Single Technology Appraisal

Liposomal cytarabine and daunorubicin for untreated acute myeloid leukaemia [ID1225]

Committee Papers

HIGHLY CONFIDENTIAL

NATIONAL INSTITUTE FOR HEALTH AND CARE EXCELLENCE

SINGLE TECHNOLOGY APPRAISAL

Liposomal cytarabine and daunorubicin for untreated acute myeloid leukaemia [ID1225]

Final Scope and Final Matrix

1. **Pre-Meeting Briefing (PMB)**
   - Part 1

2. **Company submission summary** from Jazz Pharmaceuticals

3. **Clarification letters**
   - NICE request to the company for clarification on their submission
   - Company response to NICE's request for clarification

4. **Patient group, professional group and NHS organisation submission** from:
   - Leukaemia Care
   - NCRI-ACP-RCP
   - Royal College of Pathologists and British Society for Haematology
   - NHS England

5. **Expert personal perspectives** from:
   - Professor Nigel Russell - clinical expert, nominated by NCRI-ACP-RCP – *endorsed by Dr Steven Knapper*
   - Dr Priyanka Mehta – clinical expert, nominated by Jazz Pharmaceuticals
   - Mark Sandford – patient expert, nominated by Leukaemia Care

6. **Evidence Review Group report** prepared by Centre for Reviews and Dissemination and Centre for Health Economics - York

7. **Evidence Review Group report – factual accuracy check**

8. **Evidence Review Group report – erratum**

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

NICE National Institute for Health and Care Excellence

Liposomal cytarabine and daunorubicin (CPX-351) for untreated acute myeloid leukaemia

Pre-meeting briefing

PART 1

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

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

NICE 2

Abbreviation	In full	Abbreviation	In full
AICC	Akaike information criterion (corrected)	HMA	Hypomethylating agent
AML	Acute myeloid leukaemia	HMRN	Haematological malignancy research network
AML-MRC	AML with myelodysplasia-related changes	HSCT	Haematopoietic stem cell transplantation
BIC	Bayesian information criterion	ICER	Incremental cost-effectiveness ratio
BSA	Body surface area	ITT	Intention-to-treat
CI	Confidence intervals	MDS	Myelodysplastic syndrome
CMMoL-AML	AML with antecedent chronic myelomonocytic leukaemia	OS	Overall survival
CR	Complete remission	QALY	Quality-adjusted life year
CRi	Complete remission with incomplete platelet or neutrophil recovery	t-AML	Therapy-related AML
EFS	Event-free survival	WHO	World Health Organization

NICE

3

Key issues – clinical effectiveness

- Is standard chemotherapy (3+7 daunorubicin and cytarabine) the appropriate comparator?
  - Is FLAG-Ida relevant for younger patients?
- Are the results from the trial in people aged 60-75 generalisable to adults of all ages?
- Is CPX-351 clinically effective?
- Is the follow-up period of the trial long enough to capture relevant outcomes?
- Would clinicians begin the first cycle of treatment before genetic test results are received for patients who may have *de novo* AML?

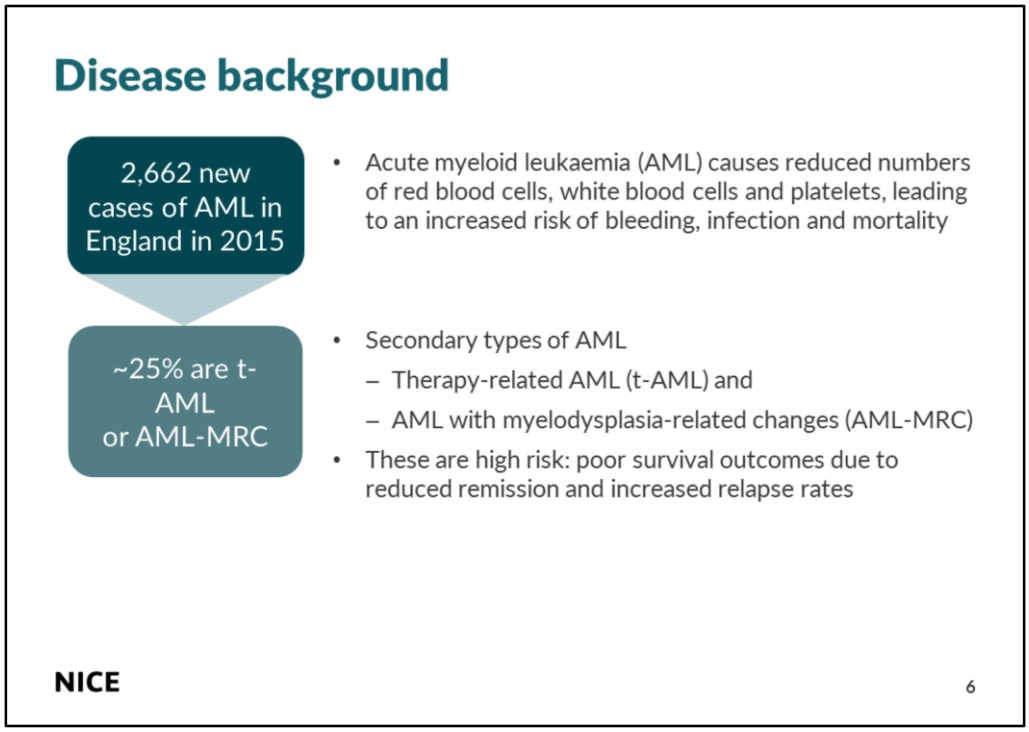
NICE 4

Key issues – cost effectiveness

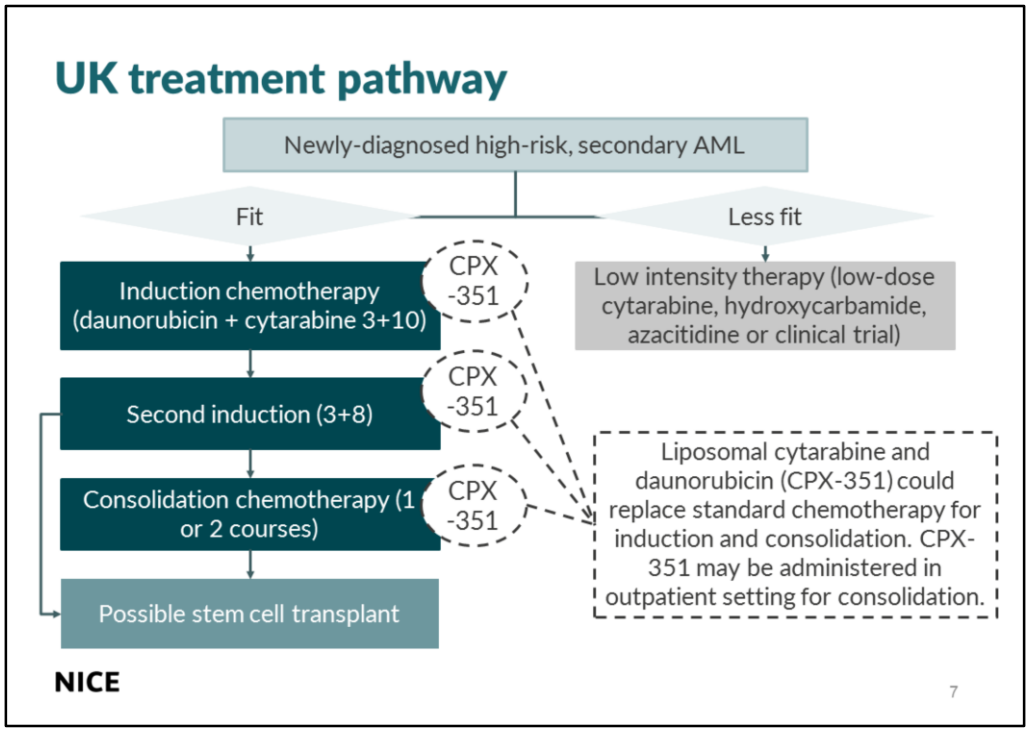
- Should post-transplant outcomes be based only on overall survival?
- What is the appropriate distribution for extrapolation of post-transplant overall survival?
- Should post-transplant mortality be higher than general population mortality?
- Should utility values in the post-transplant remission health state be adjusted for age?
- Should health-related quality of life be equivalent for both treatments in induction and consolidation?
- How should vial use be calculated?
- Should the costs of stem cell transplant include providing unrelated donor stem cells?
- Is CPX-351 cost-effective?
  - In adults of all ages?
- What is the most plausible ICER?
- Are the end of life criteria met?

NICE

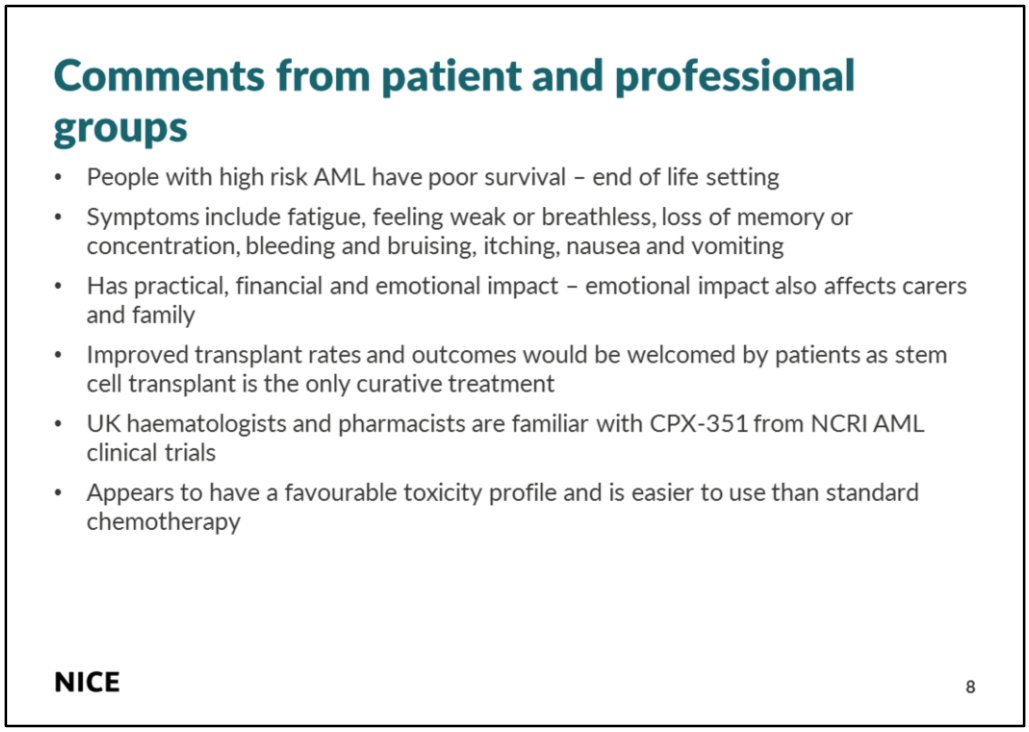
5



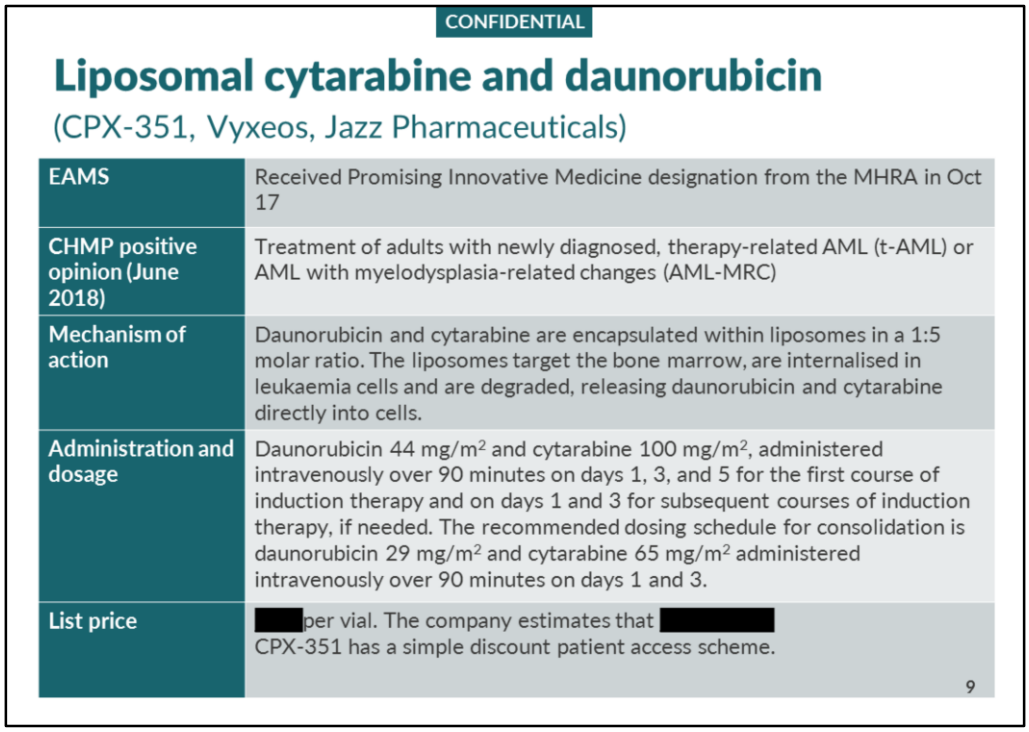
Source: company submission B1.3.1; submissions from RCPath-BSH, NCRI-ACP-RCP, and Leukaemia Care; Cancer Research UK



Source: company submission B1.3.3; submissions from RCPath-BSH and NCRI-ACP-RCP



Source: submissions from RCPath-BSH, NCRI-ACP-RCP, and Leukaemia Care



Source: company submission B1.2

Decision problem

	Final scope issued by NICE	Decision problem in the company's submission	Rationale if different
Population	People with newly diagnosed, high-risk (secondary) AML who are considered to be eligible for intensive therapy	High-risk (secondary) AML is defined by: • Therapy-related AML (t-AML) • AML with myelodysplasia related changes (AML-MRC)	In line with comments from British Society for Haematology (BSH) and diagnostic subgroups in WHO classification
Intervention	Liposomal daunorubicin and cytarabine	No change	N/A
Outcomes	Overall survival Event free survival Disease free survival Remission Health related quality of life Adverse effects of treatments	No change	N/A

10

Source: company submission B1.1

Decision problem

	Final scope issued by NICE	Decision problem in the company's submission	Rationale if different
Comparator	• Standard intensive induction and consolidation therapy • Azacitidine (for people who are not eligible for haematopoietic stem cell transplantation and have AML with 20-30% blasts and multilineage dysplasia) • Midostaurin (for people with FLT3-mutation-positive AML) (subject to ongoing NICE appraisal) • Gemtuzumab ozogamicin (subject to ongoing NICE appraisal)	Standard intensive induction and consolidation therapy (induction and consolidation with daunorubicin and cytarabine)	• Azacitidine not recommended for people eligible for HSCT; typically used as palliative therapy for people who are older and unfit • Midostaurin and gemtuzumab ozogamicin are added on to standard intensive chemotherapy, and not currently used in the NHS.

11

Source: company submission B1.1

ERG comments on decision problem

Population

- Defining population difficult in practice, particularly *de novo* AML – genetic test results available in 7-10 days: clinicians may begin the first cycle of treatment before receiving results
- Trial in company's submission included people aged 60-75
  - People older than 75 unlikely to be eligible for intensive treatment
  - 20-25% of people with high-risk AML are under 60

Comparators

- Appropriate not to consider azacitidine, midostaurin or gemtuzumab ozogamicin*
- 3+10 regimen of daunorubicin+cytarabine is standard treatment in NHS, but reasonable to consider equivalent to 3+7 regimen presented in company submission
- FLAG-Ida may be used as an alternative chemotherapy for younger patients in practice

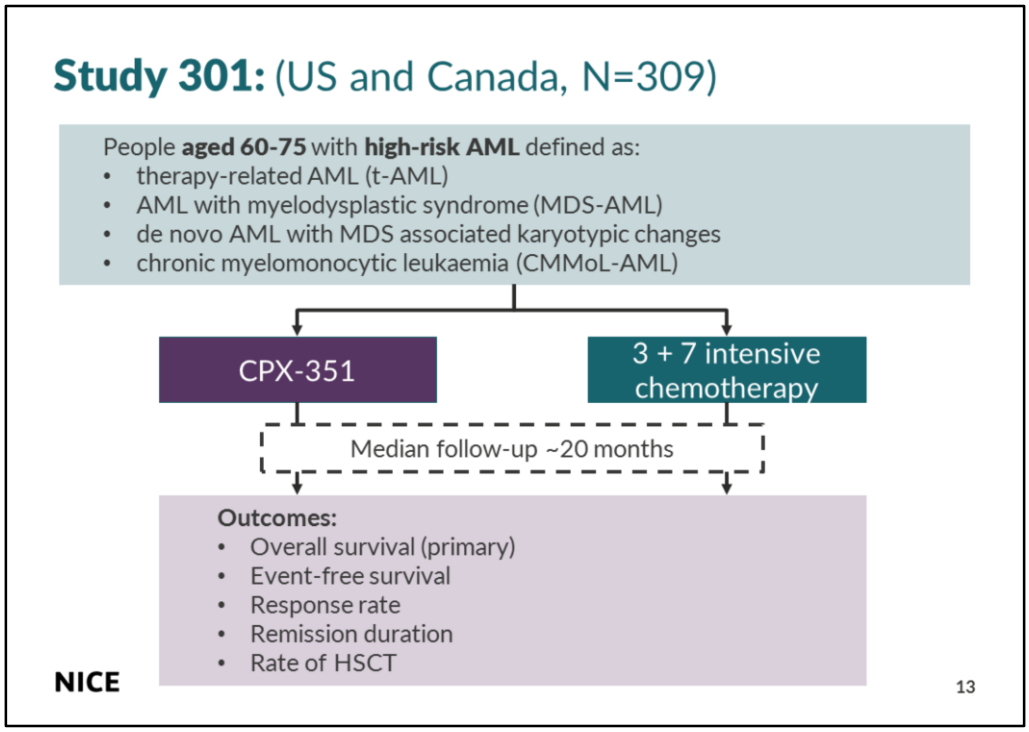
Outcomes

- Analysis was based on median follow-up 20.5 months in CPX-351 group, 21.2 months in 3+7 group - insufficient for measuring longer term post-HSCT overall survival (based on 2015 data cut)

*TA523 published June 2018: midostaurin recommended within marketing authorisation for newly diagnosed acute FLT3-mutation-positive myeloid leukaemia. Gemtuzumab ozogamicin appraisal ongoing.

12

Source: ERG report section 3



Source: company submission B2.3.1

Notes:

- Ongoing studies
    - AML18, older people with AML or high-risk MDS – CPX-351 is to be included in a 2018 amendment (estimated completion date October 2019)
    - AML19, people under 60 with AML or high-risk MDS (estimated completion date January 2021)

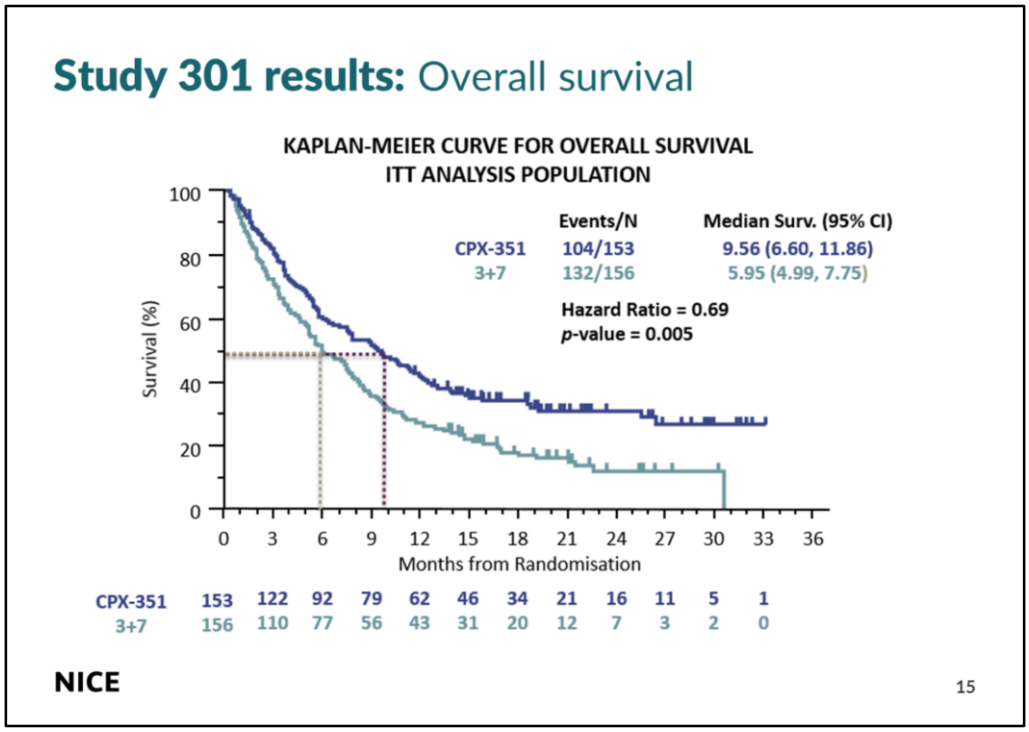
Study 301 baseline characteristics

Intention-to-treat (ITT) population

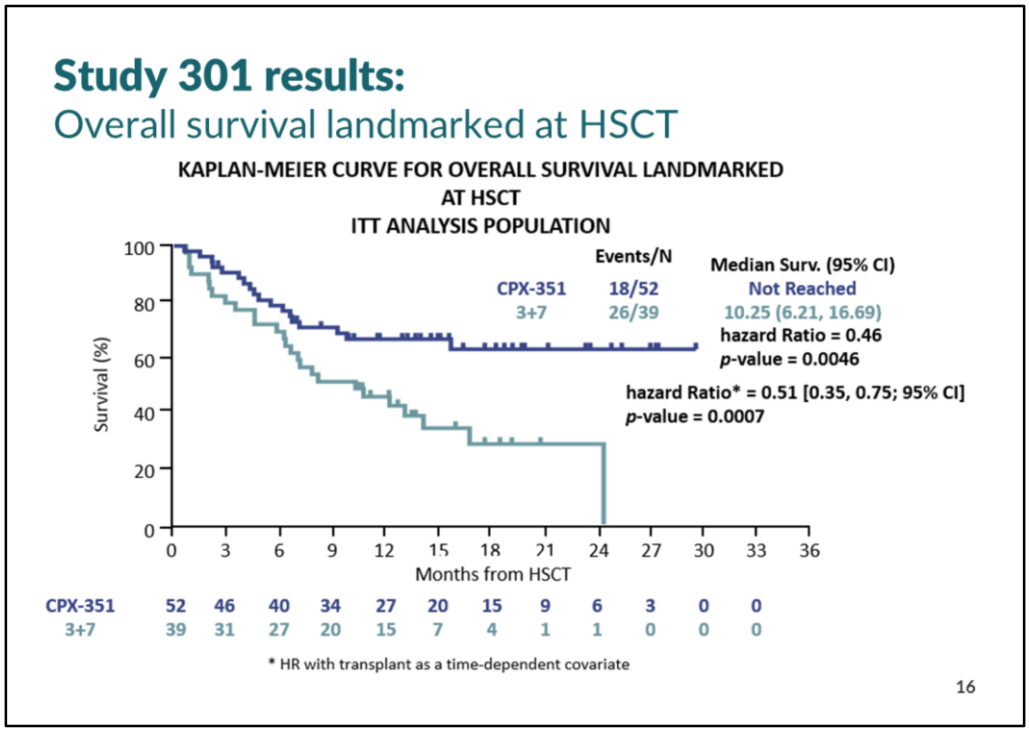
Characteristic	CPX-351 (n=153)	3+7 (n=156)
Mean age (SD), years	[redacted]	[redacted]
Age, n (%)		
60-69 years	96 (62.7)	102 (65.4)
70-75 years	57 (37.3)	54 (34.6)
Male sex, n (%)	94 (61.4)	96 (61.5)
Race, n (%)		
White	128 (83.6)	139 (89.1)
Black or African American	7 (4.6)	6 (3.8)
Asian	6 (3.9)	2 (1.3)
American Indian or Alaska native	[redacted]	[redacted]
Other	11 (7.2)	9 (5.8)
Median weight (range), kg	82.0 (49.0, 134.0)	82.7 (46.0, 136.0)
Median height (range), cm	170.2 (149.0, 198.0)	170.2 (149.0, 189.0)
Median BSA (range), m^2	2.0 (1.0, 3.0)	2.0 (1.0, 3.0)
ECOG performance group, n (%)		
PS = 0	37 (24.2)	45 (28.8)
PS = 1	101 (66.0)	89 (57.1)
PS = 2	15 (9.8)	22 (14.1)
PS ≥ 3	0	0

