA arm) received at least one follow-up therapy for AML, which may have confounded overall survival.

The investigator analyses were broadly consistent with the analyses conducted by the IRC.

[redacted] 4

	GO + DA arm	DA arm	Point estimate (95% CI)
EFS, months, median (95% CI)	[redacted]	[redacted]	[redacted]
RFS, months, median (95% CI)	[redacted]	[redacted]	[redacted]
OS, months, median (95% CI)	[redacted]	[redacted]	[redacted]
Overall response rate (CR/CRp), n (%)	[redacted]	[redacted]	[redacted]

Kaplan-Meier (KM) survival curves for EFS and OS are presented in ***[redacted] 3 and Figure 4 (Figures 6 and 10 of the CS).

Both KM curves plateau at around [redacted], suggesting no further relapse, events or mortality after around 3 to 4 years. This supports the company's claim that patients are functionally cured if there are no events within three years of treatment response; but the ERG notes that there were few patients with follow-up extending beyond three years.

[redacted]3
[redacted]

[redacted]

[redacted] 4

[redacted]

[redacted]

4.2.4.2 Analysis by cytogenetic status

The main analysis reported in Section 4.2.4.1 above includes all patients regardless of cytogenetic status. The ERG considers that to properly understand the efficacy of GO some further breakdown by cytogenetic status is required, particularly in view of the company's decision problem, specifying the use of GO for patients not known to have unfavourable cytogenetics.

Results for patients with favourable or intermediate cytogenetic profiles (excluding unknown cytogenetics) were reported in Table 14 of the CS, and are summarised in [redacted] 5 below.

[redacted] 5

	GO + DA arm (n = 94)	DA arm (n = 95)	Point estimate (95% CI)
EFS, months, median (95% CI)	[redacted]	[redacted]	[redacted]
RFS, months, median (95% CI)	[redacted]	[redacted]	[redacted]
OS months, median (95% CI)	[redacted]	[redacted]	[redacted]
Overall response rate (CR/CRp), n (%)	[redacted]	[redacted]	[redacted]

EFS and RFS were statistically significantly improved to a similar extent as the overall population in the favourable and intermediate cytogenetic subgroup, whilst OS and overall response rate favoured

the GO + DA arm, but the difference between groups was not statistically significant, consistent with the overall population results.

The CS also reports results for those patients with unfavourable cytogenetics, summarised below in ***[redacted]6. There was no benefit of GO in these patients for the outcomes of EFS, RFS or OS.

[redacted]6

In response to the ERG's request for further clarification, the company

	GO + DA arm (n = 14)	**DA arm (n = 15)**	**Point estimate (95% CI)**
EFS, months, median (95% CI)	[redacted]	[redacted]	[redacted]
RFS, months, median (95% CI)	[redacted]	[redacted]	[redacted]
OS, months, median (95% CI)	[redacted]	[redacted]	[redacted]

provided additional IRC assessment results for the outcomes EFS, RFS and OS, for each of: favourable, intermediate and unknown cytogenetic risk groups. The results are summarised below in [redacted]7,

[redacted]8 and [redacted]9, respectively. These results are broadly consistent with those for the combined favourable/intermediate group ([redacted]5).

[redacted]

[redacted]7 [redacted]

Cytogenetic profile	**GO + DA arm**	**DA arm**	**Point estimate (95% CI)**
Favourable/intermediate + unknown [redacted]	[redacted]	[redacted]	[redacted]
Favourable [redacted]	[redacted]	[redacted]	[redacted]
Intermediate [redacted]	[redacted]	[redacted]	[redacted]
Unknown [redacted]	[redacted]	[redacted]	[redacted]

[redacted]8[redacted]

Cytogenetic profile	GO + DA arm	DA arm	Point estimate (95% CI)
Favourable/intermediate + unknown [redacted]	[redacted]	[redacted]	[redacted]
Favourable [redacted]	[redacted]	[redacted]	[redacted]
Intermediate [redacted]	[redacted]	[redacted]	[redacted]
Unknown [redacted]	[redacted]	[redacted]	[redacted]

[redacted]9[redacted]

Cytogenetic profile	GO + DA arm	DA arm	Point estimate (95% CI)
Favourable/intermediate + unknown [redacted]	[redacted]	[redacted]	[redacted]
Favourable [redacted]	[redacted]	[redacted]	[redacted]
Intermediate [redacted]	[redacted]	[redacted]	[redacted]
Unknown [redacted]	[redacted]	[redacted]	[redacted]

4.2.4.3 Analysis by cytogenetic and molecular status

The company also provided additional IRC assessment results, in response to the ERG's request for further clarification, for patient subgroups based on cytogenetic and molecular risk categories, with the intermediate group divided into "intermediate-1" and "intermediate-2". It was not clear to the ERG exactly what test was used for this assessment: the CS stated that subgroup analyses were performed using risk classification based on NCCN and ELN guidelines.

The company provided data for the following groups: favourable and intermediate-1; intermediate-2 and unfavourable; intermediate-1 only; intermediate-2 only. These were provided for the outcomes EFS, RFS and OS, summarised below in [redacted], [redacted] and [redacted], respectively.

[redacted]

[redacted]

[redacted] 10

Cytogenetic/molecular profile	GO + DA arm	DA arm	Point estimate (95% CI)
Intermediate-1 [redacted]	[redacted]	[redacted]	[redacted]
Intermediate-2 [redacted]	[redacted]	[redacted]	[redacted]
Favourable and intermediate-1 [redacted]	[redacted]	[redacted]	[redacted]
Intermediate-2 and unfavourable [redacted]	[redacted]	[redacted]	[redacted]

[redacted] 11

Cytogenetic/molecular profile	GO + DA arm	DA arm	Point estimate (95% CI)
Intermediate-1 [redacted]	[redacted]	[redacted]	[redacted]
Intermediate-2 [redacted]	[redacted]	[redacted]	[redacted]
Favourable and intermediate-1 [redacted]	[redacted]	[redacted]	[redacted]
Intermediate-2 and unfavourable [redacted]	[redacted]	[redacted]	[redacted]

[redacted] 12

Cytogenetic/molecular profile	GO + DA arm	DA arm	Point estimate (95% CI)
Intermediate-1 [redacted]	[redacted]	[redacted]	[redacted]
Intermediate-2 [redacted]	[redacted]	[redacted]	[redacted]
Favourable and intermediate-1 [redacted]	[redacted]	[redacted]	[redacted]
Intermediate-2 and unfavourable [redacted]	[redacted]	[redacted]	[redacted]

4.2.4.4 Molecular status and other subgroup analyses

Further to cytogenetic profiles, subgroup analysis was also presented based on molecular testing. Molecular testing risk stratifies patients based on specific gene mutations. Current standard of care combines cytogenetic results with targeted testing for mutations in *FLT3*, *NPM1*, *CEBPA* and *KIT* to determine the prognostic subgroup.[27] The following subgroups were presented: FLT3_ITD, NPM1 status, MLL, WT1, risk based on the NCCN classification, risk based on the ELN classification, age, ECOG performance status and CD33 expression. These results were summarised in forest plots in the CS as Figures 30 and 31. However, Figure 30 (EFS) presented data from the 2011 data-cut, therefore, the ERG requested the figure to be updated using the April 2013 data cut; these forest plots are presented below as Figure 5 and Figure 6 (EFS) and *[redacted]7 7 (OS).

[redacted]

In response to the ERG's request for further clarification, the company also provided subgroup analysis results for white blood cell count <30 x 10^9/L and ≥30 x 10^9/L, for the outcome overall response rate (CR/CRp).

[redacted]

[redacted]5 [redacted]

[redacted]

[redacted]

[redacted]

[redacted]7[redacted]

[redacted]

[redacted]

[redacted]

4.2.5 Safety

The proportion of patients experiencing an adverse event (AE) was similar in the GO + DA and DA alone treatment arms ([redacted]), as shown in [redacted] 13 (Table 25 of the CS). The most common treatment-related AE was haemorrhage ([redacted]). However, the proportion of patients experiencing a serious adverse event (SAE) was higher in the GO + DA treatment arm than in the DA treatment arm ([redacted]). The most common SAE in the GO + DA arm was thrombocytopenia ([redacted]). At each treatment phase, the median time to platelet recovery was [redacted] in the GO + DA arm than the DA arm. GO was also associated with a higher proportion of patients experiencing veno-occlusive disease (VOD) ([redacted]); however, all patients who experienced a VOD event in the DA arm had received GO as subsequent therapy. More patients in the GO + DA arm were readmitted or had planned hospitalisation prolonged owing to reasons relating to AEs ([redacted]). A similar proportion of patients were admitted to intensive care units in the GO + DA and DA arms ([redacted]).

The proportion of patients who permanently discontinued treatment owing to treatment-emergent AEs (TEAEs) was higher in the GO + DA arm than in the DA arm ([redacted] This was mainly accounted for by thrombocytopenia ([redacted]) in the GO + DA arm (no discontinuations for this reason in the DA arm) and hepatobiliary disorders ([redacted]). Over [redacted] of patients in the GO + DA arm were able to receive all three fractionated doses of GO during induction therapy and just under half were able to receive GO during the two courses of consolidation therapy.

[redacted] 13

	GO + DA, (N = 131) n (%)		DA, (N = 137) n (%)	
	All-causality AEs	**Related AEs**	**All-causality AEs**	**Related AEs**
Patients with AEs	[redacted]	[redacted]	[redacted]	[redacted]
Patients with SAEs	[redacted]	[redacted]	[redacted]	[redacted]
Patients with grade 3 or 4 or severe infection AEs	[redacted]	[redacted]	[redacted]	[redacted]
Patients with fatal events	[redacted]	[redacted]	[redacted]	[redacted]
Patients who permanently discontinued study treatment owing to AEs	[redacted]	[redacted]	[redacted]	[redacted]

Treatment related deaths were slightly more frequent in the GO + DA treatment arm, as shown in [redacted] (Table 24 of the CS).

[redacted] 14

	GO + DA, (N = 131) n (%)	DA, (N = 137) n (%)
Treatment-related deaths	[redacted]	[redacted]
Cause of death		
Disease under study	[redacted]	[redacted]
Study treatment toxicity	[redacted]	[redacted]
Unknown	[redacted]	[redacted]
Other	[redacted]	[redacted]
Mechanism of death		
Disease progression or relapse	[redacted]	[redacted]
Septic shock	[redacted]	[redacted]
Infection	[redacted]	[redacted]
GVHD	[redacted]	[redacted]
Liver toxicity	[redacted]	[redacted]
Haemorrhage	[redacted]	[redacted]
Other	[redacted]	[redacted]
Treatment related death during CR/CRp[a]	[redacted]	[redacted]

During the EMA regulatory process, the population in the decision problem was extended to also include patients aged 15-17 years old. The company conducted a systematic review on safety data in

patients <18 years old treated with GO. The company state that the evaluation of the safety profile emerging from the identified paediatric studies was consistent with the known characterised safety profile in the adult population. As no details of the systematic review were presented, the ERG has not checked this evidence.

4.3 Supporting data from meta-analyses

Section D3 of the CS appendices presents data on an individual patient data (IPD) meta-analysis conducted by Pfizer and four published meta-analyses identified in a literature review. These meta-analyses are described in Sections 4.3.1 and 4.3.2, respectively.

4.3.1 Individual patient data meta-analysis

An IPD meta-analysis was conducted by Pfizer for use in regulatory submissions. This meta-analysis updated a meta-analysis by Hills et al. (described in Section 4.3.2) using data from a later time point.

4.3.1.1 Critique of the methods of the individual patient data meta-analysis

A literature search was conducted to identify RCTs of GO in addition to induction chemotherapy, compared with induction chemotherapy alone, in newly diagnosed patients aged 15 years or older with AML (either de novo or secondary) or high risk myelodysplastic syndrome (MDS). The search strategy used to identify relevant trials appears to have been appropriate.

Five RCTs were included in the meta-analysis; MRC AML15, NCRI AML16, ALFA-0701, GOELAMS AML2006IR and SWOG S0106. The IPD meta-analysis included five of the eight RCTs identified in the systematic review described in Section 4.2 of this report; the exclusion of trials AML-17, AML-18 and AML-19 from the IPD meta-analysis was appropriate; AML-17 did not include a 'no-GO' comparator group and AML-18 and AML-19 are ongoing trials.

The individual investigators and institutional sponsors of the five included trials were contacted to request individual patient data. The data collection cut-off date was April 2013 for the ALFA-0701 trial (the IA analysis was used), June 2015 for the GOELAMS AML2006IR and SWOG S0106 trials and August 2015 for the MRC AML15 and NCRI AML16 trials.

A summary of the quality of the included trials is presented in Table 77 of the CS appendices. All five trials were considered to be at low risk of selection bias, attrition bias and reporting bias; the ERG considers that the quality assessment was appropriate for these criteria.

The CS states that there was a low risk of performance bias, as the lack of blinding of patients and personnel was unlikely to have resulted in differences in care provided between the groups. The ERG considers that the lack of blinding has the potential to introduce bias, as clinicians being aware of the

treatment allocation could influence the care that the patient receives and patients' behaviour and perception of effectiveness and adverse effects may be influenced by their knowledge of the treatment received. The CS acknowledges the high risk of bias resulting from the lack of blinding for the safety outcome assessment.

The primary endpoint assessed in the IPD meta-analysis was OS, secondary endpoints were EFS, CR rate, OR rate and RFS.

A total of 3331 patients were included in the IPD meta-analysis; 1663 were randomised to the GO arm and 1668 were randomised to the no-GO arm. Most patients across the included trials received DA as induction therapy, but different induction schedules were also used. A total of 2655 patients were evaluated for cytogenetic risk, of which 2075 had favourable/intermediate risk, split evenly between the GO and no-GO treatment groups. Table 15 presents a summary of the study and participant characteristics included in the IPD meta-analysis (adapted from Tables 80-82 of the CS).

Table 15: Summary of study and participant characteristics included in the IPD meta-analysis (adapted from Tables 80-82 of CS)

Parameter	**Trial**					
	MRC AML15 n (%)	**NCRI AML16 n (%)**	**ALFA-0701 n (%)**	**GOELAMS AML 2006 IR n (%)**	**SWOG S0106 n (%)**	**All trials n (%)**
Eligibility criteria	AML, de novo or secondary, APL, usually aged < 60 years	AML, de novo, secondary, or high-risk MDS, usually aged > 60 years	De novo AML, aged 50–70 years	De novo AML, intermediate cytogenetics, aged 18–60 years	De novo AML, aged 18–60 years	
Primary endpoint	OS	OS	EFS	EFS at 3 years	DFS, CR	
Secondary endpoints	CR, Cri, RD, RFS, CIR, CIDCR, toxicity	CR, CRi, RFS, RR, DCR1, toxicity	CR, CRp, RFS, OS, toxicity	OS, CIR and CIDND at 3 years, CR, toxicity	CRi, PRem, RD, OS, RFS, toxicity	
GO dose	3 mg/m^2 on day 1	3 mg/m^2 on day 1	3 mg/m^2 on days 1, 4 and 7 (≤ 5 mg/dose)	6 mg/m^2 on day 4	6 mg/m^2 on day 4	
Number of patients	1099	1115	271	251	595	3331
Treatment arm						
GO	548 (49.9)	559 (50.1)	135 (49.8)	126 (50.2)	295 (49.6)	1663 (49.9)
No-GO	551 (50.1)	556 (49.9)	136 (50.2)	125 (49.8)	300 (50.4)	1668 (50.1)
Age, years						
15–29	150 (13.6)	0	0	25 (10.0)	80 (13.4)	255 (7.7)
30–39	156 (14.2)	0	0	31 (12.4)	95 (16)	282 (8.5)

Parameter	Trial					
	MRC AML15 n (%)	**NCRI AML16 n (%)**	**ALFA-0701 n (%)**	**GOELAMS AML 2006 IR n (%)**	**SWOG S0106 n (%)**	**All trials n (%)**
40–49	242 (22.0)	0	0	65 (25.9)	166 (27.9)	473 (14.2)
50–59	397 (36.1)	19 (1.7)	90 (33.2)	121 (48.2)	235 (39.5)	862 (25.9)
60–69	151 (13.7)	755 (67.7)	169 (62.4)	9 (3.6)	19 (3.2)	1103 (33.1)
≥ 70	3 (0.3)	341 (30.6)	12 (4.4)	0	0	356 (10.7)
Sex						
Female	514 (46.8)	444 (39.8)	137 (50.6)	108 (43.0)	282 (47.4)	1485 (44.6)
Male	585 (53.2)	671 (60.2)	134 (49.4)	143 (57.0)	313 (52.6)	1846 (55.4)
Type of AML						
De novo	1011 (92.0)	805 (72.2)	271 (100)	251 (100)	595 (100)	2933 (88.1)
Secondary	88 (8.0)	197 (17.7)	0	0	0	285 (8.6)
High-risk MDS	0	113 (10.1)	0	0	0	113 (3.4)
MRC cytogenetic risk						
Favourable	133 (12.1)	33 (3.0)	9 (3.3)	1 (0.4)	72 (12.1)	248 (7.4)
Intermediate	565 (51.4)	580 (52.0)	176 (64.9)	223 (88.8)	283 (47.6)	1827 (54.8)
Adverse	196 (17.8)	264 (23.7)	53 (19.6)	0	67 (11.3)	580 (17.4)
Unknown	205 (18.7)	238 (21.3)	33 (12.2)	27 (10.8)	173 (29.1)	676 (20.3)

Parameter	Trial					
	MRC AML15 n (%)	**NCRI AML16 n (%)**	**ALFA-0701 n (%)**	**GOELAMS AML 2006 IR n (%)**	**SWOG S0106 n (%)**	**All trials n (%)**
CD33-positive						
< 30% of blasts	127 (11.6)	239 (21.4)	37 (13.7)	28 (11.2)	0	431 (12.9)
≥ 30%	423 (38.5)	517 (46.4)	157 (57.9)	220 (87.6)	0	1317 (39.5)
< 70%	243 (22.1)	442 (39.6)	68 (25.1)	71 (28.3)	0	824 (24.7)
≥ 70%	307 (27.9)	314 (28.2)	126 (46.5)	117 (0.5)	0	924 (27.7)
Unknown	549 (50.0)	359 (32.2)	77 (28.4)	3 (1.2)	595 (100)	1583 (47.5)
Chemotherapy						
ADE/DA	630 (57.3)	773 (69.3)	271 (100)	251 (100)	595 (100)	2520 (75.7)
FLAG-Ida	469 (42.7)	0	0	0	0	469 (14.1)
DClo	0	342 (30.7)	0	0	0	342 (10.3)

4.3.1.2 Results of the individual patient data meta-analysis

Overall survival

Overall survival was statistically significantly improved in the GO arm compared with the no-GO arm ([redacted], as shown in [redacted] 8 below (Figure 26 of the CS). Overall survival was improved in four of the five trials, although the result was not statistically significant in any individual trial.

[redacted] 8 [redacted]

[redacted]

The estimated absolute improvement in OS in patients randomised to GO at 6 years was [redacted] overall pooled median overall survival was [redacted] in the GO arm and [redacted] in the no-GO arm.

[redacted]

[redacted] 9 [redacted]

[redacted]

[redacted]

[redacted] ***Event-free survival***

EFS was improved in the GO arm compared with the no-GO arm [redacted] shown in Figure 28 of the CS. In subgroup analyses, [redacted].

Relapse-free survival

RFS was evaluated in [redacted] patients in the GO arm and [redacted] patients in the no-GO arm who had achieved a response. RFS was prolonged in the GO arm compared with the no-GO arm [redacted], shown in Figure 29 of the CS. The CS states that in subgroup analyses, [redacted] Time to relapse was significantly prolonged in the GO arm compared with the no-GO arm [redacted].

Response rate

Overall response was achieved in [redacted] patients in the GO arm and [redacted] patients in the no-GO arm. The odds of achieving a response were improved with GO, but the result was not quite statistically significant [redacted]. CR was achieved in [redacted] patients in the GO arm and [redacted] patients in the no-GO arm. The odds of achieving a CR were improved with GO, but again the result was not [redacted]

In summary, the IPD meta-analysis demonstrated that the addition of GO to induction therapy improved OS, EFS and RFS. The addition of GO did not result in a statistically significant improvement in response rate. [redacted]

4.3.2 Published meta-analyses identified in the literature review

The CS states that four published meta-analyses that investigated the effect of adding GO to induction therapy in AML were identified by a systematic literature review, although no details of the methods of the systematic review were reported, so they have not been verified by the ERG. The four meta-analyses described in Section D.3.2 of the CS appendices were by Hills et al.,[29] Li et al.,[30] Loke et al.[31] and Kharfan-Dabaja et al.[32]

The meta-analysis by Hills et al. included five RCTs and a total of 3325 patients; ALFA-0701, MRC AML15, NCRI AML16, GOELAMS AML 2006 IR and SWOG S0106. The data collection cut-off date was 2011 for the ALFA-0701 trial and up to 2013 for the other trials. Overall survival and RFS were statistically significantly improved in the GO arm compared with the no-GO arm. In subgroup analyses overall survival was significantly improved in the GO arm for patients with favourable cytogenetic risk and patients with intermediate cytogenetic risk, but not for patients with adverse cytogenetic risk. The addition of GO to induction chemotherapy did not improve the rate of CR.[29]

The meta-analysis by Li et al. included five RCTs and a total of 3596 patients: ALFA-0701, MRC AML15, SWOG S0106, NCRI/University Hospital of Wales AML16 and Leukaemia Research Fund AML14 and NCRI AML16. Median follow-up ranged from 20 months for the ALFA-0701 trial to 49 months for the SWOG S0106 trial. Overall survival, RFS and relapse rate were statistically significantly improved in the GO arm compared with the no-GO arm. However, in subgroup analyses the improvement in overall survival was only statistically significant in patients with favourable cytogenetic risk, the improvement in the intermediate cytogenetic risk group did not reach statistical significance and patients in the unfavourable group did not experience a survival benefit. The addition of GO to induction chemotherapy did not improve the rate of CR.[30]

The meta-analysis by Loke et al. included eleven RCTs of GO in patients with AML, with a total of 7138 patients, although the review was not restricted to trials using GO as part of induction therapy. Overall survival was not statistically significantly different between the GO and no-GO arms, even when only the seven RCTs of induction therapy were combined (ALFA-0701, MRC AML15, NCRI AML16, GOELAMS AML 2006 IR, SWOG S0106, EORTC-Gruppo Italiano Malattie Ematologiche dell'Adulto [GIMEMA] AML17 and Children's Oncology Group trial AML0531). However, in subgroup analyses overall survival was improved in the GO arm for patients with favourable cytogenetic risk (HR: 0.46, 95% CI: 0.29-0.73), but not for patients with intermediate or adverse cytogenetics. Rate of resistant disease (HR: 0.77, 95% CI: 0.67-0.9) and cumulative incidence of relapse (HR: 0.86, 95% CI: 0.79-0.93) were reduced in the GO arm. The addition of GO to induction chemotherapy did not improve the rate of CR and there was no evidence that CD33 positivity influenced OS in patients in the GO arm.[31]

The meta-analysis by Kharfan-Dabaja et al. included seven RCTs comparing GO plus conventional chemotherapy with conventional chemotherapy alone in patients with newly diagnosed AML, with a total of 3943 patients: ALFA-0701, MRC AML15, NCRI AML16, GOELAMS AML 2006 IR, SWOG S0106, EORTC-GIMEMA AML17 and the German AML Intergroup trial. Overall survival was statistically significantly improved in the GO arm for the low/intermediate cytogenetic risk group (HR: 0.83, 95% CI: 0.70-0.99), but not for the high risk group or overall group. RFS (reported in six trials) and EFS (reported in three trials) were also statistically significantly improved in the GO arm. The addition of GO to induction chemotherapy did not improve the rate of CR.[32]

In summary, the meta-analyses demonstrated that the addition of GO to induction therapy significantly improved RFS. Overall survival was improved in patients with favourable cytogenetics (results were inconsistent between meta-analyses for the overall population and also the intermediate risk group). The addition of GO did not improve the rate of CR in any of the meta-analyses. The CS emphasises the strengths of the IPD meta-analysis described in Section 4.3.1, compared with the summary meta-analyses presented in this section.

4.4 Conclusions of the clinical effectiveness section

The CS evaluation of GO was primarily based on one reasonably good quality RCT; the ALFA-0701 trial, which compared GO + DA versus DA alone in patients aged 50-70 with previously untreated de novo AML. The anticipated marketing authorisation is patients aged 15 years or over with CD33-positive AML. Therefore, the age range in the trial was narrower than the anticipated marketing authorisation, although the majority of patients with AML are over the age of 50 and the ERG's clinical advisor did not consider that GO would work differently in patients under the age of 50 to

those over the age of 50. The restriction to patients with CD33-positive AML appears clinically appropriate, in view of the mechanism of action of GO and subgroup analysis results indicated that GO appears to be more effective in patients with a higher proportion of CD33-positive blasts than those with a lower proportion of CD33-positive blasts.

The ALFA-0701 trial demonstrated that GO + DA was effective at improving EFS and RFS, compared with DA alone, in the overall patient population ([redacted]). Whilst the response rate appeared better in the GO + DA arm ([redacted]), the result did not reach statistical significance, suggesting that duration of remission is extended with GO, rather than the proportion of patients achieving a remission. Overall survival appeared to be better in patients who received GO + DA, although the difference between treatment groups did not reach statistical significance ([redacted]). This lack of significance may have occurred because some patients received follow-up treatments, with some patients in the DA arm receiving GO as follow-up therapy.

EFS and RFS results were statistically significantly improved in the GO + DA arm for the subgroup of patients with favourable/intermediate cytogenetic risk, to a similar extent as the overall population, with OS and response rate also improved in the GO + DA arm, but not reaching statistical significance, which was consistent with the overall results. However, for patients with an unfavourable cytogenetic profile, OS and RFS outcomes appeared to be worse in the GO + DA arm, compared with the DA arm, whilst EFS results were similar ([redacted]).

Additional analyses provided to the ERG on request showed that the benefit seen in patients with an intermediate-1 cytogenetic and molecular risk profile was not found in patients with an intermediate-2 cytogenetic and molecular risk profile, suggesting potentially important heterogeneity in the broader 'intermediate' cytogenetic subgroup.

Whilst the proportion of patients experiencing an AE was similar between treatment groups, the proportion of patients reporting a SAE was higher in the GO + DA arm than the DA arm [redacted]. The most common SAE in the GO + DA arm was thrombocytopenia

[redacted]. All [redacted] patients who experienced VOD had received GO. A higher proportion of patients in the GO + DA arm discontinued treatment because of an adverse event than in the DA arm [redacted].

The CS also presented the results of an IPD meta-analysis, conducted by Pfizer for use in regulatory submissions.[28] Overall survival was statistically significantly improved in the GO + DA arm for the overall population in the IPD meta-analysis [redacted] and for the subpopulation of patients with a favourable or intermediate cytogenetic profile [redacted], but not for patients with an unfavourable cytogenetic profile [redacted], supporting the evidence from the ALFA-0701 trial. However, the IPD meta-analysis included patients with de novo or secondary AML or high-risk myelodysplastic syndrome, which is a broader population than the anticipated licence, therefore, results may not be entirely generalisable to patients eligible for GO in clinical practice.

The ALFA-0701 trial and the IPD meta-analysis did not report quality of life outcomes.

5 Cost Effectiveness

This section focuses on the economic evidence submitted by the company and the additional information provided in response to the points for clarification. The submission was subject to a critical review on the basis of the company's report and by direct examination of the electronic version of the economic 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 [33] (Appendix Section 11.1).

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

5.1.1 Searches

The CS described the search strategies used to identify relevant economic modelling studies of pharmacological treatments for AML patients. The searches were described in Section B.3.1 of the submission and full search strategies were presented in Appendix G.

The appropriate databases used for the cost effectiveness systematic literature review were searched. Additional searches of conference websites were conducted to identify information from 2015 to 2017. These are reported in Section G.2 of the CS. The search strategies used in MEDLINE, Embase, EconLIT and the Cochrane Library databases are fully reproduced in Tables 93 to 98 of the CS and the date conducted and number of records identified is given.

5.1.2 Inclusion/exclusion criteria used for study selection

Details of the inclusion and exclusion criteria used in the cost-effectiveness review were reported in the CS (Table 99, CS). The ERG considers that the inclusion/exclusion criteria used in the cost-effectiveness review were reasonable.

5.1.3 Studies included and excluded in the cost effectiveness review

Nine of the articles identified in the cost-effectiveness review were evaluations of treatments for AML. Four of these were from a UK payer perspective and subsequently included in the final review. Two of these were previous submissions to NICE, appraising azacitidine in two different patient populations [34, 35].

5.1.4 Conclusions of the cost effectiveness review

Four cost-effectiveness studies were identified and considered relevant for the cost-effectiveness review. However, none of these evaluated the cost-effectiveness of GO. The *de novo* cost-effectiveness analysis reported in the CS is, therefore, the only source of evidence which directly informs the decision problem.

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

An overall summary of the company's approach, and signposts to the relevant sections in the company's submission, are reported in Table 16 below.

Table 16 Summary of the company economic analysis

	Approach	Source / Justification	Signpost (location in company submission)
Model	Cost-effectiveness (cost-utility) analysis using a semi-Markov state-transition model	To provide flexibility to incorporate the transition probabilities for the HSCT and functionally cured health states, and to adjust survival for HSCT and non-HSCT patients.	Section B.3.2.2 pg. 99
States and events	The model contains the following health states: induction therapy, CR/CRp on consolidation therapy, CR/CRp off-treatment, relapse on salvage therapy, relapse on non-curative therapy, refractory on salvage therapy, refractory on non-curative therapy, HSCT, Post-HSCT without GVHD, post-HSCT with GVHD, functionally cured, and death.	Model health states were chosen based on the clinical pathway, clinical expert opinion, the nature of the available data, and previous AML models.	Section B.3.2.2 pg. 99 to 101
Comparators	Gemtuzumab ozogamicin (GO) in combination with standard intensive chemotherapy, consisting of daunorubicin and cytarabine (DA), was compared to DA alone.	DA was considered to be the most appropriate comparator to GO+DA, reflecting the standard intensive chemotherapy regimen used in routine clinical practice outside of a clinical trial setting. The other comparators included in the final NICE scope were not considered relevant for de-novo AML or to represent the current standard of care.	Section B.1.1 pg. 11 to 13, B.3.2.4 pg. 102
Subgroups	The company base-case is based on a subgroup of the licensed population – patients not known to have unfavourable cytogenetics.	No additional subgroup analysis was undertaken as the base-case already focused on a subgroup of the licensed population.	Section B.3.9 pg. 151

	Approach	Source / Justification	Signpost (location in company submission)
	Results were also presented separately for the entire licensed population.		
Treatment effectiveness	Clinical outcomes included were response (CR/CRp), RFS and OS, cure fraction, probability of HSCT, post-HSCT survival These data were taken from the ALFA-0701 study. OS was stratified by response status. Parametric models were fitted to RFS (CR/CRp only) and OS (stratified by response status) to extrapolate beyond the end of trial follow-up. Response and RFS endpoints were based on the blinded IRC assessment. RFS and OS analyses were based on the reference data of 30 April 2013.	Data from a subgroup (patients not known to have unfavourable cytogenetics) of the ALFA-0701 study were used to inform clinical inputs of the model. The ALFA-0701 study is the only RCT that has compared GO+DA using the licensed fractionated dosing regimen. The blinded IRC analyses were chosen to address any possible bias by local investigators due to the open-label design. These were stated by the company to be the outcomes considered most appropriate by EMA.	Section B.3.3 pg. 103 to 122
Adverse events	Grade 3 and 4 treatment-related adverse events that occurred in at least 1% of patients were included. GVHD as a consequence of HSCT was also included. .	Adverse event rates were taken from the ALFA-0701 trial. Incidence of GVHD was sourced from external literature; Battipaglia (2017)[36] and NHS England (2017)[37].	Section B.3.3.10 pg. 122 to 123
Health related quality of life	No HRQoL data was collected in ALFA-0701 and health state utilities were sourced from a systematic literature review. A vignette study undertaken by the company provided an alternate set of utility values that were used in a scenario analysis.	Functionally cured patients were assumed to have quality of life equal to that of the age-matched general population (Ara & Brazier) [38]. The remaining health state utilities were sourced from NICE TA399 [35]. Adverse event disutilities were sourced from external literature including: NICE TA399 [35], the appraisal of defibrotide by the SMC for VOD [39] and Kurosawa (2014) [40].	Section B.3.4.3 pg. 124 Section B.3.4.4 pg. 124 to 127

	Approach	Source / Justification	Signpost (location in company submission)
Resource utilisation and costs	These comprised: drug acquisition, health state costs (monitoring and management), terminal care costs, HCST and GVHD treatment costs, and treatment of adverse effects.	Drug acquisition costs for GO were based on the confidential list price. Drug acquisition costs for DA and subsequent lines of chemotherapy were sourced from BNF [41] and eMit [42]. Unit costs for health state and adverse events were taken from NHS reference costs (2016) [43] and PSSRU (2016) [44]. The treatment cost of VOD also includes data sourced from NHS England Commissioning reports [45]. Terminal care costs were sourced from Addicott & Dewar (2008) [46]. HSCT costs were sourced from a NHS Blood and Transplant analysis (2014) [47], and the GVHD cost from Esperou (2004) [48]. Resource use items were based on those used in ALFA-0701 and elicited clinical expert opinion.	Section B.3.5 pg.127 to 140
Discount rates	Costs and benefits were discounted at 3.5% per annum	In accordance with the NICE reference case.	Section B.3.2.2 pg. 101
Sensitivity analysis	Probabilistic sensitivity analysis was performed. Deterministic univariate probabilistic analysis was performed on a series of model parameters. A series of scenario analyses were also performed.	In accordance with the NICE reference case.	Section B.3.8 pg. 143 to 150

5.2.1 Model structure

The CS presented a *de novo* semi-Markov cohort state-transition model to estimate the cost-effectiveness of GO+DA compared with DA alone in a subgroup of the licensed population: adult de-novo AML patients not known to have unfavourable cytogenetics.

Cost-effectiveness was assessed over a lifetime time horizon of 40 years. The cycle length used in the model was one month, which was considered to be sufficiently granular to accurately capture model costs and outcomes. A half cycle correction was applied to costs and QALYs.

The model health states were identified and validated through a preference elicitation study undertaken undertaken by the company. The states are described in Figure 10 Model structure diagram (CS, Figure 11)

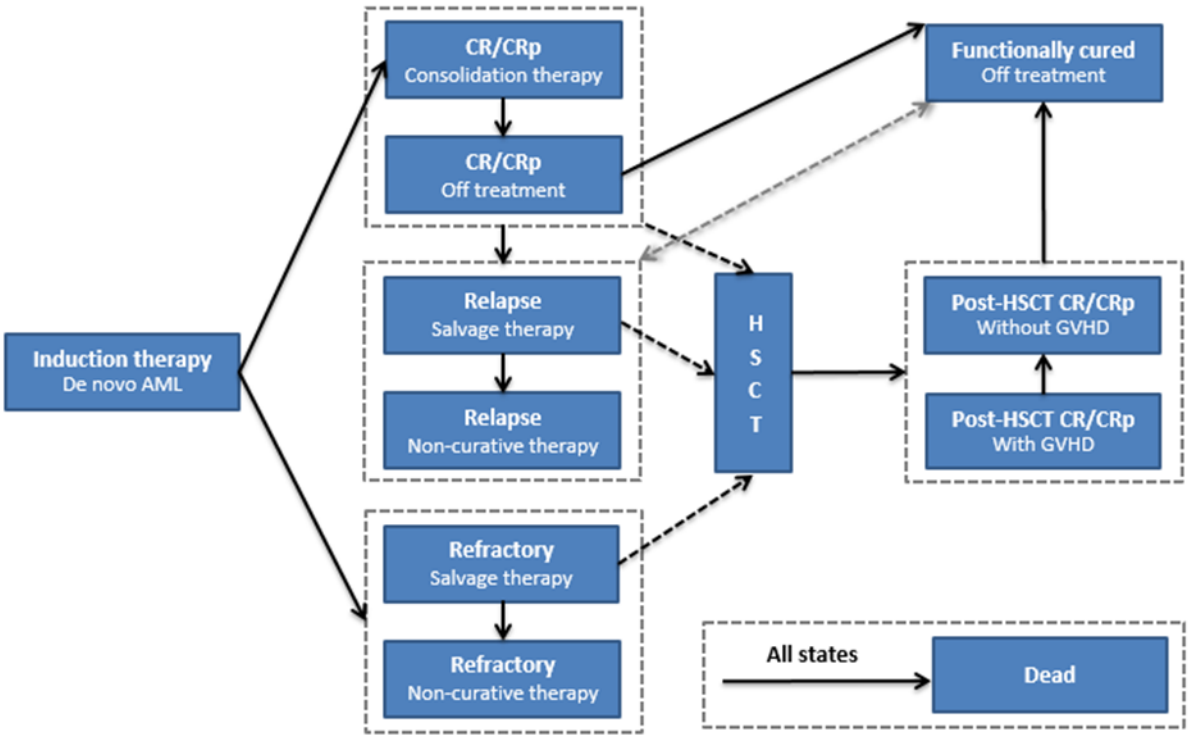


Table 17. The model structure and transitions are depicted in Figure 10. The CS reported that the design of the model structure was informed by the clinical pathway, clinical expert input and previous AML models.

Figure 10 Model structure diagram (CS, Figure 11)

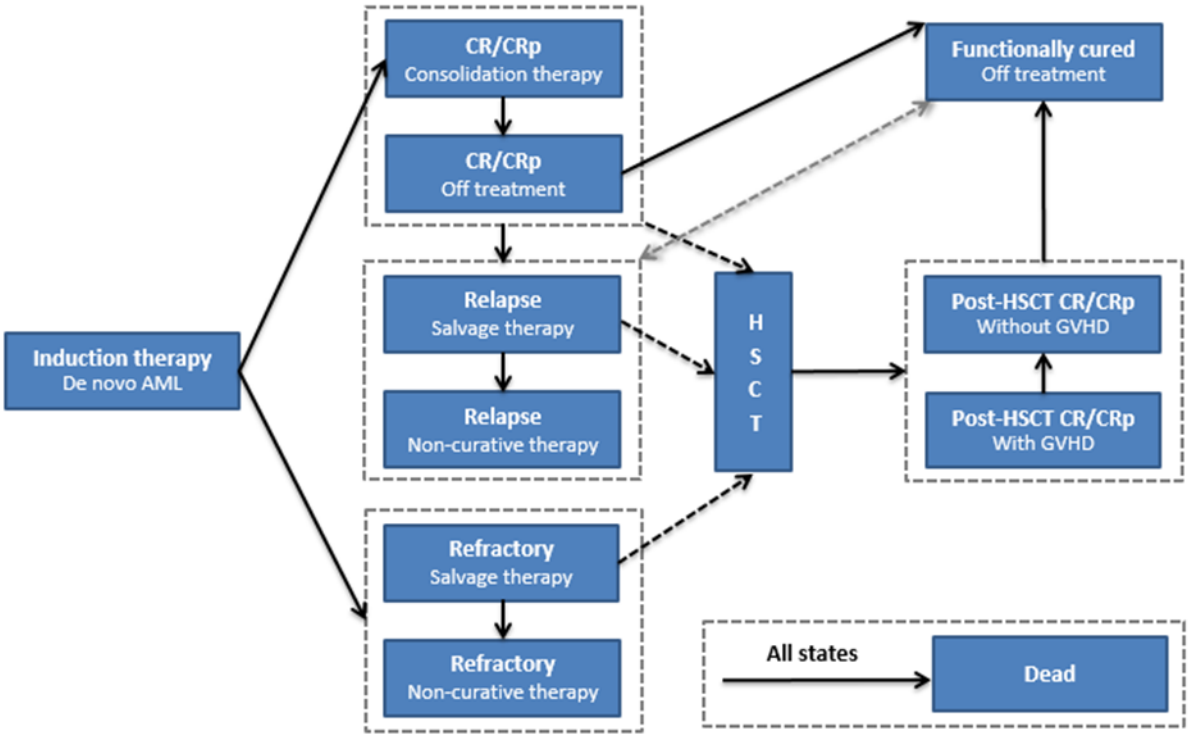


Table 17 Model health states (adapted from CS Appendix M.1, Table 130)

Health state	Description
Induction therapy	• Initial period of treatment with GO + DA or DA alone prior to determination of response status • Duration of up to two cycles
CR/CRp (consolidation)	• Consolidation treatment for patients who attain CR or CRp following induction therapy • Duration up to two cycles
CR/CRp (off-treatment)	• Period when CR or CRp patients have stopped treatment after completing induction or consolidation therapy and have regular follow-up visits • Duration: patients remain in this health state until they experience relapse or death, have an HSCT, or become 'functionally cured'
Refractory (salvage therapy)	• Refractory patients (those who failed induction therapy) who are receiving treatment with high-intensity salvage therapy • Duration of 1.5 cycles in base-case model
Refractory (non-curative therapy)	• Refractory patients on non-curative therapy, who are not eligible for salvage therapy • Refractory patients on non-curative therapy (BSC only) who failed subsequent salvage therapy in the previous cycles • Duration: until death or receive HSCT
HSCT	• Period of HSCT procedure and recovery when patients remain hospitalized • Duration: one cycle
Post-HSCT CR/CRp (without GVHD)	• Period after HSCT procedure and prior to becoming 'functionally cured', without GVHD (based on defined period of time) • Duration: In health state until functionally cured (after 60 months) or death
Post-HSCT CR/CRp (with GVHD)	• Period after HSCT procedure and prior to becoming 'functionally cured', with GVHD • Duration: in health state for pre-specified amount of time, reflecting duration of GVHD
Functionally cured	• Long-term disease-free survival (CR/CRp) with no planned follow-up • Duration: in health state until death, as modelled with adjusted general population mortality
Relapse (salvage therapy)	• Treatment with high-intensity salvage therapy for patients with relapse following an initial CR/CRp with induction therapy • Duration of 1.5 cycles in base-case model
Relapse (non-curative therapy)	• Patients with relapse on non-curative therapy, who are not eligible for salvage therapy • Patients with relapse on non-curative therapy (BSC only) who failed salvage therapy in the previous cycles • Duration: until death or receive HSCT
Dead	• Dead
GO: gemtuzumab ozogamicin; DA: daunorubicin and cytarabine; AML: acute myeloid leukaemia; CR: complete remission; CRp: complete remission with incomplete platelet recovery; BSC: best supportive care; HSCT: haematopoietic stem cell transplant; GVHD: graft versus host disease	

Induction therapy

Patients entered the model in the induction therapy health state and received one or two induction courses, depending on their initial response to treatment. After two cycles in the model, response status was assessed with patients classified as either responders (CR/CRp) or refractory to induction therapy. Response status determined subsequent pathways, with patients transitioning to either the CR/CRp state (consolidation therapy or off-treatment) or refractory states (salvage or non-curative therapy).

Deaths during the induction period were estimated using the underlying parametric OS curves for refractory patients and CR/CRp patients.

Of those patients who achieved CR/CRp, a proportion received one cycle and a further proportion received two cycles of consolidation therapy. The proportions were based on data from the ALFA-0701 trial, reflecting treatment discontinuation due to disease progression or adverse events.

CR/CRp: Off-treatment

CR/CRp patients who did not receive consolidation therapy transitioned to the CR/CRp: off-treatment state after the induction period (2 cycles). CR/CRp patients who received consolidation therapy entered this health state after the consolidation period. Patients remained in this health state until transitioning to the HSCT, relapse or death states. After 5 years, all patients remaining in this health state transitioned to the functionally cured health state.

Refractory (salvage therapy)

A proportion ([redacted]) of the patients who were refractory to the induction treatments was assumed to receive a second-line salvage therapy for a mean of 1.5 cycles. The aim of salvage therapy was to attain second-line CR/CRp and bridge to a potentially curative HSCT. Following salvage therapy, patients transitioned to either the HSCT, refractory (non-curative) or death states.

Refractory (non-curative therapy)

Refractory patients who did not receive salvage therapy were assumed to receive non-curative therapy, which consisted of either low-intensity chemotherapy or best-supportive care (BSC). Refractory patients were also permitted to transition to this health state after failure of salvage therapy (i.e. those who do not bridge to HSCT). Patients remained in this health state until death.

Relapse (salvage therapy)

Patients who initially achieved CR/CRp were allowed to transition to the relapse state over time (salvage and non-curative therapies). Patients transitioning to relapse (salvage therapy state) were assumed to receive 1.5 cycles of salvage therapy.

Relapse (non-curative therapy)

Relapsed patients who did not receive salvage therapy transitioned to the relapse (non-curative therapy state), consisting of treatment with either low-intensity chemotherapy or BSC. Relapsed patients could also enter this health state after failure of salvage therapy to bridge to HSCT. Patients remained in this health state until either HSCT or death.

Response was not explicitly modelled in either of the relapse states. However, HSCTs were assumed to have been undertaken in patients who received prior salvage therapy and had achieved CR/CRp.

HSCT

Patients could receive HSCT from a number of health states at different (fixed) time points, including from "CR/CRp off-treatment" at [redacted], from "refractory" at [redacted] and from "relapse" at [redacted], [redacted], [redacted] and [redacted]. The time points were based on the average time to HSCT for each of the states. These timings were based on calendar time (i.e. from the point of randomisation) and not time in state. The use of calendar time implies that any relapses which occur after [redacted] will not receive HSCT.

Patients entering this health state from the refractory and relapse health states were assumed to have received prior salvage therapy and achieved CR/CRp. Patients remained in the HSCT state for one cycle, corresponding to the time that the company assumed that patients would spend hospitalised in an isolation room. Following this cycle, patients transitioned to one of the post-HSCT states (with or without GVHD) or death.

HSCT: CR/CRp (with GVHD)

Separate sub-states within the main HSCT state were used to capture the incidence and duration of acute and chronic GVHD.

A proportion of patients experienced GVHD (acute or chronic) after HSCT. The company stated that acute GVHD was assumed to last 2.5 months. Chronic GVHD was assumed to occur 6 months after HSCT and last for 9 months. Incidence rates of chronic and acute GVHD were applied to the number of patients surviving HSCT to estimate patients in this health state at the time points described.

Following the respective periods of GVHD, patients remained in the main HSCT health until transitioning to death (for period up to 60 months after HSCT). Patients remaining in this health state at 60 months after HSCT transitioned to the functionally cured health state.

HSCT: CR/CRp (without GVHD)

Patients who received HSCT and did not experience either chronic or acute GVHD, entered this state until transitioning to death (for period up to 60 months after HSCT). Patients remaining in this health state at 60 months after HSCT transitioned to the functionally cured health state.

Functionally cured

The CS reported that clinicians considered patients to be "functionally cured" (i.e. have long-term disease-free survival) in clinical practice after remaining in complete remission for 3 to 5 years. In the model, CR/CRp patients could enter the functionally cured health state if they remained alive and relapse-free at 5-years. Patients from other states could enter if they remained alive at 5-years after receiving an HSCT as a refractory, relapse or CR/CRp patient.

Patients in the functionally cured health state were assumed to be no longer at risk of relapse or mortality due to AML. However, longer-term excess mortality due other causes (e.g. higher risk of secondary cancers, cardiac events, etc.) was assumed for the remainder of the model time horizon by applying a hazard ratio (HR) to general population mortality rates.

ERG comment

The proposed model structure is complex and was challenging to critique given the difficulties in determining the actual flow of patients through the model. As part of the clarification questions, the company was requested to provide a clearer description of the assumptions and to explain the advantage of the state-transition model compared to a simpler and more conventional partitioned survival analysis (PartSA) model [49].

The company responded that the key benefit of GO was improved RFS (as opposed to differences in the initial induction success rates) and that patients receiving GO remained in CR/CRp for longer, delaying or avoiding subsequent relapses. The company noted that the proposed structure allowed for: separate survival analyses to be undertaken based on induction success or failure (i.e. CR/CRp or refractory); differentiation between relapse and refractory patients as well as capturing the impact of subsequent events (e.g. HSCT). The company argued that neither a conventional PartSA model (three-state model - relapse-free, relapse and death) nor a dual-partitioned survival analysis model (i.e. six-state model with separate states and transitions applied depending on induction success or failure)

would accurately capture the accrual of costs and QALYs and the impact of subsequent events such as HSCT.

PartSA and state-transition modelling differ fundamentally in their underlying structural assumptions. In the PartSA approach, each survival endpoint (e.g. RFS, OS) is modelled independently of the other endpoints included within the model. The differences in survival outcomes (e.g. RFS, OS) predicted between treatments are determined based on the differences in the survival curves chosen for these independently modelled events. In contrast, the state-transition modelling approach uses explicit structural links between health states, such that differences in survival outcomes are determined by the combined effect of each treatment on individual health states and the structural relationships assumed between these states.

To limit some of the additional structural complexity of the state transition model (and associated programming), the company included a series of separate sub-states. For example, relapse was included as a sub-state within the CR/CRp health state and HSCT was included as a sub-state within the CR/CRp, relapse and refractory health states. These sub-states allowed tracking of patients within the main states to inform cost and HRQoL estimates. However, the specific prognostic impact of these sub-states on subsequent transitions was not explicitly modelled, and the same survival probabilities were applied to the main states (CR/CRp and refractory) and the sub-states within these.

Complexity was also reduced by including additional structural assumptions for the HSCT state. This is most evident in the decision to use calendar time (i.e. time from randomisation) rather than time in state (i.e. time from relapse) as well as absolute probabilities at fixed time points. These assumptions serve two purposes: (i) reducing the complexity of the final model by avoiding the need for separate states to capture patient 'history' (e.g. whether HSCT was undertaken before or after relapse for CR/CRp patients and time since relapse); and (ii) ensuring the internal validity of the model predictions (i.e. the model predictions for HSCT rates and overall survival will be the same as those observed in the ALFA-0701 trial over the same time periods).

While the ERG acknowledges the importance of internal validity, it should also be noted that the absence of an explicit structural link in the proposed model between relapse and HSCT and mortality rates significantly limits the flexibility of the model to fully reflect longer term uncertainties. The lack of an explicit structural link between key states also results in the same independence assumption between clinical events that underpins the PartSA approach. As a result, the ERG considers that the proposed model provides limited (if any) advantage over a simpler PartSA approach. Equally, the ERG does not consider that the model confers any specific disadvantage, putting aside issues of transparency and programming complexity.

The ERG considers that judgements concerning the appropriateness of the proposed model structure (and external validity) can largely be simplified to the 'functional cure' assumption. The cost-effectiveness of GO+DA is driven by the longer term QALY benefits (and cost-offsets) attributed to the difference in the proportion of patients assumed to be functionally cured at 5-years. The route by which patients become functionally cured (i.e. whether the pathways are explicitly or implicitly modelled) appears less critical than assessing the validity of the functionally cured assumption itself.

The assumptions relating to functional cure and the robustness of the data to support these are discussed in detail in subsequent sections.

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

Table 18 summarises the economic submission and the ERG's assessment of whether the *de novo* evaluation meets NICE's reference case and other methodological recommendations.

Table 18 Features of the de novo analysis

Elements of the economic evaluation	Reference Case	Included in submission	Comment on whether de novo evaluation meets requirements of NICE reference case
Comparator(s)	The comparator specified in the NICE final scope and CS is "Established clinical management without gemtuzumab ozogamacin". The NICE final scope lists the most commonly used chemotherapies (amsacrine, cytarabine, daunorubicin, etoposide, fludarabine, idarubicin, mitoxantrone, thioguanine and midostaurin)	Yes	The chemotherapy comparator was restricted to daunorubicin and cytarabine regimens. Specifically, daunorubicin and cytarabine (3+7 regimen) in the induction phase, followed by daunorubicin and intermediate-dose cytarabine, in the consolidation phase. Although the comparator was more restrictive than the NICE scope, the ERG considered that the chosen comparator adequately represents the standard of care outside of a clinical trial setting.
Type of economic evaluation	Cost-effectiveness analysis.	Yes	
Perspective on costs	NHS and personal and social services	Yes	NHS and PSS costs have been taken into account.
Perspective on outcomes	All health effects on individuals.	Yes	QALY benefits to treated individuals were considered.
Time horizon	Sufficient to reflect any differences in costs or outcomes between the technologies being compared.	Yes	40 years is equivalent to a life-time horizon.
Synthesis of evidence on outcomes	Systematic review.	Yes	The evaluation uses clinical evidence from ALFA-0701. Results from an IPD meta-analysis (including ALFA-0701) were reported but not used in the economic model due to the different (and unlicensed) dosing regimens in the other trials.
Measure of health effects	QALYs.	Yes	Utility values were used directly from the published literature for all health states. Utility values for the health states of chemotherapy, remission, HSCT, relapse and refractory were based on from published utilities (QLQ-C30 data mapped into EQ-5D values). The utility value for the functionally cured health state was based on UK EQ-5D population norms. Utility values obtained from a vignette study were explored in a scenario analysis.
Source of data for measurement of HRQoL	Reported directly by patients and/or caregivers.	Yes	
Source of preference data for valuation of changes in HRQoL	Representative sample of the public.	Yes	
Discount rate	Annual rate of 3.5% on both costs and health effects.	Yes	Costs and benefits were discounted at 3.5% per annum.
Equity weighting	An additional QALY has the same weight regardless of the other characteristics of the individuals receiving the health benefit.	Yes	No special weighting undertaken.
Sensitivity analysis	Probabilistic sensitivity analysis	Yes	

5.2.3 Population

The population included in the company's decision problem and economic model comprises adult patients not known to have unfavourable cytogenetics, with previously untreated, de novo AML. As previously stated in Section 3.1, this is a subpopulation within the overall population covered by the anticipated EMA marketing authorisation [23].

The license amendments regarding age and the restriction to CD33-positivity were proposed by EMA after the company submitted to NICE. Following notification of these changes, the company was requested to provide additional commentary and/or evidence (including any proposed revisions to the cost-effectiveness model) to support their initial submission.

The company stated that the clinical pathway and management of de-novo AML in patients aged 15-17 years was broadly similar to the population included in their economic model and that adolescent patients respond better to treatment. The company also noted the small number of adolescent cases per year (1%). The company concluded that the efficacy data from ALFA-0701 was generalisable to adolescent patients. Given the small numbers of adolescent cases, the company argued that the cost-effectiveness estimates would almost identical (and potentially even lower) if patients aged 15-17 years were also included. Therefore the company proposed no further amendments to the previously submitted economic model.

The ERG agrees that the small numbers of patients means that the ICER is unlikely to be significantly affected and that the ALFA-0701 trial appears sufficiently generalisable to the population who would be eligible for treatment with GO + DA in routine clinical practice. However, the ERG does not consider that the ICER would necessarily be lower, as younger patients may also respond better to DA alone.

In terms of the restriction to CD33-positive patients, the company reported that there would be no additional resource implications (e.g. additional testing). The company also noted that the majority of patients with de novo AML will express CD33. CD33 expression on AML blasts was determined in [redacted] of patients in the ALFA-0701 trial. Of the patients assessed, few patients ([redacted] were reported to have low CD33 expression (less than 30% of blasts). As a result, the company did not propose any changes to the clinical effectiveness or cost inputs in the model since the majority of patients with de novo AML express CD33.

The ERG notes that these clarifications were submitted at a very late stage in the process which did not leave sufficient time to fully explore any uncertainties and potential implications. However, the ERG considers that the arguments made by the company appear reasonable and concludes that the

impact of age is likely to be negligible in terms of the ICER results and for CD33 positive patients the base-case ICER may be slightly conservative.

As previously highlighted in Section 3.1, the company's decision problem population further excluded patients known to have unfavourable cytogenetics from within the broader marketing authorisation. The CS justified this restriction based on the subgroup results of the ALFA-0701 study and clinical advice that these patients would not be treated with GO plus intensive chemotherapy in clinical practice. The clinical advisor to the ERG also agreed with this view. Patients with unfavourable cytogenetics constituted around 21.0% of the total ALFA-0701 trial population (Table 19).

Table 19 Cytogenetic risk stratification of patients in ALFA-0701

Risk stratification	Favourable	Intermediate	Unfavourable	Unknown
GO+DA group	2.2%	67.4%	20.0%	10.4%
DA group	4.4%	65.4%	22.1%	8.1%
GO, gemtuzumab ozogamicin; DA, daunorubicin and cytarabine				

The main baseline characteristics of the population the company presented as their base-case analysis is summarised in Table 20.

Table 20 Mean baseline characteristics (population excluding known unfavourable cytogenetics)

Characteristic	GO + DA (N=108)	DA (N=106)	Total (N=214)
Age (years)	[redacted]	[redacted]	[redacted]
Gender (% male)	[redacted]	[redacted]	[redacted]
Body surface area (m^2)	[redacted]	[redacted]	[redacted]
Weight (kg)	[redacted]	[redacted]	[redacted]
GO, gemtuzumab ozogamicin; DA, daunorubicin and cytarabine			

The ERG agrees with the company's decision to exclude patients with known unfavourable cytogenetics and their rationale for focusing on the specific subpopulation where GO+DA provides clear clinical benefit and optimises cost-effectiveness.

The subgroup results requested by the ERG and summarised in Section 4.2.4.3, indicate there may be other subgroups (e.g. intermediate-2 patients) where clinical benefit also appears to be unclear and hence further optimisation in cost-effectiveness may be appropriate. The ERG does not consider that

the company have adequately explored any remaining heterogeneity within the proposed population and possible implications for clinical and cost-effectiveness.

Heterogeneity in subgroup with unknown cytogenetics

The base-case population included patients with unknown cytogenetics. The company's rationale for including these patients reflected both statistical (i.e. to increase the sample size and the robustness of statistical modelling) and clinical considerations (i.e. patients may need to be treated immediately depending on the severity of their symptoms rather than waiting for further confirmatory tests).

The ERG considers that the company did not sufficiently justify the inclusion of the subgroup with unknown cytogenetics and/or attempt to fully explore the implications of alternative assumptions. The ERG notes that the decision to include these patients appears appropriate based on the ALFA-0701 results and the additional subgroup results presented by the company. However, the ERG identified important difference between the findings from the ALFA-0701 trial and the results of the IPD meta-analysis for the subgroup with unknown cytogenetics. The ERG notes that the authors of the IPD meta-analysis concluded that "as an individual group we could not see them benefiting from GO" (HR for OS = 1.01 [0.84-1.22] [29]).

In response to the points for clarification related to this specific subgroup, the company stated that only the ALFA-0701 study included a dosing regimen for GO in the IPD meta-analysis that is expected to be recommended by the EMA. Additionally the company stated that there were differences between trial populations in the IPD meta-analysis (i.e. not all studies included only *de novo* patients). While the ERG acknowledges these differences, significant uncertainties remain regarding the appropriate management of this subgroup. The ERG considers that the differences reported across studies and implications for cost-effectiveness have not been fully addressed by the company.

The ERG also requested that the company exclude this population from the cost-effectiveness result to further consider the impact of including/excluding these patients on the incremental cost-effectiveness ratio (ICER).

[redacted]

The reason for the contradictory findings from the ALFA-0701 and the IPD meta-analysis remains unclear to the ERG. One plausible explanation might relate to heterogeneity within the subgroup

according to the reason for the unknown cytogenetic result (e.g. whether the test was not undertaken versus the test result analysis was unsuccessful in determining cytogenetic status).

The ERG's clinical advisor indicated that the majority of patients would have a known cytogenetic result prior to starting GO. However, they also noted that there would be a small proportion of patients (approximately 5%) for whom cytogenetic risk is not known at the point at which a decision to start treating with GO would be made. The majority of these patients were considered to be those for whom the cytogenetic test was unsuccessful. For these patients, the ERG's clinical advisor suggested that she would potentially start GO therapy and undertake further tests (e.g. molecular). If these additional tests subsequently indicated unfavourable prognostic features the ERG's clinical advisor stated that she would stop GO and continue with conventional chemotherapy backbone regimen.

Heterogeneity in the subgroup with intermediate cytogenetics

The intermediate group of patients is the largest subgroup in the ALFA-0701 trial. The ERG considers there to be heterogeneity within this group with regards their underlying genetic biomarkers and the resulting survival rates. The additional subgroup results requested by the ERG and previously discussed in the clinical effectiveness section also suggest important heterogeneity in the clinical effectiveness of GO in the intermediate group. The majority of survival benefit appeared to be conferred within the intermediate-1 subgroup.

The ERG's clinical advisor reported that while cytogenetic testing is routinely undertaken for AML patients in clinical practice, molecular testing is not at present. Hence, the value and appropriateness of further exploring heterogeneity within the intermediate group may be debatable if further risk stratification is not considered practical and/or feasible within routine clinical practice. However, the ERG was also aware that the ALFA-0701 trial design specified that patients classified as favourable or intermediate-1 (i.e. based on cytogenetic and molecular tests) were not considered for HSCT in first complete remission, whereas patients with intermediate-2 or unfavourable risk who experienced a CR were considered for transplant. The ERG considered that further subgroup analysis was justified to also explore the possible confounding effect of HSCT within the ALFA-0701 trial.

The ERG's clinical advisor also reported that HSCT is not routinely considered for all intermediate patients in the UK and is reserved for higher-risk patients, indicating that some clinicians are already stratifying the intermediate population based on other considerations (including additional molecular test results) in guiding decisions regarding eligibility for HSCT.

The ERG requested data on the intermediate-1 and intermediate-2 patients to undertake additional further exploratory analyses for these subgroups. The results of these are presented in Section 6.

5.2.4 Interventions and comparators

5.2.4.1 First-line therapy

The company's economic model evaluated GO+DA compared to DA alone. Treatment consisted of an induction phase on which all patients start, and a consolidation phase which a proportion of patients receive as further treatment after induction therapy. The dosing and administration schedule of GO and DA in the model reflected that of the ALFA-0701 trial.

The fractionated dose of GO given in ALFA-0701 and applied in the model reflects the expected marketing authorisation license (3mg/m^2 on days 1, 4 and 7 for a total maximum dose of 5mg). GO is not administered during second induction therapy. A second induction course was permitted in the trial and included in the model, if more than 5% leukaemic blasts persisted at the Day 15 bone marrow aspiration (BMA). Second induction in the ALFA-0701 trial was not permitted to be initiated later than Day 22 (page 34 of CSR [50]).

For patients experiencing a complete remission following induction, up to 2 consolidation courses were provided. GO is provided on Day 1 of each consolidation course, using a dose of 3mg/m^2.

Table 21 provides a summary of the dosing schedules applied in the model.