14

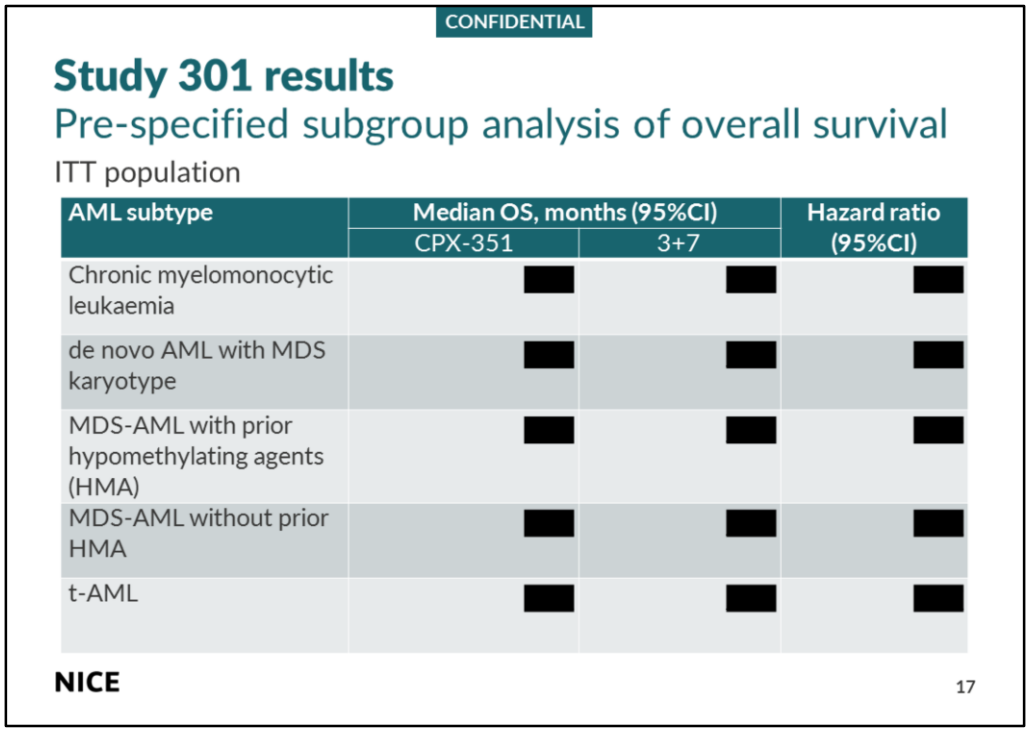
Source: company submission B2.3.3 table 8



Source: company submission figure 4



Source: company submission figure 7



Source: company submission appendix E, table 7

Notes:

- The study was not powered to evaluate differences between subgroups
- No tests for interaction performed

Study 301 results

Response rate

- CR: Complete remission
- CRi: Complete remission with incomplete platelet or neutrophil recovery

Endpoint, n (%)	CPX-351 (n=153)	3+7 (n=156)	Odds ratio (95% confidence intervals)
CR+CRi	73 (47.7)	52 (33.3)	1.77 (1.11, 2.81), p=0.016
CR	57 (37.3)	40 (25.6)	1.69 (1.03, 2.78), p=0.040
CRi	16 (10.5)	12 (7.7)	Not reported
No response	80 (52.3)	104 (66.7)	Not reported

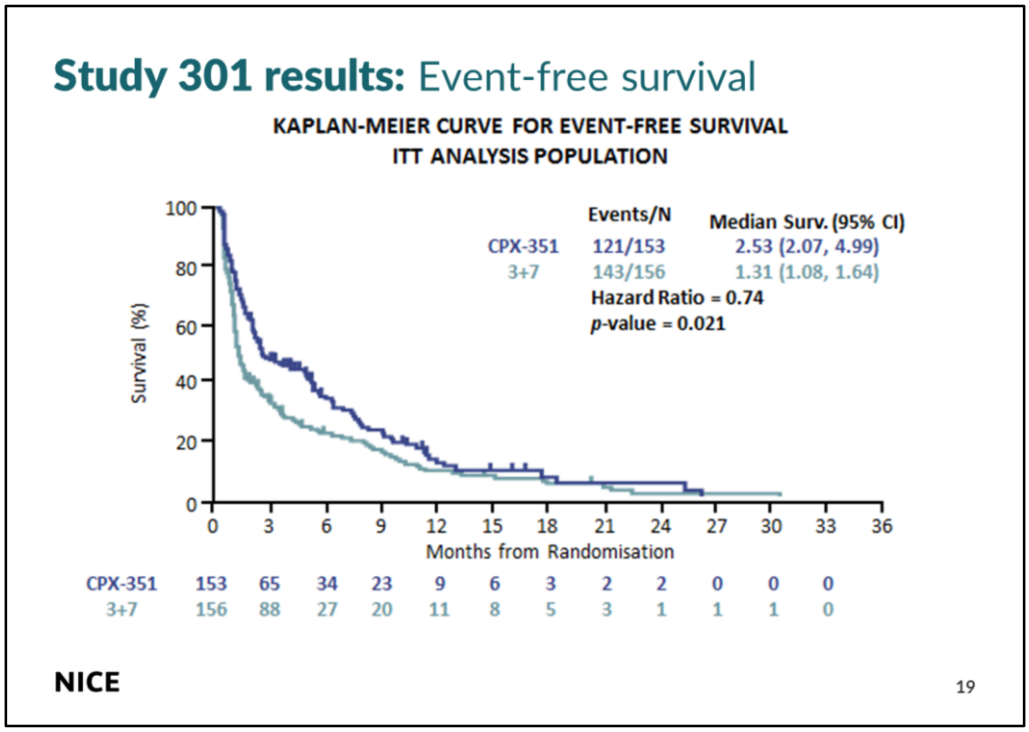
Rate of stem cell transplant

	CPX-351 (n=153)	3+7 (n=156)	Odds ratio (95% confidence intervals)
Patients having HSCT	52 (34.0%)	39 (25.0%)	1.54 (0.92, 2.56)

NICE

18

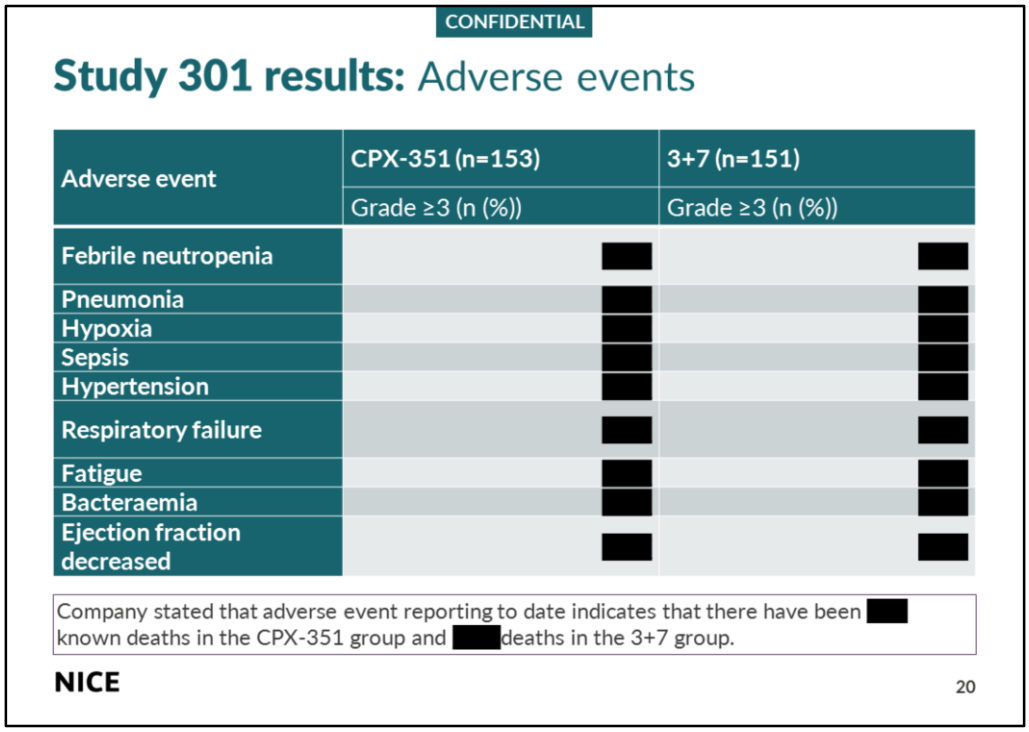
Source: company submission B.2.6.2, table 11; ERG report 4.2.3.1



Source: company submission B2.6.2, figure 6

Notes

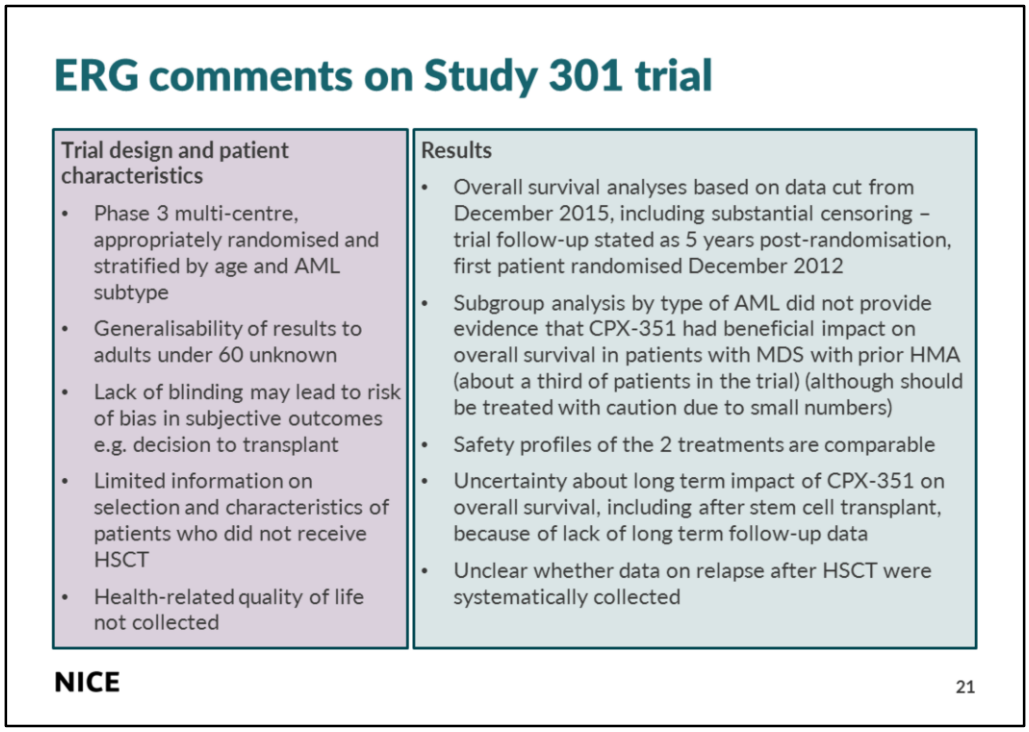
- An event was defined as the time from study randomisation to the date of induction treatment failure, relapse from either complete remission (CR) or complete remission with incomplete platelet or neutrophil recovery (CRi), or death



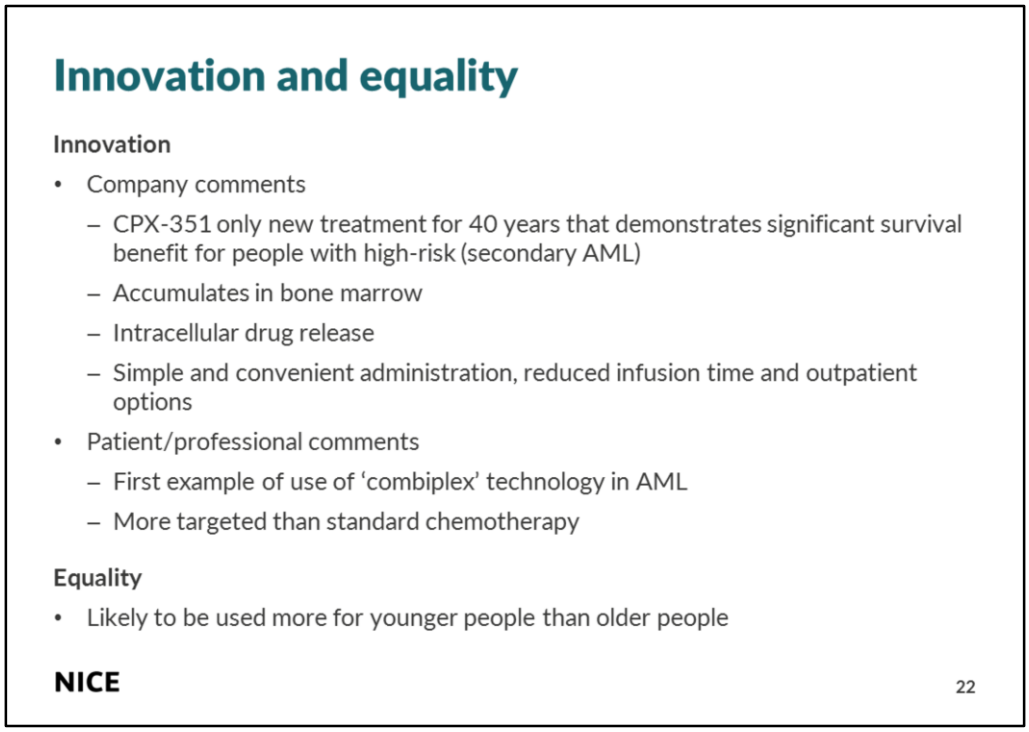
Source: company submission B2.10, table 16; response to clarification question A1iii; ERG report section 4.3.2.1

Notes:

- Adverse events were recorded from the start of infusion of day 1 to the last day of the treatment period. Serious adverse events were also collected up to 30 days after treatment completion.



Source: ERG report section 4.2.2

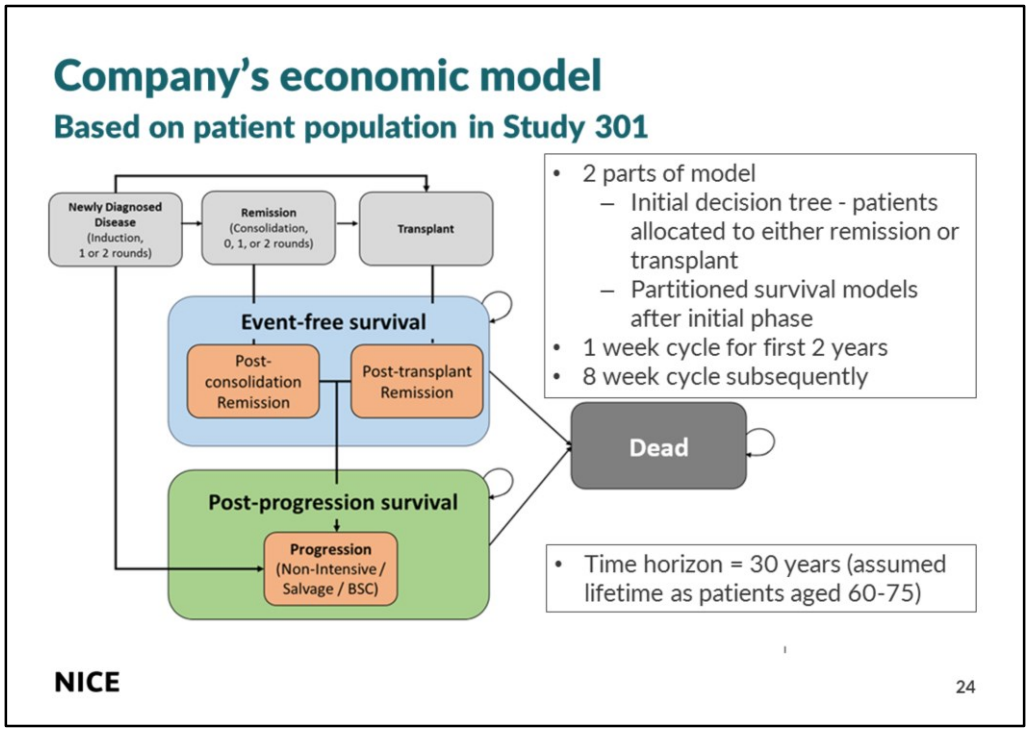


Source: company submission B2.12; submissions from RCPath-BSH, NCRI-ACP-RCP, and Leukaemia Care

Cost effectiveness

NICE

23



Source: Company submission, figure 11

Treatment effectiveness in the company's model

- Company conducted post-hoc multivariate logistic regression analyses to estimate probability of clinical pathway outcomes
- For patients where response not achieved, parametric curves applied from start of the model
- For patients where a response was achieved, a time-shift was applied, where no mortality occurs from start of model until time at which event-free and overall survival start to be tracked

Patient group	EFS and OS tracked from:
Response + transplant	Time of transplant
Response, no transplant	Last consolidation therapy

- Survival data from Study 301 extrapolated over model time horizon
  - Parametric survival curves modelled separately per treatment group
- Where modelled OS suggested lower mortality than general population, mortality based on Office for National Statistics data instead of Study 301 estimate
- Higher rate of mortality post-transplant explored in scenario analysis

NICE 25

Source: company submission section B3.3.2, B3.3.3

Extrapolation of survival in the model

Company's choice of parametric curves

Patient group	Survival curve	CPX-351 distribution	3+7 distribution
Response, no transplant	Overall survival	Log-logistic	Log-logistic
	Event-free survival	Weibull	Weibull
Response + transplant	Overall survival	Gompertz*	Gompertz
	Event-free survival	Lognormal**	Lognormal**
No response	Overall survival	Lognormal	
	Time to transplant/death	Lognormal	
	Time to transplant/relapse/death	Generalised gamma	

*See next slide

**Company and ERG agree that results based on these analyses are unreliable due to small patient numbers

NICE

26

Source: ERG report section 5.2.6.3

CONFIDENTIAL

Post-transplant overall survival extrapolation for people who had a complete response

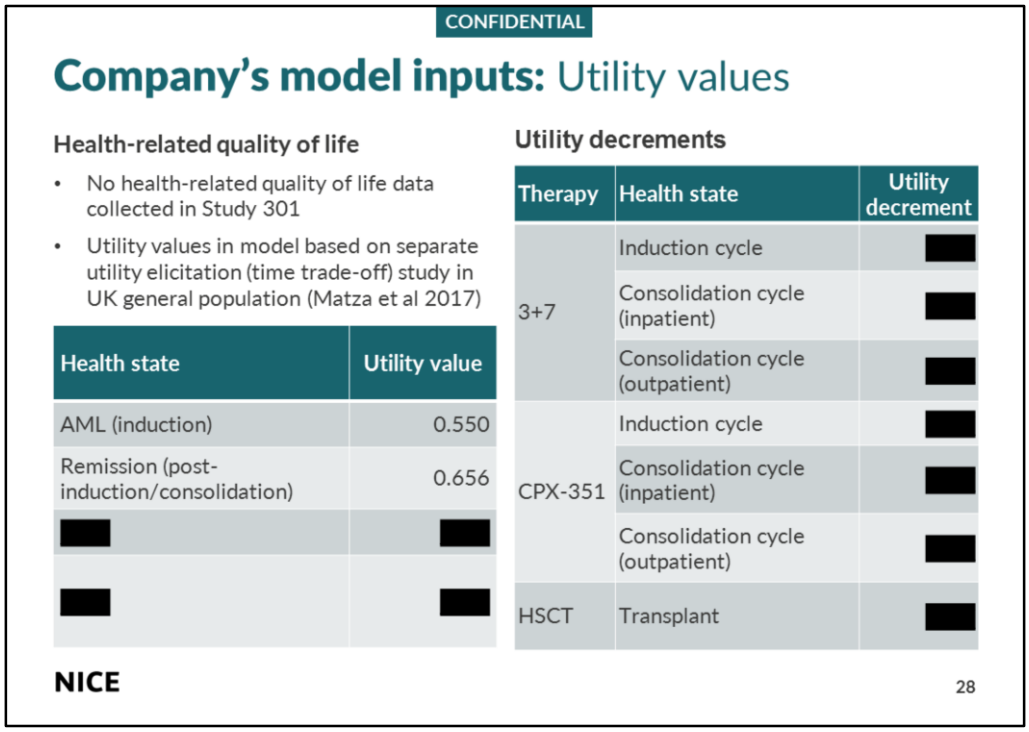
CPX-351

- AIC and BIC statistics [redacted]
- Company chose Gompertz because best fit to trial data and based on clinical plausibility
- CPX-351: n=52
- 3+7: n=39

NICE

27

Source: Company submission appendix M6 figure 170, ERG report section 5.2.6.3



Source: company submission B3.4.5

Notes

- Disutility for adverse events included in the treatment-related decrements

Company's model inputs:
Costs and resource use

Cost/Resource	Source
Drug costs	eMIT, British National Formulary
Administration costs	NHS National Schedule of Reference Costs 2016-17
Transplant costs	UK Stem Cell Strategic Forum Wang et al. 2011 NHS National Schedule of Reference Costs 2016-17
Treatment monitoring	Resource use: US NCCN Guidelines for Acute Myeloid Leukemia Costs: NHS National Schedule of Reference Costs 2016-17
Post-consolidation monitoring	Mahmoud et al. 2012
Post-transplant monitoring costs	UK Stem Cell Strategic Forum
Adverse event costs	NHS National Schedule of Reference Costs 2016-17

NICE 29

Source: Company submission table 48

ERG report section 5.2.8

CONFIDENTIAL

Company's base case cost-effectiveness results

List price analysis

Treatment	Total costs (£)	Total QALYs	Incremental costs (£)	Inc QALYs	ICER (£/QALY)
3+7	[redacted]	[redacted]	-	-	-
CPX-351	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]

Analysis with patient access scheme (PAS) for CPX-351 only

Treatment	Total costs (£)	Total QALYs	Incremental costs (£)	Incremental QALYs	ICER incremental (£/QALY)
3+7	[redacted]	[redacted]	-	-	-
CPX-351	[redacted]	[redacted]	[redacted]	[redacted]	46,631

Azacitidine is included in the model as a subsequent treatment. See part 2 for analysis including PAS for azacitidine.