Table 21 Dosing regimens of induction therapy treatments

	GO+DA	**DA alone**
Induction therapy		
First induction	GO: 3mg/m^2 on Days 1, 4 and 7 DNR: 60 mg/m^2/day on Day 1 to Day 3 AraC: 200 mg/m^2/day on Day 1 to Day 7	DNR: 60 mg/m^2/day on Day 1 to Day 3 AraC: 200 mg/m^2/day on Day 1 to Day 7
Second induction	DNR: 35 mg/m^2/day on Day 1 to Day 3 AraC: 1000 mg/m^2/12 hours on Day 1 to Day 3	DNR: 35 mg/m^2/day on Day 1 to Day 3 AraC: 1000 mg/m^2/12 hours on Day 1 to Day 3
Consolidation therapy		
First consolidation	GO: 3mg/m^2 on Day 1 DNR: 60 mg/m^2/day on Day 1 AraC: 1000 mg/m^2/12 hours on Day 1 to Day 4	DNR: 60 mg/m^2/day on Day 1 AraC: 1000 mg/m^2/12 hours on Day 1 to Day 4
Second consolidation	GO: 3mg/m^2 on Day 1 DNR: 60 mg/m^2/day on Day 1 and Day 2 AraC: 1000 mg/m^2/day on Day 1 to Day 4	DNR: 60 mg/m^2/day on Day 1 and Day 2 AraC: 1000 mg/m^2/day on Day 1 to Day 4
GO, gemtuzumab ozogamicin; DA, daunorubicin and cytarabine; DNR, daunorubicin; AraC, cytarabine		

As previously highlighted in the decision problem section (Section 3.2 and 3.3), there are differences between the licensed dosage of GO and the dosing currently used within UK clinical trials (AML18 and AML19). There also exist potential differences in the number of courses of consolidation therapy used in clinical practice for conventional chemotherapy regimens. Despite these differences, the ERG considered the dosing regimen assumed for the GO+DA to be appropriate and aligned with the anticipated marketing authorisation and that the comparator included in the model was likely to be sufficiently representative of the current standard of care outside of a clinical trial setting.

5.2.4.2 Subsequent lines of therapy

Following a relapse or a lack of response to induction therapy, patients in the model could receive salvage therapy or non-curative therapy. Salvage therapy consisted of FLAG-Ida, given for 1.5 cycles (the midpoint of 1-2 cycles, the schedule typically provided as per the clinician's discretion). The proportion of relapsed or refractory patients receiving salvage therapy was assumed to be [redacted], based on clinical advice received by the company.

Patients not fit enough to receive salvage therapy could receive non-curative therapy. This consisted of low dose chemotherapy and best supportive care (BSC). The company assumed that this consisted of hydroxycarbamide, low-dose cytarabine and azacitadine, used in a [redacted] ratio. Non-curative therapy was received after salvage therapy (until some patients received HSCT), consisting of hydroxycarbamide only. The company applied standard UK treatment schedules for these therapies (

Table 22).

Table 22 Dosing regimens of treatments provided in subsequent lines of therapy

Therapy	Dose	Frequency (doses per month)
Salvage therapy		
Fludarabine	30mg/m^2	5 doses
Cytarabine	2000mg/m^2	5 doses
G-CSF	263µg	7 doses
Idarubicin	8mg/m^2	3 doses
Non-curative therapy		
Hydroxycarbamide	1000mg	28 doses
Cytarabine	40mg	10 doses
Azacitidine	75mg/m^2	7 doses
G-CSF: granulocyte-colony stimulating factor		

ERG comment

The ERG considers that the use of FLAG-Ida to represent the second line of chemotherapy (salvage therapy) to be reasonable and reflective of UK clinical practice. The ERG notes that the proportion of patients receiving subsequent lines of therapy (and the specific types of therapies) in the model was based on clinical expert opinion. The ERG was unclear of the reason for using assumptions given that data on subsequent therapies were collected within the ALFA-0701 trial and reported in the CSR[50]. The ERG notes that the way this information was reported in the CSR made it difficult to assess the validity of the company's assumptions. However, despite the lack of justification provided by the company, the ERG does not consider that these assumptions are likely to be a significant driver of cost-effectiveness.

A proportion of patients in the DA group of the ALFA-0701 trial also received a dose of GO as part of the compassionate programme ([redacted] of patients in the overall population of the chemotherapy arm). These costs were not accounted for in the model even though the survival data was not adjusted for cross-over between arms (Section 5.2.6). The approach used by the company is therefore potentially conservative by excluding the costs and not adjusting for the potential impact on OS. However, the ERG also notes that [redacted] of patients in the GO+DA arm also received GO as part of a subsequent therapy. Although these were treated in a similar manner (i.e. costs were excluded without any adjustment), some uncertainty remains concerning whether patients might be re-treated with GO in routine clinical practice and whether these costs should have been included for GO+DA.

5.2.5 Perspective, time horizon and discounting

The economic model adopted a National Health Service (NHS) and Personal Social Services (PSS) perspective in accordance with the NICE reference case.

The time horizon used in the economic model was 40 years. The CS stated that this was sufficient to capture lifetime costs and benefits. The costs and benefits in the model were discounted at an annual rate of 3.5%, as per the NICE reference case.

Implementation of discounting in the economic model was carried out on an annual basis, such that all costs and benefits incurred with any given year are discounted by the same amount regardless of whether they occur at the start or the end of that year.

ERG comment

A 40-year time-horizon appears to be appropriate based on the average age assumed and the potential curative assumptions employed. The average age (and distribution) in the model is based on patients in the ALFA-0701 trial ([redacted]) such that the probability of patients still alive at 40 years is considered sufficiently small to represent a lifetime horizon. The ERG acknowledges that a longer horizon would be required for younger patients.

The ERG considers that discounting on a per cycle basis is more accurate as it more closely reflects the actual time at which benefits and costs occur. The impact of this issue on the ICER is, however, likely to be minimal and therefore is not explored further as part of ERG exploratory analysis.

5.2.6 Treatment effectiveness and extrapolation

The CS provided a description of the clinical data used in the model. These were based on ALFA-0701, and included: response to first-line treatment, overall survival, relapse-free survival, probability of HSCT, the cure rate of HSCT, and the incidence rate of treatment-emergent adverse events. Additionally, the model estimated mortality in the functionally cured population, which was based on general population mortality [51] and adjusted to account for excess mortality in AML survivors.

5.2.6.1 Response to first-line treatment

Response to first-line treatment was modelled as the probability of achieving CR or CRp after induction therapy. Response data were pooled across treatment arms and justified in the CS based on the lack of any statistically significant difference and clinical advice that GO was not expected to affect the initial response outcomes. The main effect of GO was therefore assumed to be in terms of the additional durability of the response outcomes.

The base-case analysis used probabilities based on response status assessed by the independent review committee (IRC). The company provided the corresponding data for investigator-assessed response (IA) in their clarification response for the base-case population and the additional subgroups requested by the ERG.

These response outcomes for the base-case population are summarised in Table 23.

Table 23 Response status for the favourable/intermediate/unknown subgroup (adapted from CS, Table 34)

	GO + DA (N = 108)	**DA (N = 106)**	**Pooled (N = 214)**
IRC assessed (base case analysis)			
CR + CRp, n (%)	[redacted]	[redacted]	[redacted]
Induction failure, n (%)	[redacted]	[redacted]	[redacted]
Investigator assessed			
CR + CRp, n (%)	[redacted]	[redacted]	[redacted]
Induction failure, n (%)	[redacted]	[redacted]	[redacted]

CR, complete remission; CRp, complete remission with incomplete platelet recovery; DA, daunorubicin and cytarabine; GO, gemtuzumab ozogamicin; IRC, independent review committee

ERG comment

The ERG did not consider that it was necessary or appropriate to pool response data. Even if the differences in response rates are not considered clinically important, any differences should still be considered and the lack of statistical significance should be reflected in the distributions assigned to these parameters. Also, subsequent treatment decisions (e.g. rates of consolidation therapy, decision to proceed to HSCT) in the ALFA-0701 trial will have been based on the individual response rates.

While the ERG acknowledges the arguments made by the company in relation to the decision to use IRC as opposed to IA analyses, the ERG considers that that the initial treatment costs of the induction and consolidation therapies should be based on the IA response outcomes. IA outcomes more appropriately reflect the actual treatment decisions and resource use incurred within the trial.

5.2.6.2 Survival analysis

The company fitted parametric survival curves for RFS and OS using the patient-level data from the ALFA-0701 trial to extrapolate over the model time horizon. The OS analysis was stratified by response status due to the heterogeneity in survival outcomes between CR/CRp and refractory patients. The parametric survival curves were converted to monthly health state transition probabilities.

The economic model was based on data collected up to the April 2013 data cut-off, and analysis of RFS was based on IRC assessment. The company acknowledged that the IA analysis may be more reflective of clinical practice, but elected to use the IRC analyses as they were "generated according to regulatory requirements" and addressed any possible bias introduced by the local investigators as a result of the open-label design of the study. The company reported good agreement between IRC and investigator analyses: the results are not presented for the cytogenetic population used in the company base-case analysis, but the two assessments in the mITT analysis provided similar HRs of RFS ([redacted] for investigator and [redacted] for IRC).

Survival models

Models were selected based on statistical tests and clinical plausibility. The company explored a range of alternative approaches, including convention single parametric models, spline-based models and mixture cure models (MCM).

A summary of the survival functions used in the company base-case model are presented in Table 24. The model predictions based on the selected survival functions were reported to have been validated by external experts (two clinicians and a statistical expert).

Table 24 Summary of survival functions in the company base-case analysis (adapted from CS, Table 38)

Endpoint	GO + DA	DA	Pooled
RFS (CR or CRp)	MCM log-normal Cure rate: [redacted]	MCM log-normal Cure rate: [redacted]	—
OS (CR or CRp)	MCM log-normal Cure rate: [redacted]	MCM log-normal Cure rate: [redacted]	—
OS (refractory)	—	—	Gompertz Cure rate: n/a

CR, complete remission; CRp, complete remission with incomplete platelet recovery; GO, gemtuzumab ozogamicin; OS, overall survival; RFS, relapse-free survival; MCM, mixed cure model

Based on inspection of the Kaplan-Meier (KM) curves of OS in CR/CRp and of RFS, the company identified a plateau occurring between [redacted]. In order to appropriately capture the plateau in the curve, the company investigated the use of more complex survival models (including flexible spline modelling and mixture cure models), as well as standard parametric models. The CS stated that the more complex survival models were considered to provide an advantage in capturing the plateau observed in the KM-curves.

In situations where a proportion of patients experience long-term durable remissions for their illness, there can be significant heterogeneity in survival data. Standard parametric (and spline models) group all patients together and provide a single prediction of survival for the entire group. In contrast, the MCM approach assumes that the group comprises a mixture of patients who are cured and patients who are not cured. The heterogeneity in survival data is reflected in the MCM by estimating the probability a patient is cured (the 'cure fraction') and associated survival predictions. Patients who are cured are assumed to be subject to background (non-AML) mortality only. Patients who are not cured are subject to the background mortality and to additional mortality from AML. The additional mortality from AML is modelled using conventional parametric functions (Weibull, lognormal, generalised gamma). The CS cited references and considered MCM to be established statistical practice in the AML disease area.

[redacted] receiving DA alone, subsequently received GO as a follow-up therapy through a compassionate use programme. Although the company undertook a feasibility assessment to determine whether formal adjustment methods could be applied, they concluded that the assumptions

required did not hold and that any adjustments would be unreliable. While the company noted that that the compassionate use programme could bias efficacy in favour of the comparator arm, there remains significant uncertainty surrounding the possible magnitude of this.

5.2.6.3 Survival models

Overall survival – CR/CRp

The curve selected by the company for their base-case analysis, the MCM lognormal curve, had the best fit according to AIC/BIC statistics and provided the most conservative estimates of overall survival. Figure 11 summarises the predicted OS estimates for the alternative MCM functions and Table 25 reports the predicted cure fractions for each MCM.

Figure 11 Predicted OS[CR]

Table 25 Cure fractions - overall survival in CR/CRp patients

MCM distribution	GO+DA group	DA group
MCM Weibull	[redacted]	[redacted]
MCM lognormal	[redacted]	[redacted]
MCM generalised gamma	[redacted]	[redacted]

Overall survival – refractory

OS for refractory patients was pooled in the base case, since the company were advised by clinical experts that GO would not affect OS for refractory patients.

A number of distributions were fit to survival data pooled from refractory patients in both treatment groups, including the standard six parametric distributions and three spline-based models with one knot. The curve selected by the company for their base-case analysis, the Gompertz curve, had the best fit according to AIC/BIC statistics, and the company also considered that it had the best visual fit, stating that the spline-based models resulted in late-occurring plateaus.

Figure 12 OS in refractory patients (CS, Figure 16)

Relapse-free survival – CR/CRp

The curve selected by the company for their base-case analysis, the MCM lognormal curve, had the best fit according to AIC/BIC statistics. The company considered that the MCM Weibull and MCM lognormal provided a similar visual fit, but that the lognormal "provides the best fit to the plateau" and was considered by the company to best capture the benefits of GO+DA.

Figure 13 RFS (CS, Figure 14

Table 26 Cure fraction of RFS in CR/CRp patients

MCM distribution	GO+DA group	DA group
MCM Weibull	[redacted]	[redacted]
MCM lognormal	[redacted]	[redacted]
MCM generalised gamma	[redacted]	[redacted]

ERG comment

Overall, the ERG considered the approach to curve fitting and the rationale for selecting distributions to be appropriately justified. Uncertainties surrounding the choice of survival functions were also explored using a range of alternative functions within separate scenarios.

Although the alternative MCM distributions reported different estimates of the absolute cure fraction for each group, the difference in the cure fraction between the groups was broadly similar for both the MCM lognormal and Weibull functions for both EFS and OS. This is important because it is the difference between the groups in the probability of long-term survival which is the main driver of QALY differences and the ICER estimates.

The ERG felt that the company should have provided more explanation regarding the larger differences estimated using the MCM generalised gamma distribution and the clinical plausibility of the higher cure fractions for OS compared to EFS. The ERG considers that the differences observed using the generalised gamma distribution may be due to difficulties in achieving convergence. However, the clinical plausibility of the higher cure fraction estimated for OS compared to EFS was not discussed in the CS. The difference in the cure fractions for EFS and OS suggests that either: (i) there are a significant number of patients who become functionally cured following relapse (potentially due to subsequent therapies and/or HSCT) or (ii) the data may not be sufficiently mature to robustly estimate the cure fraction for OS.

Despite the potential concerns regarding the difference in the absolute cure fractions reported for EFS and OS, the ERG notes that the differences in the cure fraction between the groups is similar for both endpoints. This provides additional reassurance regarding the robustness of the ICER results. As previously stated, it is the difference in the cure rates between the groups (as opposed to the absolute cure rates) this is most critical for determining the appropriateness of the ICER results.

For the base-case population, the ERG considers that the choice of survival function appears less critical than the assumptions which are subsequently applied to long-term survivors regarding potential excess morbidity (i.e. HRQoL assumptions) and mortality. However, the ERG also notes that there remains significant heterogeneity in outcomes within the base-case population which hasn't been fully explored in the CS. These issues are further explored in the following sections and additional ERG exploratory analyses are also provided in Section 6.

5.2.6.4 Mortality in the cured population

To capture the excess mortality (relative to the general population) for functionally cured patients at five years, the company applied a hazard ratio of [redacted] to the general population mortality rates. The company argued that functionally cured patients remained at higher risk of other health conditions which may increase mortality rates above that of the general population, including secondary or relapsed cancer, late complications following an HSCT, or cardiovascular disease following an anthracycline (such as daunorubicin and idarubicin).

The company undertook an analysis of pooled survival data from UK AML trials 10 to 16, restricted de novo AML patients to the intermediate and favourable cytogenetic subgroup (a total of [redacted] intermediate and [redacted] favourable patients) aged 50 to 70, using survival curves conditional on surviving the first five years. The hazard ratio was estimated by calculating the ratio of the means of

the annual mortality rates, from five years after AML diagnosis for AML patients and of those matched to the mean age of the analysis from the general population.

The excess mortality hazard ratio was applied after the cure point (at five years), and was applied to patients considered to be cured after consolidation therapy as well as HSCT.

Table 27 Excess mortality for long-term AML survivors versus the general population

Time since AML diagnosis	Mean age	Proportion alive – AML patients	Mortality rate – AML patients	Mortality rate – general population	HR per cycle
5	[redacted]	[redacted]		0.010	
6	[redacted]	[redacted]	[redacted]	0.011	[redacted]
7	[redacted]	[redacted]	[redacted]	0.012	[redacted]
8	[redacted]	[redacted]	[redacted]	0.013	[redacted]
9	[redacted]	[redacted]	[redacted]	0.015	[redacted]
10	[redacted]	[redacted]	[redacted]	0.016	[redacted]
11	[redacted]	[redacted]	[redacted]	0.018	[redacted]
12	[redacted]	[redacted]	[redacted]	0.020	[redacted]
13	[redacted]	[redacted]	[redacted]	0.023	[redacted]
14	[redacted]	[redacted]	[redacted]	0.026	[redacted]
15	[redacted]	[redacted]	[redacted]	0.028	[redacted]
16	[redacted]	[redacted]	[redacted]	0.031	[redacted]
17	[redacted]	[redacted]	[redacted]	0.034	[redacted]
18	[redacted]	[redacted]	[redacted]	0.038	[redacted]
19	[redacted]	[redacted]	[redacted]	0.043	[redacted]
20	[redacted]	[redacted]	[redacted]	0.049	[redacted]
Source: Company analysis of AML10-16 patient data Source: ONS [51]. Weighted rates based on ALFA-0701 baseline age and gender					

ERG comment

The decision to adjust general population mortality rates in AML survivors appears supported by clinical evidence and the ERG is generally satisfied with the manner in which it is implemented. However, some uncertainties remain regarding the estimation of the adjustment factor (the hazard ratio).

Firstly, the number of patients at risk in the analysis of AML10-16 trial data was not reported and therefore it was difficult to determine how robust the estimates of mortality are in the later years. These values may be based on small patient numbers.

Secondly, the HR per cycle appears higher in the years immediately following year 5 before settling into a more consistent pattern. The higher HR in the earlier years may indicate that surviving patients are still at risk of AML-related relapse and associated mortality, suggesting that 5-years may be too early to establish that patients are functionally cured. Equally, it should be noted that the dataset includes all patients who are still alive at year 5, which may include patients who have relapsed and are still at risk of AML related mortality. Since the company model only applies the assumption of functional cure to patients from particular states (e.g. CR/CRp without relapse and for post-HSCT survivors), 5-years may still be a reasonable time to implement the excess-mortality adjustment.

Finally, in some years, the probability of death was higher in the general population than in the AML survivors, which does not seem plausible. This might be just due to the small sample size and the play of chance. However, the ERG considers that a further adjustment appears appropriate, such that the mortality rate is set equal to the general population mortality rate in instances when the observed mortality rate is reported to be lower than the general population. The impact of the adjustment is explored by the ERG in Section 6.

5.2.6.5 Hematopoietic stem-cell transplantation

Patients were able to receive HSCT from three health states in the model: from the CR/CRp health state, the refractory health state, and the relapse health state. The probabilities of receiving HSCT were estimated from data of patients receiving HSCT in the ALFA-0701 trial, excluding those with unfavourable cytogenetics. The probabilities are summarised in Table 28.

Table 28 Annual probability of HSCT (CS, Table 39)

Probability of HSCT	Subpopulation			Timing (months)
	GO + DA	DA	Pooled	
From CR or CRp (without relapse)				
Total, %	NA	NA	[redacted]	[redacted]
From Refractory				
Total, %	NA	NA	[redacted]	[redacted]
From Relapse				
Year 1, %	[redacted]	[redacted]	[redacted]	[redacted]
Year 2, %	[redacted]	[redacted]	[redacted]	[redacted]
Year 3, %	[redacted]	[redacted]	[redacted]	[redacted]
Year 4, %	[redacted]	[redacted]	[redacted]	[redacted]
Year 5, %	[redacted]	[redacted]	[redacted]	[redacted]

DA, daunorubicin and cytarabine; GO, gemtuzumab ozogamicin; HSCT, hematopoietic stem-cell transplant; CR, complete remission; CRp. Complete remission with incomplete platelet recovery

ERG comment

The ERG previously highlighted that the company limited the complexity of the model by including additional structural assumptions for the HSCT state and using calendar time (i.e. time from randomisation) rather than time in state (i.e. time from relapse) as well as absolute probabilities at fixed time points. These assumptions also ensured that the model predicted identical HSCT rates as observed in the ALFA-0701 trial.

HSCT rates are one of the central drivers of cost-differences in the model. The reduced proportion of patients receiving HSCT with GO+DA provides an important potential cost-offset in the model. The main uncertainty is whether the data from ALFA-0701 can be considered sufficiently mature to provide an accurate estimate the long-term difference in HSCT rates between the two treatment arms. The ERG's concern is whether possible differences in time at risk for relapse patients between the treatment arms (i.e. patients relapsing earlier with DA alone may have a longer time at risk for HSCT) could lead to an over-estimate of the differences in the HSCT rates between the groups predicted over a longer time period. The ERG requested additional data and Kaplan-Meier curves during the clarification stage.

The data provided suggested no obvious bias or differences in the time at risk. However, the ERG notes that some of the cost-offsets for HSCT are predicated on the functional cure assumption (i.e. that those patients who have not relapsed by five years will never relapse and hence will not require HSCT at some point in the future). The absence of any structural link to the rates of HSCT limits the ability of the ERG to further assess the potential impact of this source of uncertainty.

5.2.6.6 Mortality post-HSCT

After an HSCT, patients were assumed to remain in the post-HSCT CR/CRp health state until death. The OS curve for these patients was based on the underlying OS curve for the health state that the patient was in, until the point at which the proportion alive reached the estimated cure rate of HSCT. After this point, the mortality rate was assumed to be that of the adjusted general population. Hence, until the cure rate for HSCT is reached, both HSCT and non-HSCT patients were subject to the same rate of OS, at which point the survival of the two groups in the model diverged.

The cure rate for HSCT was estimated to be [redacted] based on an analysis of post-HSCT overall survival pooled from all HSCT data in the ALFA-0701 trial. Data was pooled from both arms of the trial and for HSCTs with all health states, since clinical experts advised the company that OS would

be similar regardless of what point in the treatment pathway HSCT occurred (i.e. after refractory disease, after first remission or after relapse). No details were provided on the methods of this analysis and it is unclear whether the [redacted] was estimated via a visual inspection of the KM survival curve or from a formal statistical analysis.

ERG comment

The ERG's clinical advisor considered that the HSCT cure rate would be different depending on whether HSCT was received after relapse or after first CR/CRp. Disease characteristics are often different after second CR following relapse, with patients having higher levels of comorbidities and greater organ damage. HSCT patients after first CR/CRp would therefore be expected to have a higher survival rate. The ERG subsequently requested additional data from the company regarding the number of patients receiving HSCT (B15 in the points for clarification) and their survival outcomes following induction failure, before relapse and after relapse.

While visual inspection of the Kaplan-Meier graphs suggested differences in the cure rate across these separate groups, the ERG notes that these analyses are based on small patient numbers and thus considers the decision to pool across these groups to be reasonable.

5.2.6.7 Safety

Related to first-line therapy

The company analysis modelled the impact of grade 3 to 4 treatment-emergent adverse events (AEs) relating to GO and DA. Incidence rates were estimated from ALFA-0701 trial data. The model included AEs that occurred in at least 1% of patients in either arm of the trial. These are summarised in Table 29.

Table 29 Incidence of treatment-emergent Grade 3+ adverse events for first-line AML therapies (CS, Table 40)

Adverse event	GO + DA	DA
Skin toxicity	[redacted]	[redacted]
Mucosal toxicity	[redacted]	[redacted]
Pain	[redacted]	[redacted]
Nausea, vomiting, and diarrhoea	[redacted]	[redacted]
Pulmonary toxicity	[redacted]	[redacted]
Cardiac rhythm disorder	[redacted]	[redacted]
Other cardiac toxicity	[redacted]	[redacted]
Central neurological toxicity	[redacted]	[redacted]
Peripheral neurological toxicity	[redacted]	[redacted]
Infections	[redacted]	[redacted]
Haemorrhage	[redacted]	[redacted]
Veno-occlusive disease	[redacted]	[redacted]
GO: gemtuzumab ozogamicin, DA: daunorubicin + cytarabine		

Related to HSCT

The model also captures the impact of acute and chronic GVHD as a consequence of HSCT. The incidence rate of GVHD was extracted from two sources. Acute GVHD was assumed to occur within the first 100 days post-HCT, and had an associated incidence rate of 15% and duration of 2.5 months. Chronic GVHD was assumed to occur 6 months after HSCT, with an associated incidence rate of 5.5% and duration of 9 months. The incidence rates, duration and evidence sources are summarised in Table 30.

Table 30 Incidence rates and duration of GVHD

Event	Incidence rate	Source	Duration	Source
Acute GVHD	15%	Battipaglia (2017)	2.5 months	Clinical opinion
Chronic GVHD	5.5%	NHS England (2017)	9 months	Clinical opinion

ERG comment

[redacted] (Table 29, CS). As a result, the ERG considers that these costs should not have been assigned to DA alone. The impact of excluding these cases are considered as part of the ERG exploratory analysis presented in Section 6.

5.2.7 Health related quality of life

5.2.7.1 Health state utilities

The ALFA-0701 trial did not collect HRQoL evidence from the trial participants. The company therefore undertook a systematic literature review of utility studies that reported relevant health state values. The review did not identify any publications reporting data on HRQoL that met the NICE reference case in full. However, six publications that met the search inclusion criteria were identified.

In the base-case analysis, the company incorporated utility values from a number of sources.

Table 31 provides a summary of the utility values used within the model for the base-case analysis and tested in scenario analyses.

Table 31 Summary of health state utility values

Source	Chemotherapy treatment*	Consolidation therapy	HSCT	GVHD (post-HSCT)	CR/CRp off-treatment	Relapse	Refractory	Functionally cured
Values used in base–case analysis								
TA399	0.66	0.66	0.66	0.67*	0.74	0.57	0.57	0.82*
Values used in scenario analysis								
TA399	0.72	0.72	0.72	0.67*	0.77	0.62	0.62	0.82*
Pfizer TTO	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
Pfizer VAS	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
*Applied to patients in induction, salvage and non-curative *Varied per cycle, based on mean patient age at each time point, from Ara & Brazier *Source Kurosawa 2014 [40]								

The utility values assigned to several states were based on a previous NICE technology appraisal for azacitidine (TA399) [35] and clinical expert opinion was used to map these values to health states in the model where health state descriptions were not aligned. These utility values were calculated from EORTC QLQ-C30 data collected in the azacitidine pivotal trial, which enrolled patients over the age of 65 with *de novo* or secondary AML with >30% bone marrow blasts who were not eligible for HSCT. Two mapping algorithms were identified and used in TA399: the company in the present appraisal elected to use the values estimated with the McKenzie & Van der Pol algorithm that provided utility values closer to the values in their preference elicitation study. A company-presented scenario analysis with the alternative set of utility values resulted in very little change to the ICER.

The company made a number of assumptions in order to map utility values from TA399 to the health states in their analysis. Quality of life for patients on intensive chemotherapy (including induction therapy, consolidation therapy and salvage therapy) was assumed to be 0.657. Differences between arms for patients on GO+DA and on DA were captured by estimating the utility decrement associated with the different safety profiles. In TA399, this was the value estimated from patients in non-remission (stable disease or partial remission).

The company also assumed that patients in the month after having an HSCT would have a similar quality of life to those on intensive chemotherapy, and the same utility value (0.657) was applied in this health state.

Patients in remission (CR or CRp) and off-treatment had a utility value of 0.740, mapped from the health state utility value for remission patients (CR or CRi) in TA399. The utility value for relapsed patients and refractory patients receiving non-curative treatment (including either best supportive care

or low-intensity chemotherapy) was assumed to be 0.568, estimated from patients in TA399 who were post-progression or who had relapsed.

Patients in the functionally cured health state were assumed to have quality of life consistent with patients of the same age in the general population. Age- and gender-matched utilities were estimated from a formula estimated by Ara & Brazier, who used EQ-5D data and UK preference rates to determine quality of life according to a number of patient covariates [38]. These values were estimated on a per-cycle basis to allow the model to capture the gradual decline in HRQoL associated with ageing. Using the formula for a patient aged 61 results in a utility value of 0.820.

The utility value for patients with GVHD following HSCT was 0.67, and was taken from a published economic analysis of HSCT patients in first remission from AML [40] which was identified in the SLR. The reference for this value was not provided by the company and therefore the utility value could not be verified. The economic analysis did not differentiate between chronic and acute GVHD.

The company also conducted a separate preference elicitation study, which recruited 125 participants from the general population. A series of vignettes were developed, corresponding to each health state in the analysis. Each participant had a one-to-one face-to-face interview and was asked to value a range of health state descriptions using the time-trade-off and visual analogue techniques. However, the company chose not to use the resulting values from this study in their base-case analysis, stating that utilities from the studies identified in the review more closely aligned with the NICE reference case. A separate sensitivity analysis was conducted using the values estimated from the elicitation study.

ERG comment

In the absence of direct HRQoL data available in ALFA-0701, the ERG considered the approach used by the company to be reasonable and appropriately justified. The only exception to this was the separate assumption made that functionally cured patients experience the same HRQoL as the general population. This assumption results in a marked jump in the HRQoL estimates at 5-years for functionally cured patients. The use of general population quality of life was not considered internally consistent with the excess mortality applied for functionally cured patients to OS. Given that functionally cured patients are assumed to be at higher mortality risk than the general population, the ERG concluded that it would appear reasonable to assume that functionally cured patients would also have lower quality of life than that of the general population. Alternative assumptions have been explored by the ERG in Section 6.

5.2.7.2 Adverse event disutilities

The company incorporated disutilities for the key grade 3 to 4 adverse events. The disutility for VOD applied in the model was 0.208 and applied for a mean duration of 26.8 days. This was based on the SMC evaluation of defibrotide in which it was assumed that quality of life with severe VOD following a HSCT was equivalent to that of acute liver failure prior to a transplant [39].

For the remaining Grade 3-4 events, the company applied a mean utility decrement of 0.024, applied for a duration of one day. This was based on the disutility estimates presented in the NICE appraisal for azacitadine (TA399), where trial-based EORTC QLQ-C30 data for patients who were hospitalised with and without grade 3 or higher TEAEs was mapped to a mean disutility value [35].

ERG comment

The ERG had some minor concerns with the adverse event disutility values, but given the relatively low incident rate of AEs, these were felt to have minimal impact and were not explored further. As with the health state utility values, the disutility values were estimated from a different patient population and may not be fully representative of the potential disutility experienced by patients in this appraisal. There was also felt to be some uncertainty regarding the VOD disutility, which was obtained from an SMC appraisal of defibrotide for the treatment of VOD after HSCT. The CS stated that VOD after GO is typically less severe which would imply that the VOD disutility used in the model may be too high. This could result in a bias against GO+DA, but the impact would be very small given the low incidence rate of VOD.

5.2.8 Resources and costs

The CS provided a description of the resource use and costs incurred over time. These included: drug acquisition costs, drug administration costs, HSCT costs, costs related to the health states, costs associated with adverse events and costs related to terminal cancer patients that were applied at the end of the patient's life.

The company conducted a systematic review to identify published evidence regarding the resource use and costs associated with the management of patients with AML. The company found two studies relating to the UK setting, but these provided average lifetime costs for AML patients. The company considered that the costs informing the azacitidine appraisal to NICE were not appropriate for use in this analysis was they were a different population that includes secondary AML cases who were ineligible for HSCT. As such, there were no data found that was considered by the company to be able to inform the economic model. Resource use estimates in the company's model health states were therefore based on recommendations from their clinical key opinion leaders.

5.2.8.1 Costs of first-line therapy

The dosing of first-line (induction and consolidation) therapies was based on data from the ALFA-0701 trial, and reflects the expected marketing authorisation for GO. The drug acquisition cost for GO was based on the confidential list price. The cost per 5mg vial of GO was [redacted]. Unit costs of comparator therapies were sourced from appropriate, national resources (BNF and eMIT) [41, 42]. Cytarabine is available in a number of different vial sizes, and the company used the unit cost of cytarabine that resulted in the lowest mean cost per mg.

Table 32 Drug acquisition costs of first-line therapies (CS, Table 42)

Drug	Pack price	Source
Gemtuzumab ozogamicin (5-mg vial)	[redacted]	Pfizer
Daunorubicin (20-mg vial)	£65.00	BNF (2017)
Cytarabine (2000 mg/5 mL solution, 5 vials)	£6.60	DoH (eMiT) (2017)

The cost of each dose is estimated in the model based on mean patient body surface area (BSA) and the mean dose. Mean BSA in the ALFA-0701 trial was [redacted]. The company assumed that there would be no drug wastage in their drug cost calculations, stating that this was supported by advice from clinical experts. A separate scenario including drug wastage found the impact to be minor. An alternative method was also explored which calculated mean vial usage accounted for the distribution of BSA and allowing for vial wastage.

The cost per cycle is the estimated by combining the mean cost per dose and the expected number of doses per treatment cycle. Each induction cycle and each consolidation cycle was assumed to last for one month.

The proportion of patients who received each course of treatment in the ALFA-0701 study was applied in the model to account for treatment discontinuation. This was based on patient data for the overall population in ALFA-0701 (including those with unfavourable cytogenetics who were excluded from the efficacy analysis). In the company base-case analysis, it was assumed that the proportion of patients receiving each course of therapy would be broadly similar in each treatment group, and so a pooled estimated was used.

Allocation to consolidation therapy in the trial was based on investigator assessment of response (CR/CRp), and the proportion of patients receiving consolidation therapy in the model was subsequently adjusted to be consistent with the IRC-assessed efficacy inputs. The adjustment involved down-weighting the proportion who received consolidation therapy in the trial. The weight was estimated as the relative difference in proportion of patients who attained CR/CRp in IRC versus

investigator assessments (IRC-assessed induction success / investigator-assessed induction success, Table 27 of the CSR).

Table 33 Patients receiving each course of treatment

Proportion of randomised patients	GO + DA group	DA group	Pooled (base-case analysis)
Induction course 1	[redacted]	[redacted]	[redacted]
Induction course 2	[redacted]	[redacted]	[redacted]
Consolidation course 1 (modelled)*	[redacted]	[redacted]	[redacted]
Consolidation course 2 (modelled)*	[redacted]	[redacted]	[redacted]
Consolidation course 1 (ALFA-0701)	[redacted]	[redacted]	[redacted]
Consolidation course 2 (ALFA-0701)	[redacted]	[redacted]	[redacted]
GO, gemtuzumab ozogamicin; DA, daunorubicin and cytarabine *Downweighted using response status cross classification			

Dose reductions, due to adverse events, were not accounted for in the model. The CSR provides details of the proportion of patients receiving the expected prescribed dose in the trial (Table 33, CSR [50]): this was generally high ([redacted]) and so the exclusion of this factor was not felt to impact substantially on the model.

These treatments are administered in an inpatient setting, and administration costs were assumed to be captured within the elective inpatient cost applied as part of the health state cost for induction and consolidation therapy.

ERG comment

The ERG had two main concerns around the estimation of costs relating to first-line therapy, relating to the adjustment of the courses of treatment and pooling of data from both treatment groups.

The ERG does not consider it appropriate to adjust the courses of induction and consolidation treatment using the approach undertaken by the company. The ERG acknowledges that this creates a more aligned approach with the clinical efficacy data which were based on IRC assessment, however the unadjusted rates are a more accurate reflection of the drug costs that would be incurred in clinical practice, as these would be determined by the treating clinician assessing the patient in a similar manner to the local investigator in the trial.

Similarly to the response rate (Section 5.2.6.1), the ERG also felt that the decision to pool data from both treatment groups was unnecessary. There do appear some differences between arms, even if they are small, and the ERG felt it important that the model capture any potential differences.

These concerns are explored further in Section 6.

The assumption to not include drug wastage and that the administration costs would be captured within the health state costs were both considered reasonable. While the inclusion of drug wastage is a more accurate assumption, the company's analysis provides a more conservative estimate of cost-effectiveness, as the costs of DA are increased proportionally more than those of GO when wastage is taken into account.

5.2.8.2 Costs of subsequent lines of therapy

Subsequent lines of therapy in the model consisted of salvage therapy (assumed to be FLAG-Ida) and non-curative therapy (comprising hydroxycarbamide, low-dose cytarabine, and azacitidine) and best-supportive care (hydroxycarbamide).

Table 34 Drug acquisition costs of subsequent lines of therapy (CS, Table 43)

Drug	Pack price	Source
Salvage therapy		
Fludarabine (50 mg/2 mL concentrate, 1 vial)	£26.08	DoH (eMiT) (2017)
Cytarabine (2000 mg/5 mL solution, 5 vials)	£6.60	DoH (eMiT) (2017)
G-CSF (filgrastim) (30 million units/0.5 mL solution, 5 vials)	£49.30	BNF (2017)
Idarubicin (5-mg powder for solution, 1 vial)	£87.36	BNF (2017)
Non-curative therapies		
Low-dose cytarabine (100-mg vial)	£4.70	DoH (eMiT) (2017)
Hydroxycarbamide (100 capsules)	£8.83	DoH (eMiT) (2017)
Azacitidine (100-mg powder for suspension)	£321.00	BNF (2017)

The mean cost per dose and cost per cycle was estimated in the same manner as that of first line therapies, as described in the section above.

The duration of non-curative treatment was estimated from restricted mean survival time estimates (RMST) from the ALFA-0701, and adjusted for time assumed to be spent in the terminal care period (two cycles). RMST for relapsed and refractory patients who did not receive HSCT were estimated from data pooled from both arms, and was [redacted] for refractory patients and [redacted] for relapsed patients.

ERG comment

The ERG agreed that the choice of subsequent therapy was appropriate for UK clinical practice. There were, however, some minor issues with the use of RMST to estimate the duration of non-curative treatment. These costs were applied as a one-off cost to patients as they entered the refractory and relapse health states, and were not explicitly based on the time in the health state (i.e. not linked to any survival estimates). They calculations in the model also did not appear to adjust for the proportion that went on to receive HSCT, which would overestimate costs for DA alone due to the higher rate of HSCT in this arm.

5.2.8.3 Health state costs

Health state costs were incorporated into the company's model to account for the monitoring and management of patients with AML that did not specifically relate to systemic therapy. The same level of resource use was applied for these states to GO+DA and DA alone, with the exception of blood products (i.e. GO was not associated with any additional resources beyond those relating to the treatment of adverse events). Functionally cured patients were not assumed to incur any additional AML-related resources.

Health state costs included:

- Inpatient and outpatient attendances;
- Consultant haematologist;
- Specialist nurse;
- Disease monitoring tests (including bone marrow aspirate, full blood count and biochemistry profiles);
- Supportive therapies, consisting of prophylactic use of antifungals and antibiotics;
- Blood products, including red blood cells and platelets.

The number of units of each resource used was informed by clinical expert opinion. Table 44 in the CS presents the unit cost of each resource and resource utilisation per health state.

Resource utilisation of blood products for the individual treatment arms was estimated from ALFA-0701, and was applied within the induction, consolidation, salvage therapy and non-curative therapy health states (Table 45 of the CS). GO+DA was associated with [redacted] use of both red blood cells and platelets in all cycles (except second consolidation for RBC). For those in the non-curative health state, no data was available from ALFA-0701 and so the company used resource use data from the technology appraisal of azacitidine to inform this parameter [35]. Unit costs were extracted from national sources [43].

ERG comment

The ERG considered the majority of health state costs to be broadly reasonable, and included all important resource use items. The assumption that GO would not require any additional resource use above that required for DA appeared reasonable for most of the health states, although there are concerns that this might not be the case for hospitalisations during induction or consolidation therapy. Data on hospitalisations were collected in the ALFA-0701 trial (Table 43, CSR [50]), which indicated that GO+DA patients may experience longer time spent in hospital: GO+DA patients had a median of [redacted]) in hospital, compared with [redacted]) for DA. The ERG asked the company to clarify why these data were not included in the analysis. This was subsequently justified by the company on the basis that the data were not available in a format useful for the model, as the data in the CSR were not presented separately for each course of treatment, and also included hospitalisations due to adverse events and worsening of AML. The ERG could, therefore, not explore this issue further.

The assumption that the functionally cured health state had no associated costs was thought to be generally reasonable, as patients in remission are not expected to be monitored in perpetuity. However, the ERG notes that AML survivors are at greater risk of cardiovascular disease and secondary cancers, as reflected in the adjustment for long-term mortality for these patients, and hence some of the assumed cost-offsets may not be fully realised.

5.2.8.4 HSCT costs

The cost of HSCT was obtained from a NHS Blood and Transplant analysis (2014) [47]. The cost consisted of a one-off cost of the procedure, and monthly costs thereafter up to two years, including complications and follow-ups. Costs after the procedure were applied to those remaining alive after the HSCT procedure.

The model also takes into account transplant-related acute and chronic GVHD complications. An inflated cost of £26,889 was identified from a French publication, due to a lack of UK-specific data [48]. This was applied in addition to the background HSCT costs, and the same unit cost was applied for both acute and chronic GVHD. The cost includes 6 months of follow-up resources, and was applied in the model as a one-off cost.

Table 35 Unit costs associated with HSCT

Unit	Unit cost	Source
HSCT procedure and recovery period	£60,892 (applied as a one-off as patients enter the HSCT health state)	NHS Blood and Transplant analysis
0 to 6 months after the HSCT	£4,891 per month	NHS Blood and Transplant analysis
6 to 2 12 months after the HSCT	£3,360 per month	NHS Blood and Transplant analysis
12 to 24 months after the HSCT	£1,212 per month	NHS Blood and Transplant analysis
GVHD	£26,889 per month	Esperou et al., 2004

ERG comment

The ERG notes that HSCT costs were obtained from a costing study conducted in the Netherlands between 1994 and 1999 [52]. The HSCT process has changed substantially in the intervening period, and that inflating these costs to 2017 may not accurately reflect the current costs of HSCT. As such, NHS reference costs may provide a more appropriate cost of the procedure itself. There is a wide variation in HSCT costs provided in NHS reference costs (from £17,344 for an autologous transplant to £38,336 for an allogeneic transplant from an unspecified donor). They also are substantially lower than the unit cost used by the company. Given the uncertainty in the HSCT costs, the impact of alternative costs was explored in Section 6.

The ERG is also unclear on whether the NHS Blood and Transplant costs included treatment for patients who developed GVHD. If this was the case, the addition of the GVHD health state costs would be double counting. Overestimating HSCT costs would bias the model in favour of GO+DA as fewer of these patients had an HSCT. Given this uncertainty, the ERG explored the impact of both the inclusion and exclusion of these additional GVHD costs.

5.2.8.5 Adverse event costs

The model incorporated a weighted total AE cost, which was estimated from the unit cost of each event and weighted by the proportion of patients estimated to experience that event over the course of first-line treatment. This cost was applied as a one-off cost in the first cycle of the model. No AEs were associated with subsequent treatment costs. Unit costs were extracted from NHS Reference Costs where available [43].

For VOD, the cost was estimated on the recommended diagnosis and treatment of VOD reported within the AML 17 trial. This incorporated a cost of endoscopic ultrasound examination, and a course of defibrotide (10 mg/kg per day, for 7 days). The cost of defibrotide was estimated using mean body

weight data from ALFA-0701 ([redacted]), the expected list price of £365 per 200mg vial from NHS England Commissioning reports [45], and estimated using the same method as the other drugs (assuming no vial wastage).

Table 36 Cost of adverse events (adapted from CS, Table 46)

Adverse Event	Cost	Source (HRG Code)*
Skin toxicity	£1,586	Department of Health (2016) (JD07A-K)
Mucosal toxicity	£1,493	Department of Health (2016) (FZ91A-M)
Pain	£1,009	Department of Health (2016) (WH08A-B)
Nausea, vomiting, and diarrhoea	£1,493	Department of Health (2016) (FZ91A-M)
Pulmonary toxicity	£1,527	Department of Health (2016) (DZ20D-F)
Cardiac rhythm disorder	£997	Department of Health (2016) (EB07A-E)
Other cardiac toxicity	£1,713	Department of Health (2016) (EB14A-E)
Central neurological toxicity	£389	Department of Health (2016) (VC12Z)
Peripheral neurological toxicity	£389	Department of Health (2016) (VC12Z)
Infections	£1,938	Department of Health (2016) (WH07A-G)
Haemorrhage	£1,251	Department of Health (2016) (SA02G-J)
Venous occlusive disease	£9,452.78	NHS England Commissioning reports
	£611.79 (Doppler ultrasound)	Department of Health (2016) (GB13Z)
*Where multiple codes, weighted average estimated from HRG activity and unit cost per code		

ERG comment

The ERG is generally satisfied with the approach to implement the AE-related costs for first-line therapy. However, the ERG considered that patients experiencing VOD would also require inpatient treatment extending beyond the standard stay for treatment with GO due to the associated high mortality risk. There is some uncertainty whether the additional inpatient treatment is already captured in the length of stay assumptions. The impact of including an additional hospitalisation cost was explored in Section 6.

5.2.8.6 Mortality costs

The model incorporates a terminal care cost of £6,659 (Addicott & Dewar (2008) [46], inflating to 2015/16 values using the indices reported in PSSRU [44]), constituting the care for the last two months of life for those patients receiving non-curative therapy (including best supportive care). This reflects the increase in acute sector and community sector resources used in the final weeks of life and included additional GP visits, time with a district nurse, and acute hospital care. This was applied as a

one-off cost as patients transitioned into the refractory health state and to patients entering the relapse health state.

ERG comment

The ERG considered that the source of the mortality costs was generally appropriate. There were some minor concerns around how these costs were applied in the model. These costs appeared to be applied to all refractory and relapsed patients, regardless of whether they went onto receive HSCT and subsequent cure. As costs are applied when patients enter the health state, they are not discounted appropriately, but this is unlikely to make much difference to the ICER as life expectancy is low in these patients. Given the small impact judged to the overall results, these costs were not explored further by the ERG.

5.2.9 Cost effectiveness results

5.2.9.1 Base-case results

Cost effectiveness results

Table 37 presents the results of the company base-case analysis. The model results found GO+DA to be more costly (mean cost difference of [redacted]), but also more effective (mean gain of [redacted] QALYs). The estimated deterministic ICER for GO+DA compared with DA was £12,251 per QALY.

The company undertook a probabilistic sensitivity analysis (PSA) to explore and quantify uncertainty in the outcomes of the analysis. Probabilistic results were estimated from 1,000 iterations of the model, with values for key parameters sampled stochastically from assigned distributions to each parameter. The probabilistic ICER estimated by the company was £13,600 per QALY. The probabilistic results were relatively similar to those estimated in the deterministic base-case analysis although slightly higher.

Table 37 Results of the company base-case analysis, base-case population (excluding unfavourable cytogenetics) (CS, Table 48, p. 143)

Technologies	Total costs (£)	Total QALYs	Incremental costs (£)	Incremental QALYs	ICER incremental (£/QALY)
Deterministic results					
GO + DA	[redacted]	[redacted]	[redacted]	[redacted]	£12,251
DA	[redacted]	[redacted]			
Probabilistic results					
GO + DA	[redacted]	[redacted]	[redacted]	[redacted]	£13,600
DA	[redacted]	[redacted]			
CS, company submission; ICER, incremental cost-effectiveness ratio; QALYs, quality-adjusted life years; GO, gemtuzumab ozogamicin; DA, daunorubicin + cytarabine					

The largest cost savings in the GO+DA arm were due to fewer HSCTs and fewer patients receiving non-curative therapy after relapse, however these were not sufficient to offset the additional treatment costs of GO+DA compared to DA alone. The majority of QALYs gained were observed in the functionally cured health state and in the CR/CRp off-treatment health state (the two health states associated with the highest utility values).

5.2.9.2 Sensitivity analysis

Deterministic sensitivity analysis

The CS presented the results of a variety of one-way deterministic sensitivity analyses (DSA) to identify the key drivers of the cost-effectiveness model.

Parameters included in the DSA were those relating to: patient characteristics (age, body surface area), quality of life (HS utility values, disutility values associated with adverse events), unit costs (health state costs, HSCT cost, AE costs), proportion receiving each course of chemotherapy, resource use (usage of supportive therapies and blood products, frequency of monitoring and visits to health care professionals), probability of HSCT, time of HSCT, duration of GVHD, and incidence and duration of adverse events. The company varied each parameter value by ±10% and reported the subsequent absolute change to the ICER.

The company presented a tornado diagram depicting the results of the DSA (Figure 20, CS). This displayed little variation in the results due to parameter variation, and the change in the ICER was never greater than £1,200. The tornado diagram demonstrated that the parameters with the largest influence on the ICER are HSCT probabilities from relapse in years 1 and 2 for the DA group, and the RMST for relapsed patients.

Probabilistic sensitivity analysis

Figure 14 presents the cost-effectiveness acceptability curve (CEAC). At a willingness-to-pay threshold of £20,000, the probability of GO + DA being cost-effective was [redacted]. At a threshold of £30,000, the probability rose to [redacted].

Figure 14 Cost-effectiveness acceptability curve (CS, Fig 19, p. 145)

[redacted]

5.2.9.3 Scenario analysis

The company presented a range of scenario analysis within their base case analysis. These analyses focused on the use of alternative survival functions, health state utility weights and disutilities for adverse events, the use of data from individual arms instead of pooled data to estimate certain parameters (response status, mean RMST, HSCT probabilities OS for refractory patients, and the number of treatment courses received), and alternative values for the excess mortality of long-term AML survivors.

Across the set of scenarios exploring the alternative survival functions for RFS and OS (for CR or CRp patients), the ICER varied between £6,821 (best fitting standard parametric functions) and £12,233 (MCM Weibull) per QALY.

The ERG notes that the main drivers of the cost-effectiveness estimates are the difference in costs (e.g. initial treatment costs and subsequent HSCT) predicted over the period of follow-up of the ALFA-0701 trial and the difference in survival after the follow-up. While the alternative survival models (MCM, spline-based and conventional parametric) result in large differences in the terms of the absolute LYG and QALY estimates, they all predict very similar between group differences. This

is because the majority of functions generate similar predictions of the difference in the proportion of patients who become long-term survivors. While this is done explicitly in the MCM models using the cure fraction approach, many of the more conventional parametric functions are doing this implicitly with a shallowing of the hazard function which appears to be converging to a similar between group differences in the proportion of patients who experience long term survival. As a result, the mean difference in costs and QALYs used to estimate the ICER appear robust and relatively stable across the majority of functions.

Across all the scenarios, the ICER estimate varied between £6,821 (best fitting standard parametric functions for RFS and OS for CR/CRp) and £20,334 per QALY (when the RMST for relapse patients was based on individual treatment arms).

5.2.9.4 Subgroup analysis

In the CS appendix, the company presented cost-effectiveness results for the full indication (including those with unfavourable cytogenetic risk profile). The ICER is higher in the whole patient population (£20,457, compared with £12,251 for the favourable/intermediate and unknown cytogenetic population), since the effect of GO is lower in patients with unfavourable cytogenetics.

Table 38 Results of the analysis, all patient population (CS Appendix, Table 172)

Technologies	Total costs (£)	Total QALYs	Incremental costs (£)	Incremental QALYs	ICER incremental (£/QALY)
GO + DA	[redacted]	[redacted]	[redacted]	[redacted]	20,457
DA	[redacted]	[redacted]			

CS, company submission; ICER, incremental cost-effectiveness ratio; QALYs, quality-adjusted life years; GO, gemtuzumab ozogamicin; DA, daunorubicin + cytarabine

The mean probabilistic ICER was £21,999 (95% confidence credible interval [CrI]: [redacted] for costs; [redacted] for QALYs), and GO+DA had a [redacted] probability of being cost-effective at a willingness-to-pay threshold of £30,000. The probability was less than [redacted] when the willingness-to-pay threshold was £20,000 per QALY. The DSA showed that the results in this population were sensitive to the same input parameters as the cytogenetic population. The greatest increase in the ICER was observed when HSCT probabilities from relapse were pooled, providing an ICER of £30,206.

5.2.10 Model validation and face validity check

A comparison of modelled and trial-based relapse-free survival and overall survival in the pooled patient cytogenetic population was presented by the company for purposes of model validation. Modelled survival appears to represent the clinical data well throughout the trial period, reflecting the goodness of fit of the selected survival models to the Kaplan-Meier trial data.

The company also provided details of the external validation provided by external experts.

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. However, the ERG identified a number of key uncertainties. The main concerns identified by the ERG include:

1. *The complexity of the state-transition modelling approach*

The ERG considers that the state-transition modelling approach introduces unnecessary complexity compared to a simpler partitioned survival analysis (PartSA) approach. A series of structural assumptions (e.g. sub-states and HSCT) were imposed which ultimately resulted in the same independence assumption between clinical events (EFS and OS) that underpin the PartSA approach. While the state-transition model created significant challenges for the ERG in terms of identifying and following key assumptions, the predictions for the company base-case demonstrated internal validity and consistency with the ALFA-0701 trial results.

The ERG considers that judgements concerning the appropriateness of the proposed model structure (and external validity) can potentially be simplified to the validity of the 'functional cure' assumption.

2. *Important clinical heterogeneity remains within the population proposed by the company*

The ERG agrees with the company's decision to exclude patients with known unfavourable cytogenetics and their rationale for focusing on the specific subpopulation where GO+DA provides clear clinical benefit and optimises cost-effectiveness. However, the ERG does not believe that the company has sufficiently addressed the heterogeneity in the subgroup of patients with unknown cytogenetics and within the intermediate population.

3. *The use of IRC rather than IA response outcomes and pooling*

The ERG did not consider that it was necessary or appropriate to pool response data. Even if the differences in response rates are not considered clinically important, any differences should still be

considered and the lack of statistical significance should be reflected in the distributions assigned to these parameters

The ERG also considers that that the initial treatment costs of the induction and consolidation therapies should be based directly on the IA response outcomes, rather than attempting to adjust the IRC response outcomes in the manner proposed by the company.

4. *The differences in the cure fraction reported between: (i) the alternative MCMs and (ii) ES and OS.*

The ERG highlighted differences in the predicted cure fractions reported across the alternative MCM distributions and between EFS and OS that had not been explained and/or justified by the company. Despite these potential concerns, the ERG concluded that the differences in the cure fraction between the treatments appeared sufficiently robust for the base-case population. However, the ERG considers that there remains significant heterogeneity within the base-case population which may have important implications concerning the difference in the cure fraction for further subgroups within the overall population.

5. *The lack of explicit structural link between relapse and HSCT*

The final model did not include an explicit structural link between relapse and HSCT. The ERG highlights that the important cost-offsets assumed for HSCT are predicated on the functional cure assumption (i.e. that those patients who have not relapsed by 5-years will never relapse and hence will not require HSCT at some point in the future).

6. *The functionally cured assumption*

The ERG considers that there remains significant uncertainty surrounding the long-term morbidity and survival of functionally cured patients. Although the model included an adjustment for excess mortality, the ERG identified additional uncertainties in the estimate of the HR due to the small number of patients on which this was based. The use of general population quality of life was also not internally consistent with the excess mortality applied for functionally cured patients for OS. Given that functionally cured patients are assumed to be at higher mortality risk than the general population, the ERG concluded that it would appear reasonable to assume that functionally cured patients would have lower quality of life than that of the general population.

7. The potential confounding effect of compassionate use of GO

[redacted] receiving DA alone, subsequently received GO as a follow-up therapy through a compassionate use programme. Although the company undertook a feasibility assessment to determine whether formal adjustment methods could be applied, they concluded that the assumptions required did not hold and that any adjustments would be unreliable. While the company noted that that the compassionate use programme could bias efficacy in favour of the comparator arm, there remains significant uncertainty surrounding the possible magnitude of this.

8. Uncertainties surrounding the costing assumptions for VOD and HSCT

The ERG identified a number of uncertainties surrounding the costing assumptions for VOD (specifically the inclusion of VOD for DA alone) and for HSCT.

Given the importance of these issues, additional analyses requested by the ERG from the company and independently undertaken by the ERG are presented in Section 6 considering the potential impact of uncertainty in VOD & HSCT costing on the cost-effectiveness results.

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

6.1 Overview

This section details the ERG's further exploration of the assumptions and uncertainties raised in the review and critique of the company's cost-effectiveness analysis, presented in Section 5. This section is organised in four parts. Section 6.2 details the impact of correction of errors identified in ERG's validation of the executable model and other amendments to the company base-case analysis. Section 6.3 details a series of scenario analyses exploring the robustness of the cost-effectiveness results to specific assumptions and additional uncertainties identified by the ERG. These analyses were conducted within the company corrected base-case analysis. The scenario analyses presented in Section 6.3 focus on exploring the following issues and uncertainties:

- The number of induction and consolidation courses received;
- Rate of response to treatment;
- Treatment costs associated with HSCT and VOD;
- Inclusion of VOD events;
- Quality of life in functionally cured patients;
- Excess mortality risk in functionally cured patients;
- Impact of treatment in different patient populations.

In Section 6.4, the ERG alternative base-case is presented based on a combination of the exploratory analyses presented in Section 6.3. Further exploratory analysis is also presented in Section 6.5, exploring the impact of GO in different patient populations in the context of the ERG alternative base-case. Section 6.6 presents a brief conclusion summarising the ERG's additional analyses.

Due to time constraints, ICERs based on the deterministic analysis are presented throughout this section with the exception of the ERG alternative base-case.

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

A small number of errors were identified by the ERG in the company model. These errors were minor and their correction did not have a large impact on the model results. The ERG also identified a source of data used by the company which was not the most up to date available.

The amendments made by the ERG are as follows:

- Inconsistencies in the data source for mortality: calculations for the survival models made reference to two different sources of general population mortality data in the model (one for the UK and one for England & Wales). The ERG incorporated the more recently published mortality data for England & Wales for the survival analysis, and the mortality data for the UK for the mortality HR calculations [53].
- Discrepancy for HSCT probabilities after relapse: these calculations were based on an arm-specific denominator (the number of patients achieving CR/CRp), while the model uses pooled data for CR response in the base case. The ERG amendment involved changing the calculations to reflect the actual number of patients achieved CR/CRp in the model.
- Patients who did not receive the second cycle of induction therapy in the second cycle of the model were considered equivalent to those off-treatment for HRQoL purposes and did not have any associated costs that cycle. The ERG applied the cost associated with the off-treatment health state to these patients in that cycle.
- Estimation of the proportion of refractory patients receiving salvage therapy: [redacted] of patients received the first cycle of salvage therapy, and [redacted] of these salvage patients then went on to receive the subsequent cycles (i.e. these patients were double adjusted). This was corrected by the ERG so that all refractory patients receiving the first cycle of salvage therapy also received the subsequent cycles of salvage therapy

Table 39 presents the results of the ERG corrections to the company model: the ICER increase by approximately 10.7% from £12,251 to £13,561 per QALY.

Table 39 Results of the ERG-corrected company base-case model

Technologies	Total costs	Total QALYs	Incremental costs	Incremental QALYs	ICER incremental (£/QALY)
Company base-case					
GO + DA	[redacted]	[redacted]	[redacted]	[redacted]	£12,251
DA	[redacted]	[redacted]	-	-	-
Company base-case (including ERG corrections)					
GO + DA	[redacted]	[redacted]	[redacted]	[redacted]	£13,561
DA	[redacted]	[redacted]	-	-	-

ERG, Evidence Review Group; ICER, incremental cost-effectiveness ratio; QALYs, quality-adjusted life years; GO, gemtuzumab ozogamicin; DA, daunorubicin + cytarabine
Based on deterministic analysis

6.3 Additional ERG analyses

6.3.1 Re-estimation of induction/consolidation courses and response rate

In the company base-case analysis, the number of courses of treatment and the rate of response were estimated from trial data pooled from both treatment groups, assuming no difference in treatment received and treatment response between GO+DA and GO patients (Section 5.2.8.1). In this section, the ERG explores the impact of alternative assumptions, where arm-specific rates were modelled.

Treatment courses – impact on resource use

The ERG explored the impact of modelling individual rates of courses of treatment, which did not include the company adjustment to align with the IRC-assessed clinical data. These values, presented in Table 40, were felt to be more appropriate as they reflected actual treatment decisions in the trial and would be more reflective of clinical practice. As subsequent treatment decisions in the ALFA-0701 trial were based on the individual response rates, this analysis explores the implications on the resource use relating to treatment costs (changes to outcomes were modelled in the subsequent analysis).

Table 40 Courses of induction and consolidation therapy

	Company modelled values		**ERG modelled values**	
Proportion of patients	**GO + DA group**	**DA group**	**GO + DA group**	**DA group**
Induction course 1	[redacted]	[redacted]	[redacted]	[redacted]
Induction course 2	[redacted]	[redacted]	[redacted]	[redacted]
Consolidation course 1	[redacted]	[redacted]	[redacted]	[redacted]
Consolidation course 2	[redacted]	[redacted]	[redacted]	[redacted]

ERG, Evidence Review Group; GO, gemtuzumab ozogamicin; DA, daunorubicin and cytarabine

As presented in Table 41, the results of this analysis are associated with higher incremental costs and a higher ICER of £14,249 per QALY. Although GO+DA was associated with a lower number of courses of treatment than DA, the potential cost savings were not offset by the additional cost of GO, resulting in the higher overall costs in this scenario.

Table 41 Results of ERG exploratory analysis on courses of treatment and response rate

	Total costs	Total QALYs	Incremental costs	Incremental QALYs	ICER
Company base-case (including ERG corrections)					
GO + DA	[redacted]	[redacted]	[redacted]	[redacted]	£13,561
DA	[redacted]	[redacted]	-	-	-
Scenario: Courses of treatment based on unpooled investigator-assessed data					
GO + DA	[redacted]	[redacted]	[redacted]	[redacted]	£14,249
DA	[redacted]	[redacted]	-	-	-

ERG, Evidence Review Group; QALYs, quality-adjusted life year; ICER, incremental cost-effectiveness ratio; GO, gemtuzumab ozogamicin; DA, daunorubicin and cytarabine
Based on deterministic analysis

Treatment response rate – impact on outcomes

The ERG explored the impact of using individual rates of response (based on unpooled ALFA-0701 trial data), instead of using pooled rates as per the company base-case analysis. Rates of response are presented in Table 23 in Section 5.2.6.1: GO+DA was associated with a response rate of [redacted] and DA with a response rate of [redacted]. While the rates were similar between arms, the ERG considered that any observed differences should still be captured. Decision to allocate to subsequent therapies in the ALFA-0701 trial was based on the individual response rates, each associated with different long-term prognosis. As such, this analysis explores the implications on patient outcomes captured through the use of individual response rates.

The impact of using alternative assumptions for treatment response was that the ICER was reduced to £10,526 (Table 43) as a result of the higher response rate for GO+DA patients. With fewer patients subsequently entering the refractory health states with a lower quality of life, the number of QALYs gained increased with GO+DA.