NICE

30

Source: company submission tables 51 and 52

CONFIDENTIAL

ERG corrected deterministic sensitivity analysis of company's model

NICE 31

Source: ERG report figure 19

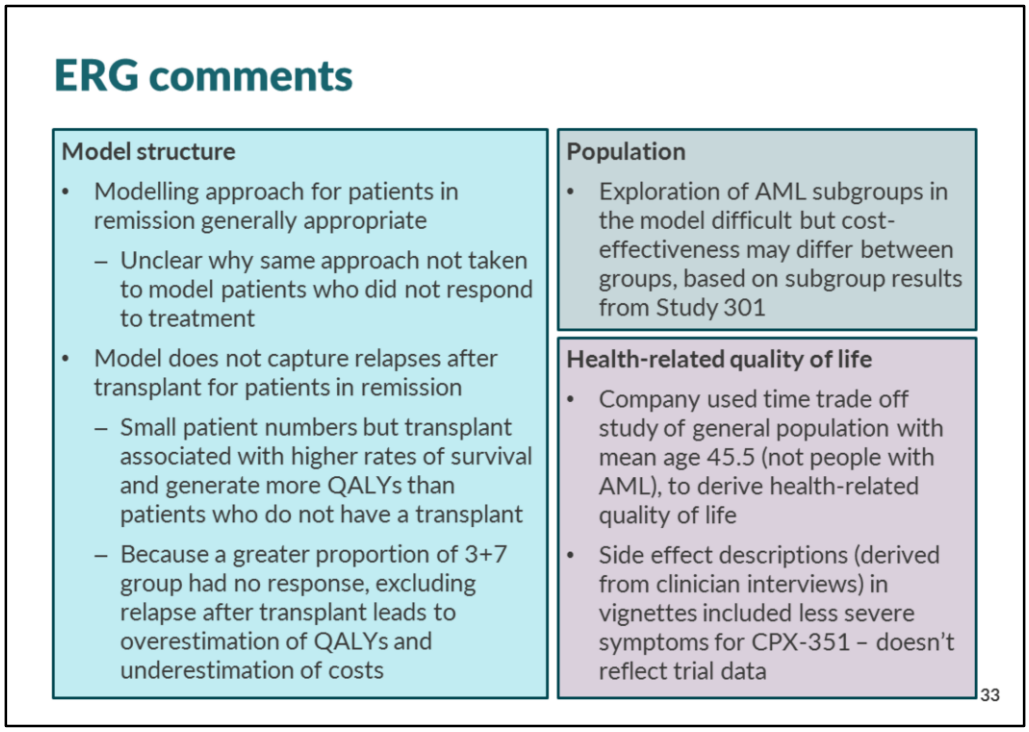
CONFIDENTIAL

Company's scenario analyses: list price

Scenario	Incremental costs (£)	Incremental QALYs	ICER (£/QALY)
Company base case	[redacted]	[redacted]	[redacted]
30% of modelled cohort aged <60 years	[redacted]	[redacted]	[redacted]
10% of modelled cohort aged <60 years	[redacted]	[redacted]	[redacted]
Comparator based on UK 3+10 regimen	[redacted]	[redacted]	[redacted]
Alter percentage of patients treated with CPX-351 who receive second induction and consolidation therapy as outpatients (patients receiving consolidation as outpatients changed from 50 to 100%)	[redacted]	[redacted]	[redacted]
5 year time horizon	[redacted]	[redacted]	[redacted]
10 year time horizon	[redacted]	[redacted]	[redacted]
Utilities from Hensen et al. 2017	[redacted]	[redacted]	[redacted]
Mortality after HSCT 2.3 times general population (included in ERG base case)	[redacted]	[redacted]	[redacted]
No transplant for non-responders	[redacted]	[redacted]	[redacted]

32

Source: company submission tables 60 to 68



Source: ERG report sections 5.2.1, 5.2.3, 5.2.7 and 5.2.8

ERG comments

Treatment effectiveness

- Data from 2015 datacut of Study 301 used – although follow-up is continuing for 5 years post-randomisation, no newer datacut available
- Potential lack of clinical plausibility of post-transplant overall survival for CPX-351 – higher than reported in literature, not justified clinically, level of censoring of trial data between 1 and 2 years leads to uncertainty in extrapolation
- Some inconsistencies in inclusion of stratification variables – ERG concerned that adjusting by variables added uncertainty due to sparsity of data in subgroups
  - Categories of AML not used consistently in analysis – original 5 categories are sometimes grouped together into 2 categories, due to low patient numbers
  - Prevents exploring the impact of CPX-351 within AML subtypes
- Company scenario analysis adjusted general population mortality after stem cell transplant by standardised mortality ratio (SMR) of 2.34 (derived from Martin et al., 2010) – ERG unclear how SMR derived from this study but has face validity and ERG used in base case

NICE 34

Source: ERG report section 5.2.6

CONFIDENTIAL

ERG scenario analyses

Including patients under 60 years of age

- Company used Dumas study to estimate relative response rate in younger patients in scenario analyses - did not include patients under age of 60, so ERG considered that Study 301 could have been used to conduct similar analysis
  - ERG scenario analysis used lower odds ratio in younger population than company, based on company's method of calculating difference between age groups 60-69 and 70-75 in Study 301 (OR of [redacted] for post-induction remission)
- ERG also adjusted drug acquisition costs on clinical advice that patients under 60 more likely to receive intermediate-dose cytarabine in consolidation phase

Technologies	Total costs	Total QALYs	Incremental costs	Incremental QALYs	ICER (£/QALY)
Company base-case (including ERG corrections)					
3+7	[redacted]	[redacted]	-	-	-
CPX-351	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
Scenario: 10% of patients under 60 years					
3+7	[redacted]	[redacted]	-	-	-
CPX-351	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
Scenario: 30% of patients under 60 years					
3+7	[redacted]	[redacted]	-	-	-
CPX-351	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]

35

Source: ERG report section 5.2.3, 6.3.1, table 36

CONFIDENTIAL

ERG scenario analyses

Post-transplant overall survival for CPX-351 [1]

- Predicted post-transplant overall survival from Gompertz model plateaued at [redacted] survival rate after ~2yrs (median overall survival [redacted] weeks for 3+7 group)
  - Suggests [redacted] of patients can be cured – trial data not mature enough to justify after 2 years
- Common treatment effect for both groups explored, although predicted survival deviates from observed clinical data

- ERG explored uncertainty around choice of distribution using model averaging approach – each model assigned a weight that represents how well that distribution predicts the lifetime survival of the modelled cohort, compared to other distributions

Weights used	1 year survival	2 year survival	5 year survival
AICC	[redacted]	[redacted]	[redacted]
BIC	[redacted]	[redacted]	[redacted]

NICE

36

Source: ERG report section 5.2.6.3, 6.3.2, figure 21

Notes

- Model averaging approach methods from Jackson et al., Hettle et al.

CONFIDENTIAL

ERG scenario analyses

*Included in ERG alternative base case

Post-transplant overall survival for CPX-351 [2]

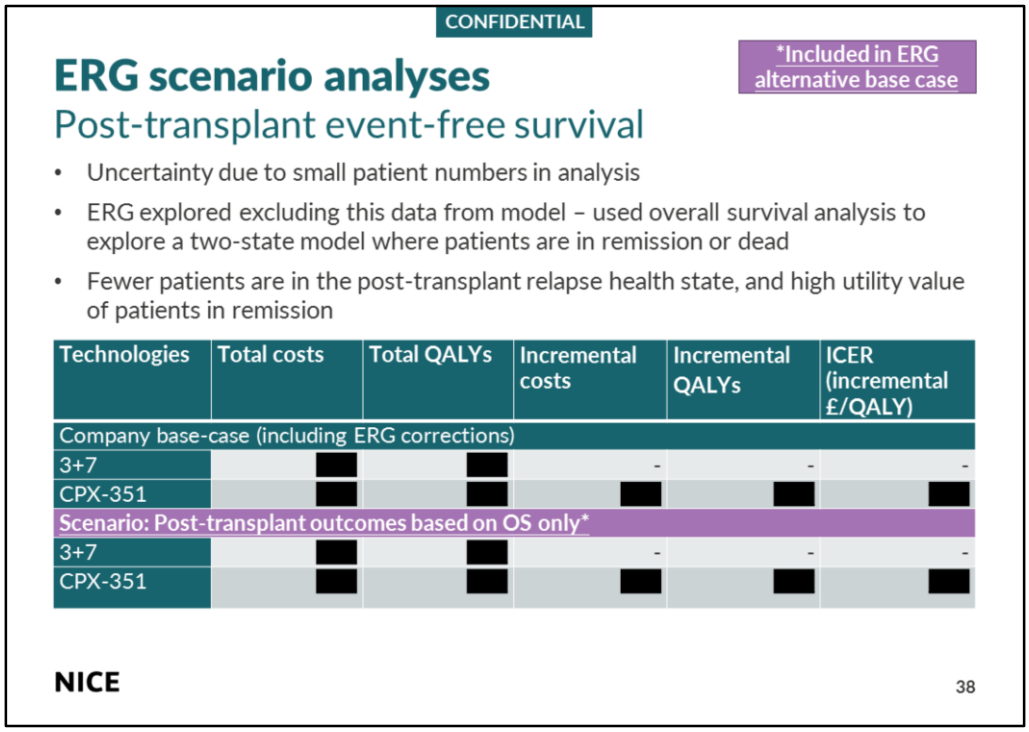
Scenario (post-transplant overall survival)	Tech-nology	Total costs	Total QALYs	Inc. costs	Inc. QALYs	ICER (£/QALY)
Company base-case (including ERG corrections) - Gompertz	3+7	[redacted]	[redacted]	-	-	-
	CPX-351	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
Weibull for CPX-351	3+7	[redacted]	[redacted]	-	-	-
	CPX-351	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
Log-logistic for CPX-351	3+7	[redacted]	[redacted]	-	-	-
	CPX-351	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
Log-normal for CPX-351	3+7	[redacted]	[redacted]	-	-	-
	CPX-351	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
Exponential for CPX-351	3+7	[redacted]	[redacted]	-	-	-
	CPX-351	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
Generalised gamma for CPX-351	3+7	[redacted]	[redacted]	-	-	-
	CPX-351	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
Combining treatment arms (Gompertz distribution)	3+7	[redacted]	[redacted]	-	-	-
	CPX-351	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
Weighted CPX-351 curve (weighted average of Gompertz, Weibull, log-normal, log-logistic and exponential using AICC weights) *	3+7	[redacted]	[redacted]	-	-	-
	CPX-351	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
Weighted CPX-351 curve (weighted average of Gompertz, Weibull, log-normal, log-logistic and exponential using BIC weights)	3+7	[redacted]	[redacted]	-	-	-
	CPX-351	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]

37

Source: ERG report section 5.2.6.3, 6.3.2, figure 21

Notes

- Model averaging approach methods from Jackson et al., Hettle et al.



Source: ERG report section 5.2.6.3, 6.3.3, table 38

CONFIDENTIAL

ERG scenario analyses

Post-transplant remission utility

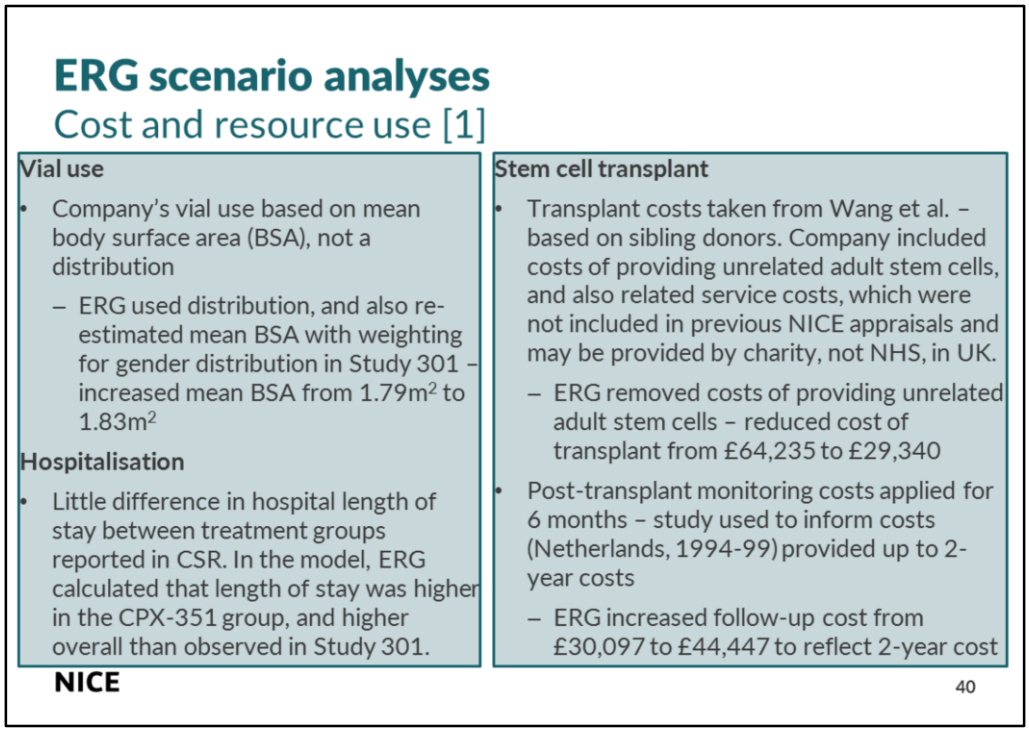
*Included in ERG alternative base case

- Company's utility value for post-transplant remission health state (0.86) is similar or higher than general population
- Utility values not adjusted for aging
- Treatment-related disutilities based on clinician experience not evidence in trial and didn't include impact of more serious AEs included in model

Scenario	Technology	Total costs	Total QALYs	Inc. costs	Inc. QALYs	ICER (£/QALY)
Company base-case (with ERG corrections)	*3+7*	[redacted]	[redacted]	-	-	-
	CPX-351	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
Utility value of 0.75 for durable remission*	3+7	[redacted]	[redacted]	-	-	-
	CPX-351	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
General population utility value (0.79) for durable remission	3+7	[redacted]	[redacted]	-	-	-
	CPX-351	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
Post-transplant remission utility, adjusted for aging*	3+7	[redacted]	[redacted]	-	-	-
	CPX-351	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
Mean on-treatment utility during induction and consolidation treatment*	3+7	[redacted]	[redacted]	-	-	-
	CPX-351	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]

39

Source: ERG report section 5.2.7, 6.3.4, table 39



Source: ERG report section 5.2.8, 6.3.5

CONFIDENTIAL

*Included in ERG alternative base case

ERG scenario analyses

Cost and resource use [2]

	Total costs	Total QALYs	Incremental costs	Incremental QALYs	ICER (incremental £/QALY)
Company base-case (with ERG corrections)					
3+7	[redacted]	[redacted]	-	-	-
CPX-351	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
Scenario: Vial usage reflecting the distribution of body surface area*					
3+7	[redacted]	[redacted]	-	-	-
CPX-351	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
Scenario: Reduced number of hospital days during consolidation period*					
3+7	[redacted]	[redacted]	-	-	-
CPX-351	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
Scenario: Alternative cost of transplant*					
3+7	[redacted]	[redacted]	-	-	-
CPX-351	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]

41

Source: ERG report section 6.3.5, table 41

CONFIDENTIAL

Summary of ERG's alternative base case

Scenario	Inc. costs (£)	Inc. QALYs	ICER (£/QALY)
Company's base case, corrected	[redacted]	[redacted]	[redacted]
1. Post-transplant outcomes based on overall survival only	[redacted]	[redacted]	[redacted]
2. Post-transplant overall survival based on analysis weighted by goodness-of-fit, based on AICC weights	[redacted]	[redacted]	[redacted]
3. Adjusting post-transplant general population mortality (company scenario analysis, ***ICER based on uncorrected company base case***)	[redacted]	[redacted]	[redacted]
4. Utility value of 0.75 for durable remission	[redacted]	[redacted]	[redacted]
5. Utilities in the post-transplant remission health state adjusted for age	[redacted]	[redacted]	[redacted]
6. Equivalent quality of life for both treatment groups in induction and consolidation	[redacted]	[redacted]	[redacted]
7. Vial use based on distribution of body surface area, and mean re-weighted based on gender distribution in Study 301	[redacted]	[redacted]	[redacted]
8. Reduced number of hospital days in consolidation	[redacted]	[redacted]	[redacted]
9. Excluding costs of providing unrelated donor stem cells	[redacted]	[redacted]	[redacted]
ERG's base case (1-9)	[redacted]	[redacted]	[redacted]

42

CONFIDENTIAL

ERG alternative base case results and further scenario analyses

	Total costs	Total QALYs	Incremental costs	Incremental QALYs	ICER (£/QALY)
Company base-case (with ERG corrections)					
3+7	[redacted]	[redacted]	-	-	-
CPX-351	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
ERG alternative base-case					
3+7	[redacted]	[redacted]	-	-	-
CPX-351	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]

Scenario	Incremental costs	Incremental QALYs	ICER (£/QALY)
Most favourable post-transplant OS for CPX-351 (Gompertz)	[redacted]	[redacted]	[redacted]
Least favourable post-transplant OS for CPX-351 (exponential)	[redacted]	[redacted]	[redacted]
10% patients under the age of 60	[redacted]	[redacted]	[redacted]
30% patients under the age of 60	[redacted]	[redacted]	[redacted]

NICE

43

Source: ERG report section 6.4, table 42 and section 6.5, table 44

CONFIDENTIAL

End of life considerations

Criterion	Data source	Indication	Age	Overall survival	
				Median (months)	Mean
Short life expectancy, normally < 24 months	Study 301	High risk AML	60-75	5.95	-
	Company's economic model (undiscounted life years)	High risk AML	60-75	-	[redacted]
	ERG alternative base case	High risk AML	60-75	-	[redacted]
	HMRN Yorkshire registry	t-AML and AML-MRC	All	~3-4	-
	Swedish registry	t-AML	<55	14	-
			55-74	9	-
			≥75	8	-
Extension to life, normally of a mean value of ≥ 3 months				Increase with CPX-351	
				Median (months)	Mean
	Study 301			3.61	-
	Company's economic model (undiscounted life years) – total population			-	[redacted]
	ERG alternative base case model			-	[redacted]
	ERG scenario for post-transplant OS for CPX-351 (least optimistic estimate)			-	[redacted]

44

Source: company submission table 22

Authors

Kirsty Pitt
Technical Lead

Alex Filby
Technical Adviser

with input from the Lead Team (Natalie Hallas, John Hampson and David Chandler)

NICE

45

NATIONAL INSTITUTE FOR HEALTH AND CARE EXCELLENCE

Single technology appraisal

CPX-351 for untreated high-risk (secondary) acute myeloid leukaemia (ID1225)

Document A

Company evidence submission summary for committee

Jazz Pharmaceuticals Ltd. confirm that all information in the submission summary is an accurate summary or replication of evidence in the main submission and accompanying appendices and that wherever possible a cross reference to the original source is provided.

August 2018

File name	Version	Contains confidential information	Date
ID1225_Evidencesummary [redacted]_V4	4.0	No	07/08/2018

Contents

Tables and figures

Submission summary

A.1 Health condition

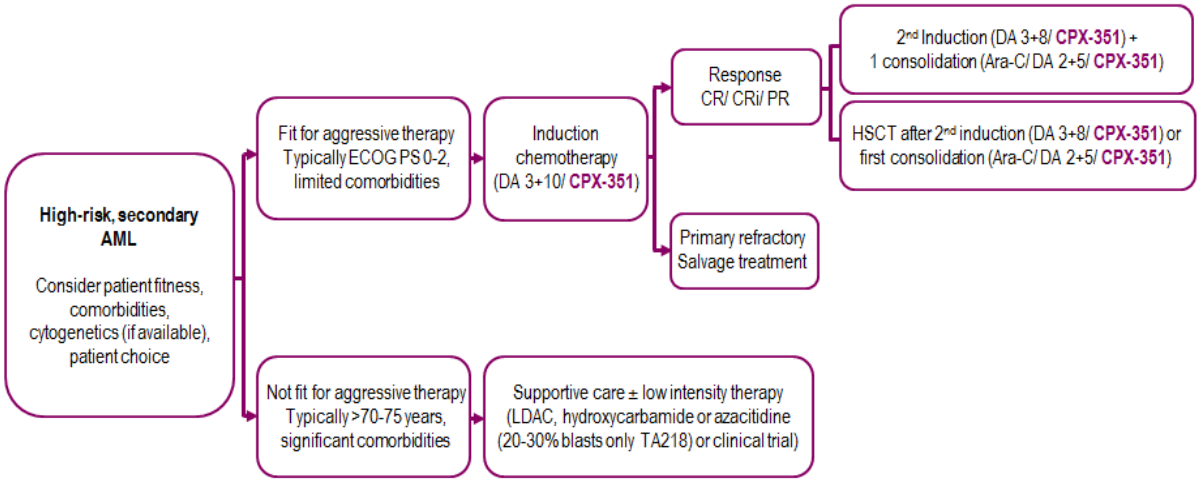
Acute myeloid leukaemia (AML) is a rare clonal disorder of haematopoietic progenitor cells, whereby a population of leukaemic stem cells is thought to give rise to the proliferation of abnormal myeloid precursor cells (blasts), which fail to differentiate (1, 2). The accumulation of leukaemic blasts in the bone marrow and the suppression of normal haematopoiesis lead to neutropaenia, anaemia, and thrombocytopaenia. If untreated, patients may die of infections, complications of infections (e.g. sepsis and multisystem organ failure), or bleeding events (typically central nervous system, respiratory, or gastrointestinal bleeding), usually in a matter of weeks (3).

CPX-351 is proposed to be indicated for the treatment of newly diagnosed adults with high-risk (secondary) AML. Patients with high-risk AML include those with therapy-related AML (t-AML) or AML with myelodysplasia-related changes (AML-MRC). High-risk (secondary) AML is particularly difficult to treat as it conflates two different factors that contribute to poor survival: adverse characteristics of the disease itself, and risk emanating from patient specific factors (4, 5). Adverse characteristics of the AML disease itself are those that reduce tumour sensitivity to treatment, resulting in either failure to achieve response or early relapse. These adverse characteristics include poor risk cytogenetics, antecedent haematologic disorders, and prior exposure to chemotherapy or radiation therapy. Patient specific factors are those that reduce patient fitness, resilience and tolerance for intensive chemotherapy or haematopoietic stem cell transplant (HSCT). These factors are often evaluated by age, comorbid conditions and overall performance status.

A.2 Clinical pathway of care

The current treatment paradigm for high-risk (secondary) AML along with the proposed positioning of CPX-351 is summarised in Figure 1.

Figure 1: Clinical Context of CPX-351 in the Current Treatment Paradigm for High-risk (secondary) AML



DA 3+10, 3+8 or 2+5 refers to the number of days of drug delivery for daunorubicin and cytarabine, respectively; Ara-C, cytarabine; CR, complete remission; CRi, complete remission with incomplete blood count recovery; PR, partial response; ECOG PS, Eastern Cooperative Oncology Group performance status; LDAC, low dose cytarabine

A.3 Equality considerations

Use of CPX-351 is unlikely to raise any equality issues. As detailed in section B1.3.4, the prognosis of patients with high-risk AML is poor, and there is a great degree of similarity in unmet medical need for patients of all ages with high-risk (secondary) AML.

A.4 The technology

Table 1: Technology being appraised – B.1.2 (page 15)

UK approved name and brand name	CPX-351; Brand name: Vyxeos™
Mechanism of action	CPX-351 is a combination of the antineoplastic drugs daunorubicin and cytarabine encapsulated in CombiPlex® liposomes for intravenous (IV) administration. The unique composition of CPX-351 liposomes imparts remarkable stability, drug delivery and drug release properties (6). These distinctive features differentiate CPX-351 liposomes from conventional antineoplastic liposomal products, and confer unique pharmacological advantages that enable targeted delivery and deployment of a synergistic fixed molar ratio of the drug cargo. Daunorubicin and cytarabine are encapsulated within the advanced CombiPlex® liposomes at a fixed 1:5 molar ratio, a ratio which has

	been shown to exhibit synergy in promoting leukaemia cell death in both *in vitro* and *in vivo* models. As the CPX-351 liposomes are 'solid phase', stable liposomes, encapsulation of daunorubicin and cytarabine into liposomes results in high retention of drug cargo and low *in vivo* release rates as they circulate in plasma, and effectively changes the pharmacokinetic profiles of the two drugs. Correspondingly, CPX-351 exhibits a prolonged plasma half-life following IV infusion in both animals and humans, with greater than 99% of the daunorubicin and cytarabine in the plasma remaining encapsulated within the liposomes. Encapsulation of the active moieties within specialised liposomes also affects drug delivery to target tissues. Based on data in animals, CPX-351 liposomes target and persist in the bone marrow, where they are preferentially taken up intact by leukaemia cells in an active engulfment process. After internalisation by leukaemia cells, CPX-351 liposomes undergo degradation, releasing daunorubicin and cytarabine within the intracellular environment, enabling the drugs to exert their 1:5 synergistic antineoplastic activities directly at the tumour target. Targeted, fixed ratio drug delivery with CPX-351 introduces the possibility of enhancing anti-tumour efficacy over both the free drug cocktail as well as other delivery systems in which the importance of drug ratios is neglected. Furthermore, CPX-351 also introduces the possibility of superior efficacy with a lower cumulative dose of daunorubicin and cytarabine relative to standard intensive chemotherapy regimen. Taken together, these attributes can translate to improved clinical outcomes following intensive cytotoxic therapy with a curative intent, especially in high-risk patients for whom there is an unmet need.
Marketing authorisation Status	On 28 June 2018, the Committee for Medicinal Products for Human Use (CHMP) adopted a positive opinion, recommending the granting of a marketing authorisation for the medicinal product Vyxeos, (CPX-351). In accordance with published EC guidelines, marketing authorisation is expected around 4th September 2018 In September 2017, the CHMP granted CPX-351 accelerated assessment, which is designed for products of major interest for public health and therapeutic innovation. CPX-351 was granted orphan drug status by both the EMA (EU/3/11/942; January 2012) and the US FDA (August 2016) for the treatment of AML. In October, 2017, CPX-351 received Promising Innovative Medicine (PIM) designation from the Medicines and Healthcare Products Regulatory Agency in the United Kingdom. On August 3rd 2017, CPX-351 was approved by the U.S. Food and Drug Administration (FDA), for the treatment of adults with high-risk AML i.e. t-AML and AML-MRC.

Indications and any restriction(s) as described in the summary of product characteristics	CPX-351 (Vyxeos) is indicated for the treatment of adults with newly diagnosed, therapy-related acute myeloid leukaemia (t-AML) or AML with myelodysplasia-related changes (AML-MRC) Contraindications, special warnings and precautions for use are listed as per the draft SmPC (see Appendix C)
Method of administration and dosage	CPX-351 is administered by intravenous infusion over 90 minutes. The injection is supplied in a single-patient-use vial. Each 50 ml vial contains 44 mg of daunorubicin and 100 mg of cytarabine. The concentration of the reconstituted solution is 44 mg/20 mL (2.2 mg/mL daunorubicin) and 100 mg/20 mL (5 mg/mL cytarabine). CPX-351 dosing is based on the patient's body surface area (BSA) according to the schedule below. **<u>Recommended dosing schedule for induction of remission</u>** Daunorubicin 44 mg/m^2 and cytarabine 100 mg/m^2, administered intravenously over 90 minutes on days 1, 3, and 5 for the first course of induction therapy and on days 1 and 3 for subsequent courses of induction therapy, if needed. A subsequent course of induction may be administered in patients who do not show disease progression or unacceptable toxicity. The attainment of a normal-appearing bone marrow may require more than one induction course. Evaluation of the bone marrow following recovery from the previous course of induction therapy determines whether a further course of induction is required. **<u>Recommended dosing schedule for consolidation</u>** Administer the first consolidation cycle 5 to 8 weeks after the start of the last induction. The recommended dosing schedule of CPX-351 is daunorubicin 29 mg/m^2 and cytarabine 65 mg/m^2 administered intravenously over 90 minutes on days 1 and 3. Consolidation therapy is recommended for patients achieving remission who have recovered to absolute neutrophil count (ANC) greater than 500/µL and the platelet count has recovered to greater than 50,000/µL in the absence of unacceptable toxicity. A subsequent course of consolidation may be administered in patients who do not show disease progression or unacceptable toxicity.
Additional tests or investigations	No additional tests or investigations are required for CPX-351 versus standard induction chemotherapy.
List price and average cost of a course of treatment	[redacted] Based on experience from the randomised Phase III pivotal study, and based on standard practice in England, it is estimated that a 1.8m^2 patient will require on average, [redacted] vials of CPX-351. Average cost of a course of treatment is estimated to be

	[redacted] per patient.
Patient access scheme (if applicable)	[redacted]

A.5 Decision problem and NICE reference case

The submission covers the technology's full marketing authorisation for this indication.

The company submission differs from the final NICE scope and the NICE reference case.

Table 2: The decision problem – B.1.1 (pages 12-14)

	Final scope issued by NICE/reference case	Decision problem addressed in the company submission	Rationale if different from the final NICE scope
Population	People with newly diagnosed, high-risk AML who are considered to be eligible for intensive therapy	People with newly diagnosed, high-risk (secondary) AML who are considered to be eligible for intensive therapy. High-risk (secondary) AML is defined by: Therapy-related AML (t-AML) AML with myelodysplasia related changes (AML-MRC)	Company has provided specificity in line with comments from British Society for Haematology (BSH). t-AML and AML-MRC are diagnostic subgroups in the 2016 World Health Organization (WHO) classification (7).
Intervention	Liposomal daunorubicin and cytarabine	CPX-351 is a combination of the antineoplastic drugs daunorubicin and cytarabine encapsulated in CombiPlex® liposomes for IV administration.	CPX-351 is the first dual-drug advanced liposomal formulation. The unique composition of CPX-351 liposomes imparts remarkable stability, drug delivery and drug release properties. These distinctive features differentiate CPX-351 liposomes from conventional liposomal products, and confer unique pharmacological advantages that enable targeted delivery and deployment of a synergistic fixed molar ratio of the drug cargo directly to leukaemia cells for a prolonged period of time.
Comparator(s)	Standard intensive induction and consolidation therapy	For people who are considered to be fit for intensive chemotherapy, the standard therapy in the UK is induction and	The submission contains reference to the comparators outlined in the NICE scope.

	Final scope issued by NICE/reference case	Decision problem addressed in the company submission	Rationale if different from the final NICE scope
	Azacitidine (for people who are not eligible for haematopoietic stem cell transplantation and have AML with 20-30% blasts and multilineage dysplasia) Midostaurin (for people with FLT3-mutation-positive AML) (subject to ongoing NICE appraisal) Gemtuzumab ozogamicin (subject to ongoing NICE appraisal)	consolidation with a daunorubicin and cytarabine regimen.	After validation with clinical experts, the following products are not considered to be relevant comparators to CPX-351 for the purpose of the evaluation: Azacitidine is not recommended for patients eligible for HSCT and is typically used in older, unfit AML patients as a palliative therapy. This is a different population to CPX-351 and is therefore not considered a relevant comparator. Midostaurin is not currently used in the NHS and is indicated as an add-on therapy to standard intensive chemotherapy. Midostaurin is therefore not considered a relevant comparator. Gemtuzumab ozogamicin is not currently used in the NHS. As the indication (*de novo* AML) and usage (add-on therapy) differ from the intended use of CPX-351, gemtuzumab ozogamicin is not considered a relevant comparator.
Equity considerations		No equity or equality issues are anticipated.	

A.6 Clinical effectiveness evidence

The primary evidence of efficacy and safety of CPX-351 in newly diagnosed patients with AML is derived from Study 301, a randomised, pivotal Phase III study. Importantly, the pivotal Study 301 is the first randomised study in AML patients to demonstrate a superior survival benefit over the 3+7 regimen. Overall survival and best response rates (CR, CR + CRi) all showed statistically significant improvement in the CPX-351 treatment group relative to the standard 3+7 regimen. Treatment with CPX-351 also enabled a greater proportion of subjects to receive HSCT, with superior outcomes following transplant, including significantly longer overall survival.

Table 3: Clinical effectiveness evidence

Study title	**CLTR0310-301**
Study design	Multicentre, open-label, randomised, parallel-arm, standard therapy-controlled, Phase III
Population	Patients aged 60 to 75 years with untreated high-risk (secondary) AML
Intervention(s)	CPX-351
Comparator(s)	Daunorubicin plus cytarabine (3+7)
Outcomes specified in the decision problem	• **Overall survival (OS), measured from the date of randomisation to death.** • **Response rate, defined as the number of patients who achieved CR or CRi during the treatment phase divided by the total number of patients in the corresponding treatment group.** • **Event-free survival (EFS), defined as the time from study randomisation to the date of induction treatment failure (persistent disease), relapse from complete remission (CR) or complete remission with incomplete platelet or neutrophil recovery (CRi), or death.** • Remission duration, measured from the date of achievement of a remission (CR or CRi) until the date of relapse or death. • Rate of transfer to HSCT after induction treatment was recorded.
Reference to section in submission	B.2.2. (page 34)

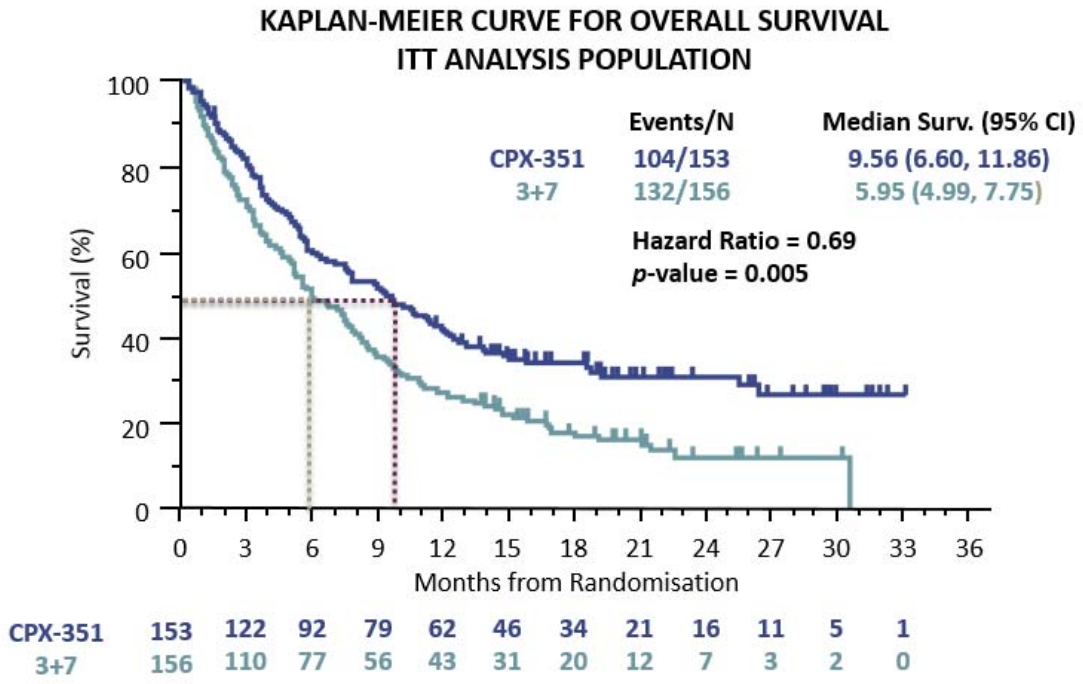
A.7 Key results of the clinical effectiveness evidence

A.7.1 *Overall survival*

The primary endpoint was OS measured from the date of randomisation to death from any cause. CPX-351 demonstrated superior improvement in OS in the

intention-to-treat (ITT) population compared with the standard of care 3+7 treatment regimen (Figure 2). The median survival for the CPX-351 treatment group was 9.56 months compared with 5.95 months for the 3+7 treatment group (Hazard ratio = 0.69, 95% CI = 0.52, 0.90, stratified (one-sided) log-rank test p = 0.003; stratified (two-sided) log-rank test p = 0.005).

Figure 2: Study 301: Kaplan-Meier Curve for Overall Survival, ITT Population



CI, confidence interval; ITT, intent-to-treat; OS, overall survival
Source: (8)

A.7.2 *Response rate: CR; CR + CRi*

A significantly greater proportion of subjects in the CPX-351 treatment group achieved a CR compared with subjects in the 3+7 treatment group (37.3% vs 25.6%, respectively; stratified two-sided p = 0.040). In addition, the proportion of subjects with a CR + CRi was greater in the CPX-351 treatment group than in the 3+7 treatment group (47.7% vs 33.3%, respectively; stratified two-sided p = 0.016).

More patients treated with CPX-351 achieved remission after one induction cycle than with 3+7 (CR: 47/105 [45%] vs 28/100 [28%]; CR+CRi: 58/105 [55%] vs 34/100 [34%], respectively). Remission rates after two induction cycles were similar between

treatment arms (CR: 10/48 [21%] vs 12/51 [24%]; CR+CRi: 15/48 [31%] vs 18/51 [35%], respectively).

A.7.3 *Event-free survival*

EFS was a secondary endpoint in Study 301 and was calculated from the date of randomisation to the date that persistent disease was documented or the date of relapse after CR or death. Subjects in the CPX-351 treatment group demonstrated superior EFS compared with subjects in the 3+7 treatment group. The median EFS was 2.53 months and 1.31 months for subjects in the CPX-351 and 3+7 treatment groups, respectively (1-sided p = 0.011; 2-sided p = 0.011).

A.7.4 *Remission duration*

No clinically significant difference was observed in remission duration between subjects with CR + CRi in the CPX-351 treatment group compared with the 3+7 treatment group. The median remission duration in the CPX-351 treatment group was 6.93 months vs 6.11 months in the 3+7 treatment group (Hazard ratio = 0.77, 1-sided p = 0.147; 2-sided p = 0.294).

A.7.5 *Rate of achieving morphologic leukaemia-free state*

Morphologic leukaemia-free state was achieved by a significantly greater proportion of subjects in the CPX-351 treatment group than the 3+7 treatment group (69.0% vs 55.5%, respectively; p = 0.017).

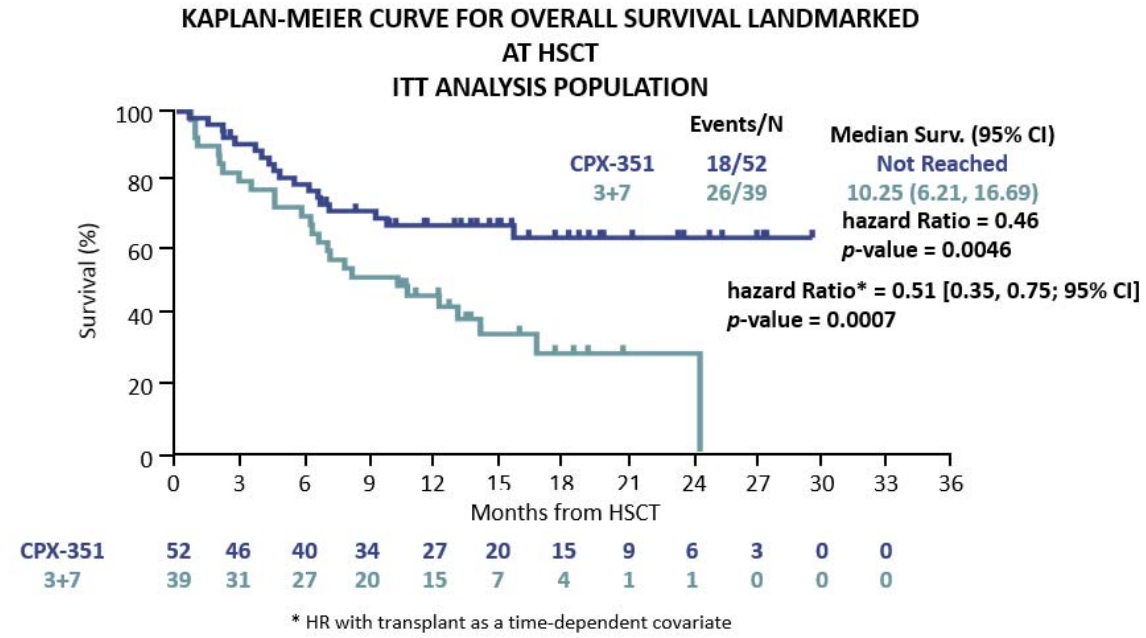
A.7.6 *Patients achieving transplant*

Patients achieving CR are not assumed to be completely free of AML and post-remission treatment with consolidation or HSCT is administered to prevent relapse. However, many patients, despite achieving CR, are deemed not fit enough for HSCT and thereby excluded from this potentially curative treatment (9, 10). Thus, the goal of treatment is to use cytotoxic therapy to induce a CR whilst maintaining patient fitness to enable HSCT.

Treatment with CPX-351 enabled a greater proportion of patients to go on to achieve a HSCT. A total of 52 (34%) patients in the CPX-351 arm received a HSCT compared with 39 (25%) patients treated with 3+7. Outcomes after HSCT strongly favoured subjects in the CPX-351 arm. A Kaplan-Meier analysis of the subjects who

received a transplant, landmarked at the time of stem cell transplant, showed that median survival was not reached in the CPX-351 treatment group, whereas the median survival in the 3+7 treatment group was 10.25 months (HR=0.46, p = 0.009 [1-sided]); Figure 3).

Figure 3: Kaplan-Meier Curve for Overall Survival Landmarked at Stem Cell Transplant



CT, haematopoietic stem cell transplant; Source: Study 301 (11)

A.8 Evidence synthesis

The key evidence that supports the submission for CPX-351 is derived from the large, statistically powered, randomised, controlled, pivotal Phase III study. Although the comparison used in the clinical study was the 3+7 standard intensive chemotherapy regimen, evidence suggests that in UK practice, the combination of 3+10 is widely used.

Published results indicate that the additional cytarabine received in the UK 3+10 schedule does not translate into an overall improvement in efficacy versus 3+7. This is consistent with UK clinical advisory board feedback, which supports equivalence of the two schedules and also the British Committee for Standards in Haematology (BCSH) AML guidelines, which recommend that patients not eligible or unwilling to

participate in the NCRI studies should be offered daunorubicin and cytarabine 3+10 or 3+7 intensive chemotherapy (12).

In addition, qualitative research was conducted to demonstrate comparativeness of treatment options.

To determine a relative treatment effect between CPX-351 and 3+10, an indirect treatment comparison (ITC) between treatments (3+7, 3+10 and CPX-351) was attempted. The efforts began with a systematic literature review (SLR) to identify and assess the available data and progressed to the evaluation of the feasibility to compare the effects of those treatments regimens in the population of untreated high-risk (secondary) AML patients aged 60-75 years

Search results revealed limited clinical evidence to allow for comparisons other than that between CPX-351 and 3+7. Besides the sponsor's studies on CPX-351 and 3+7, the only other identified study (13) compared two 3+7 regimens with doses of daunorubicin (45mg/m^2 vs 90mg/m^2) that differ to that used in the 3+7 arm of the 301 study (60mg/m^2)(14-17). Due to various reasons explained in detail in Section B, the effect modifiers were reasoned to introduce too great a degree of heterogeneity into an indirect comparison, and it was concluded that this could obscure the true relative treatment effect between CPX-351 and 3+7.

Figure 4: Relationship between treatment arms in the studies identified through a systematic review of the literature

A qualitative assessment of the indirect treatment effect between CPX-351, 3+7 and 3+10 was conducted. In this pragmatic review, a selection of studies with 3+7 or 3+10 treatment arms that did not report results for the specific population studied in the pivotal trial were examined (for which the principal reason for exclusion was a failure to report results exclusively for the secondary AML sub-population).