Table 42 ERG exploratory analysis on rate of response

	Total costs	Total QALYs	Incremental costs	Incremental QALYs	ICER
Company base-case (including ERG corrections)					
GO + DA	[redacted]	[redacted]	[redacted]	[redacted]	£13,561
DA	[redacted]	[redacted]	-	-	-
Scenario: Rate of response to treatment for individual arms					
GO + DA	[redacted]	[redacted]	[redacted]	[redacted]	£10,526
DA	[redacted]	[redacted]	-	-	-

ERG, Evidence Review Group; QALYs, quality-adjusted life year; ICER, incremental cost-effectiveness ratio; GO, gemtuzumab ozogamicin; DA, daunorubicin and cytarabine
Based on deterministic analysis

6.3.2 Treatment of HSCT and VOD

The ERG explored a number of issues relating to the modelling of HSCT resource use and of VOD events relating to treatment.

The ERG was concerned that the costs used for HSCT were based on an old, non-UK dataset [52]. For the initial costs of HSCT, the ERG applied a weighted cost obtained from NHS reference costs for an allogeneic stem cell transplant, using the costs associated with the procedure in adults to be consistent with the other parameters in the analysis [43]. These costs were significantly lower than those from Blood & Transplant (£60,892 vs £31,628)[47]. The inflated costs from NHS Blood & Therapeutics used by the company were used to capture the ongoing resource use.

A scenario explored the impact of removing the GVHD-specific health state costs. It was unclear whether the GVHD cost would be captured within the HSCT health state costs. While not explicitly stated by Van Agthoven et al, the ERG considered it likely that the costs of managing GVHD were already captured given the magnitude of the follow-up costs in the study [52]. If this was the case, the cost of GVHD would be double counted in the model.

The ERG also explored the impact of excluding VOD cases in the DA arm, since these were considered to be related to the patients who crossed over and received a dose of GO in the ALFA-0701 trial (discussed in Section 5.2.6.7).

There was also some uncertainty whether the cost of VOD treatment was fully captured in the model. The ERG explored the impact of including additional hospitalisation costs, applied for the duration as the disutility associated with this event (26.8 days). A longer length of inpatient stay for GO during this period would suggest that these costs were captured implicitly; however, this would not be the

case if pooled courses of treatment were applied in the model. The ERG felt that even if unpooled courses of induction and consolidation treatment were modelled then the additional impact of VOD would not be captured, since DA is associated with greater levels of hospitalisations during this period overall.

Results of these analyses are presented in Table 43. The impact of these analyses is relatively minor, although the impact of reducing the cost of HSCT resulted in the ICER increasing from £13,561 to £16,003 per QALY. This is due to the higher rate of HSCT in the DA group; thus reducing the HSCT unit cost will reduce costs in the DA arm to a greater degree, leading to lower incremental costs and a higher ICER.

Table 43 Results of the ERG exploratory analysis with alternative assumptions for HSCT and VOD

	Total costs	Total QALYs	Incremental costs	Incremental QALYs	ICER
Company base-case (including ERG corrections)					
GO + DA	[redacted]	[redacted]	[redacted]	[redacted]	£13,561
DA	[redacted]	[redacted]	-	-	-
Scenario: Alternative HSCT costs					
GO + DA	[redacted]	[redacted]	[redacted]	[redacted]	£16,003
DA	[redacted]	[redacted]	-	-	-
Scenario: Exclusion of additional GVHD costs					
GO + DA	[redacted]	[redacted]	[redacted]	[redacted]	£14,020
DA	[redacted]	[redacted]	-	-	-
Scenario: Exclusion of VOD events in the DA alone group					
GO + DA	[redacted]	[redacted]	[redacted]	[redacted]	£13,704
DA	[redacted]	[redacted]	-	-	-
Scenario: Inclusion of hospital costs for the treatment of VOD					
GO + DA	[redacted]	[redacted]	[redacted]	[redacted]	£13,733
DA	[redacted]	[redacted]	-	-	-

ERG, Evidence Review Group; QALYs, quality-adjusted life year; ICER, incremental cost-effectiveness ratio; GO, gemtuzumab ozogamicin; DA, daunorubicin and cytarabine; HSCT, hematopoietic stem cell transplant; GVHD, graft versus host disease; VOD, venous occlusive disease
Based on deterministic analysis

6.3.3 Quality of life in functionally cured patients

The ERG considered that the assumption around quality of life in patients who are functionally cured was not sufficiently justified (as discussed in Section 5.2.7.1). These patients experienced the same HRQoL as the general population, which resulted in a marked jump in the HRQoL estimates at five

years. It was also not considered internally consistent with the excess mortality applied for functionally cured patients to OS.

The ERG, therefore, explored a scenario where functionally cured patients would have lower quality of life than that of the general population. The utility associated with the CR/CRp off-treatment patients health state of 0.74 was applied to these patients [34]. A second scenario also applied the utility value for CR/CRp: off-treatment to functionally cured patients, which was further adjusted for aging using the values from Ara & Brazier used by the company in their analysis [38].

Results of these analyses are presented in Table 44. Both scenarios were associated with lower QALYs due to the lower utility values. The scenarios also had a higher ICER: with a greater amount of patients in the GO+DA group achieving functional cure, a lower utility value reduces the overall QALY gain for these patients more than in the DA group.

Table 44 Results of the ERG exploratory analysis with alternative values for functionally cured patients

	Total costs	Total QALYs	Incremental costs	Incremental QALYs	ICER
Company base-case (with ERG corrections)					
GO + DA	[redacted]	[redacted]	[redacted]	[redacted]	£13,561
DA	[redacted]	[redacted]	-	-	-
Scenario: Alternative utility values for functionally cured patients					
GO + DA	[redacted]	[redacted]	[redacted]	[redacted]	£13,878
DA	[redacted]	[redacted]	-	-	-
Scenario: Alternative utility values for functionally cured patients, adjusted for aging					
GO + DA	[redacted]	[redacted]	[redacted]	[redacted]	£15,279
DA	[redacted]	[redacted]	-	-	-

ERG, Evidence Review Group; QALYs, quality-adjusted life year; ICER, incremental cost-effectiveness ratio; GO, gemtuzumab ozogamicin; DA, daunorubicin and cytarabine
Based on deterministic analysis

6.3.4 Excess mortality in functionally cured patients

As noted in Section 5.2.6.4, the ERG had some concerns with the estimation of the hazard ratio to model excess mortality in functionally cured patients, in some years the probability of death of AML survivors was higher in the general population. The hazard ratio was re-calculated by the ERG by adjusting the calculations so that the higher of the AML mortality rate and the general population mortality rate was used, and the hazard ratio subsequently increased from [redacted] to [redacted] (based on the mortality data used in the calculations originally).

The results of the exploratory analysis are presented in Table 45. The amendment of this parameter results in a modest increase to the ICER, as a result of increased mortality and fewer QALYs accrued in the functionally cured patients.

Table 45: Results of the ERG exploratory analysis with alternative hazard ratio for survival

	Total costs	Total QALYs	Incremental costs	Incremental QALYs	ICER
Company base-case (with ERG corrections)					
GO + DA	[redacted]	[redacted]	[redacted]	[redacted]	£13,561
DA	[redacted]	[redacted]	-	-	-
Scenario: ERG-estimated HR for long-term survival					
GO + DA	[redacted]	[redacted]	[redacted]	[redacted]	£14,337
DA	[redacted]	[redacted]	-	-	-

ERG, Evidence Review Group; HR, hazard ratio; QALYs, quality-adjusted life year; ICER, incremental cost-effectiveness ratio; GO, gemtuzumab ozogamicin; DA, daunorubicin and cytarabine
Based on deterministic analysis

6.4 ERG alternative base-case

Table 46 presents the results of the ERG alternative base-case analysis. These incorporate a number of changes to key model parameters and assumptions, which were previously explored individually in Section 6.3.

The ERG alternative base-case analysis includes the following changes to the company base-case analysis:

- The number of induction and consolidation therapy courses reflecting what was provided in the ALFA-0701 trial (i.e. based on investigator assessment and unpooled data);
- Arm-specific rate of response to treatment (i.e. based on unpooled trial data);
- The initial cost of HSCT estimated from NHS Reference Costs;
- Removal of VOD events in the DA treatment group;
- Exclusion of GVHD-specific costs;
- Inclusion of hospital costs for the treatment of VOD;
- Quality of life in functionally cured patients based on the utility value for off-treatment CR patients, and further adjusted for age;
- Long-term mortality in functionally cured patients adjusted for excess mortality using the ERG-calculated hazard ratio.

These results reflect the population included in the company base-case, including those with favourable, intermediate and unknown cytogenetics. The cost-effectiveness of GO+DA in different subgroups, under the ERG new base case assumptions, is explored in Section 6.5.

Under the ERG's alternative set of assumptions, the deterministic ICER for GO+DA versus DA alone is £16,910 per QALY, and the probabilistic ICER is £17,956. The ERG notes that the probabilistic ICER is the most relevant to inform decisions based on cost-effectiveness, and is referred to as the key ICER for the ERG alternative base-case analysis elsewhere in this report.

Table 46 Results of the ERG alternative base-case analysis

	Total costs	Total QALYs	Incremental costs	Incremental QALYs	ICER
Deterministic results					
GO + DA	[redacted]	[redacted]	[redacted]	[redacted]	£16,910
DA	[redacted]	[redacted]	-	-	-
Probabilistic results					
GO + DA	[redacted]	[redacted]	[redacted]	[redacted]	£17,956
DA	[redacted]	[redacted]	-	-	-

ERG, Evidence Review Group; QALYs, quality-adjusted life year; ICER, incremental cost-effectiveness ratio; GO, gemtuzumab ozogamicin; DA, daunorubicin and cytarabine

6.5 Subgroup analysis

The heterogeneity reported between the cytogenetic subgroups included within the base-case population was not sufficiently explored in the CS. The main issues and uncertainties raised by the ERG in Section 5 were:

(i) The ERG did not consider that the company had fully justified the inclusion of patient with unknown cytogenetic results and explained the differences in the findings between the ALFA-0701 trial and the IPD meta-analysis for this specific subgroup;

(ii) The intermediate population is the largest subgroup in the ALFA-0701 trial. The potential impact of heterogeneity between the results of this subgroup and other subgroups included within the base-case population was not explored in the CS;

(iii) The ERG also noted the heterogeneity within the intermediate group with regards underlying genetic biomarkers, indicating potential variability in outcomes between individual patients which might be explained by additional molecular testing and further risk-stratification.

For the first issue, the ERG considers the decision to include the patients with unknown cytogenetics appears appropriate based on the ALFA-0701 results. When the company excluded the subgroup of patients with unknown cytogenetic results, the base-case ICER [redacted] per QALY (response to question B3 in response to PFCs), compared to £12,251 when they were included. This suggests that GO was [redacted]. The reason for the apparent contradictory findings for this subgroup in the ALFA-0701 trial and the IPD meta-analysis remain unclear to the ERG. The ERG acknowledges that these differences may be explained by the different trial populations and dosing regimens. It is also possible that that this might be due to differences across the studies in the underlying reasons for the unknown cytogenetic result (i.e. test not undertaken vs. unsuccessful test). Unfortunately, due to the absence of further information on the reasons for the unknown result, the ERG is unable to further explore this issue.

A series of exploratory subgroup analysis were conducted by the ERG to explore the impact of heterogeneity between the subgroups included in the base-case population and possible variability within the intermediate group. These analyses explored the impact in different subgroups based on cytogenetic results alone as well as subgroups based on cytogenetic and molecular results.

6.5.1 Data inputs

The ERG requested additional subgroup analysis results (response, RFS and OS and statistical cure rates) for the subgroups summarised in Table 47. The clinical data provided by the company was previously summarised and discussed in Section 4.

Table 47 Summary of addition subgroups requested by the ERG

Cytogenetic subgroups	Cytogenetic and molecular subgroups
Favourable/intermediate (excluding unknown)	Favourable and intermediate-1 (excluding unknown)
Favourable only	Intermediate-2 and unfavourable (excluding unknown)
Intermediate only	Intermediate-1 only
Unknown only	Intermediate-2 only

[redacted], the company reported that they were unable to fit MCM models to the following subgroups:

(i) [redacted] Cytogenetic and molecular subgroups: intermediate-2 only subgroup (n=24).

Table 48 to

Table 50 summarise the response and cure fraction results (RFS and OS) for these subgroups. The summary of RFS and OS focuses on the cure fraction estimates and the difference predicted in the cure rates between the subgroups, since this has previously been identified as the main driver of survival and QALY benefits. To aid interpretation, the results are also summarised for the populations considered in the CS base-case (favourable, intermediate and unknown) and all-patients (entire ALFA-0701 population).

Across the majority of populations and subgroups, the log-normal MCM model was considered the best-fit for RFS and OS, and only these results are presented in the summary tables below. The only exception was the intermediate-2 and unfavourable subgroup, where the best fitting model differed for GO+DA (lognormal for RFS and Weibull for OS) and DA alone (Weibull for RFS and lognormal for OS).

Table 48 Response rate for subgroups

Populations considered by company	**GO+DA**	**DA**
Base-case (favourable, intermediate and unknown)	[redacted]	[redacted]
All patients	[redacted]	[redacted]
Additional cytogenetic subgroups (ERG exploratory)	**GO+DA**	**DA**
Favourable and intermediate	[redacted]	[redacted]
Intermediate only	[redacted]	[redacted]
Additional cytogenetic and molecular subgroup (ERG exploratory)	**GO+DA**	**DA**
Favourable and intermediate-1	[redacted]	[redacted]
Intermediate-2 and unfavourable	[redacted]	[redacted]
Intermediate-1 only	[redacted]	[redacted]

*incorrectly reported by company in the response to points for clarification
Source: Company response to PFCs (estimated from data provided to B5 and B6)
ERG; evidence review group; GO, gemtuzumab ozogamicin; DA, daunorubicin and cytarabine; NR, not reported

Table 49 Survival data for subgroups – cure fractions for RFS (CR/CRp)

Subgroup	GO+DA	n[CR]	DA	n[CR]	Difference in cure rate
Populations considered by company					
Base-case (favourable, intermediate and unknown)	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
All patients	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
Additional cytogenetic subgroups (ERG exploratory)					
Favourable and intermediate	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
Intermediate only	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
Additional cytogenetic and molecular subgroup (ERG exploratory)					
Favourable and intermediate-1	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
Intermediate-2 and unfavourable (*lognormal MCM*)	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
Intermediate-2 and unfavourable (*Weibull MCM*)	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
Intermediate-1 only	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]

RFS, relapse-free survival; ERG; evidence review group; GO, gemtuzumab ozogamicin; DA, daunorubicin and cytarabine; CR, complete remission; MCM, mixture cure model

Table 50 Survival data for subgroups – cure fraction for OS (CR/CRp)

Subgroup	GO+DA	n[CR]	DA	n[CR]	Difference in cure rate
Populations considered by company					
Base-case (favourable, intermediate and unknown)	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
All patients	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
Additional cytogenetic subgroups (ERG exploratory)					
Favourable and intermediate	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
Intermediate only	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
Additional cytogenetic and molecular subgroup (ERG exploratory)					
Favourable and intermediate-1	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
Intermediate-2 and unfavourable (l*ognormal MCM*)	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
Intermediate-2 and unfavourable (*Weibull MCM*)	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
Intermediate-1 only	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]

OS, overall survival; ERG; evidence review group; GO, gemtuzumab ozogamicin; DA, daunorubicin and cytarabine; CR, complete remission; MCM, mixture cure model

The results appear broadly consistent across the majority of subgroups defined based on cytogenetic results only and based on cytogenetic and molecular results. Differences are most evident in the predicted differences in the cure fractions for EFS ([redacted] and for OS ([redacted]), indicating the heterogeneity between particular subgroups. The differences in the cure fractions are most apparent in the intermediate-2 and unfavourable subgroup. The results for EFS and OS for this subgroup appear sensitive to the specific MCM model applied and lead to inconsistent findings for EFS and OS. As a result, the MCM models do not appear sufficiently robust to be used for this subgroup.

The impact on the ICER of the heterogeneity in outcomes across the subgroups was explored by the ERG using a series of exploratory subgroup analyses. These analyses used subgroup specific differences for the following outcomes (response, EFS[CR/CRp] and OS[CR/CRp]). Table 51 summarises the differences in patient characteristics (age, gender, BSA and weight) assumed across the subgroups.

Table 51 Patient characteristics in subgroups

	All patients	Intermediate, favourable and unknown	Intermediate and favourable*	Intermediate**
Age (years)	[redacted]	[redacted]	[redacted]	[redacted]
Female (%)	[redacted]	[redacted]	[redacted]	[redacted]
Body surface area (m^2)	[redacted]	[redacted]	[redacted]	[redacted]
Weight (kg)	[redacted]	[redacted]	[redacted]	[redacted]
* Applied to intermediate-1 and favourable patients ** Applied to intermediate-1 patients				

Due to the limited data available to the ERG, it was not possible to alter several parameters for the exploratory subgroup analyses (e.g. courses of treatment, adverse events, HSCT probabilities, OS in refractory patients). Hence, the exploratory analyses presented by the ERG should be seen as indicative only. In addition, due to the incorrect reporting of response rate for the intermediate-2 and unfavourable subgroup and the lack of robustness in the cure fraction results, the ERG was not able to consider this subgroup further.

6.5.2 Results

Table 52 summarises the ICER results across the various populations and subgroups based on the ERG's alternative assumptions. The ICER of GO+DA versus DA alone varied between £16,343 (intermediate-1 only, defined by cytogenetic and molecular test) and £31,709 (intermediate only, defined by cytogenetic results only).

Table 52 Subgroup analysis (based on ERG alternative assumptions)

	Total costs	Total QALYs	Incremental costs	Incremental QALYs	ICER
Populations considered by company					
Base-case (Favourable, intermediate and unknown)					
GO + DA	[redacted]	[redacted]	[redacted]	[redacted]	£16,910
DA	[redacted]	[redacted]	-	-	-
All patients					
GO + DA	[redacted]	[redacted]	[redacted]	[redacted]	£25,941
DA	[redacted]	[redacted]	-	-	-
Additional cytogenetic subgroups (ERG exploratory)					
Favourable and intermediate					
GO + DA	[redacted]	[redacted]	[redacted]	[redacted]	£24,581
DA	[redacted]	[redacted]	-	-	-
Intermediate					
GO + DA	[redacted]	[redacted]	[redacted]	[redacted]	£31,709
DA	[redacted]	[redacted]	-	-	-
Additional cytogenetic and molecular subgroup (ERG exploratory)					
Favourable and intermediate-1					
GO + DA	[redacted]	[redacted]	[redacted]	[redacted]	£17,614
DA	[redacted]	[redacted]	-	-	-
Intermediate-1 only					
GO + DA	[redacted]	[redacted]	[redacted]	[redacted]	£16,343
DA	[redacted]	[redacted]	-	-	-

ERG, Evidence Review Group; QALYs, quality-adjusted life year; ICER, incremental cost-effectiveness ratio; GO, gemtuzumab ozogamicin; DA, daunorubicin and cytarabine
Based on deterministic analysis

Although the exploratory analyses can only be considered indicative, the findings suggest that there may be value in further risk stratification to that proposed by the company. The ERG notes that the ICER in the intermediate cytogenetic subgroup is £31,709 per QALY, indicating that the lower ICER reported for the company's base case population is being driven by the higher effect evident in the favourable and unknown patients.

Further risk stratification using genetic and molecular testing appears to produce a clearer separation between subgroups. The clinical and economic value of GO+DA appears largely confined to the favourable and intermediate-1 population defined by cytogenetic and molecular tests. While small

numbers precluded the company from undertaking cure fraction models for the intermediate-2 population, the results presented in Section 4 for EFS and OS suggest limited clinical benefit is evident in this subgroup.

6.6 Conclusions from ERG analyses

The ERG has presented a number of additional analyses. These analyses were carried out in a number of stages. The first stage addressed a number of minor calculation errors in the company's revised model. The impact of these changes was to reduce the ICER by a small amount from £12,251 per QALY to £13,561 per QALY.

Using the corrected and updated model, the ERG then presented a number of analyses considering a range of issues raised in Section 5. These scenario analyses addressed the following issues:

- The number of induction and consolidation courses received;
- Rate of response to treatment;
- Treatment costs associated with HSCT and VOD;
- Inclusion of VOD events;
- Quality of life in functionally cured patients;
- Excess mortality risk in functionally cured patients.

The majority of these changes resulted in an increase to the ICER, with the exception of using individual rates of response to treatment, although the scenarios were not associated with substantial differences to the ICER. The scenarios associated with the greatest impact on cost-effectiveness outcomes related to changes made by the ERG to the HSCT costs, the quality of life in functionality cured patients and to the use of individual rates of response. This exploration of alternative modelling assumptions and parameter values was concluded with the ERG presenting a base-case with a preferred set of assumptions.

The ERG alternative base-case, based on a probabilistic analysis, estimated GO+DA to be more costly (cost difference [redacted]) and more effective ([redacted] QALY gain) compared with DA, and suggests that the ICER for GO+DA compared with DA is £17,956 per QALY.

A series of exploratory subgroup analysis were conducted by the ERG to explore the impact of heterogeneity between the subgroups included in the base-case population and possible variability within the intermediate group. These analyses explored the impact in different subgroups based on cytogenetic results alone as well as subgroups based on cytogenetic and molecular results. The ICER

of GO+DA versus DA alone varied between £16,343 (intermediate-1 only, defined by cytogenetic and molecular test) and £31,709 (intermediate only, defined by cytogenetic results only).

The ERG concludes that further risk stratification using genetic and molecular testing may provide a clearer separation between subgroups. The results suggest that clinical and economic value of GO+DA appears largely confined to the favourable and intermediate-1 population, defined by cytogenetic and molecular tests. However, these findings can only be considered indicative due to data limitations. Uncertainties also remain concerning the practicality and feasibility of introducing additional risk stratification within routine clinical practice.

7 End of life

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

8 Submissions from practitioner and patient groups

A submission was received from Dr Andrew Goddard, RCP Registrar, on behalf of NCRI-ACP-RCP. Dr Goddard described the aim of treatment for AML; the aim of intensive AML therapy is curative, firstly by achieving remission and then by giving further chemotherapy (with or without the addition of allogeneic stem cell transplant) to prevent relapse. The care pathway for patients with AML is generally well-defined; patients are treated in larger centres, usually these are centres that participate in NCRI AML trials. Dr Goddard stated that the technology would have very little direct impact on the current pathway of care, but that there may be greater onus placed on rapid turnaround of cytogenetic analysis prior to starting chemotherapy, as trials of GO to date have shown that the 20% of patients with adverse risk disease derive no benefit from the addition of GO. If GO is to be restricted to patients with favourable or intermediate risk cytogenetics, the more rapid turnaround of these tests may create an administrative/workload challenge for genetics laboratories. Haematologists and pharmacy departments already have significant experience of using GO from NCRI AML clinical trials over the last 10-15 years. Dr Goddard discussed some of the results of the IPD meta-analysis by Hills et al.,[29] which was updated by Pfizer and is described in Section 4.3.1 of this report, and the ALFA-0701 trial.[26]

A submission was received from Zack Pemberton-Whiteley, Campaigns and Advocacy Director at Leukaemia Care. The results of a Leukaemia Care patient experience survey 'Living with Leukaemia' were used to inform the submission. Mr Whiteley described AML symptoms and the rapidly progressing nature of the condition; many patients are diagnosed via emergency presentation and have to start treatment quickly. There is usually very little time to take in information and start to cope with it. Mr Whiteley stated that AML has extremely poor outcomes and high unmet need, with little progress in decades. 80% of AML patients reported that they would be willing to experience additional side effects for a more effective treatment.

9 Overall conclusions

9.1 Clinical effectiveness

Evidence from one reasonably good quality RCT demonstrates that GO is effective at improving EFS and RFS in patients aged 50-70 years receiving GO + DA compared with those receiving DA alone. Overall survival and response rate also appeared to be improved with GO, although results were not statistically significant. Subgroup analysis results, according to cytogenetic risk profile, indicated that patients with favourable/intermediate cytogenetic risk benefited to a similar degree as the overall population. However, for patients with unfavourable cytogenetics, outcomes appeared to be worse in the GO + DA arm, compared with the DA arm. The CS restricted the decision problem population to patients without known unfavourable cytogenetic risk; which appears appropriate in view of these results. Whilst the proportion of patients experiencing an adverse event was similar between treatment groups, the proportion of patients reporting a serious adverse event, and the proportion of patients who permanently discontinued treatment because of an adverse event, was higher in the GO + DA arm than the DA arm.

The anticipated marketing authorisation for GO is for the treatment of patients age 15 years or above with previously untreated, de novo CD33-positive AML. Therefore, the age range in the trial was narrower than the anticipated marketing authorisation, although the majority of patients with AML are over the age of 50 and the ERG's clinical advisor did not consider that GO would work differently in patients under the age of 50 to those over the age of 50. The restriction to patients with CD33-positive AML appears clinically appropriate, in view of the mechanism of action of GO and subgroup analysis results indicated that GO appears to be more effective in patients with a higher proportion of CD33-positive blasts than those with a lower proportion of CD33-positive blasts.

An IPD meta-analysis, presented as supporting evidence, demonstrated that OS was statistically significantly improved for patients who received GO + DA, compared with those who received DA alone. However, the IPD meta-analysis included patients with de novo or secondary AML or high-risk myelodysplastic syndrome, which is a broader population than the anticipated licence. Therefore, the results may not be entirely generalisable to patients eligible for GO in clinical practice.

Neither the ALFA-0701 trial nor the IPD meta-analysis presented health related quality of life results.

9.2 Cost-effectiveness

The economic evidence presented by the company primarily consisted of a *de novo* model. The company's model used a semi-Markov cohort state-transition approach which directly used the time-to-event data from the ALFA-0701 trial to determine the patient transitions between the health states. The company found GO+DA to be more costly (cost difference of £[redacted]) and more effective ([redacted]) compared with DA alone. The deterministic base-case ICER was £12,251 and the mean probabilistic ICER was £13,600 per QALY.

The ERG considers that the economic analysis presented by the company addressed the decision problem specified in NICE's scope; however, there were some areas of uncertainty that the ERG did not feel were fully explored or able to be captured. The ERG's key concerns related to the structure of the model; while providing predictions that aligned with the clinical trial, the model did contained a number of structural limitations which did not allow uncertainty in a number of key parameters to be fully captured. The ERG was unable to fully address all the identified issues with the company's model, but was able to carry out a number of analyses using assumptions and data inputs it believes are more plausible than those used in the company's base-case analysis. The ERG alternative base-case analysis estimated GO+DA to be more costly ([redacted]) and more effective ([redacted] QALY gain) compared with DA alone, and suggests that the ICER for GO+DA compared with DA is £17,956 per QALY.

The ERG also considered that there was heterogeneity within the base-case population that was not fully explored by the company. This was between the subgroups included in the base-case population and possible variability within the intermediate group. These may have important implications concerning the difference in the cure fraction for further subgroups within the overall population. A series of exploratory subgroup analyses conducted by the ERG provided indicative results for these subgroups, suggesting that clinical and economic value of GO+DA appears confined to the favourable and intermediate-1 population, defined by cytogenetic and molecular tests. Uncertainties also remain concerning the practicality and feasibility of introducing additional risk stratification within routine clinical practice.

9.3 Implications for research

There are two ongoing studies of GO; AML18 and AML19. The dosing schedule of GO used in these trials does not match the dose in the anticipated marketing authorisation. This may have implications for dosing in practice, if GO is approved, as UK clinicians are currently using the AML18 and AML19 doses, rather than the dose used in the ALFA-0701 trial, which is the dose in the anticipated marketing authorisation.

There is no evidence in GO in patients aged under 50 or over 70 in the licensed dose. Further RCT evidence for GO in these patients is required. Further exploration of the impact of GO in different cytogenetic and molecular risk subgroups may be warranted.

10 References

1. Cancer Research UK. *About acute myeloid leukaemia (AML)*. [cited 2017 26th October]. Available from: http://www.cancerresearchuk.org/about-cancer/acute-myeloid-leukaemia-aml.
2. Duncan N. Adult myeloid leukaemias pathogenesis, clinical features and classification. *Clinical Pharmacist* 2010;2:117-21.
3. De Kouchkovsky I, Abdul-Hay M. Acute myeloid leukemia: a comprehensive review and 2016 update. *Blood Cancer Journal* 2016;6:e441.
4. Office for National Statistics. *Cancer Registration Statistics, England, 2015*. 2015. [cited 2018 23 August]. Available from: https://www.ons.gov.uk/peoplepopulationandcommunity/healthandsocialcare/conditionsanddiseases/datasets/cancerregistrationstatisticscancerregistrationstatisticsengland (accessed on 23 August 2017).
5. Cancer Research UK. *Acute myeloid leukaemia (AML) incidence statistics*. [cited 2017 26th October]. Available from: http://www.cancerresearchuk.org/health-professional/cancer-statistics/statistics-by-cancer-type/leukaemia-aml/incidence?_ga=2.56878186.1169686380.1516030870-1880561788.1516030870.
6. Cancer Research UK. *Acute myeloid leukaemia (AML) statistics*. [cited 2017 26th October]. Available from: http://www.cancerresearchuk.org/health-professional/cancer-statistics/statistics-by-cancer-type/leukaemia-aml?_ga=2.56151275.1169686380.1516030870-1880561788.1516030870.
7. Hulegardh E, Nilsson C, Lazarevic V, Garelius H, Antunovic P, Rangert Derolf A, et al. Characterization and prognostic features of secondary acute myeloid leukemia in a population-based setting: a report from the Swedish Acute Leukemia Registry. *Am J Hematol* 2015;90:208-14. Available from: https://www.ncbi.nlm.nih.gov/pubmed/25421221
8. NHS. University College London Hospital. *What is acute myeloid leukaemia (AML)?* [cited 2017 26th October]. Available from: https://www.uclh.nhs.uk/OurServices/ServiceA-Z/Cancer/CBD/AML/Pages/WhatisAML.aspx.
9. American Cancer Society. *Signs and symptoms of acute myeloid leukemia*. [cited 2017 11th October]. Available from: https://www.cancer.org/cancer/acute-myeloid-leukemia/detection-diagnosis-staging/signs-symptoms.html.
10. Bloodwise. *Acute myeloid leukaemia (AML) symptoms and diagnosis*. 2017. [cited 2018 15th january]. Available from: https://bloodwise.org.uk/info-support/acute-myeloid-leukaemia/symptoms.
11. Kumar C. Genetic abnormalities and challenges in the treatment of acute myeloid leukemia. *Genes Cancer* 2011;2:95-107.
12. Public Health England. *National Cancer Intelligence Network. Trends in incidence and outcome for haematological cancers in England: 2001-2010*. London: PHE; 2014. Available from: http://www.ncin.org.uk/cancer_type_and_topic_specific_work/cancer_type_specific_work/haematological_cancers/#
13. Duncan N. Adult myeloid leukaemias: current and future treatments. *Clinical Pharmacist* 2010;2:122–27.
14. British Committee for Standards in Haematology, Milligan D, Grimwade D, Cullis J, Bond L, Swirsky D, et al. Guidelines on the management of acute myeloid leukaemia in adults. *Br J Haematol* 2006;135:450-74.
15. Laszlo GS, Estey EH, Walter RB. The past and future of CD33 as therapeutic target in acute myeloid leukemia. *Blood Rev* 2014;28:143-53. Available from: https://www.ncbi.nlm.nih.gov/pubmed/24809231
16. Jilani I, Estey E, Huh Y, Joe Y, Manshouri T, Yared M, et al. Differences in CD33 intensity between various myeloid neoplasms. *Am J Clin Pathol* 2002;118:560-6.
17. Liersch R, Muller-Tidow C, Berdel W, Krug U. Prognostic factors for acute myeloid leukaemia in adults – biological significance and clinical use. *Br J Haematol* 2014;165:17-38.