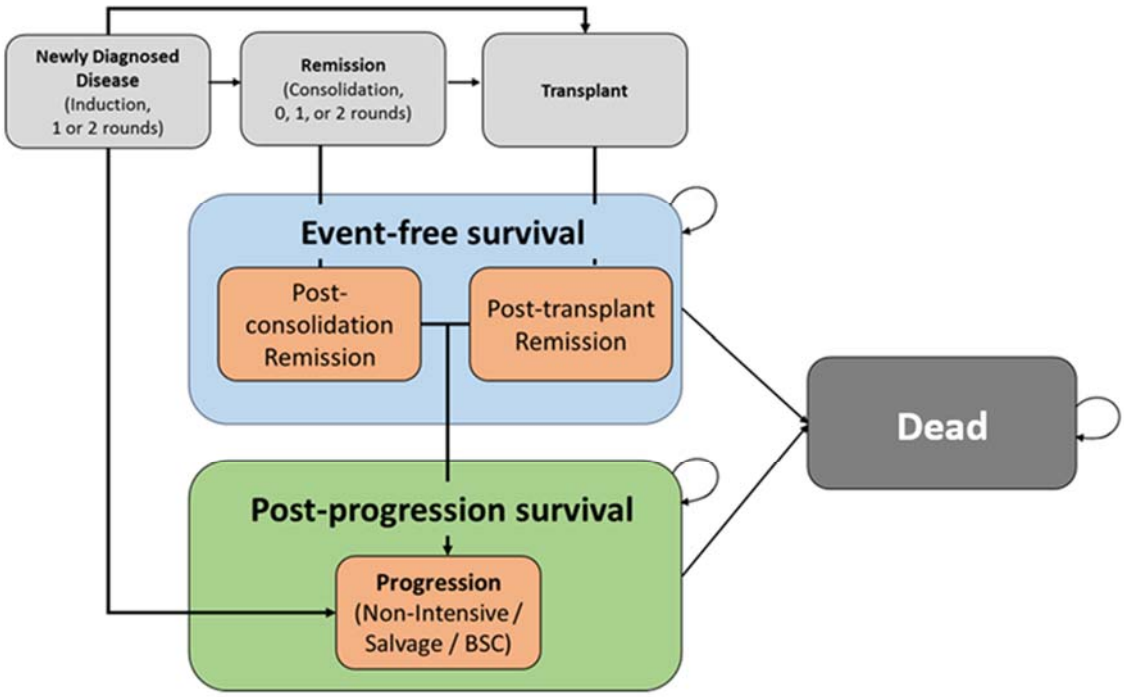
Results from the quantitative and qualitative research conclude that the identified evidence does not allow robust comparison of survival between 3+7 and 3+10 in the target population. Differences between interventions studied and patient characteristics preclude meaningful comparisons of efficacy as the evidence base for 3+10 in this population appears to be very limited. However, as described above, both UK guidelines and clinical advisory board feedback, indicate that the additional cytarabine received in the UK 3+10 schedule does not translate into an overall improvement in efficacy versus 3+7 (18).

A.9 Key clinical issues

- Study 301 was open-label in design because the unique colour of CPX-351 enabled visible identification of differences between the treatment administration regimens, rendering blinding impossible.
- Study 301 included a North America-based AML population; hence, no UK-specific AML population was studied.
- The 3+7 control arm is the international standard intensive chemotherapy regimen while UK practice uses a 3+10 schedule.
- No health-related quality of life (HRQL) measurements were collected.
- No testing of minimal residual disease was performed.

Overview of the economic analysis

Figure 5: Model diagram – B.3.2.3 (page 86)



BSC, best supportive care

A.10 Incorporating clinical evidence into the model

A survival-based cohort model was developed to predict disease progression, health benefits, and costs of treatment with CPX-351 versus treatment with the 3+7

regimen in the patient population by tracking their outcomes during movement through a series of health states over a lifetime time horizon. Patient characteristics for the modelled cohort were taken from the patient population in the Phase III clinical trial. To implement data from the study, *post-hoc* statistical analyses were conducted using individual patient-level data from the trial. Multivariate logistic regression analyses (binomial or multinomial, as appropriate) were performed to estimate the probability of clinical pathway outcomes. These results were used to estimate the proportion of patients that would follow each pathway in the cost-effectiveness model.

Patients who achieve response, OS and EFS are not tracked until a certain amount of time has elapsed. In effect, a horizontal time-shift occurs, during which no mortality occurs for patients from the start of the model until the time at which their EFS and OS starts to be tracked. For patients who respond and receive a transplant, EFS and OS are tracked starting at the time of transplant. Similarly, for patients who respond but do not receive a transplant, EFS and OS are tracked starting at the time of last consolidation therapy, including patients who receive no consolidation but do respond and do not receive HSCT. Multivariate linear regressions are used to calculate these time-shifts for each predetermined subgroup run through the model. Note that for patients who do not achieve response, no time-shifts should be considered in relation to their survival; therefore, the parametric curves assessing the successive composite endpoints (OS, OS or HSCT, and OS or HSCT or progression) are applied from the beginning of the model.

In order to define patient survival beyond the trial period and estimate costs and health benefits over the model time horizon, survival data (OS and EFS) from the pivotal Phase III trial CLTR0310-301 (11) were extrapolated. Extrapolation was conducted by exploring standard parametric fits to the Kaplan-Meier data from the study, including Weibull, log-normal, log-logistic, exponential, generalised gamma, and Gompertz according to the methodology outlined in the NICE DSU document (19). The Akaike information criterion (AIC) and Bayesian information criterion (BIC) were used to choose among the different model fittings. Once fitted, the resulting parametric survival curves were modelled separately per treatment arm and extended by integrating the following predictors: treatment, AML type, number of

rounds of induction therapy, and number of rounds of consolidation (when applicable), as data permitted.

The model also captures the possibility of experiencing adverse events (AE) with each treatment, rates for which were sourced from the Phase III trial CLTR0310-301 (11). Grade 3-5 AEs with at least 5% frequency of occurrence in at least one treatment arm in the trial were included in the model. AEs were assumed to occur in Year 1 due to limitations in data availability and AE costs were applied as a one-off cost.

Age- and sex-specific mortality rates for the general UK population were also calculated for each period. In any period and for any group where modelled OS suggested lower mortality than the general population, all-cause age- and gender-adjusted mortality based on the UK Office of National Statistics was used instead of the study based estimate (20). To assess the impact of post-transplant mortality on the cost-effectiveness results, a scenario adapting the general mortality based on the findings of Martin *et al*., 2010 (21) was performed whereby the general mortality was adjusted by means of a hazard ratio equal to 2.3.

A.11 Key model assumptions and inputs

Key model assumptions are as follows:

- The model was run based on subgroups defined by patient characteristics, number of rounds of induction therapy (1 or 2), and number of rounds of consolidation therapy (0, 1, or 2).
- Linear regressions were used to calculate the times at which OS and EFS start to be tracked for each subgroup above, accounting for the time spent in induction and consolidation regimens and in transplant.
- For patients who respond and receive a transplant, OS and EFS were tracked from the time of transplant.

- For patients who do not achieve response (i.e., induction failure), OS and EFS were tracked from the beginning of the model.
- Background mortality was applied when its probability exceeded the computed probability of death from the parametric OS equations.

Further assumptions are listed in Table 4.

Table 4: Key model assumptions and inputs

Model input and cross reference	Source/assumption	Justification
Tracking overall survival (OS) B.3.6.2. (page 128)	Separate OS + EFS curves were used according to therapy received, remission status and whether transplant was achieved.	These assumptions were based on data from the Phase III study in CPX-351.
Dosing B.3.6.2. (page 128)	Patients receive full doses of drugs over the durations of active treatment. Dose reduction was not considered.	Dose reduction was not necessary for standard chemotherapy doses (22).
Adverse events B.3.6.2. (page 129)	AEs were included if grade 3 to 5 and occurred in ≥5% of patients in at least one treatment arm.	Only clinically significant adverse events were included.
	AEs were assumed to occur in year 1 and their costs were applied as a one-off cost.	This is due to limitations in available data about when events occurred.
Healthcare resource use B.3.6.2. (page 129)	Types of follow-up procedures are based on NCCN guidelines for AML and clinical expert opinion.	Lack of resource use data available in current UK and European AML guidelines.
	50% of administrations of consolidation therapy for CPX-351 occurred in an outpatient setting.	Data from the Phase III study in CPX-351.
	All costs associated with the management of adverse events were incurred in an inpatient setting.	Data from the Phase III study in CPX-351.
HRQL B.3.6.2. (page 129)	Utility values used in the model were based on a time-trade-off study that was conducted in members of the UK general population.	The pivotal clinical trial did not collect HRQL data.

EFS, event-free survival; HRQL, health-related quality of life; OS, overall survival

A.12 Base-case results

Table 5: Base-case results (deterministic) – B.3.7 (page 133)

Treatment	Total costs (£)	Total LYG	Total QALYs	Incremental costs (£)	Incremental LYG	Incremental QALYs	ICER versus baseline (£/QALY)	ICER incremental (£/QALY)
3+7	[redacted]	[redacted]	[redacted]	-	-	-	-	-
CPX-351	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]

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

Table 6: Base-case results with PAS (deterministic) – Table 52 (page 133)

Treatment	Total costs (£)	Total LYs	Total QALYs	Incremental costs (£)	Incremental LYG	Incremental QALYs	ICER versus baseline (£/QALY)	ICER incremental (£/QALY)
3+7	[redacted]	[redacted]	[redacted]	-	-	-	-	-
CPX-351	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	46,631	46,631

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

A.13 Probabilistic sensitivity analysis

Table 7: Base-case results (probabilistic) – B.3.8 (page 138)

Treatment	Total mean costs (£)	Total mean LYG	Total mean QALYs	Mean incremental costs (£)	Mean incremental LYG	Mean incremental QALYs	Mean ICER versus baseline (£/QALY)	Mean ICER incremental (£/QALY)
3+7	84,019	[redacted]	[redacted]	-	-	-	-	-
CPX-351	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]

Figure 6: Scatterplot of probabilistic results – B.3.8 (page 139)

A.14 Key sensitivity and scenario analyses

Figure 7: Tornado diagram – B.3.8 (page 143)

Table 8: Key scenario analyses

Scenario and cross reference	Scenario detail	Brief rationale	ICER (impact on base-case ICER) [£/QALY]
Base-case ICER (£/QALY)			[redacted]
Extrapolation of modelled outcomes to younger patient population (i.e. mean age 54) (Section B3.6.3, Table 50, page 129, section B.3.8, Table 60, page 143)	Extension of model to younger population – various percentages of modelled cohort <60 years of age	In the case of an age-agnostic marketing authorisation, younger patients with high-risk (secondary) AML would be eligible for treatment with CPX-351. To assess the cost-effectiveness of treating younger patients with CPX-351, first a SLR was conducted to identify relative effect estimates comparing younger versus older populations with previously untreated high-risk (secondary) AML patients. Effect of age on response is not treatment-related; and 2) the baseline patient characteristics and treatment experience of patients above 60 are similar to patients below 60.	30% of modelled cohort aged <60 years: [redacted]
Alter percentage of patients treated with	In second induction, 50% of CPX-351 patients receive	CPX-351 could be administered as an outpatient and therefore, the	[redacted]

Scenario and cross reference	Scenario detail	Brief rationale	ICER (impact on base-case ICER) [£/QALY]
CPX-351 in outpatient setting (Section B3.6.3, Table 50, page 129, section B.3.8, Table 63, page 144)	therapy in an outpatient setting and 100% of CPX-351 patients receive consolidation in an outpatient setting.	assumption was made for second induction and consolidation.	
Shorter time horizon (Section B3.6.3, Table 50, page 129, section B.3.8, Tables 64 and 65, page 144)	Estimating ICER at 5-year and 10-year time horizons	Shorter time horizons were tested to assess the impact of the long-term survival extrapolation.	5-year time horizon: [redacted] 10-year time horizon: [redacted]
Population mortality post-HSCT (Section B3.6.3, Table 50, page 129, section B.3.8, Table 67, page 145)	Adjust general population mortality for HSCT patients 2.3-fold	Background mortality increased 2.3-fold in patients post-transplant based on Martin *et al.*, 2010 (21).	[redacted]

A.15 Innovation

CPX-351 is the only new AML treatment for 40 years with evidence demonstrating a significant survival benefit in patients with high-risk (secondary) AML compared with existing high-intensity chemotherapy regimens. CPX-351 advances AML treatment through the use of a novel drug delivery technology that reformulates daunorubicin and cytarabine at a synergistic fixed molar ratio within stable liposomes to produce pharmacologic advantages that enhance efficacy without increasing toxicity. The liposomes are directly and preferentially internalised by malignant myeloblasts, potentially bypassing membrane associated efflux pumps that are a major source of chemotherapy resistance. The stability of the liposomes results in their persistence in the circulation, which extends the potential exposure of AML blasts to chemotherapy far beyond that of conventional cytarabine and daunorubicin regimens. Moreover, the liposomes are confined to the vascular space after administration, reducing drug exposure to normal tissues. The net result is an improvement in efficacy marked by increases in survival and response rate, in a population of patients expected to have short survival, with a higher proportion of patients referred to transplant for potentially curative post remission treatment. Furthermore, CPX-351 could transform administrative practice by allowing patients to spend less time in the hospital and more time continuing with their daily lives compared with 3+10 standard intensive chemotherapy.

For further information see the section on innovation in the main submission: B.2.12 (pages 70-73).

A.16 'End-of-Life' criteria

Table 9: 'End-of-Life' criteria – B.2.13.3 (pages 78-80)

Criterion	Data available
The treatment is indicated for patients with a short life expectancy, normally less than 24 months	Survival results from the CLTR0310-301 Phase III trial show: 1.The median OS for the CPX-351 group was 9.56 months compared with 5.95 months for the 3+7 group (HR=0.69, 95% CI: 0.52, 0.90, 2-sided log-rank test p=0.005). 2. The modelled mean life expectancy from the cost-effectiveness analysis shows that the mean survival to be [redacted]

Criterion	Data available
	3. Registry results from the UK, Sweden and Denmark consistently show a median OS for secondary AML <1 year. This poor survival is seen regardless of age at diagnosis.
There is sufficient evidence to indicate that the treatment offers an extension to life, normally of at least an additional 3 months, compared with current NHS treatment	Survival results from the CLTR0310-301 Phase III trial show: The difference in median OS for the CPX-351 group was 3.61 months compared with the 3+7 group (HR=0.69, 95% CI: 0.52, 0.90, 2-sided log-rank test p=0.005). In AML, HSCT offers the potential for cure and the survival benefit from CPX-351 versus 3+7 was associated with more patients receiving HSCT and better OS post HSCT (HR=0.46, 95% CI: 0.24, 0.89, p=0.0007). In contrast to the median OS which will not fully capture this benefit for CPX-351 vs 3+7, the modelled mean life expectancy from the trial will include survival benefits due to cure: 1 The modelled mean life expectancy from the cost-effectiveness analysis for the entire patient population shows that the difference to be [redacted] 2 The modelled mean life expectancy from the cost-effectiveness analysis for the transplanted patient population shows that the difference to be [redacted]

AML, acute myeloid leukaemia; CI, confidence interval; HSCT, haematopoietic stem cell transplant; HR, hazard ratio; MDS, myelodysplastic syndrome; NHS, National Health Service; OS, overall survival.

A.17 Budget impact

Table 10: Budget impact – Budget impact analysis (page 11, 24)

	Company estimate	Cross Reference
Number of people in England, Wales, NI who would have treatment	[redacted] incident patient population (all ages) [redacted] over 60s	Budget impact analysis, Eligible population: (Table 5 Page 11)
Average treatment cost per person	Based on experience from the Phase III pivotal study, and based on standard practice in England, it is estimated that a 1.8m^2 patient will require [redacted] vials. Average cost of a course of treatment is estimated to be [redacted]	Study 301 data used in cost effectiveness model
Estimated annual budget impact on the NHS in England	Year 1: £2,424 Year 2: £6,152 Year 3: £8,113 Year 4: £8,551 Year 5: £8,553 (All values are in 1000's)	Budget impact analysis, Estimated annual budget impact assuming adults of all ages: (Table 23, page 24)

A.18 Interpretation and conclusions of the evidence

The introduction of CPX-351 into the treatment pathway within UK, for people with high-risk (secondary) AML, presents opportunity for a small population with high unmet need and a corresponding high mortality.

- Due to its underlying biology, high-risk AML is associated with chemo-resistance and outcomes are universally poor regardless of age (23). Survival is considerably less than 2 years (16).

- High-risk AML has the worst outcomes of all the AML diagnostic subtypes (24) and therefore represents an area of significant unmet medical need.

- Existing standard intensive chemotherapy regimens typically require daily infusions for up to ten days (25, 26), in a hospital inpatient setting, thereby presenting a substantial burden to patients, carers and healthcare systems.

As demonstrated, in study 301**,** CPX-351 improved outcomes versus standard intensive chemotherapy in patients with high-risk AML:

- CPX-351 demonstrated a significant improvement in overall survival (median 9.56 vs 5.95 months, respectively; p=0.005) (16).

- CPX-351 was associated with significantly higher remission rates (p=0.016) and a decreased risk of 60-day mortality (16).

- More patients treated with CPX-351 received a haematopoietic stem cell transplant (HSCT) (34% vs 25%), with improved survival post-HSCT (27).

- The rate of adverse events per patient year was lower with CPX-351, giving patients increased chance of benefit without an increase in toxicity (28).

CPX-351 brings additional value to patients, carers and the NHS.

- The 90-minute infusion of CPX-351 allows for outpatient administration, providing the opportunity to free up hospital beds and reduce costs compared with inpatient-only standard intensive chemotherapy regimens (29).

- CPX-351 meets the ‘End-of-Life’ criteria, by improving survival vs standard intensive chemotherapy for greater than 3 months in a high-risk AML population, where survival is considerably less than 2 years (8).

Due to the small identifiable patient population, uptake of CPX-351 can be managed within a well-defined budget and be shown to be a cost-effective use of NHS resource.

References

1. Estey E, Dohner H. Acute myeloid leukaemia. Lancet. 2006;368(9550):1894-907.
2. Robak T, Wierzbowska A. Current and emerging therapies for acute myeloid leukemia. Clin Ther. 2009;31 Pt 2:2349-70.
3. Li X, Xu S, Tan Y, Chen J. The effects of idarubicin versus other anthracyclines for induction therapy of patients with newly diagnosed leukaemia. Cochrane Database Syst Rev. 2015(6):CD010432.
4. Appelbaum FR, Gundacker H, Head DR, Slovak ML, Willman CL, Godwin JE, et al. Age and acute myeloid leukemia. Blood. 2006;107(9):3481-5.
5. Kantarjian H, O'Brien S, Cortes J, Giles F, Faderl S, Jabbour E, et al. Results of intensive chemotherapy in 998 patients age 65 years or older with acute myeloid leukemia or high-risk myelodysplastic syndrome: predictive prognostic models for outcome. Cancer. 2006;106(5):1090-8.
6. Tardi P, Wan CP, Mayer L. Passive and semi-active targeting of bone marrow and leukemia cells using anionic low cholesterol liposomes. J Drug Target. 2016;24(9):797-804.
7. Arber DA, Orazi A, Hasserjian R, Thiele J, Borowitz MJ, Le Beau MM, et al. The 2016 revision to the World Health Organization classification of myeloid neoplasms and acute leukemia. Blood. 2016;127(20):2391-405.
8. Lancet JE, Uy G, Cortes JE, Newell LF, Lin TL, al. e. Final results of a phase III randomized trial of VYXEOS™ (CPX-351) vs 7 + 3 in older patients with newly diagnosed high-risk (secondary) AML [abstract/oral presentation], ASCO Annual Meeting (June 2016). J Clin Oncol. 2016;34(15_suppl (May 20 Supplement)):7000.
9. Döhner H, Estey E, Grimwade D, Amadori S, Appelbaum FR, Büchner T, et al. Diagnosis and management of AML in adults: 2017 ELN recommendations from an international expert panel. Blood. 2017;129(4):424-47.
10. Cornelissen JJ, van Putten WL, Verdonck LF, Theobald M, Jacky E, Daenen SM, et al. Results of a HOVON/SAKK donor versus no-donor analysis of myeloablative HLA-identical sibling stem cell transplantation in first remission acute myeloid leukemia in young and middle-aged adults: benefits for whom? Blood. 2007;109(9):3658-66.
11. Jazz Pharmaceuticals Inc. Clinical Study Report: Study CLTR0310-301. Phase III, Multicenter, Randomized, Trial of CPX-351 (Cytarabine: Daunorubicin) Liposome Injection versus Cytarabine and Daunorubicin in Patients 60 -75 years of Age with Untreated High Risk (Secondary) AML. 2017 2/24/2017. Report No.: V1.0.
12. Milligan DW, Grimwade D, Cullis JO, Bond L, Swirsky D, Craddock C, et al. British Committee for Standards in Haematology: Guidelines on the management of acute myeloid leukaemia in adults. Br J Haematol. 2006;135(4):450-74.
13. Cheson BD, Bennett JM, Kopecky KJ, Buchner T, Willman CL, Estey EH, 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(24):4642-9.
14. Lancet JE, Rizzieri D, Schiller GJ, Stuart RK, Kolitz JE, Solomon SR, et al. Overall survival with CPX-351 versus 7+3 in older adults with newly diagnosed,

therapy-related acute myeloid leukemia: Subgroup analysis of a phase 3 study. Haematologica. 2017;102:214-5.
15. Lancet JE, Rizzieri D, Schiller GJ, Stuart RK, Kolitz JE, Solomon SR, et al. Overall survival (OS) with CPX-351 versus 7+3 in older adults with newly diagnosed, therapy-related acute myeloid leukemia (tAML): Subgroup analysis of a phase III study. Journal of Clinical Oncology Conference. 2017;35(15 Supplement 1).
16. Lancet JE, Uy GL, Cortes JE, Newell LF, Lin TL, Ritchie EK, et al. Final results of a phase III randomized trial of CPX-351 versus 7+3 in older patients with newly diagnosed high risk (secondary) AML. Journal of Clinical Oncology Conference. 2016;34(no pagination).
17. Medeiros BC, Lancet JE, Cortes JE, Newell LF, Lin TL, Ritchie EK, et al. Analysis of efficacy by age for patients aged 60-75 with untreated secondary acute myeloid leukemia (AML) treated with CPX-351 liposome injection versus conventional cytarabine and daunorubicin in a phase III trial. Blood Conference: 58th Annual Meeting of the American Society of Hematology, ASH. 2016;128(22).
18. Preisler H, Davis RB, Kirshner J, Dupre E, Richards Fd, Hoagland HC, et al. Comparison of three remission induction regimens and two postinduction strategies for the treatment of acute nonlymphocytic leukemia: a cancer and leukemia group B study. Blood. 1987;69(5):1441.
19. Latimer N. Survival analysis for economic evaluations alongside clinical trials - extrapolation with patient-level data 2011. Contract No.: 14.
20. Office of National Statistics (ONS). National life tables, UK: 2014 to 2016. 2017.
21. Martin PJ, Counts GW, Jr., Appelbaum FR, Lee SJ, Sanders JE, Deeg HJ, et al. Life expectancy in patients surviving more than 5 years after hematopoietic cell transplantation. J Clin Oncol. 2010;28(6):1011-6.
22. Ofran Y, Tallman MS, Rowe JM. How I treat acute myeloid leukemia presenting with preexisting comorbidities. Blood. 2016;128(4):488-96.
23. 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(3):208-14.
24. Roman E, Smith A, Appleton S, Crouch S, Kelly R, Kinsey S, et al. Myeloid malignancies in the real-world: Occurrence, progression and survival in the UK's population-based Haematological Malignancy Research Network 2004-15. Cancer Epidemiol. 2016;42:186-98.
25. Burnett AK, Russell NH, Hills RK, Kell J, Cavenagh J, Kjeldsen L, et al. A randomized comparison of daunorubicin 90 mg/m2 vs 60 mg/m2 in AML induction: results from the UK NCRI AML17 trial in 1206 patients. Blood. 2015;125(25):3878-85.
26. Burnett AK, Russell NH, Hills RK, Kell J, Nielsen OJ, Dennis M, et al. A comparison of clofarabine with ara-C, each in combination with daunorubicin as induction treatment in older patients with acute myeloid leukaemia. Leukemia. 2017;31(2):310-7.
27. Lancet JE, Hoering A, Uy GL, Cortes JE, Newell LF, Lin TL, et al. Survival following allogeneic hematopoietic cell transplantation in older high-risk acute myeloid leukemia patients initially treated with CPX-351 liposome injection versus standard cytarabine and daunorubicin: Subgroup analysis of a large phase III trial.