18. Byrd J, Mrozek K, Dodge R, Carroll A, Edwards C, Arthur D, et al. Pretreatment cytogenetic abnormalities are predictive of induction success, cumulative incidence of relapse, and overall survival in adult patients with de novo acute myeloid leukemia: results from Cancer and Leukemia Group B (CALGB 8461). *Blood* 2002;100:4325-36.
19. Grimwade D, Hills R, Moorman A, Walker H, Chatters S, Goldstone A, et al. Refinement of cytogenetic classification in acute myeloid leukemia: determination of prognostic significance of rare recurring chromosomal abnormalities among 5876 younger adult patients treated in the United Kingdom Medical Research Council trials. *Blood* 2010;116:354-65.
20. Dohner H, Estey E, Amadori S, Appelbaum F, Büchner T, Burnett A, et al. Diagnosis and management of acute myeloid leukemia in adults: recommendations from an international expert panel, on behalf of the European LeukemiaNet. *Blood* 2010;115:453-74.
21. Döhner H, Estey E, Grimwade D, Amadori S, Appelbaum F, Büchner T, et al. Diagnosis and management of AML in adults: 2017 ELN recommendations from an international expert panel. *Blood* 2017;129:424-47.
22. Sekeres M, Elson P, Kalaycio M, Advani A, Copelan E, Faderl S, et al. Time from diagnosis to treatment initiation predicts survival in younger, but not older, acute myeloid leukemia patients. *Blood* 2009;113:28-36.
23. European Medicines Agency. *Mylotarg: gemtuzumab ozogamicin. EMA/CHMP/85476/2018*. Committee for Medicinal Products for Human Use (CHMP); 2018. [cited 2018 23rd February]. Available from: http://www.ema.europa.eu/docs/en_GB/document_library/Summary_of_opinion_-_Initial_authorisation/human/004204/WC500244277.pdf.
24. Gardin C, Chevret S, Pautas C, Turlure P, Raffoux E, Thomas X, et al. Superior long-term outcome with idarubicin compared with high-dose daunorubicin in patients with acute myeloid leukemia age 50 years and older. *J Clin Oncol* 2013;31:321-7. Available from: https://www.ncbi.nlm.nih.gov/pubmed/23248249
25. Minden M, Wei A. Revisiting late relapses in acute myeloid leukemia. *Leukemia & Lymphoma* 2001;51:735-6.
26. Castaigne S, Pautas C, Terré C, Raffoux E, Bordessoule D, Bastie J-N, et al. Effect of gemtuzumab ozogamicin on survival of adult patients with de-novo acute myeloid leukaemia (ALFA-0701): a randomised, open-label, phase 3 study. *Lancet* 2012;379:1508-16.
27. Yohe S. Molecular genetic markers in acute myeloid leukemia. *Journal of Clinical Medicine* 2015;4:460-78.
28. Pfizer. *A patient-level meta-analysis of studies utilizing intensive induction chemotherapy with or withour mylotarg (gemtuzumab ozogamicin) in patients with previously untreated acute myeloid leukemia. PF-05208747*.
29. Hills RK, Castaigne S, Appelbaum FR, Delaunay J, Petersdorf S, Othus M, et al. Addition of gemtuzumab ozogamicin to induction chemotherapy in adult patients with acute myeloid leukaemia: a meta-analysis of individual patient data from randomised controlled trials. *The Lancet Oncology* 2014;15:986-96.
30. Li X, Xu S, Qin D, Tan Y, Gong Q, Chen J. Effect of adding gemtuzumab ozogamicin to induction chemotherapy for newly diagnosed acute myeloid leukemia: a meta-analysis of prospective randomized phase III trials. *Ann Oncol* 2014;25:455-61.
31. Loke J, Khan J, Wilson J, Craddock C, Wheatley K. Mylotarg has potent anti-leukaemic effect: a systematic review and meta-analysis of anti-CD33 antibody treatment in acute myeloid leukaemia. *Ann Hematol* 2015;94:361-73.
32. Kharfan-Dabaja M, Hamadani M, Reljic T, Pyngolil R, Komrokji R, Lancet J, et al. Gemtuzumab ozogamicin for treatment of newly diagnosed acute myeloid leukaemia: a systematic review and meta-analysis. *Br J Haematol* 2013;163:315-25.
33. Drummond M. *Methods for the economic evaluation of health care programmes*. 2nd ed. Oxford: Oxford University Press; 1997.
34. National Institute for Health and Care Excellence. *Azacitidine for the treatment of myelodysplastic syndromes, chronic myelomonocytic leukaemia and acute myeloid leukaemia*. London: NICE; 2011. Available from: https://www.nice.org.uk/guidance/ta218

35. National Institute for Health and Care Excellence. *Azacitidine for treating acute myeloid leukaemia with more than 30% bone marrow blasts*. London: NICE; 2016. Available from: https://www.nice.org.uk/guidance/ta399
36. Battipaglia G, Labopin M, Candoni A, Fanin R, El Cheikh J, Blaise D, et al. Risk of sinusoidal obstruction syndrome in allogeneic stem cell transplantation after prior gemtuzumab ozogamicin treatment: a retrospective study from the Acute Leukemia Working Party of the EBMT. *Bone Marrow Transplant* 2017;52:592-99. Available from: https://www.ncbi.nlm.nih.gov/pubmed/28092357
37. NHS England Specialised Services Clinical Reference Group for Blood and Marrow Transplantation. *Clinical Commissioning Policy: Treatments for Graft versus Host Disease (GvHD) following Haematopoietic Stem Cell Transplantation*: NHS England; 2017. Available from: https://www.england.nhs.uk/publication/treatments-for-graft-versus-host-disease-gvhd-following-haematopoietic/
38. Ara R, Brazier JE. Populating an economic model with health state utility values: moving toward better practice. *Value Health* 2010;13:509-18. Available from: https://www.ncbi.nlm.nih.gov/pubmed/20230546
39. Scottish Medicines Consortium. *defibrotide (Defitelio). Advice*. 2014. [cited 2017 7th December]. Available from: https://www.scottishmedicines.org.uk/SMC_Advice/Advice/967_14_defibrotide_Defitelio.
40. Kurosawa S, Yamaguchi H, Yamaguchi T, Fukunaga K, Yui S, Kanamori H, et al. Decision analysis of allogeneic hematopoietic stem cell transplantation versus chemotherapy in cytogenetically standard-risk acute myeloid leukemia in first complete remission: the impact of FLT3-ITD profile. *Blood* 2014;124:1221.
41. *British National Formulary*. [cited 2017 5th November]. Available from: https://www.bnf.org.
42. Department of Health. *Drugs and pharmacuetical electronic market information (eMit)*. 2011. [cited 2017 6th December]. Available from: https://www.gov.uk/government/publications/drugs-and-pharmaceutical-electronic-market-information-emit.
43. Department of Health. *Reference Costs 2015-16*. London: DoH; 2016. Available from: https://www.gov.uk/government/publications/nhs-reference-costs-2015-to-2016
44. Curtis L, Burns A. *Unit Costs of Health and Social Care 2016*. Canterbury, Kent: Personal Social Services Research Unit. University of Kent; 2016. Available from: https://www.pssru.ac.uk/pub/uc/uc2016/full.pdf?uc=2016-full
45. NHS England Clinical Reference Group for Blood and Marrow Transplantation. *Clinical Commissioning Policy: Use of defibrotide in severe venoocclusive disease following stem cell transplant*. In: NHS England; 2015. Available from: https://www.england.nhs.uk/commissioning/wp-content/uploads/sites/12/2015/01/b04-use-defibrotide.pdf
46. Addicott R, Dewar S. *Improving choice at end of life: a descrptive analysis of the impact and costs of the Marie Curie Delivering Choice programme in Lincolnshire*. London: Kings Fund. Available from: https://www.kingsfund.org.uk/sites/default/files/improving-choice-end-of-life-descriptive-analysis-impact-costs-marie-curie-choice-programme-lincolnshire-rachael-addicot-steve-dewar-april-2008.pdf
47. NHS Blood and Transplant. *Unrelated donor stem cell transplantation in the UK. Effective affordable, sustainable. A report from the UK Stem Cell Strategy Oversight Committee*; 2014. Available from: http://docplayer.net/7404866-Unrelated-donor-stem-cell-transplantation-in-the-uk.html
48. Esperou Hln, Brunot A, Roudot-Thoraval Fo, Buzyn A, Dhedin N, Rio B, et al. Predicting the Costs of Allogeneic Sibling Stem-Cell Transplantation: Results from a Prospective, Multicenter, French Study. *Transplantation* 2004;77:1854-58.
49. Woods B, Sideris E, Palmer S, Latimer N, Soares M. *Partitioned survival analysis for decision modelling in health care: a critical review. 2017 [Available from http://www.nicedsu.org.uk]*. York: Centre for Health Economics, University of York and School of Health and Related Research, University of Sheffield; 2017. Available from: http://nicedsu.org.uk/wp-content/uploads/2017/06/Partitioned-Survival-Analysis-final-report.pdf

50. Pfizer. *Full Clinical Study Report for Protocol WS936568 (ALFA-0701 [MyloFrance 3]), 15 March 2016*.
51. Office for National Statistics. *National life tables, UK: 2013 – 2015*. 2016. [cited 2018 21st March]. Available from: https://www.ons.gov.uk/releases/nationallifeexpectancytablesuk2013to2015.
52. van Agthoven M, Groot M, Verdonck L, Löwenberg B, Schattenberg A, Oudshoorn M, et al. Cost analysis of HLA-identical sibling and voluntary unrelated allogeneic bone marrow and peripheral blood stem cell transplantation in adults with acute myelocytic leukaemia or acute lymphoblastic leukaemia. *Bone Marrow Transplant* 2002;30:243-51.
53. Office for National Statistics. *National life tables: United Kingdom*. ONS; 2017. [cited 2018 21st March]. Available from: https://www.ons.gov.uk/peoplepopulationandcommunity/birthsdeathsandmarriages/lifeexpectancies/datasets/nationallifetablesunitedkingdomreferencetables.

11 Appendices

11.1 Economic analysis quality checklist

Table 53 summarises the results of the Drummond checklist applied to the company cost effectiveness submission.

Table 53 Quality checklist for the company model

Company submission		Reviewer's judgment	Notes
Study design			
1	Was the research question stated?	Yes	The decision problem was described in detail
2	Was the economic importance of the research question stated?	Yes	Yes
3	Was/were the viewpoint(s) of the analysis clearly stated and justified?	Yes	NHS and Personal Social Services, as required to meet the NICE reference case
4	Was a rationale reported for the choice of the alternative programmes or interventions compared?	Yes	Intervention and comparator described in the decision problem section of the CS
5	Were the alternatives being compared clearly described?	Yes	Description provided of the dosing regimen and number of courses provided
6	Was the form of economic evaluation stated?	Yes	Cost-effectiveness analysis
7	Was the choice of form of economic evaluation justified in relation to the questions addressed?	Yes	To capture all relevant outcomes and meet the NICE reference case
Data collection			
8	Was/were the source(s) of effectiveness estimates used stated?	Yes	Pivotal trial for GO (ALFA-0701)
9	Were details of the design and results of the effectiveness study given (if based on a single study)?	Yes	ALFA-0701 methodology described in clinical effectiveness section
10	Were details of the methods of synthesis or meta-analysis of estimates given (if based on an overview of several effectiveness studies)?	NA	Only one study with the licensed dose of GO has been completed, meta-analysis not possible or necessary
11	Were the primary outcome measure(s) for the economic evaluation clearly stated?	Yes	The ICER (cost per QALY) was reported
12	Were the methods used to value health states and other benefits stated?	Yes	Benefits captured with QALYs
13	Were the details of the subjects from whom valuations were obtained given?	No	Limited details of the patients in the utility elicitation study were provided in the CS, although a reference to the source was provided.
14	Were productivity changes (if included) reported separately?	NA	Not relevant to decision problem
15	Was the relevance of productivity changes to the study question discussed?	Yes	The CS acknowledged that indirect costs were often production losses due to premature mortality
16	Were quantities of resources reported separately from their unit cost?	Yes	Unit costs and resource use for each item were presented

17	Were the methods for the estimation of quantities and unit costs described?	Yes	The CS provides a description of how resource use and cost data was modelled
18	Were currency and price data recorded?	Yes	Unit costs were reported in GBP. Unit costs, their sources and the year in which they were published were described.
19	Were details of price adjustments for inflation or currency conversion given?	No	Neither the conversion rate nor the inflation factor were reported
20	Were details of any model used given?	Yes	A detailed description of the economic model, with a model schematic and details of health states, were provided
21	Was there a justification for the choice of model used and the key parameters on which it was based?	Yes	The model structure was determined by the clinical pathway, the nature of the available data, previous AML models and clinical expert input. Details of the clinical expert input were not provided. A detailed description and a summary of all key parameters was provided.
Analysis and interpretation of the results			
22	Was time horizon of cost and benefits stated?	Yes	Stated and justified in the CS
23	Was the discount rate stated?	Yes	Stated for both cost and benefits
24	Was the choice of rate justified?	Yes	In accordance with the NICE reference case
25	Was an explanation given if cost or benefits were not discounted?	NA	NA
26	Were the details of statistical test(s) and confidence intervals given for stochastic data?	Yes	A description of the standard error and the probability distribution for the parameters included in the PSA were provided in an Appendix to the CS
27	Was the approach to sensitivity analysis described?	Yes	A deterministic and probabilistic sensitivity analysis was undertaken
28	Was the choice of variables for sensitivity analysis justified?	No	Rationale for inclusion of variables in the probabilistic and deterministic sensitivity analysis was not stated. The scenario analyses tested “key model drivers and areas of uncertainty”
29	Were the ranges over which the parameters were varied stated?	Yes	In the CS, it was stated that parameters were varied by +/- 10%
30	Were relevant alternatives compared? (i.e. Were appropriate comparisons made when conducting the incremental analysis?)	Yes	The comparator intervention, DA, was relevant and well-justified
31	Was an incremental analysis reported?	Yes	The company presented incremental costs and QALYs, and the ICER
32	Were major outcomes presented in a disaggregated as well as aggregated form?	Yes	These were provided in an Appendix to the CS
33	Was the answer to the study question given?	Yes	The CS provided an estimate of cost-effectiveness for the intervention within the described decision problem

34	Did conclusions follow from the data reported?	Yes	The interpretation of the results was consistent with the data reported
35	Were conclusions accompanied by the appropriate caveats?	Yes	The company provided a description of the limitations with the analysis
36	Were generalizability issues addressed?	Yes	Relevance to UK clinical practice and to those who could potentially use the technology were discussed

National Institute for Health and Care Excellence
Centre for Health Technology Evaluation

Pro-forma Response

ERG report

Gemtuzumab ozogamiacin for untreated acute myeloid leukaemia [ID982]

You are asked to check the ERG report from Centre for Reviews and Dissemination and Centre for Health Economics – York to ensure there are no factual inaccuracies contained within it.

If you do identify any factual inaccuracies you must inform NICE by **5pm on Wednesday 4 April 2018** using the below proforma comments table. All factual errors will be highlighted in a report and presented to the Appraisal Committee and will subsequently be published on the NICE website with the committee papers.

The proforma document should act as a method of detailing any inaccuracies found and how and why they should be corrected.

4th April 2018

Dear Meindert,

Pfizer would like to thank the York ERG for their thorough review of the gemtuzumab ozogamiacin for untreated acute myeloid leukaemia submission and welcomes the opportunity to check the ERG report to ensure there are no factual inaccuracies.

We have structured our response using the proforma document provided as follows:

- Issue 1, Typos or incorrectly reported values (entire ERG report)
- Issue 2, Increase clarity of sentences/descriptions (Sections 1-4 of ERG report)
- Issue 3, Increase clarity of sentences/descriptions (Sections 5-9 of ERG report)
- Issue 4, Confidentiality marking (entire ERG report)

Please do not hesitate to contact us should you require further information.

With best wishes,

[redacted]

Issue 1 Typos or incorrectly reported values (entire ERG report)

Description of problem	Description of proposed amendment	Justification for amendment	ERG response
On page 64, Table 16 (Sensitivity analysis), the ERG report states: *"Probabilistic sensitivity analysis was performed. Deterministic univariate **probabilistic** analysis was performed on a series of model parameters. A series of scenario analyses were also performed."*	Please replace "probabilistic" with "sensitivity"	Typo. No impact on clinical and cost-effectiveness conclusions presented.	Text has been amended as per the company suggestion
On page 80 and page 74 (table 20) please correct the average model entry age so that it matches base-case model (favourable /intermediate/unknown): [redacted]	These references should be correct to the following: [redacted]	So average entry age matches base-case population model entry age. No impact on clinical and cost-effectiveness conclusions presented.	The mean age reported on page 80 has been updated so that it reports the value for the combined population, as suggested by the company. Table 20 on page 74 summarises the mean age for each arm individually and combined (to two decimal places). The ERG has included an additional footnote in the table clarifying that data from the combined arms is used in the model.
On page 98 of the ERG report, It states: *"RMST for relapsed and refractory patients who did not receive HSCT were estimated from data pooled*	Please correct these pooled values for the basecase model to: *"RMST for relapsed and refractory patients who did not receive HSCT were estimated from data pooled from both arms, and was [redacted] for refractory patients and*	So the RMST values are correct and match base-case model. No impact on clinical and cost-effectiveness conclusions presented.	Text has been corrected to report the cytogenetic subgroup value, rather than the overall population value

from both arms, and was [redacted] for refractory patients and [redacted] for relapsed patients."	[redacted] *for relapsed patients."*		
On page 125 the ERG report states: *"The impact of these changes was to reduce the ICER by a small amount from £12,251 per QALY to £13,561 per QALY."*	*"The impact of these changes was to **increase** the ICER by a small amount from £12,251 per QALY to £13,561 per QALY."*	Typo. No impact on clinical and cost-effectiveness conclusions presented.	Text has been amended as per the company suggestion
Page 93; Table 31: The footnotes are all denoted by "*" despite being relevant for different data.	Ensure the footnotes match the corresponding data.	Correct vague footnotes to table. No impact on clinical and cost-effectiveness conclusions presented.	Footnotes have been amended so that each are denoted numerically
Page 42, table 6: The following sample sizes are incorrect for the unfavourable cytogenetic patients: GO+DA ([redacted]), DA ([redacted])	Please correct to the following: GO+DA ([redacted]), DA ([redacted])	So the samples sizes are correct. No impact on clinical and cost-effectiveness conclusions presented.	Sample sizes corrected on page 42, table 6 for EFS and OS result. Sample size of 14 and 15 is reported for RFS (table 87 of CS).

Issue 2 Increase clarity of sentences/descriptions (Sections 1-4 of ERG report)

Description of problem	Description of proposed amendment	Justification for amendment	ERG response
On page 14 and 60, the ERG report states: *"All [redacted] patients who experienced veno-occlusive disease (VOD) had received GO"* Please clarify to avoid confusion	Suggest changing sentence to following: *"All [redacted] patients who experienced veno-occlusive disease (VOD) had received GO. This included [redacted] patients in the DA arm who received GO during follow-up as part of the compassionate use program. They experienced VOD more than 28 days after*	Increase clarity. No impact on clinical and cost-effectiveness conclusions presented.	Text amended on pages 14 and 60 for clarity.

(see CSR page 156-157).	*receiving their last dose of GO ([redacted]).”*		
On page 15 (and page 59), the ERG report states: [redacted] This should be changed to avoid confusion (please see page 98 and Figure 5 in CSR; or figure 5 in ERG report).	Suggest changing to the following to be more specific: [redacted] Suggest also adding a caveat based on IPD meta-analysis (see meta-analysis CSR table 11, p65; figure 11, p74): [redacted]	Increase accuracy when reporting ALFA sub-group analyses. No impact on clinical and cost-effectiveness conclusions presented.	A modified version of the amendment (see below) has been added to page 15, and a shorter version to page 59 [redacted]
On page 15 the ERG report states: *“The IPD meta-analysis, presented as supporting evidence, included patients aged 15 years or older with newly diagnosed AML (either de novo or secondary), or high-risk myelodysplastic syndrome (MDS), which is a broader population than that defined in the decision problem or the anticipated licence.”*	Suggest changing for added detail to: *“The IPD meta-analysis, presented as supporting evidence, included patients aged 15 years or older with newly diagnosed AML (either de novo or secondary), or high-risk myelodysplastic syndrome (MDS), which is a broader population than that defined in the decision problem or the anticipated licence. However, the number of MDS patients corresponds to [redacted] of the all trial meta analyses sample.”*	Increase accuracy about disease status of meta analyses population. No impact on clinical and cost-effectiveness conclusions presented.	Text amended on pages 15 and 26.

Also on page 26 of the ERG report: *"The CS also included an individual patient data (IPD) meta-analysis as supporting evidence in Appendix D.3.1 of the CS. Patients included in the IPD meta-analysis were aged 15 years or older with newly diagnosed AML (either de novo or secondary), or high-risk myelodysplastic syndrome (MDS), which is a broader population than that defined in the decision problem or the anticipated licence."* Should give proportion of MDS patients in meta-analyses (presented in Table 81, page 98 of the Pfizer submission appendix) for added clarity.			
On page 36 the ERG report states: *"The anticipated marketing authorisation for GO specifies patients with CD33-positive AML, whilst the trial also included AML patients who were not CD33-positive, therefore, a small proportion of patients in the trial would not be eligible for GO + DA in clinical practice."* Clarify that in clinical practice	Suggested change: *"The anticipated marketing authorisation for GO specifies patients with CD33-positive AML, whilst the trial also included AML patients who were not CD33-positive, therefore, a small proportion of patients in the trial would not be eligible for GO + DA in clinical practice.* ***However, clinical expert opinion suggested that in current clinical practice patients are usually treated irrespective of their CD33 status."***	Increase clarity of sentence. No impact on clinical and cost-effectiveness conclusions presented.	This is not a factual inaccuracy and is not based on the clinical effectiveness evidence presented. Therefore, amendment not made.

patients are often treated irrespective of CD33 status.			
On page 39 it is stated: [redacted] Please add “3 years” to give context to [redacted]	Please adapt to: [redacted]	Increase clarity of sentence. No impact on clinical and cost-effectiveness conclusions presented.	Text amended on page 39 for clarity.
In table 5 of the ERG report, row 4 (response rates) it is not specified that this is IRC and not IA data. In table 6 of the ERG report please specify that the reported endpoints are based on IRC data. In table 7, 8, 9, 10, 11 and 12 of the ERG report the data cut-off should be stated and that the first column is based on IRC data.	Please specify that these are IRC response rates and data where applicable.	Increase accuracy. No impact on clinical and cost-effectiveness conclusions presented.	Table 5 and Table 6 amended for accuracy. ‘IRC analysis’ has been added to the headings of Tables 7-12, however, data cut-off was not provided in the company response to the ERG’s points for clarification. We have also added a sentence to Section 4.2.4 (page 39): *All tables in this section report the results of the IRC analysis.*
On page 43 of the ERG report it states: *“It was not clear to the ERG exactly*	These classifications were based on ELN guidelines, please adapt.	Increase accuracy. No impact on clinical and cost-effectiveness conclusions presented.	Text amended on page 43 for accuracy, so it now reads: *These classifications were*

what test was used for this assessment: the CS stated that subgroup analyses were performed using risk classification based on NCCN and ELN guidelines." *See the heading of Tables 14.2.2.11 to Table 14.2.10.33 sent in response to NICE clarification questions.*			*based on ELN guidelines.*
On page 43/44 of the ERG report it states: [redacted]	Please account for powering and sample size. Suggested change: [redacted]	Increase clarity of sentence. No impact on clinical and cost-effectiveness conclusions presented	Not a factual inaccuracy.
On page 57 the ERG report states: *"Overall survival and RFS were statistically significantly improved in the GO arm compared with the no-GO arm. In subgroup analyses overall survival was significantly improved in the GO arm for patients with favourable cytogenetic risk and patients with intermediate cytogenetic risk, but not for patients*	Suggested change: *"Overall survival and RFS were statistically significantly improved in the GO arm compared with the no-GO arm. In subgroup analyses overall survival was significantly improved in the GO arm for patients with favourable cytogenetic risk and patients with intermediate cytogenetic risk, but not for patients with adverse cytogenetic risk.* ***The addition of GO to induction chemotherapy***	Increase clarity of sentence. No impact on clinical and cost-effectiveness conclusions presented	The text on pages 57 and 58 has been amended to clarify that the addition of GO did not significantly improve the rate of CR (consistent with the wording in the Hills et al., Lancet Oncology publication).

with adverse cytogenetic risk. ***The addition of GO to induction chemotherapy did not improve the rate of CR."*** *There was some improvement in point estimate CR but this was not statistically significant (.OR 0.91 (0.77-1.07) p=0.3).* *A similar statement is made in the line at the end of the next paragraph.*	***did improve the rate of CR but the difference was not statistically significant."***		

Issue 3 Increase clarity of sentences/descriptions (Sections 5-9 of ERG report)

Description of problem	Description of proposed amendment	Justification for amendment	ERG response
On page 63, Table 16 (treatment effectiveness), the ERG report states: *"Clinical outcomes included were response (CR/CRp), RFS and OS, cure fraction, probability of HSCT,* ***post-HSCT survival****"*	Post-HSCT survival was not explicitly used in the model. Post-HSCT cure fraction was used to adjust OS. Suggest deleting "*post-HSCT survival"*	Increase accuracy. No impact on clinical and cost-effectiveness conclusions presented.	"Post-HSCT survival" has been replaced with "HSCT cure fraction"
On page 63, Table 16 (HRQoL), the ERG report states: *"A vignette study undertaken by the company provided an alternate set of utility values that were used in a scenario analysis."*	Two sets of values (TTO & VAS) were provided and used in scenario analysis. Suggest updating to: *"A vignette study undertaken by the company provided two alternate sets of utility values that were used in scenario analysis."*	Increase accuracy. No impact on clinical and cost-effectiveness conclusions presented.	Text amended based on company suggestion
On page 63, Table 16 (HRQoL), the ERG report states:	The Kurosawa estimate was sourced from the 2016 publication, not 2014:	Incorrect reference & increase accuracy.	The ERG noted that Kurosawa (2016) is a

"Functionally cured patients were assumed to have quality of life equal to that of the age-matched general population (Ara & Brazier). The remaining health state utilities were sourced from NICE TA399. *Adverse event disutility's were sourced from external literature including: NICE TA399, the appraisal of defibrotide by the SMC for VOD and* ***Kurosawa (2014).****"*	Kurosawa, S.,Yamaguchi, H.,Yamaguchi, T.,Fukunaga, K.,Yui, S.,Wakita, et al. Decision Analysis of Postremission Therapy in Cytogenetically Intermediate-Risk Acute Myeloid Leukemia: The Impact of FLT3 Internal Tandem Duplication, Nucleophosmin, and CCAAT/Enhancer Binding Protein Alpha. Biology of Blood and Marrow Transplantation. 2016. 6(22):1125-1132. The estimate was used for GVHD but this was captured as a health-state utility not a disutility. Suggest updating to: *"Functionally cured patients were assumed to have quality of life equal to that of the age-matched general population (Ara & Brazier). The remaining health state utilities were sourced from NICE TA399, except for post-HSCT CR/CRp with GVHD which was sourced from Kurosawa (2016).* *All adverse event disutility's were sourced from NICE TA399, except for VOD which was sourced from the appraisal of defibrotide by the SMC"*	No impact on clinical and cost-effectiveness conclusions presented.	published decision analysis and economic model, and is not the primary study in which the utility value was estimated. Kurosawa (2015) appears to be the primary study on which the utility value was based (as referenced within Kurosawa (2016)). However, the company justified the use of the source citing Kurosawa (2014) in Table 41 on 126 of the CS). Given the ambiguity and the lack of access to the 2014 publication, the ERG agrees to change the reference to Kurosawa (2016). The text has been updated using the company's suggestion.
On page 64, Table 16 (RU & Costs), the ERG report states: *"These comprised: drug acquisition, health state costs (monitoring and management), terminal care costs, HCST and GVHD treatment costs,*	Suggested change: *"These comprised: drug acquisition, health state costs* ***(including inpatient and outpatient attendances, disease monitoring and management)*** *HCST and GVHD treatment costs, and treatment of*	Increase accuracy. No impact on clinical and cost-effectiveness conclusions presented.	Additional items (inpatient and outpatient attendances) included in the paragraph on page 64

and treatment of adverse effects."	*adverse effects."*		
On page 66, Table 17 (Refractory: non-curative therapy), the ERG report states: *"Refractory patients on non-curative therapy (BSC only) who failed **subsequent** salvage therapy in the previous cycles"*	Please delete "subsequent"	Increase accuracy. No impact on clinical and cost-effectiveness conclusions presented.	Text amended as per the company suggestion
On page 68 (Relapse: non-curative therapy) the ERG report states: *"Response was not explicitly modelled in either of the relapse states. However, HSCTs were assumed to have been undertaken in patients who received prior salvage therapy and had achieved CR/CRp."*	Response to second-line treatment (CR2) from relapse and refractory health states was not explicitly modelled because no data were available from the ALFA-0701 study. Suggested change: *"Response was not explicitly modelled in either of the relapse states because this outcome was not included in the* ALFA-0701 study*. However, HSCTs were assumed to have been undertaken in patients who received prior salvage therapy and had achieved CR/CRp."*	Acknowledge data limitations and increase accuracy. No impact on clinical and cost-effectiveness conclusions presented.	Not a factual error.
On page 68 (HSCT) the ERG report states: [redacted]	The time points were based on the average time to HSCT for the CR/CRp and refractory states, but not for relapse [redacted] Suggested change: [redacted]	Increase accuracy. No impact on clinical and cost-effectiveness conclusions presented.	Additional text included in the paragraph on page 68, based on a modified version of the company's suggested statements. Change made: Patients could receive HSCT

[redacted]	[redacted]		from a number of health states at different (fixed) time points, including from "CR/CRp off-treatment" at six months, from "refractory" at [redacted] and from "relapse" at [redacted], [redacted], [redacted], [redacted] and [redacted]. For the CR/CRp and refractory states, the time points were based on the average time to HSCT. For relapse, the mid-point of each year was assumed for years 2 to 5. These timings were based on calendar time (i.e. from the point of randomisation) and not time in state. The use of calendar time implies that any relapses which occur after [redacted] would not receive HSCT.
On page 69 (functionally cured) the ERG report states: *"Patients in the functionally cured health state were assumed to be no longer at risk of relapse or mortality due to AML.* *However, longer-term excess mortality due other causes (e.g.*	Survival and relapse for the functionally cured state were still determined by the underlying OS and RFS curves (which had cure factored in and excess mortality HR applied at 5 years). Although there was not a structural link between the functionally cured and relapse states, CR/CRp patients could still relapse after 5 years – it was just that all patients in CR/CRp, as determined by the curves, were assumed to be functionally	Increase accuracy. No impact on clinical and cost-effectiveness conclusions presented.	Paragraph has been amended using the company's suggested text

higher risk of secondary cancers, cardiac events, etc.) was assumed for the remainder of the model time horizon by applying a hazard ratio (HR) to general population mortality rates"	cured after 5 years. Suggested change: *"Patients in CR/CRp after 5 years transitioned to the functionally cured health state. Patients were assumed to be no longer at risk of mortality due to AML. However, longer-term excess mortality due to other causes (e.g. higher risk of secondary cancers, cardiac events, etc.) was assumed for the remainder of the model time horizon by applying a hazard ratio (HR) to general population mortality rates"*		
On page 87 the ERG report states: *"The difference in the cure fractions for EFS and OS suggests that either: (i) there are a significant number of patients who become functionally cured following relapse (potentially due to subsequent therapies and/or HSCT) or (ii) the data may not be sufficiently mature to robustly estimate the cure fraction for OS."* Please clarify by adding model/trace predictions about proportion of relapsers that become functionally cured.	Suggested change: *"The difference in the cure fractions for EFS and OS suggests that either:* *(i) there are a significant number of patients who become functionally cured following relapse (potentially due to subsequent therapies and/or HSCT) .* ***However, the proportion of patients that the model predicts are relapsers who do not receive a HSCT but enter the functionally cured state at 5 years is small and balanced between arms ([redacted] for GO and [redacted] for GO+DA).*** *(ii) or (ii) the data may not be sufficiently mature to robustly estimate the cure fraction for OS.* ***Although clinical opinion suggested the***	Increase clarity. No impact on clinical and cost-effectiveness conclusions presented.	Additional text has been included in the subsequent paragraph, stating the ERGs considerations on these concerns, based on the company's suggestion. Regarding the second point, the ERG has described evidence earlier in the report (Section 3.4) that suggests that functional cure is more robustly estimated from longer-term survival data. The ERG does not consider the addition of the statement suggested by the company to be reasonable on this basis.

	establishment of a cured proportion by 3-5 years*".*		
On page 90 the ERG report states: *"The data provided suggested no obvious bias or differences in the time at risk. However, the ERG notes that some of the cost-offsets for HSCT are predicated on the functional cure assumption* ***(i.e. that those patients who have not relapsed by five years will never relapse and hence will not require HSCT at some point in the future)****. The absence of any structural link to the rates of HSCT limits the ability of the ERG to further assess the potential impact of this source of uncertainty."*	Per above, although there was not a structural link between the functionally cured and relapse states, CR/CRp (and functionally cured) patients could still relapse after 5 years – it was just that all patients in CR/CRp, as determined by the curves, were assumed to be functionally cured after 5 years. Therefore, it was the absence of an explicit structural link between the relapse state and HSCT rates that limited the ability to assess assumptions about HSCT rates beyond the trial time horizon (although no HSCTs occurred in year 5). Suggest re-wording.	Increase accuracy. No impact on clinical and cost-effectiveness conclusions presented.	Not a factual error. The issue regarding the lack of a structural link between relapse and HSCT has also been discussed extensively in previous sections.
On page 93 the ERG report states: *"Quality of life for patients on intensive chemotherapy (including induction therapy, consolidation therapy and salvage therapy) was assumed to be 0.657. Differences between arms for patients on GO+DA and on DA were captured by estimating the utility decrement associated with the different safety profiles.* ***In TA399, this was the value estimated from patients in non-remission (stable disease or partial remission)."***	Bold text is referring to the intensive chemotherapy utility estimate (0.657) not the disutility; could be misinterpreted. Suggested change: *"In TA399, this utility value (0.657) was estimated from patients in non-remission (stable disease or partial remission)."*	Increased accuracy No impact on clinical and cost-effectiveness conclusions presented.	Paragraph text has been amended on page 93.