Blood Conference: 58th Annual Meeting of the American Society of Hematology, ASH. 2016;128(22).
28. Medeiros BC, Lancet JE, Cortes JE, Newell LF, Lin TL, Ritchie EK, et al. Rates of Adverse Events per Patient-Year in a Randomized, Phase III Study of CPX-351 Versus 7+ 3 in Older Adults with Newly Diagnosed, Treatment-Related Acute Myeloid Leukemia (AML) or AML with Myelodysplasia-Related Changes. Am Soc Hematology; 2017.
29. Kolitz JE, Strickland SA, Cortes J, Hogge D, Lancet JE, Goldberg SL, et al. Efficacy by consolidation administration site: Subgroup analysis of a phase III study of CPX-351 versus 7+3 in older adults with newly diagnosed, high-risk acute myeloid leukemia (AML). J Clin Oncol. 2017;35(no. 15_suppl):7036.

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

Liposomal cytarabine and daunorubicin for untreated acute myeloid leukaemia ID1225

Dear Jazz Pharmaceuticals

The Evidence Review Group, Centre for Reviews and Dissemination and Centre for Health Economics, University of York, and the technical team at NICE have looked at the submission received on 26 April 2018 from Jazz Pharmaceuticals. 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 **Tuesday 5 June 2018.** Your response and any supporting documents should be uploaded to NICE Docs/Appraisals [embed NICE DOCS LINK].

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 [redacted] in turquoise, and all information submitted as [redacted] 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

Alex Filby
Technical Adviser – Appraisals
Centre for Health Technology Evaluation

NICE National Institute for Health and Care Excellence

Level 1A
City Tower
Manchester
M1 4BT
United Kingdom

+44 (0)300 323 0140

On behalf of:
Frances Sutcliffe
Associate Director – Appraisals
Centre for Health Technology Evaluation

Section A: Clarification on effectiveness data

Study 301

A1. **Priority question:** The data utilised in the analyses presented in the company submission and used in the model appears to be based on what is now a relatively old data cut dated December 2015. On page 38 it is stated that trial follow up is designed to continue for 5 years after randomisation.

(i) Please confirm whether follow up is continuing and whether a more recent data cut is available.

(ii) If a more recent data cut is available, please update the survival analyses (overall survival (OS) and event-free survival (EFS)), i.e. update figures 4-7 with events/N, median values, hazard ratio (HR), 95% confidence intervals (CIs) and p-values.

(iii) Please also state how many trial participants are currently known to be alive and the duration of their survival.

A2. **Priority question:** Please provide Kaplan-Meier curves for EFS from time of haematopoietic stem cell transplantation (HSCT) among patients who received HSCT for the intention-to-treat (ITT) analysis population, with events/N, median values, HR, HR with transplant as time-dependent covariate, 95% CIs and p-values (as per company submission figure 7).

A3. **Priority question:** Please present remission duration data (as described in question A2) separately for patients who had a HSCT and patients who did not receive HSCT, using the most up to date data available.

A4. Imputation of missing data for the primary endpoint (OS) is described in table 9 of the company submission. Please specify how many patients had (a) a missing month and day imputed and (b) a missing day imputed, in each arm.

A5. **Priority question:** Please state what criteria were used to assess eligibility for HSCT in the CLTR0310-301 trial. Please comment on whether it can be considered generalisable to the UK.

NICE National Institute for Health and Care Excellence

Level 1A
City Tower
Manchester
M1 4BT
United Kingdom

+44 (0)300 323 0140

A6. **Priority question:** Were there any patients eligible for HSCT who did not receive it? If so, please provide reasons for not receiving HSCT (e.g. lack of available donor), as well as characteristics and the number and percentage of patients as per table 12 in the company submission.

A7. Please provide results of relevant statistical significance tests comparing characteristics of patients undergoing HSCT, in an additional column to company submission table 12.

A8. **Priority question:** Please provide details of any dose reductions or delays, including the number of patients and reason for the dose reduction/delay, by treatment group.

A9. The company submission states "The higher observed rate of serious [adverse events] AEs may be due, in part, to the greater proportion of patients in the CPX-351 arm who received consolidation in the outpatient setting compared with the 3+7 arm, because a move to the hospital setting is one of the criteria for classifying an AE as serious." However, the clinical study report (CSR) (section 9.5.4.3) states that [redacted]). The Medeiros poster (reference 8 in company submission) also shows a higher median rate of serious treatment-emergent adverse events (TEAEs) per patient-year in the treatment group, suggesting that this difference may not be explained by differences in treatment duration either. Please provide any other clinically plausible justifications for this difference in observed rates of serious adverse events (SAEs).

A10. Please clarify whether the adverse events reported in the submission (tables 15-18) are only those that occurred during the treatment phase of the trial. Please provide details of grade 3-5 adverse events that occurred during the follow-up phase of the trial, by treatment group (separately for patients who received HSCT and patients who did not receive HSCT).

A11. **Priority question:** Please provide further details of 'adverse events of special interest', which are only briefly summarised in the text (page 62-63), i.e. present tables similar to table 16 for 'infection-related adverse events', 'bleeding-related adverse events' and 'cardiac adverse events'.

A12. Please define what constitutes 'other' reasons for withdrawal from treatment in the patient disposition (figure 2) of the company submission appendices.

A13. Please confirm whether there were any patients in the 3+7 arm of the trial who subsequently received treatment with CPX-351 (i.e. any crossover).

A14. **Priority question:** Please state whether there are any other subtypes of high-risk acute myeloid leukaemia (AML) that would be eligible for CPX-351 in practice (under

Level 1A
City Tower
Manchester
M1 4BT
United Kingdom

+44 (0)300 323 0140

the anticipated marketing authorisation), which were not included in CLTR0310-301 (e.g. acute promyelocytic leukaemia). If so, please explain why they were not included in the trial.

Systematic review

A15. **Priority question:** Please justify restricting the systematic review inclusion criteria to patients aged 60-75 years, when the anticipated marketing authorisation is 'adults' (no age restriction)?

A16. **Priority question:** Were there any CPX-351 trials that were excluded only because they included participants under 60 years, but would otherwise have met the review inclusion criteria? If so, please provide references.

Section B: Clarification on cost-effectiveness data

General

B1. Please provide the "data on file" reference for the key opinion leader (KOL) advisory board (page 90 of the company submission).

B2. **Priority question**. Please clarify what "Pathways A to H" refer to (introduced on page 214 of the appendix to the company submission). Please provide a diagram of the pathway referred to in appendix M.

Model structure

B3. Please provide additional clarity on how patients whose disease did not respond to treatment were modelled, with regard to the following points:

(i) Please clarify how patients whose disease did not respond to treatment who had a transplant were implemented in the model.

(ii) From inspection of the executable model, it appears that the time of transplant in patients whose disease did not respond to treatment was based on an analysis of time to HSCT or OS, while the time of transplant in patients whose disease did respond to treatment was based on the mean time of transplant. Please confirm that this is the case, and comment on why two different methods were used. What was the mean time for treatment non-response to transplant?

(iii) Please provide information on the treatments received by this group while waiting for the transplant in the 301 trial, and whether this is generalisable to UK clinical practice.

NICE National Institute for Health and Care Excellence

Level 1A
City Tower
Manchester
M1 4BT
United Kingdom

+44 (0)300 323 0140

Clinical data

B4. The company submission describes a hazard ratio estimated from data reported by Martin (2010) that was applied to general population mortality in a scenario analysis (page 100).

Please provide additional information, including a description and example calculations if necessary, of how the 30% reduction in life expectancy reported by Martin (2010) was used to estimate the hazard ratio.

B5. **Priority question**. In the model it is assumed that there is a different post-HSCT OS curve for patients receiving CPX-351 than for patients those receiving 3+7. This is one of the most important assumptions in the model and a key driver of the model because patients receiving CPX-351 experience significantly improved survival post-HSCT, as demonstrated by figure 13, page 97 of the company submission. However, there is little difference in EFS post-transplant between the two treatment groups (figure 16, page 99 of the company submission).

(i) Please provide an interpretation of these clinical data, and comment on the plausibility of CPX-531 providing an OS but not an EFS benefit after transplant.

(ii) Please provide an additional scenario in the executable model based on an analysis of post-transplantation OS that combines data from both treatment groups (i.e. no treatment effect on OS post-transplant). Please include all parametric distributions, using the later data cut of the CLTR0310-301 trial (see question A1), if available.

B6. **Priority question:** Please provide Kaplan–Meier curves (OS and EFS, with descriptive statistics presented as per figure 7) for the following groups of patients:

(i) Patients who had a complete remission (CR)/complete remission with incomplete platelet or neutrophil recovery (CRi) and received transplant

(ii) Patients who had a CR/CRi and did not receive transplant

(iii) Patients who did not have a CR/CRi and received transplant

(iv) Patients who did not have a CR/CRi and did not receive transplant.

Survival analysis

B7. The categories of the AML types are not consistent for estimating the percentages of patients who follow each clinical pathway (e.g. the analysis for remission post-induction uses 5 categories of AML (analysis 2, appendix M.2.1), but the analysis for

Level 1A
City Tower
Manchester
M1 4BT
United Kingdom

+44 (0)300 323 0140

rounds of consolidation (analysis 3, appendix M.2.1) in patients who had a remission uses 2 categories of AML).

(i) Please provide justification for regrouping the categories in the different analyses, and comment on whether it was prespecified in the statistical analysis plan.

(ii) Please provide results of the logistic regression on the need for a transplant in patients who had a remission and consolidation rounds (similar to table 52 and table 53 in appendix M, pages 270-271)

(iii) Include a scenario analysis in the executable model that incorporates the results of the analysis with the 5 categories of AML.

B8. **Priority question:** The covariates used in the different regression models fitted to EFS and OS are not consistent. Specifically, in the regression models for post-consolidation progression-free survival (PFS), post-consolidation OS, and OS among patients who did not have a remission, 5 categories of AML were used, whereas in the post-HSCT EFS and post-HSCT OS regression models 2 categories were used.

(i) Please provide a justification for regrouping the categories in the different analyses, and comment on whether it was prespecified in the statistical analysis plan.

(ii) Please provide results for post-HSCT PFS and post-HSCT OS using the five categories of AML (similar to table 76 to table 81 in appendix M, page 305-313).

(iii) Please provide a scenario analysis in the executable model that incorporates post-HSCT EFS and post-HSCT OS using 5 categories of AML.

B9. The process of selecting the survival models is described in appendix M.1.7 and M.4 of the company submission, which reports that it involved both statistical and clinical considerations. The range of statistical tests of fit are extensive, and the results are comprehensively provided. However, the company submission stated that the models were also selected according to their clinical plausibility, specifically to validate projected survival after the trial period and the median estimate of survival.

Please provide additional information on this assessment, including how it was assessed, any external datasets that were used in the assessment, and the outcomes of the assessments.

B10. **Priority question**. A number of parametric distributions were explored for EFS and OS. However, a limited number of distributions were implemented in the executable model. For example, for OS among patients in Pathways G and H, one distribution

Level 1A
City Tower
Manchester
M1 4BT
United Kingdom

+44 (0)300 323 0140

(lognormal) was implemented (page 236, appendix M). The model fit statistics are very similar for lognormal, with similar visual fit.

(i) Please justify the selection of the lognormal distribution over the other distributions.

(ii) Please provide an executable model that incorporates all fitted parametric distributions for PFS and OS. Where possible this should use any updated cut of the trial data (see question A1). If this is not feasible please provide fitted parametric distributions for post-HSCT OS based on the later data cut incorporating them into the executable model as scenario analyses.

Resource use

B11. Please provide additional information on the following points, regarding the setting for administration of induction and consolidation therapy:

(i) On what basis were the assumptions about the proportion of CPX-351 infusions that could be delivered in an outpatient setting made?

(ii) The model assumes that 50% of CPX-351 patients would receive consolidation therapy as outpatients. What proportion of patients received consolidation therapy as an outpatient in the CLTR0310-301 trial?

(iii) Please provide the data on file (page 73 of the company submission) regarding the proportion of US patients who received CPX-351 in an outpatient setting.

(iv) Please provide the actual proportion of patients receiving CPX-351 as outpatients in the scenario analysis (table 63, page 144 of the company submission).

B12. Please provide additional information on the following points, regarding second line therapy:

(i) In the model, patients could only receive one additional line of therapy post progression (either salvage therapy, best supportive care or low intensity chemotherapy). The CancerMPact report (reference 89, company submission) makes reference to a number of relapses that occur in patients with AML, suggesting that the assumption of only one further line of therapy is not appropriate. Please comment on whether you might expect patients would receive more than one type or line of therapy after progression.

(ii) **Priority question**. Please provide information on the second-line therapies that were used in the CLTR0310-301 trial, differentiating (if possible) between those

Level 1A
City Tower
Manchester
M1 4BT
United Kingdom

+44 (0)300 323 0140

used in patients whose disease did not respond to treatment, patients who experienced relapse after consolidation therapy (no transplant) and patients who experienced relapse after transplant. Please provide this information by treatment group.

(iii) Please comment on any differences between second-line therapies provided in the CLTR0310-301 trial and those in the described in the CancerMPact report, and whether any differences in treatment patterns may affect patient outcomes (e.g. monitoring requirements, survival, safety profile).

B13. Please provide additional information on the following points, regarding the monitoring of patients in the model:

(i) Please provide further justification for using the US-specific National Comprehensive Cancer Network (NCCN) monitoring guidelines and comment on their generalisability to UK patients.

(ii) The NCCN guidelines do not appear to differentiate between the different types of AML, and it is plausible that elderly patients with high risk AML may be monitored at a different rate to patients with AML in general. Please comment on how patients with high-risk AML would be monitored and whether it might be at a different rate to those patients represented in the NCCN guidelines?

(iii) Please confirm the source of the estimates for monitoring patients in the CPX-351 group, as this does not appear to be extracted from the NCCN guidelines as for the 3+7 group. What is the rationale for the decreased need for blood count and chemistry panels during induction and consolidation compared with the 3+7 group?

(iv) Please clarify whether a complete blood count includes a platelet count, and if so, why platelet count is included as an additional monitoring test.

(v) The model only includes transfusions in the “best supportive care” health state. However, blood products are also recommended for use during intensive chemotherapy for AML. Please clarify why these were not included in the induction and consolidation health states.

B14. Please provide additional summary statistics for body surface area from the 301 trial (the standard error in addition to the mean value), in each treatment group separately and combined.

B15. Please incorporate any observed dose reductions requested in question A8 into the calculations on the vial usage of CPX-351 and 3+7, if appropriate.

NICE National Institute for Health and Care Excellence

Level 1A
City Tower
Manchester
M1 4BT
United Kingdom

+44 (0)300 323 0140

Quality of life

B16. In the analysis, CPX-351 is assumed to be associated with a lower treatment-related disutility than 3+7, which persists through induction and consolidation (table 33 of the company submission).

(i) Please describe the basis on which this assumption was made.

(ii) The vignettes in the utility elicitation study (appendix N of the company submission) suggest a difference in levels of fatigue, risk of infection and hair loss, with less severe symptoms associated with CPX-351. Please provide supporting clinical evidence for these assumptions (e.g. rates of these events that suggest a meaningful difference between treatments).

(iii) The vignettes also appear to suggest that consolidation therapy has a similar safety profile to induction therapy, for both 3+7 and CPX-351. Please comment on whether this is a clinically plausible assumption, and provide any supporting evidence from the CLTR0310-301 trial (e.g. rates of these events during induction therapy and during consolidation therapy).

Section C: Textual clarifications and additional points

C1. Table 16 states that 3 patients in the 3+7 group (2%) had bacteraemia, but the text states [redacted] Please clarify which figure is correct.

C2. Please clarify whether the numbers of records identified, reported in the PRISMA diagram (figure 1, appendix D, page 15), are correct. They differ from the number of records reported from the searches of each database for MEDLINE, EMBASE and CENTRAL in tables 1, 2 and 3 in Appendix D.

C3. The coefficient for consolidation therapy in post-transplant OS in the executable model (presented in the “Clinical Data” sheet cell E383) does not match the value in appendix M (table 77, page 306). Please clarify the correct estimate for this parameter.

Single technology appraisal

Liposomal cytarabine and daunorubicin for untreated acute myeloid leukaemia ID1225

Jazz Pharmaceuticals response to clarification questions

Section A: Clarification on effectiveness data

Study 301

A1. **Priority question**: The data utilised in the analyses presented in the company submission and used in the model appears to be based on what is now a relatively old data cut dated December 2015. On page 38 it is stated that trial follow up is designed to continue for 5 years after randomisation.

(i) Please confirm whether follow up is continuing and whether a more recent data cut is available.

Company Response:

Follow up is continuing for 5 years post-randomisation, but no recent data cut is available.

(ii) If a more recent data cut is available, please update the survival analyses (overall survival (OS) and event-free survival (EFS)), i.e. update figures 4-7 with events/N, median values, hazard ratio (HR), 95% confidence intervals (CIs) and p-values.

Company Response:

Follow up is continuing for 5 years post-randomisation, but no recent data cut is available

(iii) Please also state how many trial participants are currently known to be alive and the duration of their survival.

Company Response:

Data on the number of trial participants known to be alive, and duration of survival, are not available; however, adverse event reporting indicates that to date there have been [redacted] *deaths on CPX-351 and* [redacted] *on 3+7.*

A2. **Priority question**: Please provide Kaplan-Meier curves for EFS from time of haematopoietic stem cell transplantation (HSCT) among patients who received HSCT for the intention-to-treat (ITT) analysis population, with events/N, median values, HR, HR with transplant as time-dependent covariate, 95% CIs and p-values (as per company submission figure 7).

***Company Response*:**

Information about relapse after HSCT was not collected. Therefore it is not possible to conduct an EFS analysis post-HSCT.

A3. **Priority question:** Please present remission duration data (as described in question A2) separately for patients who had a HSCT and patients who did not receive HSCT, using the most up to date data available.

Company Response:

For reasons explained in A2, there is no post-HSCT relapse information.

A4. Imputation of missing data for the primary endpoint (OS) is described in table 9 of the company submission. Please specify how many patients had (a) a missing month and day imputed and (b) a missing day imputed, in each arm.

Company Response:

There was no imputation of any missing data for the primary endpoint since there were no missing dates in the death report.

A5. **Priority question:** Please state what criteria were used to assess eligibility for HSCT in the CLTR0310-301 trial. Please comment on whether it can be considered generalisable to the UK.

Company Response:

All subjects treated in Study 301 were fit for intensive chemotherapy (based on study eligibility) and had high-risk disease, and for this reason choice of transplant as post-remission therapy in CR1 is considered standard of care and a precondition for best chance of long-term survival.

Besides the above mentioned risk assessment of the disease the actual decision to proceed to transplant depends on a number of factors: a) patient's physical condition after induction therapy and resolution of any treatment-related toxicities, b) quality of response after induction, availability of a suitable donor, and c) patient preference (as some patients opt not to receive a transplant regardless of prognosis).

All aspects of this process are generalisable to normal UK clinical practice. This approach applies to all newly diagnosed patients with high-risk AML and is not specific to Study 301 and US practice.

In summary, a diagnosis of high-risk (secondary) AML in addition to being fit for intensive induction chemotherapy, and patient choice would be considered criteria to assess eligibility for HSCT and would be generalisable to UK treatment practices.

A6. **Priority question:** Were there any patients eligible for HSCT who did not receive it? If so, please provide reasons for not receiving HSCT (e.g. lack of available donor), as well as characteristics and the number and percentage of patients as per table 12 in the company submission.

Company Response:

At diagnosis, the treatment aim for all patients fulfilling the trial entry criteria was to deliver intensive chemotherapy with the ultimate objective of proceeding to HSCT. The reasons for not receiving HSCT were not captured.

A7. Please provide results of relevant statistical significance tests comparing characteristics of patients undergoing HSCT, in an additional column to company submission table 12.

Company Response:

Relevant Mantel-Haenszel Chi-Square tests and non-parametric Kruskal-Wallis tests were performed on all characteristics of patients undergoing HSCT. None of the variables had a p value <0.05.

A8. **Priority question:** Please provide details of any dose reductions or delays, including the number of patients and reason for the dose reduction/delay, by treatment group.

Company Response:

The full details of the doses and schedules of CPX-351 and 3+7 are included in Table 6 of the main submission. It was the intention of the study to treat all patients at full dose and therefore, the extent to which the dose was reduced was not captured. The protocol encouraged investigators to make up any delays. Details of dose delays are provided in the Table 1 below:

Table 1: Adverse events that led to delay of study drug

	CPX-351 (n=153) ***n (%)***	***3+7 (n=151)*** ***n (%)***
Patients with AEs (any grade)	[redacted]	[redacted]
Cardiac disorders	[redacted]	[redacted]
Atrial fibrillation	[redacted]	[redacted]

	CPX-351 (n=153) n (%)	3+7 (n=151) n (%)
Pericarditis	[redacted]	[redacted]
General disorders and administration site conditions	[redacted]	[redacted]
Localised oedema	[redacted]	[redacted]
Immune system disorders	[redacted]	[redacted]
Hypersensitivity	[redacted]	[redacted]
Infections and infestations	[redacted]	[redacted]
Enterobacter bacteraemia	[redacted]	[redacted]
Metabolism and nutrition disorders	[redacted]	[redacted]
Tumour lysis syndrome	[redacted]	[redacted]
Respiratory, thoracic and mediastinal disorders	[redacted]	[redacted]
Pulmonary haemorrhage	[redacted]	[redacted]

A patient who experienced multiple events within a system organ class or preferred term was counted once for that class and once for that preferred term

A9. The company submission states "The higher observed rate of serious [adverse events] AEs may be due, in part, to the greater proportion of patients in the CPX-351 arm who received consolidation in the outpatient setting compared with the 3+7 arm, because a move to the hospital setting is one of the criteria for classifying an AE as serious." However, the clinical study report (CSR) (section 9.5.4.3) states that AEs leading to inpatient hospitalisation and prolongation of existing hospitalisation both qualify as serious adverse events (SAEs). The Medeiros poster (reference 8 in company submission) also shows a higher median rate of serious treatment-emergent adverse events (TEAEs) per patient-year in the treatment group, suggesting that this difference may not be explained by differences in treatment duration either. Please provide any other clinically plausible justifications for this difference in observed rates of serious adverse events (SAEs).