On page 94 the ERG report states: *"The utility value for patients with GVHD following HSCT was 0.67, and was taken from a published economic analysis of HSCT patients in first remission from AML 40 which was identified in the SLR.* ***The reference for this value was not provided by the company and therefore the utility value could not be verified.*** *The economic analysis did not differentiate between chronic and acute GVHD."*	A reference is included in Table 31 in the ERG report (Kurosawa, 2014). However, this should be Kurosawa, 2016. This reference was provided in the CS (in table footnote): Kurosawa, S.,Yamaguchi, H.,Yamaguchi, T.,Fukunaga, K.,Yui, S.,Wakita, et al. Decision Analysis of Postremission Therapy in Cytogenetically Intermediate-Risk Acute Myeloid Leukemia: The Impact of FLT3 Internal Tandem Duplication, Nucleophosmin, and CCAAT/Enhancer Binding Protein Alpha. Biology of Blood and Marrow Transplantation. 2016. 6(22):1125-1132 Suggested change: *"The utility value for patients with GVHD following HSCT was 0.67, which was taken from a published economic analysis identified in the SLR (Kurosawa, 2016). The economic analysis did not differentiate between chronic and acute GVHD."*	Incorrect information No impact on clinical and cost-effectiveness conclusions presented.	The reference has been updated (see ERG response to the previous related point). The text has also been updated, excluding the statement regarding the availability of the reference as per the company's suggestion.
On page 95 the ERG report states: *"The CS provided a description of the resource use and costs incurred over time. These included: drug acquisition costs, drug administration costs, HSCT costs, costs related to the health states, costs associated with adverse events and costs related to terminal cancer patients that were applied at the end of the*	Drug acquisition costs were not explicitly included (captured within hospitalisation costs, which were part of health state costs). Suggested change: *"The CS provided a description of the resource use and costs incurred over time. These included: drug acquisition costs, HSCT costs, costs related to the health states* ***(including inpatient and outpatient attendances, disease monitoring and management)*** *costs associated with adverse*	Increased accuracy No impact on clinical and cost-effectiveness conclusions presented.	Not a factual inaccuracy: this statement is a summary and further details of the health state costs are provided in the relevant subsection.

patient's life."	*events and costs related to terminal cancer patients that were applied at the end of the patient's life."*		

Issue 4 Confidentiality marking (entire ERG report)

Description of problem	Description of proposed amendment	Justification for amendment	ERG response
Probabilities of relative cost-effectiveness (from the PSA) should be CiC marked as in the company submission: See page 16, 105 (company submission probabilities): Also page 106 (ERG model probabilities)	Please mark.	Marking should be in line with company submission.	The ERG notes that the probability of cost-effectiveness did not have any confidentiality marking in the version that it received when describing the results relative to the £30,000 per QALY threshold (page 144 of CS and page 405 of appendix). However, these probabilities have been marked CiC as requested by the company.
Please mark the following: [redacted]	Trend statements about marked data should be marked.	Marking should be in line with company submission.	Sentence on page 42 marked as AIC data.
Please mark the following on page 43/44: [redacted]	Trend statements about marked data should be marked.	Marking should be in line with company submission.	Paragraph on page 43/44 marked as AIC data.

[redacted]			
On page 45 please mark the following: [redacted]	Trend statements about marked data should be marked.	Marking should be in line with company submission.	Paragraph on page 45 marked as AIC data.
On page 68 please mark the following (at least the values): [redacted]	Please mark.	Marking should be in line with company submission.	Values in the paragraph on page 68 marked as AIC data.

[redacted]			

trial is the pivotal study used to support the EMA marketing authorisation. The remaining seven RCTs did not use a dose or dosing schedule of GO that is expected to be approved by the EMA; therefore, only the ALFA-0701 trial is the primary source of clinical effectiveness evidence for the appraisal.

ALFA-0701 was a phase 3 multicentre open-label RCT undertaken at 26 haematology centres in France. Patients aged 50-70 years of age were randomised to GO + DA or DA alone.

The ALFA-0701 trial demonstrated that GO + DA was effective at improving event-free survival (EFS) by approximately [redacted] and relapse-free survival (RFS) by approximately [redacted], compared with DA alone, in the overall patient population. Whilst overall survival (OS) and response rate appeared better in the GO + DA arm, these results did not reach statistical significance [redacted].

EFS and RFS results were statistically significantly improved in the GO + DA arm for the subgroup of patients with favourable/intermediate cytogenetic risk, to a similar extent as the overall population, with OS and response rate also improved in the GO + DA arm, but not reaching statistical significance, consistent with the overall results. However, for patients with an unfavourable cytogenetic profile, OS [redacted] and RFS [redacted] outcomes appeared to be worse in the GO + DA arm, compared with the DA arm, whilst EFS results were similar [redacted].

Additional analyses provided to the ERG on request showed that the benefit seen in patients with an intermediate-1 cytogenetic and molecular risk profile was not found in patients with an intermediate-2 cytogenetic and molecular risk profile, suggesting potentially important heterogeneity in the broader 'intermediate' cytogenetic subgroup.

Whilst the proportion of patients experiencing an adverse event was similar between treatment groups, the proportion of patients experiencing a serious adverse event was higher in the GO + DA arm than the DA arm [redacted]. The most common serious adverse event in the GO + DA arm was thrombocytopenia [redacted]. All [redacted] patients who experienced veno-occlusive disease (VOD) had received GO. This included [redacted] patients in the DA arm who received GO during follow-up as part of the compassionate use program. They experienced VOD more than 28 days after receiving their last dose of GO ([redacted]). A higher proportion of patients in the GO + DA arm permanently discontinued treatment because of an adverse event than in the DA arm [redacted].

the GO arm for the overall population in the IPD meta-analysis [redacted] and for the subpopulation of patients with a favourable or intermediate cytogenetic profile [redacted], but not for patients with an unfavourable cytogenetic profile [redacted], supporting the evidence from the ALFA-0701 trial.

The ALFA-0701 trial and the IPD meta-analysis did not report quality of life outcomes.

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

The evidence for the clinical effectiveness of GO + DA is based on one reasonably good quality open-label RCT and the results are likely to be reliable. The ALFA-0701 trial included GO + DA at the recommended dose, which the company states is in line with the anticipated marketing authorisation.

The ALFA-0701 trial included patients aged 50-70 years with previously untreated de novo AML. This is a sub-population of the patient population described in the final NICE scope, and the anticipated marketing authorisation, which specifies patients age 15 years and above. The clinical advisor to the ERG stated that GO is unlikely to work differently in patients under 50 years old, and that GO is unlikely to be used extensively in patients over 70 years old, due to the poorer prognosis of older patients. The majority of patients diagnosed with AML are over 50 years of age, therefore, the ALFA-0701 trial is likely to be reflective of the majority of patients who would be eligible for treatment with GO + DA in clinical practice.

The ALFA-0701 trial included patients with AML, regardless of CD33-positivity, so patients in the trial who were not CD33-positive would not be eligible for GO + DA in clinical practice under the anticipated licence.

[redacted]

The IPD meta-analysis, presented as supporting evidence, included patients aged 15 years or older with newly diagnosed AML (either de novo or secondary), or high-risk myelodysplastic syndrome (MDS), which is a broader population than that defined in the decision problem or the anticipated licence. The number of MDS patients corresponds to [redacted] of the all trial meta-analysis sample. Therefore, the results may not be entirely generalisable to patients eligible for GO + DA in clinical practice.

including UK based economic evaluations which were used to inform model parameters in the analysis, but did not identify any relevant economic assessments of GO.

The cost effectiveness of GO+DA compared with DA alone was informed by an economic evaluation conducted by the company. The population included in the company's decision problem and economic model comprises patients not known to have unfavourable cytogenetics. The company's model used a semi-Markov cohort state-transition model. The model used the time-to-event data from the ALFA-0701 clinical trial to determine the movement of patients between the health states. The model structure consisted of the following health states: (i) induction therapy; (ii) complete remission (CR/CRp): consolidation therapy; (iii) CR/CRp: off-treatment; (iv) refractory (receiving salvage therapy); (v) refractory (receiving non-curative therapy); (vi) hematopoietic stem cell transplant (HSCT); (vii) post-HSCT CR/CRp (without graft versus host disease (GVHD)); (viii) post-HSCT CR/CRp (with GVHD); (ix) relapse (receiving salvage therapy), (x) relapse (receiving non-curative therapy); (xi) functionally cured; and (xii) death. Patients could receive (hematopoietic stem cell transplant) HSCT after achieving complete remission (CR/CRp) after induction therapy, after receiving salvage therapy if they were refractory to first-line therapy, and after receiving salvage therapy if they experienced a relapse.

The efficacy data, treatment and comparator dosage, proportions of patients receiving induction and consolidation therapy, rates of HSCT, adverse event rates and patient characteristics (age, gender, weight, body surface area) used in the economic model were sourced from the ALFA-0701 clinical trial, with the remaining inputs informed by studies identified in the cost-effectiveness review and other sources. Overall survival and relapse-free survival was estimated using mixture cure models fit to Kaplan-Meier data from the ALFA-0701 trial. Patients remaining alive or relapse-free at 60 months after induction therapy or after HSCT were assumed to be cured and experienced general population mortality adjusted to reflect excess mortality in AML survivors. Probabilities of HSCT were estimated from ALFA-0701 trial data and were applied at distinct time points during the first 60 months in the model.

In the base-case analysis of patients excluding those with known, unfavourable cytogenetics, the company found GO+DA to be more costly (cost difference of £[redacted]) and more effective ([redacted] QALY gain) compared with DA alone. The deterministic base-case incremental cost-effectiveness ratio (ICER) was £12,251 per QALY, and the mean probabilistic ICER was £13,600 per quality-adjusted life year (QALY). The predicted probability that GO+DA was cost-effective compared with DA alone was [redacted]% at a cost-effectiveness threshold of £20,000 per QALY and [redacted]% at a cost-effectiveness threshold of £30,000 per QALY. The company reported that the most influential

marketing authorisation for GO in 'combination therapy with daunorubicin (DNR) and cytarabine (AraC) for the treatment of patients age 15 years and above with previously untreated, de novo CD33-positive acute myeloid leukaemia (AML), except acute promyelocytic leukaemia (APL)'.[23]

The ALFA-0701 trial included patients aged 50-70 years with previously untreated de novo AML. Therefore, this is a sub-population of the patient population described in the final NICE scope, in which no age restriction was applied, and the anticipated marketing authorisation. The clinical advisor to the ERG stated that GO is unlikely to work differently in patients under 50 years old, and that GO is unlikely to be used extensively in patients over 70 years old, due to the poorer prognosis of older patients, who are more likely to have comorbidities and be less able to tolerate intensive chemotherapy. The majority of patients diagnosed with AML are over 50 years of age, therefore, the population included in the ALFA-0701 trial is likely to be reflective of the majority of patients who would be eligible for treatment with GO + DA in clinical practice.

The ALFA-0701 trial included patients with AML, regardless of CD33-positivity, therefore, patients in the trial who were not CD33-positive would not be eligible for GO + DA in clinical practice under the anticipated licence.

The CS also included an individual patient data (IPD) meta-analysis as supporting evidence in Appendix D.3.1 of the CS. Patients included in the IPD meta-analysis were aged 15 years or older with newly diagnosed AML (either de novo or secondary), or high-risk myelodysplastic syndrome (MDS), which is a broader population than that defined in the decision problem or the anticipated licence. The number of MDS patients corresponds to [redacted] of the all trial meta-analysis sample.

3.2 Intervention

The intervention specified in the NICE scope is GO in combination with chemotherapy. The intervention addressed in the CS is GO in combination with daunorubicin plus cytarabine (DA; intensive chemotherapy). GO does not currently have EMA marketing authorisation for the indication in this submission. On 22 February 2018 the CHMP adopted a positive opinion, recommending the granting of a marketing authorisation for GO in combination with DA.[23] The anticipated date of EMA approval is in Q2 2018.

The CS states that GO is administered as an intravenous infusion in combination with DA. The recommended dose is 3 mg/m^2/dose (up to a maximum of 5 mg/dose) infused over 2 hours on days 1, 4 and 7 of the induction therapy course and as a 3 mg/m^2/dose (up to a maximum of 5 mg/dose) infused over 2 hours on day 1 of consolidation therapy, for up to two cycles. The company states that this is in line with the anticipated marketing authorisation. This dose was used in the ALFA-0701

terms of data assessment, although further treatment decisions made in practice were based on the unblinded investigator analysis. However, results were broadly similar between IRC and IA, as shown in Table 13 of the CS. Attrition was similar between treatment arms and there is no evidence that additional outcomes were measured and not reported.

Overall the ERG believes the trial was well conducted and has a low risk of bias, up to the limits of its open-label design.

4.2.4 Summary of the results of the ALFA-0701 trial

Throughout this section, we focus on the results from the mITT population using the data cut-off of 30 April 2013. The April 2013 data-cut is preferred as additional outcomes are captured and therefore the estimate produced is more conservative than that produced using the earlier 1 August 2011 cut-off. Data using the 1 August 2011 cut-off are summarised in Appendix D.2.3 of the CS.

The CS reported both results according to the IRC assessment, and the assessment made by the original investigators (IA). In this section we focus on the IRC assessment, as these results are less likely to be influenced by lack of blinding, and were similar, but generally slightly more conservative than those of the IA. All tables in this section report the results of the IRC analysis.

4.2.4.1 Efficacy results

Key clinical effectiveness results are presented in Section B.2.6 of the CS. The overall summary of efficacy end-points in ALFA-0701 are presented in **Error! Reference source not found.**.

[redacted] **Error! Reference source not found.** The median EFS in the GO + DA arm was [redacted] compared to [redacted] in the DA arm. Relapse free survival (RFS) also showed a statistically significant improvement, with a median RFS in the GO + DA arm of [redacted] compared with [redacted] in the DA arm.

Results for OS and overall response rate (CR + CRp) favoured the GO + DA arm, but the difference between treatment groups was not statistically significant. Patient reported outcomes and health related quality of life (HRQoL) were not assessed in the ALFA-0701 trial. [redacted] patients in the DA arm received GO as follow-up therapy through a compassionate use programme, which may have confounded overall survival results. Similarly, most patients ([redacted] in GO + DA arm and

[redacted] in DA arm) received at least one follow-up therapy for AML, which may have confounded overall surviva

[redacted] 1

[redacted]

[redacted]

4.2.4.2 Analysis by cytogenetic status

The main analysis reported in Section 4.2.4.1 above includes all patients regardless of cytogenetic status. The ERG considers that to properly understand the efficacy of GO some further breakdown by cytogenetic status is required, particularly in view of the company's decision problem, specifying the use of GO for patients not known to have unfavourable cytogenetics.

Results for patients with favourable or intermediate cytogenetic profiles (excluding unknown cytogenetics) were reported in Table 14 of the CS, and are summarised in [redacted]1 below.

[redacted]1

	GO + DA arm (n = 94)	DA arm (n = 95)	Point estimate (95% CI)
EFS, months, median (95% CI)	[redacted]	[redacted]	[redacted]
RFS, months, median (95% CI)	[redacted]	[redacted]	[redacted]
OS months, median (95% CI)	[redacted]	[redacted]	[redacted]
Overall response rate (CR/CRp), n (%) [IRC analysis data]	[redacted]	[redacted]	[redacted]

EFS and RFS were statistically significantly improved to a similar extent as the overall population in the favourable and intermediate cytogenetic subgroup, whilst OS and overall response rate favoured the GO + DA arm, but the difference between groups was not statistically significant, consistent with the overall population results.

The CS also reports results for those patients with unfavourable cytogenetics, summarised below in ***[redacted]2. There was no benefit of GO in these patients for the outcomes of EFS, RFS or OS.

[redacted] 2 [redacted]

	GO + DA arm ([redacted])	DA arm ([redacted])	Point estimate (95% CI)
EFS, months, median (95% CI)	[redacted]	[redacted]	[redacted]
RFS, months, median (95% CI) [redacted]	[redacted]	[redacted]	[redacted]
OS, months, median (95% CI)	[redacted]	[redacted]	[redacted]

[redacted] In response to the ERG's request for further clarification, the company provided additional IRC assessment results for the outcomes EFS, RFS and OS, for each of: favourable, intermediate and unknown cytogenetic risk groups. The results are summarised below in

[redacted] 3, [redacted] 4 and [redacted] 5, respectively. These results are broadly consistent with those for the combined favourable/intermediate group ([redacted] 1).

[redacted]

[redacted] 3 [redacted]

Cytogenetic profile	GO + DA arm	DA arm	Point estimate (95% CI)
Favourable/intermediate + unknown [redacted]	[redacted]	[redacted]	[redacted]
Favourable [redacted]	[redacted]	[redacted]	[redacted]
Intermediate [redacted]	[redacted]	[redacted]	[redacted]
Unknown [redacted]	[redacted]	[redacted]	[redacted]

[redacted]

Cytogenetic profile	GO + DA arm	DA arm	Point estimate (95% CI)
Favourable/intermediate + unknown [redacted]	[redacted]	[redacted]	[redacted]
Favourable [redacted]	[redacted]	[redacted]	[redacted]
Intermediate [redacted]	[redacted]	[redacted]	[redacted]
Unknown [redacted]	[redacted]	[redacted]	[redacted]

[redacted]

Cytogenetic profile	GO + DA arm	DA arm	Point estimate (95% CI)
Favourable/intermediate + unknown [redacted]	[redacted]	[redacted]	[redacted]
Favourable [redacted]	[redacted]	[redacted]	[redacted]
Intermediate [redacted]	[redacted]	[redacted]	[redacted]
Unknown [redacted]	[redacted]	[redacted]	[redacted]

4.2.4.3 Analysis by cytogenetic and molecular status

The company also provided additional IRC assessment results, in response to the ERG's request for further clarification, for patient subgroups based on cytogenetic and molecular risk categories, with the intermediate group divided into "intermediate-1" and "intermediate-2". These classifications were based on ELN guidelines.

The company provided data for the following groups: favourable and intermediate-1; intermediate-2 and unfavourable; intermediate-1 only; intermediate-2 only. These were provided for the outcomes EFS, RFS and OS, summarised below in [redacted], [redacted] and [redacted], respectively.

[redacted]

[redacted] 6

Cytogenetic/molecular profile	**GO + DA arm**	**DA arm**	**Point estimate (95% CI)**
Intermediate-1 [redacted]	[redacted]	[redacted]	[redacted]
Intermediate-2 [redacted]	[redacted]	[redacted]	[redacted]
Favourable and intermediate-1 [redacted]	[redacted]	[redacted]	[redacted]
Intermediate-2 and unfavourable [redacted]	[redacted]	[redacted]	[redacted]

[redacted] 7

Cytogenetic/molecular profile	**GO + DA arm**	**DA arm**	**Point estimate (95% CI)**
Intermediate-1 [redacted]	[redacted]	[redacted]	[redacted]
Intermediate-2 [redacted]	[redacted]	[redacted]	[redacted]
Favourable and intermediate-1 [redacted]	[redacted]	[redacted]	[redacted]
Intermediate-2 and unfavourable [redacted]	[redacted]	[redacted]	[redacted]

[redacted] 8

Cytogenetic/molecular profile	**GO + DA arm**	**DA arm**	**Point estimate (95% CI)**
Intermediate-1 [redacted]	[redacted]	[redacted]	[redacted]
Intermediate-2 [redacted]	[redacted]	[redacted]	[redacted]
Favourable and intermediate-1 [redacted]	[redacted]	[redacted]	[redacted]
Intermediate-2 and unfavourable [redacted]	[redacted]	[redacted]	[redacted]

4.2.4.4 Molecular status and other subgroup analyses

Further to cytogenetic profiles, subgroup analysis was also presented based on molecular testing. Molecular testing risk stratifies patients based on specific gene mutations. Current standard of care combines cytogenetic results with targeted testing for mutations in *FLT3*, *NPM1*, *CEBPA* and *KIT* to determine the prognostic subgroup.[27] The following subgroups were presented: FLT3_ITD, NPM1 status, MLL, WT1, risk based on the NCCN classification, risk based on the ELN classification, age, ECOG performance status and CD33 expression. These results were summarised in forest plots in the CS as Figures 30 and 31. However, Figure 30 (EFS) presented data from the 2011 data-cut, therefore, the ERG requested the figure to be updated using the April 2013 data cut; these forest plots are presented below as Figure 5 and Figure 6 (EFS) and **Error! Reference source not found.** 7 (OS).

[redacted]

In response to the ERG's request for further clarification, the company also provided subgroup analysis results for white blood cell count <30 x 10^9/L and ≥30 x 10^9/L, for the outcome overall response rate (CR/CRp). [redacted]

[redacted]

[redacted], shown in Figure 29 of the CS. The CS states that in subgroup analyses, [redacted] Time to relapse was significantly prolonged in the GO arm compared with the no-GO arm [redacted].

Response rate

Overall response was achieved in [redacted] patients in the GO arm and [redacted] patients in the no-GO arm. The odds of achieving a response were improved with GO, but the result was not quite statistically significant [redacted] CR was achieved in [redacted] patients in the GO arm and [redacted] patients in the no-GO arm. The odds of achieving a CR were improved with GO, but again the result was not [redacted]

In summary, the IPD meta-analysis demonstrated that the addition of GO to induction therapy improved OS, EFS and RFS. The addition of GO did not result in a statistically significant improvement in response rate [redacted]

4.3.2 Published meta-analyses identified in the literature review

The CS states that four published meta-analyses that investigated the effect of adding GO to induction therapy in AML were identified by a systematic literature review, although no details of the methods of the systematic review were reported, so they have not been verified by the ERG. The four meta-analyses described in Section D.3.2 of the CS appendices were by Hills et al.,[29] Li et al.,[30] Loke et al.[31] and Kharfan-Dabaja et al.[32]

The meta-analysis by Hills et al. included five RCTs and a total of 3325 patients; ALFA-0701, MRC AML15, NCRI AML16, GOELAMS AML 2006 IR and SWOG S0106. The data collection cut-off date was 2011 for the ALFA-0701 trial and up to 2013 for the other trials. Overall survival and RFS were statistically significantly improved in the GO arm compared with the no-GO arm. In subgroup analyses overall survival was significantly improved in the GO arm for patients with favourable cytogenetic risk and patients with intermediate cytogenetic risk, but not for patients with adverse cytogenetic risk. The addition of GO to induction chemotherapy did not significantly improve the rate of CR.[29]

The meta-analysis by Li et al. included five RCTs and a total of 3596 patients: ALFA-0701, MRC AML15, SWOG S0106, NCRI/University Hospital of Wales AML16 and Leukaemia Research Fund AML14 and NCRI AML16. Median follow-up ranged from 20 months for the ALFA-0701 trial to 49

months for the SWOG S0106 trial. Overall survival, RFS and relapse rate were statistically significantly improved in the GO arm compared with the no-GO arm. However, in subgroup analyses the improvement in overall survival was only statistically significant in patients with favourable cytogenetic risk, the improvement in the intermediate cytogenetic risk group did not reach statistical significance and patients in the unfavourable group did not experience a survival benefit. The addition of GO to induction chemotherapy did not significantly improve the rate of CR.[30]

The meta-analysis by Loke et al. included eleven RCTs of GO in patients with AML, with a total of 7138 patients, although the review was not restricted to trials using GO as part of induction therapy. Overall survival was not statistically significantly different between the GO and no-GO arms, even when only the seven RCTs of induction therapy were combined (ALFA-0701, MRC AML15, NCRI AML16, GOELAMS AML 2006 IR, SWOG S0106, EORTC-Gruppo Italiano Malattie Ematologiche dell'Adulto [GIMEMA] AML17 and Children's Oncology Group trial AML0531). However, in subgroup analyses overall survival was improved in the GO arm for patients with favourable cytogenetic risk (HR: 0.46, 95% CI: 0.29-0.73), but not for patients with intermediate or adverse cytogenetics. Rate of resistant disease (HR: 0.77, 95% CI: 0.67-0.9) and cumulative incidence of relapse (HR: 0.86, 95% CI: 0.79-0.93) were reduced in the GO arm. The addition of GO to induction chemotherapy did not improve the rate of CR and there was no evidence that CD33 positivity influenced OS in patients in the GO arm.[31]

The meta-analysis by Kharfan-Dabaja et al. included seven RCTs comparing GO plus conventional chemotherapy with conventional chemotherapy alone in patients with newly diagnosed AML, with a total of 3943 patients: ALFA-0701, MRC AML15, NCRI AML16, GOELAMS AML 2006 IR, SWOG S0106, EORTC-GIMEMA AML17 and the German AML Intergroup trial. Overall survival was statistically significantly improved in the GO arm for the low/intermediate cytogenetic risk group (HR: 0.83, 95% CI: 0.70-0.99), but not for the high risk group or overall group. RFS (reported in six trials) and EFS (reported in three trials) were also statistically significantly improved in the GO arm. The addition of GO to induction chemotherapy did not improve the rate of CR.[32]

In summary, the meta-analyses demonstrated that the addition of GO to induction therapy significantly improved RFS. Overall survival was improved in patients with favourable cytogenetics (results were inconsistent between meta-analyses for the overall population and also the intermediate risk group). The addition of GO did not improve the rate of CR in any of the meta-analyses. The CS emphasises the strengths of the IPD meta-analysis described in Section 4.3.1, compared with the summary meta-analyses presented in this section.

4.4 Conclusions of the clinical effectiveness section

The CS evaluation of GO was primarily based on one reasonably good quality RCT; the ALFA-0701 trial, which compared GO + DA versus DA alone in patients aged 50-70 with previously untreated de novo AML. The anticipated marketing authorisation is patients aged 15 years or over with CD33-positive AML. Therefore, the age range in the trial was narrower than the anticipated marketing authorisation, although the majority of patients with AML are over the age of 50 and the ERG's clinical advisor did not consider that GO would work differently in patients under the age of 50 to those over the age of 50. The restriction to patients with CD33-positive AML appears clinically appropriate, in view of the mechanism of action of GO and subgroup analysis results, in which there was some trend towards GO being more effective based on a 30% and 70% CD33-positivity cut-off.

The ALFA-0701 trial demonstrated that GO + DA was effective at improving EFS and RFS, compared with DA alone, in the overall patient population ([redacted]). Whilst the response rate appeared better in the GO + DA arm ([redacted]), the result did not reach statistical significance, suggesting that duration of remission is extended with GO, rather than the proportion of patients achieving a remission. Overall survival appeared to be better in patients who received GO + DA, although the difference between treatment groups did not reach statistical significance ([redacted]). This lack of significance may have occurred because some patients received follow-up treatments, with some patients in the DA arm receiving GO as follow-up therapy.

EFS and RFS results were statistically significantly improved in the GO + DA arm for the subgroup of patients with favourable/intermediate cytogenetic risk, to a similar extent as the overall population, with OS and response rate also improved in the GO + DA arm, but not reaching statistical significance, which was consistent with the overall results. However, for patients with an unfavourable cytogenetic profile, OS and RFS outcomes appeared to be worse in the GO + DA arm, compared with the DA arm, whilst EFS results were similar ([redacted]).

Additional analyses provided to the ERG on request showed that the benefit seen in patients with an intermediate-1 cytogenetic and molecular risk profile was not found in patients with an intermediate-2

cytogenetic and molecular risk profile, suggesting potentially important heterogeneity in the broader 'intermediate' cytogenetic subgroup.

Whilst the proportion of patients experiencing an AE was similar between treatment groups, the proportion of patients reporting a SAE was higher in the GO + DA arm than the DA arm [redacted]. The most common SAE in the GO + DA arm was thrombocytopenia [redacted]. All [redacted] patients who experienced VOD had received GO. This included [redacted] patients in the DA arm who received GO during follow-up as part of the compassionate use program. They experienced VOD more than 28 days after receiving their last dose of GO ([redacted]). A higher proportion of patients in the GO + DA arm permanently discontinued treatment because of an adverse event than in the DA arm [redacted].

The CS also presented the results of an IPD meta-analysis, conducted by Pfizer for use in regulatory submissions.[28] Overall survival was statistically significantly improved in the GO + DA arm for the overall population in the IPD meta-analysis [redacted] and for the subpopulation of patients with a favourable or intermediate cytogenetic profile [redacted], but not for patients with an unfavourable cytogenetic profile [redacted], supporting the evidence from the ALFA-0701 trial. However, the IPD meta-analysis included patients with de novo or secondary AML or high-risk myelodysplastic syndrome, which is a broader population than the anticipated licence, therefore, results may not be entirely generalisable to patients eligible for GO in clinical practice.