Company Response:

The Medeiros paper analysed incidence rates per patient year at the patient level and provided summary statistics at the treatment level. To understand what appears to be a higher SAE rate in the CPX-351 arm in both the CSR and Medeiros poster, it was further explored at the treatment level using a Poisson distribution to estimate the rates of SAEs per patient year by treatment group. When analysing this data with a more conventional method, the estimates are [redacted] *SAEs per patient year for CPX-351 and 3+7 respectively. Given this and the 95% confidence interval around the difference in the estimates including zero, the company believes there is no real difference in the SAE rates between the two arms.*

A10. Please clarify whether the adverse events reported in the submission (tables 15-18) are only those that occurred during the treatment phase of the trial. Please provide details of grade 3-5 adverse events that occurred during the follow-up phase of the trail, by treatment group (separately for patients who received HSCT and patients who did not receive HSCT).

Company Response:

Per Study 301 protocol, adverse events were to be recorded in the case report form from the start of the infusion on Day 1 to the last day of the treatment period, with the exception of serious adverse events. The Study 301 adverse event tables summarised all subjects having events with onset date on or after the first infusion date. There were [redacted] *in the CPX351 arm who underwent transplant reporting Grade 3 or higher serious adverse events after the end of treatment phase. Given the sparse data, any detailed summary table for these events will not be meaningful. The events have been included in the summary tables used for submission.*

A11. **Priority question:** Please provide further details of 'adverse events of special interest', which are only briefly summarised in the text (page 62-63), i.e. present tables similar to table 16 for 'infection-related adverse events', 'bleeding-related adverse events' and 'cardiac adverse events'.

Company Response:

Grade 1-5 and grade 3-5 'infection-related' adverse events are shown in Table 2 and Table 3.

Grade 1-5 'bleeding-related' adverse events are shown in Table 4. No grade 3-5 bleeding-related adverse events occurred at a frequency of ≥5% in the safety population.

Grade 1-5 'cardiac-related' adverse events are shown in Table 5. The only grade 3-5 cardiac-related AE with ≥5% frequency in the safety population was a decreased ejection fraction, which occurred in 8 (5%) patients on both treatment arms.

Table 2: Number of subjects with grade 1-5 infectious adverse events, ≥5% safety population

	CPX-351 (n=153) n, %	3+7 (n=151) n, %
Any AE of infection	[redacted]	[redacted]
Febrile neutropenia	[redacted]	[redacted]
Chills	[redacted]	[redacted]
Pneumonia	[redacted]	[redacted]
Pyrexia	[redacted]	[redacted]
Sepsis	[redacted]	[redacted]
Cellulitis	[redacted]	[redacted]
Bacteraemia	[redacted]	[redacted]

Table 3: Number of subjects with grade 3-5 infectious adverse events, ≥5% safety population

	CPX-351 (n=153) n (%)	3+7 (n=151) n (%)
Febrile neutropenia	104 (68)	107 (71)
Pneumonia	30 (20)	22 (15)
Sepsis	14 (9)	11 (7)
Bacteraemia	15 (10)	3 (2)

Table 4: Number of subjects with grade 1-5 bleeding adverse events, ≥5% safety population

	CPX-351 (n=153) n, %	3+7 (n=151) n, %
Any bleeding-related AE	[redacted]	[redacted]
Epistaxis	[redacted]	[redacted]
Petechiae	[redacted]	[redacted]
Mouth haemorrhage	[redacted]	[redacted]
Ecchymosis	[redacted]	[redacted]
Contusion	[redacted]	[redacted]
Haematuria	[redacted]	[redacted]
Blood blister	[redacted]	[redacted]
Gingival bleeding	[redacted]	[redacted]
Haemoptysis	[redacted]	[redacted]
Conjunctival haemorrhage	[redacted]	[redacted]

Table 5: Number of subjects with grade 1-5 cardiac adverse events, ≥5% safety population

	CPX-351 (n=153) n, %	3+7 (n=151) n, %
Any cardiac AE	[redacted]	[redacted]
Tachycardia	[redacted]	[redacted]
Atrial fibrillation	[redacted]	[redacted]
Chest pain	[redacted]	[redacted]
Ejection fraction decreased	[redacted]	[redacted]
Chest discomfort	[redacted]	[redacted]

A12. Please define what constitutes 'other' reasons for withdrawal from treatment in the patient disposition (figure 2) of the company submission appendices.

Company Response:

If a patient was identified to have an "other" reason for discontinuation ([redacted]), the investigator could enter free text in the CRF to describe the reason in more detail. Full details of those listed are provided in the Table 6 below:

Table 6: "Other" reason for discontinuation

Drug	***"Other" discontinuation text***
CPX-351	[redacted]
3+7	[redacted]

A13. Please confirm whether there were any patients in the 3+7 arm of the trial who subsequently received treatment with CPX-351 (i.e. any crossover).

Company Response:

No patients in the 3+7 arm of the trial subsequently received CPX-351, as crossover was not permitted in the study.

A14. **Priority question:** Please state whether there are any other subtypes of high-risk acute myeloid leukaemia (AML) that would be eligible for CPX-351 in practice (under the anticipated marketing authorisation), which were not included in CLTR0310-301 (e.g. acute promyelocytic leukaemia). If so, please explain why they were not included in the trial.

Company Response:

CPX-351 is awaiting EMA marketing authorisation, and the anticipated EMA label is 'for the treatment of adults with newly diagnosed therapy-related AML (t-AML) or AML with myelodysplasia-related changes (AML-MRC).'

Study 301 did not enrol patients with all types of AML but rather was designed to include patients with poor-prognosis, high-risk AML. T-AML was enrolled into Study 301 and is reflected within the proposed EMA label. A diagnosis of AML-MRC can be made if there is (i) a prior history of MDS or MDS/MPN, (ii) specific myelodysplasia-related cytogenetic abnormalities or (iii) multilineage dysplasia. Patients with MPN were specifically excluded. AML patients with multilineage dysplasia could have been enrolled into Study 301 if this abnormality was present in conjunction with t-AML, a MDS-related cytogenetic abnormality or a prior history of MDS/ CMML.

Systematic review

A15. **Priority question:** Please justify restricting the systematic review inclusion criteria to patients aged 60-75 years, when the anticipated marketing authorisation is 'adults' (no age restriction)?

Company Response:

This population was chosen as it is aligned to the patient population in Study 301 (CLTR0310-301), the results of which formed the basis of the effectiveness evidence for CPX-351 in this STA submission. Including studies of patients outside this patient population would have introduced biased comparisons with CPX-351.

A broader systematic review in all adults with AML is ongoing.

A16. Priority question: Were there any CPX-351 trials that were excluded only because they included participants under 60 years, but would otherwise have met the review inclusion criteria? If so, please provide references.

Company Response:

No CPX-351 studies were excluded because they included patients under 60 years of age alone. Study 205 was excluded because it included patients under 60 years of age in the relapsed/refractory treatment setting. It was therefore excluded on grounds of both patient age and different disease state.

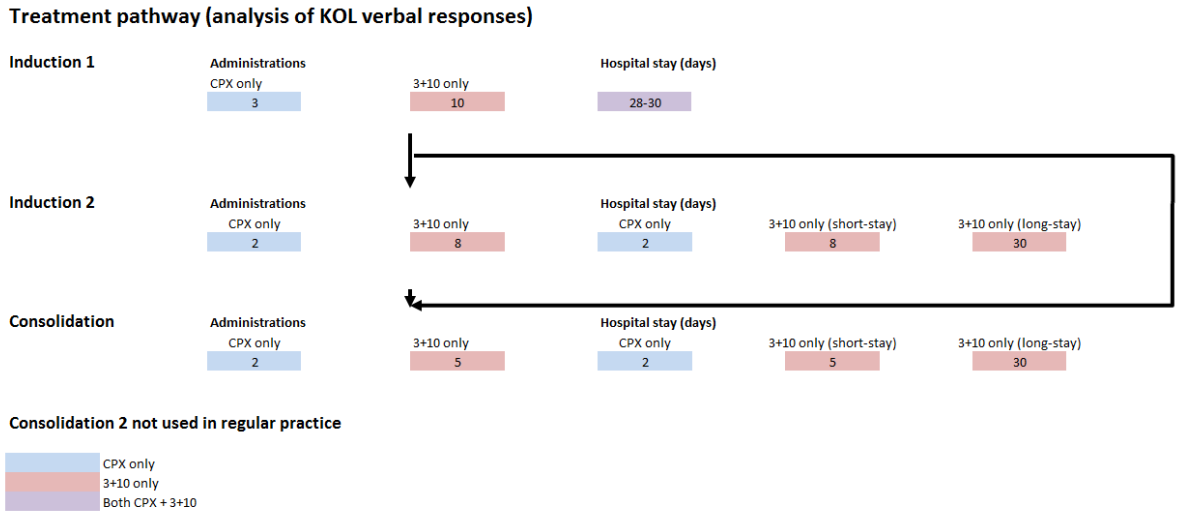
Section B: Clarification on cost-effectiveness data

General

B1. Please provide the "data on file" reference for the key opinion leader (KOL) advisory board (page 90 of the company submission).

Company Response:

The reference on page 90 of the submission refers to verbal feedback from UK AML experts at an advisory board. According to the advisory board participants, it is standard UK practice for patients to receive induction as an inpatient for around 30 days and the feedback is summarised below:



The advisors also highlighted that the inpatient stay data for daunorubicin & cytarabine (DA) 3+10 was captured and is published for the UK AML16 study. Consistent with the advisory board feedback, the mean days in hospital for DA 3+10 was 33.8 and 25.0 days during course 1 and course 2, respectively (Burnett AK, 2017).

B2. **Priority question.** Please clarify what "Pathways A to H" refer to (introduced on page 214 of the appendix to the company submission). Please provide a diagram of the pathway referred to in **appendix M.**

Company Response:

Patients were distributed among potential treatment pathways determined by post-hoc analyses of patient-level efficacy data from Study 301. Treatment pathways were defined based on response following induction and whether patients received a hematopoietic stem cell transplant (HSCT). The model tracks number of induction (1 or 2) and consolidation (0, 1,

or 2) cycles. Figure 1 shows the clinical pathways with all the steps of treatment that patients could have received in their life course after induction therapy. Each pathway is denoted by a letter and the number of corresponding patients found for each in the trial data.

At initiation of the simulation model, patients with newly diagnosed AML receive induction therapy (1 or 2 rounds). Patients who respond to treatment are considered to be in remission (A – F). Patients who respond well to induction can receive up to two rounds of consolidation and those who are sufficiently fit may receive a HSCT (A, C and E). Note that patients achieving remission post-induction may relapse after consolidation or transplantation. Those who do not achieve remission post-induction (G and H) may progress, receiving a transplant (with the model only capturing the cost impact as the trial data provided weak evidence of a difference in survival outcomes among these patients and the number of patients was balanced between the arms). Patients may die at any time.

Figure 2 illustrates how patients flow between the health states in the model.

Figure 1. Clinical pathway

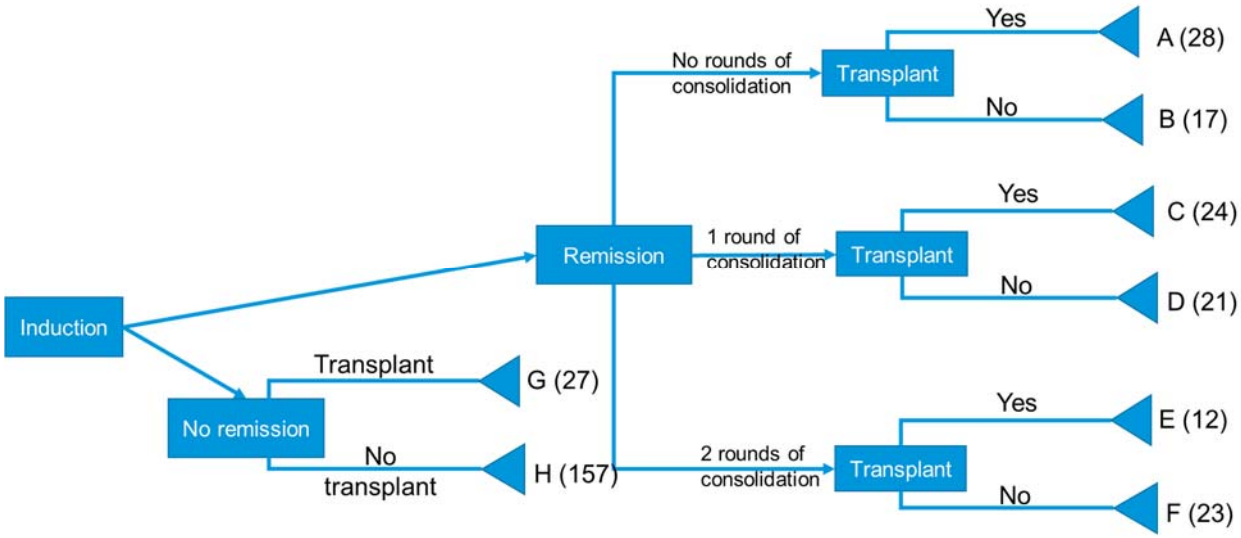
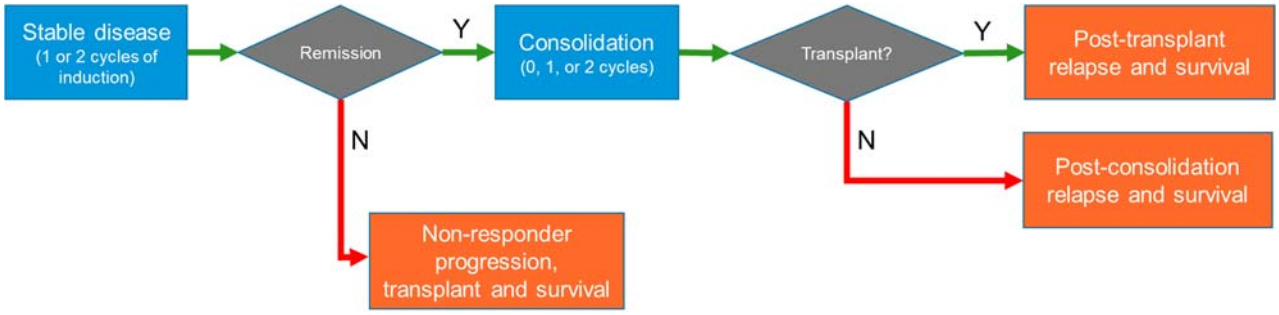


Figure 2. Modelling Framework for the Transitions between Health States



Blue boxes: no mortality assumed during these periods; duration and number of cycles based on patient characteristics and prior treatment (number of inductions and/or consolidations). Grey diamonds denote decision points; instantaneous; percent making each decision based on patient characteristics and prior treatment (number of inductions and/or consolidations). Orange boxes: Progression/relapse-free and overall survival curves were

obtained; patient characteristics, response and prior treatment (number of inductions and/or consolidations) were treated as risk factors.

Model structure

B3. Please provide additional clarity on how patients whose disease did not respond to treatment were modelled, with regard to the following points:

(i) Please clarify how patients whose disease did not respond to treatment who had a transplant were implemented in the model.

Company Response:

For patients that did not respond to treatment, parametric curves assessing the successive composite endpoints of overall survival (OS), OS or hematopoietic stem cell transplant (HSCT), and OS or HSCT or progression are applied from the beginning of the model (no time-shifts need to be considered in relation to patient survival). The proportion of patients in the transplanted health state is calculated as the proportion of patients alive (OS) minus the proportion alive and not transplanted (OS or HSCT).

(ii) From inspection of the executable model, it appears that the time of transplant in patients whose disease did not respond to treatment was based on an analysis of time to HSCT or OS, while the time of transplant in patients whose disease did respond to treatment was based on the mean time of transplant. Please confirm that this is the case, and comment on why two different methods were used. What was the mean time for treatment non-response to transplant?

Company Response:

This is correct - the time of transplant in patients whose disease did not respond to treatment was based on an analysis of time from randomisation, while the time of transplant in patients whose disease did respond to treatment was based on time from transplant. This modelling approach was pre-specified based on the information available from the definition of the patient cohorts. Specifically, by defining cohorts in terms of the treatments they received, the cohorts necessarily had 100% survival until the initiation of the last treatment in that sequence.

Modelling OS and EFS using parametric survival models by measuring time from baseline for all cohorts would have yielded inaccurate fits because such models would not explicitly acknowledge the time length required for each consolidation therapy. For example, including the number of rounds of consolidation therapy as an explanatory variable in a parametric model where time would be measured from randomisation would provide a hazard ratio or acceleration factor that would then be applied to the reference group of patients who received no consolidation therapy in order to explain time-to-event data for patients who received 1 or 2 rounds of consolidation. We believe that this modelling approach would be inappropriate because each round of consolidation therapy requires that patients survived the necessary increment of time to receive it. Using rounds of consolidation as a hazard or an acceleration would slow events, but some fraction of the cohort would still be modelled as dying before they could receive the therapies they are defined by having received. A similar argument holds for the time between consolidation therapy and HSCT. In the case of non-responders, the

decision was made based on evaluation of the data to consider survival for a single cohort, regardless of whether they received transplant (though transplant is tracked for cost accounting). For details of these data please see question B6. A key consideration was sample size – only [redacted] patients did not respond but still received transplant (Figure 1, Pathway G). For the combined cohort of all patients who did not respond, however, there is no minimum duration for which 100% must survive by definition and thus the analysis plan requires that their survival be analysed from randomisation.

The mean time to HSCT amongst non-responders in Pathway G was [redacted] weeks. (see Table 7*)*

Table 7. Mean Time to HSCT amongst non-responders who received transplant

[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]

(iii) Please provide information on the treatments received by this group while waiting for the transplant in the 301 trial, and whether this is generalisable to UK clinical practice.

Company Response:

Therapies received by non-responding patients who ultimately went to transplant include intermediate dose cytarabine (IDAC) ± idarubicin or daunorubicin or mitoxantrone, mitoxantrone + etoposide + cytarabine (MEC) and fludarabine + cytarabine + G-CSF + idarubicin (FLAG-Ida). These therapies are consistent with UK clinical practice.

According to BCSH guidelines, there are limited randomised trials for salvage treatment in AML but the mainstay of UK AML therapy is a variable-dose cytarabine-based regimen combined with other agents (Milligan DW, 2006)*. Based on experience from the UK AML15 trial, FLAG-Ida is frequently used as salvage treatment as a bridge to transplant.*

Clinical data

B4. The company submission describes a hazard ratio estimated from data reported by Martin (2010) that was applied to general population mortality in a scenario analysis (page 100).

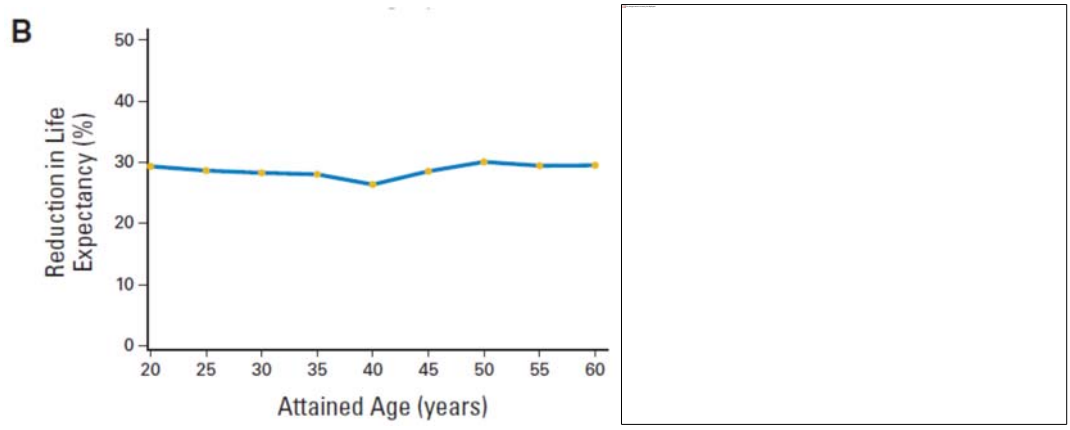
Please provide additional information, including a description and example calculations if necessary, of how the 30% reduction in life expectancy reported by Martin (2010) was used to estimate the hazard ratio.

Company Response:

Figure 3B in Martin et al. (Martin PJ, 2010) suggests a stable 30% reduction in life expectancy regardless of attained age (see reference below). This reduction was reproduced in the model by gradually increasing the SMR post-transplant until life-years after transplant were reduced by 30%, yielding the estimates provided in the submission. Because of the differences in

mortality by gender, separate SMR values were estimated for males and females. The resulting SMR values are in agreement with what might be anticipated by extrapolating the age-dependent SMR curve provided in Figure A3C of Martin et al. (Martin PJ, 2010) (see referenced graph below). These values are higher than that recently determined in the recent NICE appraisal of Mylotarg (SMR = 1.34 in a European AML population).

Figure 3B **Figure A3C**



Source: (Martin PJ, 2010)

B5. **Priority question.** In the model it is assumed that there is a different post-HSCT OS curve for patients receiving CPX-351 than for patients those receiving 3+7. This is one of the most important assumptions in the model and a key driver of the model because patients receiving CPX-351 experience significantly improved survival post-HSCT, as demonstrated by figure 13, page 97 of the company submission. However, there is little difference in EFS post-transplant between the two treatment groups (figure 16, page 99 of the company submission).

(i) Please provide an interpretation of these clinical data, and comment on the plausibility of CPX-531 providing an OS but not an EFS benefit after transplant.

Company Response:

The apparent decoupling of the EFS and OS outcomes post-transplant reflects limitations in the data, particularly around EFS, rather than an anomalous clinical finding. Only [redacted] relapses were observed prior to the date of the last examination, therefore the results from the post-HSCT EFS analysis is an unreliable indication of the true treatment effect. Given these limitations, no clinical inferences should be drawn from estimates of post-transplant EFS. Rather, EFS post-transplant was analysed exclusively for the purpose of assigning a utility score and management cost to patients following transplant. The curves used for EFS in the model are, as noted in the question, much less favourable to CPX-351 than the (much more robust) OS data would imply. This leads to a lower number of incremental QALYs accumulated post-transplant and higher excess management cost post-transplant in the CPX-351 arm versus 3+7 compared to what might have resulted from an alternative approach to the data. For example, assuming patients were no longer at risk of progression after surviving

two years following HSCT would result in those patients receiving higher utility scores ([redacted]) for their remaining lifetime.

For clarity, more details on the limitations of the EFS data post-HSCT is included here. The analysis was based on data collected close to the end-of-study follow-up. For this reason, these data were subject to misclassification and truncation. In particular, the following was observed in the EFS analysis:

[redacted]

Analysis of time to event (TTE) outcomes cannot be carried out when those take negative values. Therefore, patients with negative values for time to relapse or death were removed and the post-HSCT EFS analysis was based only on the subset of [redacted] patients with a positive value for the TTE outcome. Among these [redacted] patients, only [redacted] progressed (all subsequently died), [redacted] died (including those who progressed before dying), and [redacted] were censored.