The ALFA-0701 trial and the IPD meta-analysis did not report quality of life outcomes.

	Approach	Source / Justification	Signpost (location in company submission)
Treatment effectiveness	Clinical outcomes included were response (CR/CRp), RFS and OS, cure fraction, probability of HSCT, HSCT cure fraction. These data were taken from the ALFA-0701 study. OS was stratified by response status. Parametric models were fitted to RFS (CR/CRp only) and OS (stratified by response status) to extrapolate beyond the end of trial follow-up. Response and RFS endpoints were based on the blinded IRC assessment. RFS and OS analyses were based on the reference data of 30 April 2013.	Data from a subgroup (patients not known to have unfavourable cytogenetics) of the ALFA-0701 study were used to inform clinical inputs of the model. The ALFA-0701 study is the only RCT that has compared GO+DA using the licensed fractionated dosing regimen. The blinded IRC analyses were chosen to address any possible bias by local investigators due to the open-label design. These were stated by the company to be the outcomes considered most appropriate by EMA.	Section B.3.3 pg. 103 to 122
Adverse events	Grade 3 and 4 treatment-related adverse events that occurred in at least 1% of patients were included. GVHD as a consequence of HSCT was also included. .	Adverse event rates were taken from the ALFA-0701 trial. Incidence of GVHD was sourced from external literature; Battipaglia (2017)[36] and NHS England (2017)[37].	Section B.3.3.10 pg. 122 to 123
Health related quality of life	No HRQoL data was collected in ALFA-0701 and health state utilities were sourced from a systematic literature review. A vignette study undertaken by the company provided two alternate sets of utility values that were used in a scenario analysis.	Functionally cured patients were assumed to have quality of life equal to that of the age-matched general population (Ara & Brazier) [38]. The remaining health state utilities were sourced from NICE TA399 [35], except for post-HSCT CR/CRp with GVHD which was sourced from Kurosawa (2016) {Kurosawa, 2016 #145}. All adverse event disutilities were sourced from NICE TA399 [35], except for VOD which was sourced from the appraisal of defibrotide by the SMC [39].	Section B.3.4.3 pg. 124 Section B.3.4.4 pg. 124 to 127

	Approach	Source / Justification	Signpost (location in company submission)
Resource utilisation and costs	These comprised: drug acquisition, health state costs (including inpatient and outpatient attendances, disease monitoring and management, and terminal care costs), HSCT and GVHD treatment costs, and treatment of adverse effects	Drug acquisition costs for GO were based on the confidential list price. Drug acquisition costs for DA and subsequent lines of chemotherapy were sourced from BNF [41] and eMit [42]. Unit costs for health state and adverse events were taken from NHS reference costs (2016) [43] and PSSRU (2016) [44]. The treatment cost of VOD also includes data sourced from NHS England Commissioning reports [45]. Terminal care costs were sourced from Addicott & Dewar (2008) [46]. HSCT costs were sourced from a NHS Blood and Transplant analysis (2014) [47], and the GVHD cost from Esperou (2004) [48]. Resource use items were based on those used in ALFA-0701 and elicited clinical expert opinion.	Section B.3.5 pg.127 to 140
Discount rates	Costs and benefits were discounted at 3.5% per annum	In accordance with the NICE reference case.	Section B.3.2.2 pg. 101
Sensitivity analysis	Probabilistic sensitivity analysis was performed. Deterministic univariate sensitivity analysis was performed on a series of model parameters. A series of scenario analyses were also performed.	In accordance with the NICE reference case.	Section B.3.8 pg. 143 to 150

5.2.1 Model structure

The CS presented a *de novo* semi-Markov cohort state-transition model to estimate the cost-effectiveness of GO+DA compared with DA alone in a subgroup of the licensed population: adult de-novo AML patients not known to have unfavourable cytogenetics.

Table 9 Model health states (adapted from CS Appendix M.1, Table 130)

Health state	Description
Induction therapy	• Initial period of treatment with GO + DA or DA alone prior to determination of response status • Duration of up to two cycles
CR/CRp (consolidation)	• Consolidation treatment for patients who attain CR or CRp following induction therapy • Duration up to two cycles
CR/CRp (off-treatment)	• Period when CR or CRp patients have stopped treatment after completing induction or consolidation therapy and have regular follow-up visits • Duration: patients remain in this health state until they experience relapse or death, have an HSCT, or become 'functionally cured'
Refractory (salvage therapy)	• Refractory patients (those who failed induction therapy) who are receiving treatment with high-intensity salvage therapy • Duration of 1.5 cycles in base-case model
Refractory (non-curative therapy)	• Refractory patients on non-curative therapy, who are not eligible for salvage therapy • Refractory patients on non-curative therapy (BSC only) who failed salvage therapy in the previous cycles • Duration: until death or receive HSCT
HSCT	• Period of HSCT procedure and recovery when patients remain hospitalized • Duration: one cycle
Post-HSCT CR/CRp (without GVHD)	• Period after HSCT procedure and prior to becoming 'functionally cured', without GVHD (based on defined period of time) • Duration: In health state until functionally cured (after 60 months) or death
Post-HSCT CR/CRp (with GVHD)	• Period after HSCT procedure and prior to becoming 'functionally cured', with GVHD • Duration: in health state for pre-specified amount of time, reflecting duration of GVHD
Functionally cured	• Long-term disease-free survival (CR/CRp) with no planned follow-up • Duration: in health state until death, as modelled with adjusted general population mortality
Relapse (salvage therapy)	• Treatment with high-intensity salvage therapy for patients with relapse following an initial CR/CRp with induction therapy • Duration of 1.5 cycles in base-case model
Relapse (non-curative therapy)	• Patients with relapse on non-curative therapy, who are not eligible for salvage therapy • Patients with relapse on non-curative therapy (BSC only) who failed salvage therapy in the previous cycles • Duration: until death or receive HSCT
Dead	• Dead

GO: gemtuzumab ozogamicin; DA: daunorubicin and cytarabine; AML: acute myeloid leukaemia; CR: complete remission; CRp: complete remission with incomplete platelet recovery; BSC: best supportive care; HSCT: haematopoietic stem cell transplant; GVHD: graft versus host disease

Relapse (salvage therapy)

Patients who initially achieved CR/CRp were allowed to transition to the relapse state over time (salvage and non-curative therapies). Patients transitioning to relapse (salvage therapy state) were assumed to receive 1.5 cycles of salvage therapy.

Relapse (non-curative therapy)

Relapsed patients who did not receive salvage therapy transitioned to the relapse (non-curative therapy state), consisting of treatment with either low-intensity chemotherapy or BSC. Relapsed patients could also enter this health state after failure of salvage therapy to bridge to HSCT. Patients remained in this health state until either HSCT or death.

Response was not explicitly modelled in either of the relapse states. However, HSCTs were assumed to have been undertaken in patients who received prior salvage therapy and had achieved CR/CRp.

HSCT

Patients could receive HSCT from a number of health states at different (fixed) time points, including from "CR/CRp off-treatment" at six months, from "refractory" at [redacted] and from "relapse" at [redacted], [redacted], [redacted], [redacted] and [redacted]. For the CR/CRp and refractory states, the time points were based on the average time to HSCT. For relapse, the mid-point of each year was assumed for years 2 to 5. These timings were based on calendar time (i.e. from the point of randomisation) and not time in state. The use of calendar time implies that any relapses which occur after [redacted] would not receive HSCT.

Patients entering this health state from the refractory and relapse health states were assumed to have received prior salvage therapy and achieved CR/CRp. Patients remained in the HSCT state for one cycle, corresponding to the time that the company assumed that patients would spend hospitalised in an isolation room. Following this cycle, patients transitioned to one of the post-HSCT states (with or without GVHD) or death.

HSCT: CR/CRp (with GVHD)

Separate sub-states within the main HSCT state were used to capture the incidence and duration of acute and chronic GVHD.

A proportion of patients experienced GVHD (acute or chronic) after HSCT. The company stated that acute GVHD was assumed to last 2.5 months. Chronic GVHD was assumed to occur 6 months after

HSCT and last for 9 months. Incidence rates of chronic and acute GVHD were applied to the number of patients surviving HSCT to estimate patients in this health state at the time points described.

Following the respective periods of GVHD, patients remained in the main HSCT health until transitioning to death (for period up to 60 months after HSCT). Patients remaining in this health state at 60 months after HSCT transitioned to the functionally cured health state.

HSCT: CR/CRp (without GVHD)

Patients who received HSCT and did not experience either chronic or acute GVHD, entered this state until transitioning to death (for period up to 60 months after HSCT). Patients remaining in this health state at 60 months after HSCT transitioned to the functionally cured health state.

Functionally cured

The CS reported that clinicians considered patients to be "functionally cured" (i.e. have long-term disease-free survival) in clinical practice after remaining in complete remission for 3 to 5 years. In the model, CR/CRp patients could enter the functionally cured health state if they remained alive and relapse-free at 5-years. Patients from other states could enter if they remained alive at 5-years after receiving an HSCT as a refractory, relapse or CR/CRp patient.

Patients in the CR/CRp health state after five years transitioned to the functionally cured health state, and were assumed to be no longer at risk of mortality due to AML. However, longer-term excess mortality due other causes (e.g. higher risk of secondary cancers, cardiac events, etc.) was assumed for the remainder of the model time horizon by applying a hazard ratio (HR) to general population mortality rates.

ERG comment

The proposed model structure is complex and was challenging to critique given the difficulties in determining the actual flow of patients through the model. As part of the clarification questions, the company was requested to provide a clearer description of the assumptions and to explain the advantage of the state-transition model compared to a simpler and more conventional partitioned survival analysis (PartSA) model [50].

The company responded that the key benefit of GO was improved RFS (as opposed to differences in the initial induction success rates) and that patients receiving GO remained in CR/CRp for longer, delaying or avoiding subsequent relapses. The company noted that the proposed structure allowed for: separate survival analyses to be undertaken based on induction success or failure (i.e. CR/CRp or refractory); differentiation between relapse and refractory patients as well as capturing the impact of

impact of age is likely to be negligible in terms of the ICER results and for CD33 positive patients the base-case ICER may be slightly conservative.

As previously highlighted in Section **Error! Reference source not found.**, the company's decision problem population further excluded patients known to have unfavourable cytogenetics from within the broader marketing authorisation. The CS justified this restriction based on the subgroup results of the ALFA-0701 study and clinical advice that these patients would not be treated with GO plus intensive chemotherapy in clinical practice. The clinical advisor to the ERG also agreed with this view. Patients with unfavourable cytogenetics constituted around 21.0% of the total ALFA-0701 trial population (Table 10).

Table 10 Cytogenetic risk stratification of patients in ALFA-0701

Risk stratification	Favourable	Intermediate	Unfavourable	Unknown
GO+DA group	2.2%	67.4%	20.0%	10.4%
DA group	4.4%	65.4%	22.1%	8.1%
GO, gemtuzumab ozogamicin; DA, daunorubicin and cytarabine				

The main baseline characteristics of the population the company presented as their base-case analysis is summarised in Table 11.

Table 11 Mean baseline characteristics (population excluding known unfavourable cytogenetics)

Characteristic	GO + DA (N=108)	DA (N=106)	Total (N=214)
Age (years)	[redacted]	[redacted]	[redacted]
Gender (% male)	[redacted]	[redacted]	[redacted]
Body surface area (m^2)	[redacted]	[redacted]	[redacted]
Weight (kg)	[redacted]	[redacted]	[redacted]
GO, gemtuzumab ozogamicin; DA, daunorubicin and cytarabine Model based on mean values from the combined arms			

The ERG agrees with the company's decision to exclude patients with known unfavourable cytogenetics and their rationale for focusing on the specific subpopulation where GO+DA provides clear clinical benefit and optimises cost-effectiveness.

The subgroup results requested by the ERG and summarised in Section 4.2.4.3, indicate there may be other subgroups (e.g. intermediate-2 patients) where clinical benefit also appears to be unclear and hence further optimisation in cost-effectiveness may be appropriate. The ERG does not consider that

5.2.5 Perspective, time horizon and discounting

The economic model adopted a National Health Service (NHS) and Personal Social Services (PSS) perspective in accordance with the NICE reference case.

The time horizon used in the economic model was 40 years. The CS stated that this was sufficient to capture lifetime costs and benefits. The costs and benefits in the model were discounted at an annual rate of 3.5%, as per the NICE reference case.

Implementation of discounting in the economic model was carried out on an annual basis, such that all costs and benefits incurred with any given year are discounted by the same amount regardless of whether they occur at the start or the end of that year.

ERG comment

A 40-year time-horizon appears to be appropriate based on the average age assumed and the potential curative assumptions employed. The average age (and distribution) in the model is based on patients in the ALFA-0701 trial ([redacted]) such that the probability of patients still alive at 40 years is considered sufficiently small to represent a lifetime horizon. The ERG acknowledges that a longer horizon would be required for younger patients.

The ERG considers that discounting on a per cycle basis is more accurate as it more closely reflects the actual time at which benefits and costs occur. The impact of this issue on the ICER is, however, likely to be minimal and therefore is not explored further as part of ERG exploratory analysis.

5.2.6 Treatment effectiveness and extrapolation

The CS provided a description of the clinical data used in the model. These were based on ALFA-0701, and included: response to first-line treatment, overall survival, relapse-free survival, probability of HSCT, the cure rate of HSCT, and the incidence rate of treatment-emergent adverse events. Additionally, the model estimated mortality in the functionally cured population, which was based on general population mortality [51] and adjusted to account for excess mortality in AML survivors.

5.2.6.1 Response to first-line treatment

Response to first-line treatment was modelled as the probability of achieving CR or CRp after induction therapy. Response data were pooled across treatment arms and justified in the CS based on the lack of any statistically significant difference and clinical advice that GO was not expected to affect the initial response outcomes. The main effect of GO was therefore assumed to be in terms of the additional durability of the response outcomes.

using the generalised gamma distribution may be due to difficulties in achieving convergence. However, the clinical plausibility of the higher cure fraction estimated for OS compared to EFS was not discussed in the CS. The difference in the cure fractions for EFS and OS suggests that either: (i) there are a significant number of patients who become functionally cured following relapse (potentially due to subsequent therapies and/or HSCT) or (ii) the data may not be sufficiently mature to robustly estimate the cure fraction for OS.

Despite the potential concerns regarding the difference in the absolute cure fractions reported for EFS and OS, the ERG notes that the differences in the cure fraction between the groups is similar for both endpoints. This provides additional reassurance regarding the robustness of the ICER results. In addition, the proportion of patients that the model predicts are relapsers who do not receive a HSCT but enter the functionally cured state at 5 years is small and balanced ([redacted] for GO and [redacted] for GO+DA)**.** As previously stated, it is the difference in the cure rates between the groups (as opposed to the absolute cure rates) this is most critical for determining the appropriateness of the ICER results.

For the base-case population, the ERG considers that the choice of survival function appears less critical than the assumptions which are subsequently applied to long-term survivors regarding potential excess morbidity (i.e. HRQoL assumptions) and mortality. However, the ERG also notes that there remains significant heterogeneity in outcomes within the base-case population which hasn't been fully explored in the CS. These issues are further explored in the following sections and additional ERG exploratory analyses are also provided in Section 6.

5.2.6.4 Mortality in the cured population

To capture the excess mortality (relative to the general population) for functionally cured patients at five years, the company applied a hazard ratio of [redacted] to the general population mortality rates. The company argued that functionally cured patients remained at higher risk of other health conditions which may increase mortality rates above that of the general population, including secondary or relapsed cancer, late complications following an HSCT, or cardiovascular disease following an anthracycline (such as daunorubicin and idarubicin).

The company undertook an analysis of pooled survival data from UK AML trials 10 to 16, restricted de novo AML patients to the intermediate and favourable cytogenetic subgroup (a total of [redacted] intermediate and [redacted] favourable patients) aged 50 to 70, using survival curves conditional on surviving the first five years. The hazard ratio was estimated by calculating the ratio of the means of the annual mortality rates, from five years after AML diagnosis for AML patients and of those matched to the mean age of the analysis from the general population. .

Table 12 Summary of health state utility values

Source	Chemotherapy treatment[1]	Consolidation therapy	HSCT	GVHD (post-HSCT)	CR/CRp off-treatment	Relapse	Refractory	Functionally cured
Values used in base–case analysis								
TA399	0.66	0.66	0.66	0.67[3]	0.74	0.57	0.57	0.82[2]
Values used in scenario analysis								
TA399	0.72	0.72	0.72	0.67[3]	0.77	0.62	0.62	0.82[2]
Pfizer TTO	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
Pfizer VAS	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]

[1] Applied to patients in induction, salvage and non-curative health states
[2] Varied per cycle, based on mean patient age at each time point, from Ara & Brazier
[3] Source Kurosawa 2016 {Kurosawa, 2016 #145}

The utility values assigned to several states were based on a previous NICE technology appraisal for azacitidine (TA399) [35] and clinical expert opinion was used to map these values to health states in the model where health state descriptions were not aligned. These utility values were calculated from EORTC QLQ-C30 data collected in the azacitidine pivotal trial, which enrolled patients over the age of 65 with *de novo* or secondary AML with >30% bone marrow blasts who were not eligible for HSCT. Two mapping algorithms were identified and used in TA399: the company in the present appraisal elected to use the values estimated with the McKenzie & Van der Pol algorithm that provided utility values closer to the values in their preference elicitation study. A company-presented scenario analysis with the alternative set of utility values resulted in very little change to the ICER.

The company made a number of assumptions in order to map utility values from TA399 to the health states in their analysis. Quality of life for patients on intensive chemotherapy (including induction therapy, consolidation therapy and salvage therapy) was assumed to be 0.657. In TA399, this was the value estimated from patients in non-remission (stable disease or partial remission). Differences between arms for patients on GO+DA and on DA were captured by estimating the utility decrement associated with the different safety profiles.

The company also assumed that patients in the month after having an HSCT would have a similar quality of life to those on intensive chemotherapy, and the same utility value (0.657) was applied in this health state.

Patients in remission (CR or CRp) and off-treatment had a utility value of 0.740, mapped from the health state utility value for remission patients (CR or CRi) in TA399. The utility value for relapsed patients and refractory patients receiving non-curative treatment (including either best supportive care

or low-intensity chemotherapy) was assumed to be 0.568, estimated from patients in TA399 who were post-progression or who had relapsed.

Patients in the functionally cured health state were assumed to have quality of life consistent with patients of the same age in the general population. Age- and gender-matched utilities were estimated from a formula estimated by Ara & Brazier, who used EQ-5D data and UK preference rates to determine quality of life according to a number of patient covariates [38]. These values were estimated on a per-cycle basis to allow the model to capture the gradual decline in HRQoL associated with ageing. Using the formula for a patient aged 61 results in a utility value of 0.820.

The utility value for patients with GVHD following HSCT was 0.67, and was taken from a published economic analysis of HSCT patients in first remission from AML {Kurosawa, 2016 #145} which was identified in the SLR. The economic analysis did not differentiate between chronic and acute GVHD.

The company also conducted a separate preference elicitation study, which recruited 125 participants from the general population. A series of vignettes were developed, corresponding to each health state in the analysis. Each participant had a one-to-one face-to-face interview and was asked to value a range of health state descriptions using the time-trade-off and visual analogue techniques. However, the company chose not to use the resulting values from this study in their base-case analysis, stating that utilities from the studies identified in the review more closely aligned with the NICE reference case. A separate sensitivity analysis was conducted using the values estimated from the elicitation study.

ERG comment

In the absence of direct HRQoL data available in ALFA-0701, the ERG considered the approach used by the company to be reasonable and appropriately justified. The only exception to this was the separate assumption made that functionally cured patients experience the same HRQoL as the general population. This assumption results in a marked jump in the HRQoL estimates at 5-years for functionally cured patients. The use of general population quality of life was not considered internally consistent with the excess mortality applied for functionally cured patients to OS. Given that functionally cured patients are assumed to be at higher mortality risk than the general population, the ERG concluded that it would appear reasonable to assume that functionally cured patients would also have lower quality of life than that of the general population. Alternative assumptions have been explored by the ERG in Section **Error! Reference source not found.**.

These concerns are explored further in Section **Error! Reference source not found.**.

The assumption to not include drug wastage and that the administration costs would be captured within the health state costs were both considered reasonable. While the inclusion of drug wastage is a more accurate assumption, the company's analysis provides a more conservative estimate of cost-effectiveness, as the costs of DA are increased proportionally more than those of GO when wastage is taken into account.

5.2.7.2 Costs of subsequent lines of therapy

Subsequent lines of therapy in the model consisted of salvage therapy (assumed to be FLAG-Ida) and non-curative therapy (comprising hydroxycarbamide, low-dose cytarabine, and azacitidine) and best-supportive care (hydroxycarbamide).

Table 13 Drug acquisition costs of subsequent lines of therapy (CS, Table 43)

Drug	Pack price	Source
Salvage therapy		
Fludarabine (50 mg/2 mL concentrate, 1 vial)	£26.08	DoH (eMiT) (2017)
Cytarabine (2000 mg/5 mL solution, 5 vials)	£6.60	DoH (eMiT) (2017)
G-CSF (filgrastim) (30 million units/0.5 mL solution, 5 vials)	£49.30	BNF (2017)
Idarubicin (5-mg powder for solution, 1 vial)	£87.36	BNF (2017)
Non-curative therapies		
Low-dose cytarabine (100-mg vial)	£4.70	DoH (eMiT) (2017)
Hydroxycarbamide (100 capsules)	£8.83	DoH (eMiT) (2017)
Azacitidine (100-mg powder for suspension)	£321.00	BNF (2017)

The mean cost per dose and cost per cycle was estimated in the same manner as that of first line therapies, as described in the section above.

The duration of non-curative treatment was estimated from restricted mean survival time estimates (RMST) from the ALFA-0701, and adjusted for time assumed to be spent in the terminal care period (two cycles). RMST for relapsed and refractory patients who did not receive HSCT were estimated from data pooled from both arms, and was [redacted] for refractory patients and [redacted] for relapsed patients.

Probabilistic sensitivity analysis

Figure 2 presents the cost-effectiveness acceptability curve (CEAC). At a willingness-to-pay threshold of £20,000, the probability of GO + DA being cost-effective was [redacted]%. At a threshold of £30,000, the probability rose to [redacted]%.

Figure 2 Cost-effectiveness acceptability curve (CS, Fig 19, p. 145)

[redacted]

5.2.9.3 Scenario analysis

The company presented a range of scenario analysis within their base case analysis. These analyses focused on the use of alternative survival functions, health state utility weights and disutilities for adverse events, the use of data from individual arms instead of pooled data to estimate certain parameters (response status, mean RMST, HSCT probabilities OS for refractory patients, and the number of treatment courses received), and alternative values for the excess mortality of long-term AML survivors.

Across the set of scenarios exploring the alternative survival functions for RFS and OS (for CR or CRp patients), the ICER varied between £6,821 (best fitting standard parametric functions) and £12,233 (MCM Weibull) per QALY.

The ERG notes that the main drivers of the cost-effectiveness estimates are the difference in costs (e.g. initial treatment costs and subsequent HSCT) predicted over the period of follow-up of the ALFA-0701 trial and the difference in survival after the follow-up. While the alternative survival models (MCM, spline-based and conventional parametric) result in large differences in the terms of

the absolute LYG and QALY estimates, they all predict very similar between group differences. This is because the majority of functions generate similar predictions of the difference in the proportion of patients who become long-term survivors. While this is done explicitly in the MCM models using the cure fraction approach, many of the more conventional parametric functions are doing this implicitly with a shallowing of the hazard function which appears to be converging to a similar between group differences in the proportion of patients who experience long term survival. As a result, the mean difference in costs and QALYs used to estimate the ICER appear robust and relatively stable across the majority of functions.

Across all the scenarios, the ICER estimate varied between £6,821 (best fitting standard parametric functions for RFS and OS for CR/CRp) and £20,334 per QALY (when the RMST for relapse patients was based on individual treatment arms).

5.2.9.4 Subgroup analysis

In the CS appendix, the company presented cost-effectiveness results for the full indication (including those with unfavourable cytogenetic risk profile). The ICER is higher in the whole patient population (£20,457, compared with £12,251 for the favourable/intermediate and unknown cytogenetic population), since the effect of GO is lower in patients with unfavourable cytogenetics.

Table 14 Results of the analysis, all patient population (CS Appendix, Table 172)

Technologies	Total costs (£)	Total QALYs	Incremental costs (£)	Incremental QALYs	ICER incremental (£/QALY)
GO + DA	[redacted]	[redacted]	[redacted]	[redacted]	20,457
DA	[redacted]	[redacted]			

CS, company submission; ICER, incremental cost-effectiveness ratio; QALYs, quality-adjusted life years; GO, gemtuzumab ozogamicin; DA, daunorubicin + cytarabine

The mean probabilistic ICER was £21,999 (95% confidence credible interval [CrI]: [redacted] for costs; [redacted] for QALYs), and GO+DA had a [redacted]% probability of being cost-effective at a willingness-to-pay threshold of £30,000. The probability was less than [redacted]% when the willingness-to-pay threshold was £20,000 per QALY. The DSA showed that the results in this population were sensitive to the same input parameters as the cytogenetic population. The greatest increase in the ICER was observed when HSCT probabilities from relapse were pooled, providing an ICER of £30,206.

numbers precluded the company from undertaking cure fraction models for the intermediate-2 population, the results presented in Section 4 for EFS and OS suggest limited clinical benefit is evident in this subgroup.

6.6 Conclusions from ERG analyses

The ERG has presented a number of additional analyses. These analyses were carried out in a number of stages. The first stage addressed a number of minor calculation errors in the company's revised model. The impact of these changes was to increase the ICER by a small amount from £12,251 per QALY to £13,561 per QALY.

Using the corrected and updated model, the ERG then presented a number of analyses considering a range of issues raised in Section **Error! Reference source not found.**. These scenario analyses addressed the following issues:

- The number of induction and consolidation courses received;
- Rate of response to treatment;
- Treatment costs associated with HSCT and VOD;
- Inclusion of VOD events;
- Quality of life in functionally cured patients;
- Excess mortality risk in functionally cured patients.

The majority of these changes resulted in an increase to the ICER, with the exception of using individual rates of response to treatment, although the scenarios were not associated with substantial differences to the ICER. The scenarios associated with the greatest impact on cost-effectiveness outcomes related to changes made by the ERG to the HSCT costs, the quality of life in functionality cured patients and to the use of individual rates of response. This exploration of alternative modelling assumptions and parameter values was concluded with the ERG presenting a base-case with a preferred set of assumptions.

The ERG alternative base-case, based on a probabilistic analysis, estimated GO+DA to be more costly (cost difference [redacted]) and more effective ([redacted] QALY gain) compared with DA, and suggests that the ICER for GO+DA compared with DA is £17,956 per QALY.

A series of exploratory subgroup analysis were conducted by the ERG to explore the impact of heterogeneity between the subgroups included in the base-case population and possible variability within the intermediate group. These analyses explored the impact in different subgroups based on cytogenetic results alone as well as subgroups based on cytogenetic and molecular results. The ICER

35. National Institute for Health and Care Excellence. *Azacitidine for treating acute myeloid leukaemia with more than 30% bone marrow blasts*. London: NICE; 2016. Available from: https://www.nice.org.uk/guidance/ta399
36. Battipaglia G, Labopin M, Candoni A, Fanin R, El Cheikh J, Blaise D, et al. Risk of sinusoidal obstruction syndrome in allogeneic stem cell transplantation after prior gemtuzumab ozogamicin treatment: a retrospective study from the Acute Leukemia Working Party of the EBMT. *Bone Marrow Transplant* 2017;52:592-99. Available from: https://www.ncbi.nlm.nih.gov/pubmed/28092357
37. NHS England Specialised Services Clinical Reference Group for Blood and Marrow Transplantation. *Clinical Commissioning Policy: Treatments for Graft versus Host Disease (GvHD) following Haematopoietic Stem Cell Transplantation*: NHS England; 2017. Available from: https://www.england.nhs.uk/publication/treatments-for-graft-versus-host-disease-gvhd-following-haematopoietic/
38. Ara R, Brazier JE. Populating an economic model with health state utility values: moving toward better practice. *Value Health* 2010;13:509-18. Available from: https://www.ncbi.nlm.nih.gov/pubmed/20230546
39. Scottish Medicines Consortium. *defibrotide (Defitelio). Advice*. 2014. [cited 2017 7th December]. Available from: https://www.scottishmedicines.org.uk/SMC_Advice/Advice/967_14_defibrotide_Defitelio.
40. Kurosawa S, Yamaguchi H, Yamaguchi T, Fukunaga K, Yui S, Wakita S, et al. Decision Analysis of Postremission Therapy in Cytogenetically Intermediate-Risk Acute Myeloid Leukemia: The Impact of FLT3 Internal Tandem Duplication, Nucleophosmin, and CCAAT/Enhancer Binding Protein Alpha. *Biology of Blood and Marrow Transplantation*. 2016; 6(22):1125-1132.
41. *British National Formulary*. [cited 2017 5th November]. Available from: https://www.bnf.org.
42. Department of Health. *Drugs and pharmacuetical electronic market information (eMit)*. 2011. [cited 2017 6th December]. Available from: https://www.gov.uk/government/publications/drugs-and-pharmaceutical-electronic-market-information-emit.
43. Department of Health. *Reference Costs 2015-16*. London: DoH; 2016. Available from: https://www.gov.uk/government/publications/nhs-reference-costs-2015-to-2016
44. Curtis L, Burns A. *Unit Costs of Health and Social Care 2016*. Canterbury, Kent: Personal Social Services Research Unit. University of Kent; 2016. Available from: https://www.pssru.ac.uk/pub/uc/uc2016/full.pdf?uc=2016-full
45. NHS England Clinical Reference Group for Blood and Marrow Transplantation. *Clinical Commissioning Policy: Use of defibrotide in severe venoocclusive disease following stem cell transplant*. In: NHS England; 2015. Available from: https://www.england.nhs.uk/commissioning/wp-content/uploads/sites/12/2015/01/b04-use-defibrotide.pdf
46. Addicott R, Dewar S. *Improving choice at end of life: a descrptive analysis of the impact and costs of the Marie Curie Delivering Choice programme in Lincolnshire*. London: Kings Fund. Available from: https://www.kingsfund.org.uk/sites/default/files/improving-choice-end-of-life-descriptive-analysis-impact-costs-marie-curie-choice-programme-lincolnshire-rachael-addicot-steve-dewar-april-2008.pdf
47. NHS Blood and Transplant. *Unrelated donor stem cell transplantation in the UK. Effective affordable, sustainable. A report from the UK Stem Cell Strategy Oversight Committee*; 2014. Available from: http://docplayer.net/7404866-Unrelated-donor-stem-cell-transplantation-in-the-uk.html
48. Esperou Hln, Brunot A, Roudot-Thoraval Fo, Buzyn A, Dhedin N, Rio B, et al. Predicting the Costs of Allogeneic Sibling Stem-Cell Transplantation: Results from a Prospective, Multicenter, French Study. *Transplantation* 2004;77:1854-58.
49. Woods B, Sideris E, Palmer S, Latimer N, Soares M. *Partitioned survival analysis for decision modelling in health care: a critical review. 2017 [Available from http://www.nicedsu.org.uk]*. York: Centre for Health Economics, University of York and School of Health and Related Research, University of Sheffield; 2017. Available from: http://nicedsu.org.uk/wp-content/uploads/2017/06/Partitioned-Survival-Analysis-final-report.pdf