As the proportion of incomplete data that was excluded from this analysis was high ([redacted] people were excluded as explained above), and the total observed events ([redacted] relapses) was low, the results from the post-HSCT EFS analysis is an unreliable indication of the true treatment effect. In particular, the proportion of patients who could relapse may be underestimated due to the way information was recorded as part of the trial protocol close to the end of follow-up. On the contrary, the post-HSCT OS analysis did not suffer from this problem. Death times were recorded accurately and systematically even beyond the trial follow up. The separation of the survival curves of CPX-351 and 3+7 with respect to overall survival post-HSCT is clear and statistically significant (p=0.0120, see ***Error! Reference source not found.*** *and* ***Error! Reference source not found.*** *below).*

Figure 3. Product-Limit Survival Estimates

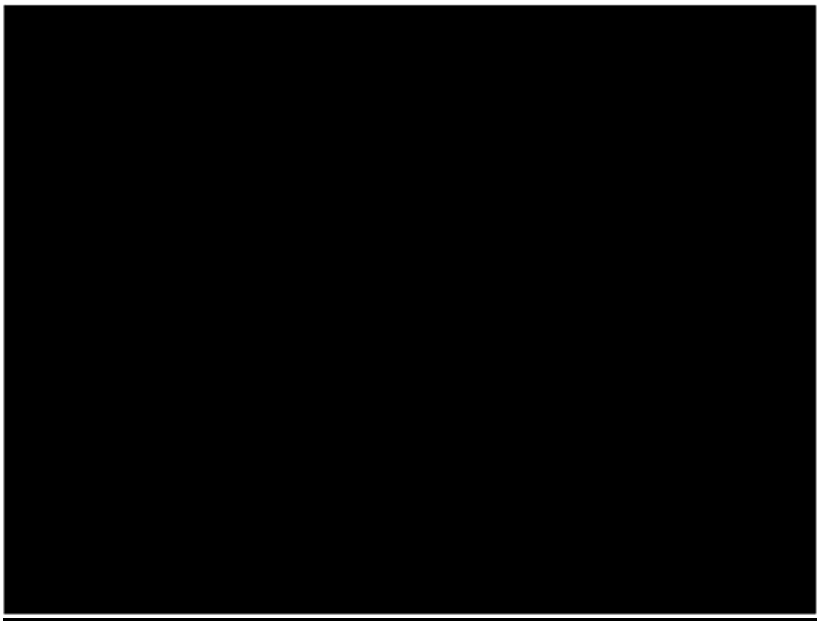


Figure 4. Product-Limit Survival Estimates (a)

(ii) Please provide an additional scenario in the executable model based on an analysis of post-transplantation OS that combines data from both treatment groups (i.e. no treatment effect on OS post-transplant). Please include all parametric distributions, using the later data cut of the CLTR0310-301 trial (see question A1), if available.

Company Response:

The version of the model provided with this response includes an option to conduct analyses assuming no difference between treatment arms in OS post-transplant. It bears emphasising, however, that this assumption contradicts the available evidence as a clear separation in the post-HSCT overall survival curves favouring CPX-351 versus 3+7 was observed, which was statistically significant (p=0.0120). Applying the assumption of a common OS curve post-transplant also yields predicted survival that deviates markedly from the observed clinical data as early as 36 weeks post-randomisation. (See Figure 5*) In keeping with the discussion above, an option has also been added to the model to use a post-HSCT EFS curve that is more reflective of the OS data. Specifically, with this option surviving patients are assumed to be relapse-free, with correspondingly higher utility and no further monitoring or management costs, beginning 2 years post-HSCT. The resulting ICER is substantially lower than the submitted base case.*

Figure 5. Predicted OS assuming no difference in post-HSCT OS between treatment arms

B6. **Priority question:** Please provide Kaplan–Meier curves (OS and EFS, with descriptive statistics presented as per figure 7) for the following groups of patients:

(i) Patients who had a complete remission (CR)/complete remission with incomplete platelet or neutrophil recovery (CRi) and received transplant

Company Response:

Please see

and Figure 7. EFS in Patients who Achieve Remission and Receive HSCT (Pathways A, C and E)

below displaying the OS and EFS survival curves in patients who achieve remission and receive a HSCT. Each figure displays the KM survival estimates with the number of patients at risk.

Figure 6. OS in Patients who Achieve Remission and Receive HSCT (Pathways A, C and E)

Figure 7. EFS in Patients who Achieve Remission and Receive HSCT (Pathways A, C and E)

(ii) Patients who had a CR/CRi and did not receive transplant

Company Response:

Please see ***Error! Reference source not found.*** *and* ***Error! Reference source not found.*** *below displaying the OS and EFS survival curves in patients who achieve remission and do not receive a HSCT. Each figure displays the KM survival estimates with the number of patients at risk.*

Figure 8. OS in Patients who Achieve Remission and Do Not Receive HSCT (Pathways B, D and F)

[redacted]

Figure 9. EFS in Patients who Achieve Remission and Do Not Receive HSCT (Pathways B, D and F)

[redacted]

(iii) Patients who did not have a CR/CRi and received transplant

Company Response:

***Please see* Figure 10. OS in Patients who do not Achieve Remission and Receive HSCT (Pathway G)**

[redacted]11 *and* ***Error! Reference source not found.*** *below displaying the OS and EFS survival curves in patients who do not achieve remission and receive a HSCT. Each figure displays the KM survival estimates with the number of patients at risk.*

Figure 10. OS in Patients who do not Achieve Remission and Receive HSCT (Pathway G)

[redacted]11[redacted]

Figure 12. EFS in Patients who do not Achieve Remission and Receive HSCT (Pathway G)

[redacted]

(iv) Patients who did not have a CR/CRi and did not receive transplant.

Company Response:

Please see

Figure 13 and ***Error! Reference source not found.*** *below displaying the OS and EFS survival curves in patients who do not achieve remission or receive a HSCT. Each figure displays the KM survival estimates with the number of patients at risk.*

OS curves in patients who do not achieve remission or receive a HSCT is similar to that of all patients who do not achieve remission, regardless of transplant status which is used in the model (combined pathways G and H, ***Error! Reference source not found.****). Using the combined G and H curve is a conservative approach for estimating the cost-effectiveness of CPX-351, as any difference in long-term survival between the treatment arms using separate survival analyses would be driven by pathway G (****Error! Reference source not found.****) and favour CPX-351, albeit in a small fraction of the patient population.*

Figure 13. OS in Patients who do not Achieve Remission or Receive HSCT (Pathway H)

Figure 14. EFS in Patients who do not Achieve Remission or Receive HSCT (Pathway H)

15

Figure 16. OS in Patients who did not achieve remission (Pathways G and H)

Survival analysis

B7 The categories of the AML types are not consistent for estimating the percentages of patients who follow each clinical pathway (e.g. the analysis for remission post-induction uses 5 categories of AML (analysis 2, appendix M.2.1), but the analysis for rounds of consolidation (analysis 3, appendix M.2.1) in patients who had a remission uses 2 categories of AML)

(i) Please provide justification for regrouping the categories in the different analyses, and comment on whether it was prespecified in the statistical analysis plan.

Company Response:

The statistical analysis plan specified a priori that all regression analyses would be adjusted for the sampling stratification variables age (60–69 vs. 70–75 years old) and AML type. Age and AML type were used in the stratified sampling process of the trial data and there is ample documentation in the literature explaining that analysing data from randomised trials without adjusting for the sampling stratification variables can lead to biased estimation of treatment effects and misleading statistical conclusions. Whenever possible, analyses were adjusted for age and AML type using the same groupings that were used in the stratification process:

- *Age:*
  - *60–69 years old*
  - *70–75 years old*
- *AML type:*
  1. *Treatment-related AML (referred to as "t AML")*
  2. *AML with documented history of MDS with prior treatment with hypomethylating agents (referred to as "MDSAML with HMA")*
  3. *AML with documented history of MDS without prior treatment with hypomethylating agents (referred to as "MDSAML without HMA")*
  4. *AML with karyotype characteristic of MDS (referred to as "denovoAML")*
  5. *AML with documented history of CMMoL (referred to as "CMMoLAML")*

Unfortunately, adjusting for AML type using the 5-level grouping listed above was not possible for a limited number of analyses due to small sample sizes. In particular, there were only 23 total patients in group 5 above (CMMoLAML). In these analyses, adjusting for AML type using 5 levels led to either convergence issues when fitting regression models or resulted in unstable models (e.g. very large regression coefficients and standard errors). For these reasons, clinical expert opinion was sought on how to merge AML type categories in a clinically meaningful way. The clinical guidance was to merge the 5 levels of AML types into 2 levels by merging AML types 1, 4 and 5 from the list above into one group and AML types 2 and 3 into a second. This reduction from 5 groups to 2 groups was not prespecified, but provided a pragmatic approach to solving the problem of adjusting for all stratification variables when data proved insufficient.

(ii) Please provide results of the logistic regression on the need for a transplant in patients who had a remission and consolidation rounds (similar to table 52 and table 53 in appendix M, pages 270-271)

(iii) Include a scenario analysis in the executable model that incorporates the results of the analysis with the 5 categories of AML.

Company Response (points ii and iii):

The results of the analysis with 5 categories of AML type yielded regression coefficients that were not meaningful in particular for the CMMoL →AML subtype and the number of patients was very low as stated earlier. As such the model could not be incorporated in the CEM. For more details, please see Table 8.

Table 8. Logistic regression

[redacted]				
[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
[redacted]				
[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
[redacted]	[redacted]	[redacted]	[redacted]	[redacted]

B8. **Priority question:** The covariates used in the different regression models fitted to EFS and OS are not consistent. Specifically, in the regression models for post-consolidation progression-free survival (PFS), post-consolidation OS, and OS among patients who did not have a remission, 5 categories of AML were used, whereas in the post-HSCT EFS and post-HSCT OS regression models 2 categories were used.

(i) Please provide a justification for regrouping the categories in the different analyses, and comment on whether it was prespecified in the statistical analysis plan.

Company Response:

Please see response to Question B7. The statistical analysis plan specified a priori that all regression analyses would be adjusted for the sampling stratification variables age and AML type, but due to small sample sizes in some subgroups, maintaining 5 levels for AML type was not possible in all analyses. AML type was reduced to 2 levels, with the choice of grouping

provided by a clinical expert. For the post-HSCT fits, in particular, sample size was limiting with no patients in the 3+7 arm with the CMMoLAML type receiving transplant following remission (see tables below).

Table 9. Number of patients in EFS post-HSCT analysis by AML type

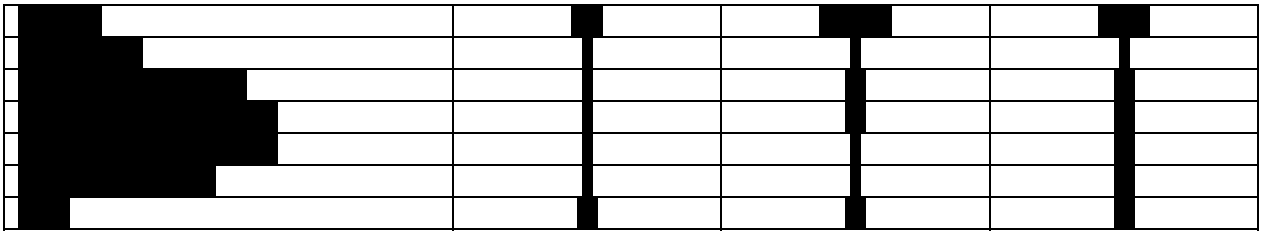
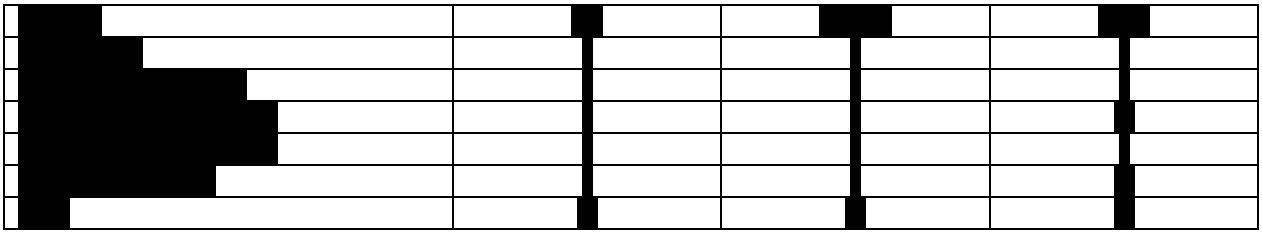


Table 10. Number of patients in OS post-HSCT analysis by AML type



(ii) Please provide results for post-HSCT PFS and post-HSCT OS using the five categories of AML (similar to table 76 to table 81 in appendix M, page 305-313).

(iii) Please provide a scenario analysis in the executable model that incorporates post-HSCT EFS and post-HSCT OS using 5 categories of AML.

Company Response (points ii and iii):

There were no patients in the CMMoL → AML category receiving 3+7. Therefore, when attempting to adjust for AML type using all five categories for the 3+7 treatment arm, there was no regression coefficient for CMMoL → AML. Furthermore, the regression coefficients for age and the constant when comparing the models where AML type was adjusted using 5 categories versus 2 categories differed by roughly a factor of 2, signalling that the model with more AML categories was unstable. As such, an analysis with 5 levels of AML included could not be incorporated in the CEM.

B9. The process of selecting the survival models is described in appendix M.1.7 and M.4 of the company submission, which reports that it involved both statistical and clinical considerations. The range of statistical tests of fit are extensive, and the results are comprehensively provided. However, the company submission stated that the models were also selected according to their clinical plausibility, specifically to validate projected survival after the trial period and the median estimate of survival.

Please provide additional information on this assessment, including how it was assessed, any external datasets that were used in the assessment, and the outcomes of the assessments.

Company Response:

The only fit for which clinical plausibility of projected survival was considered was post-HSCT OS. For patients that receive transplant, clinical expectation based on the literature is that survival will plateau among those patients surviving at least 2 years after transplant. This behaviour is also observed in the trial data, which shows a plateau in OS beginning roughly 72 weeks post-transplant, albeit with relatively small patient numbers. Literature sources consulted included Martin (2010), discussed elsewhere in these responses. More relevant sources for this population, however, were Shimoni et al (Shimoni A, 2016) and Wingard et al (Wingard JR, 2011). In Shimoni et al., among patients with AML receiving SCT and surviving to two years post-SCT, 10 year survival was roughly 74%, with no dependence on age. Similarly, Wingard et al. found overall survival of 84% among patients with AML surviving 2 years post-SCT. Comparison of the distributions fitted to the trial data for post-HSCT showed that only the Gompertz function was able to adequately capture this plateau, and so only this form was considered an appropriate, clinically plausible extrapolation of post-HSCT survival. Of note, even with this form, the modelled OS for the full CPX-351 arm still undershoots the observed KM late in the trial period.

B10. **Priority question.** A number of parametric distributions were explored for EFS and OS. However, a limited number of distributions were implemented in the executable model. For example, for OS among patients in Pathways G and H, one distribution (lognormal) was implemented (page 236, appendix M). The model fit statistics are very similar for lognormal, with similar visual fit.

(i) Please justify the selection of the lognormal distribution over the other distributions.

Company Response:

The log-normal distribution was selected based on having the best fit statistics – the AIC and BIC values for the log-normal distribution were the lowest observed (1423.663/1430.026) among the tested distributions. Predicted median OS from the log-normal model was 17.97 weeks, which was comparable to the results observed with the KM survival curve (16.71 weeks). As noted, all of the distributions tested provided similar visual fits and had similar fit statistics. Because the data are quite mature, with the observed KM OS data dropping below 10% within the trial follow-up, alternatives were not implemented in the model.

(ii) Please provide an executable model that incorporates all fitted parametric distributions for PFS and OS. Where possible this should use any updated cut of the trial data (see question A1). If this is not feasible please provide fitted parametric distributions for post-HSCT OS based on the later data cut incorporating them into the executable model as scenario analyses.

Additional clarification from the ERG:

We have prioritised the survival analyses as requested in question B10 (ii) of the PFCs. We have prioritised specific outcomes that we would like to explore further rather than specific survival models. We'd like to prioritise being able to explore the impact of different models on overall survival, and if it remains within the company's capacity to incorporate any others then EFS would also be useful.

- Highest priority: All distributions for "post-HSCT OS" (table 74 in Appendix)
- High priority: All distributions for "OS in non-responders" (Table 60 in appendix) and "post-consolidation OS" (Table 66 in appendix)
- Medium priority: All distributions for "post-consolidation EFS" (Table 70 in appendix) and "post-HSCT EFS" (Table 78 in appendix)
- Lower priority: All distributions for "Time to HSCT or death in non-responders" (Table 62 in appendix), "time to progression or HSCT or death in non-responders" (Table 64 in appendix)

Company response

The highest and high prioritised analyses have been added in the model version accompanying these responses as has an alternative post-HSCT EFS assumption. These analyses can now be selected on the Clinical Inputs tab. Please see question B4 for additional discussion.

Resource use

B11. Please provide additional information on the following points, regarding the setting for administration of induction and consolidation therapy:

(i) On what basis were the assumptions about the proportion of CPX-351 infusions that could be delivered in an outpatient setting made?

Company Response:

The assumption was based on observed setting of administration in Study 301.

(ii) The model assumes that 50% of CPX-351 patients would receive consolidation therapy as outpatients. What proportion of patients received consolidation therapy as an outpatient in the CLTR0310-301 trial?

Company Response:

The inclusion of CPX-351 administration on an outpatient basis for consolidation was based on the experience of the CLTR0310-301 trial. In the trial, approximately 50% of CPX-351 patients were discharged and received their consolidation therapy in an outpatient infusion clinic.

(iii) Please provide the data on file (page 73 of the company submission) regarding the proportion of US patients who received CPX-351 in an outpatient setting.

Company Response:

The reference on page 73 refers to the US practice of delivering CPX-351 treatment in the outpatient setting, beginning as early as the induction phase in some cases. This was taken from US post-launch quantitative market research by Naxion Research consulting performed in February to March 2018 (6-7 months post FDA market authorisation). The market research indicates that [redacted] (increases with number of doses during the induction cycle) of patients received CPX-351 induction treatment in the outpatient-setting (N=34 CPX-351 PRFs)

In second induction, [redacted] of CPX-351 patients receive therapy in an outpatient setting and 100% of CPX-351 patients receive consolidation in an outpatient setting. All 3+7 patients are treated as inpatients in this scenario. See table 50, page 132 of evidence submission.

Figure 17. Place of administration of CPX-351

(iv) Please provide the actual proportion of patients receiving CPX-351 as outpatients in the scenario analysis (table 63, page 144 of the company submission).

Company Response:

In this scenario analysis, all CPX-351 patients were assumed to receive both second induction and consolidation treatment as outpatients.

B12. Please provide additional information on the following points, regarding second line therapy:

(i) In the model, patients could only receive one additional line of therapy post progression (either salvage therapy, best supportive care or low intensity chemotherapy). The CancerMPact report (reference 89, company submission)

makes reference to a number of relapses that occur in patients with AML, suggesting that the assumption of only one further line of therapy is not appropriate. Please comment on whether you might expect patients would receive more than one type or line of therapy after progression.

Company Response:

The treatments utilised in CancerMPact report covers a broad population of AML and is not specific to secondary AML. High-risk (secondary) AML is particularly aggressive as its underlying disease biology confers resistance to therapy. Data from the two largest published registries have shown that high-risk (secondary) AML differs significantly to de novo AML, with high-risk (secondary) AML having fewer complete remissions to intensive chemotherapy and significantly inferior survival (Hulegårdh E, 2015) (Østgård LS, 2015). The older age of these patients, comorbidities, and complex disease biology will negatively impact on the risk-benefit of salvage therapies which can have an adverse side effect profile. This is consistent with results from the Yorkshire registry. Out of a total of 353 patients diagnosed with secondary AML (n=136 t-AML, n=217 AML-MRC), only 3 patients (0.8%) received any third line therapy (data courtesy of Haematological Malignancies Research Network).

(ii) **Priority question**. Please provide information on the second-line therapies that were used in the CLTR0310-301 trial, differentiating (if possible) between those used in patients whose disease did not respond to treatment, patients who experienced relapse after consolidation therapy (no transplant) and patients who experienced relapse after transplant. Please provide this information by treatment group.

Company Response:

Appendix 1 contains pooled second-line therapies for both study arms in the CLTR0310-301 trial. These data are limited as they do not distinguish between therapies utilised in patients whose disease did not respond to treatment (primary refractory) versus relapse following a complete remission and does not provide exact details of the regimens used. In addition, second line therapies for patients who experienced relapse post-transplant were not collected.

(iii) Please comment on any differences between second-line therapies provided in the CLTR0310-301 trial and those in the described in the CancerMPact report, and whether any differences in treatment patterns may affect patient outcomes (e.g. monitoring requirements, survival, safety profile).

Company Response:

In the CancerMPact report, there are a variety of second-line therapies that generally consist of cytarabine-based chemotherapies (low dose cytarabine, fludarabine + cytarabine + G-CSF + idarubicin (FLAG-Ida), cytarabine ± another chemotherapy agent) or hypomethylating agents (azacitidine or decitabine). Appendix 1 for second-line therapies utilised in the CLTR0310-301 trial does not provide granularity of the exact regimens used but is consistent with the CancerMPact report. The main second-line chemotherapy agents utilised include cytarabine [redacted] and fludarabine [redacted] with the hypomethylating agents azacitidine [redacted] and decitabine [redacted] also being employed.

Due to the limitations in the data, it is not possible to provide specific comments on how any differences between second-line therapies used in CLTR0310-301 trial versus the CancerMPact report may impact on monitoring requirements or safety. However, as described in the response to B12 part (i), patients with high-risk (secondary) AML have very poor outcomes from diagnosis and are therefore unlikely to derive a clinically meaningful survival benefit depending on which salvage regimens are employed. The company therefore do not anticipate any differences in survival regardless of second-line salvage treatments used.

B13. Please provide additional information on the following points, regarding the monitoring of patients in the model:

(i) Please provide further justification for using the US-specific National Comprehensive Cancer Network (NCCN) monitoring guidelines and comment on their generalisability to UK patients.

Company Response:

The current BCSH AML guidelines provide details on diagnosis, prognostication and treatment but do not specify monitoring tests and frequencies. In contrast, the NCCN guidelines make specific recommendations for monitoring during induction and post-remission therapy.

We have consulted with 2 UK AML experts on the NCCN monitoring recommendations and generalisability to the UK. Blood test profiles and frequency during induction and post-remission were consistent with UK practice. The only identified potential difference was in the frequency of bone marrow (BM) examinations. In the US guidelines, it is recommended to perform a BM exam 7-10 day's post intensive chemotherapy to document hypoplasia and then repeat at haematological recovery to document remission. In contrast, UK practice is to perform a BM exam only to document remission and not hypoplasia.

(ii) The NCCN guidelines do not appear to differentiate between the different types of AML, and it is plausible that elderly patients with high risk AML may be monitored at a different rate to patients with AML in general. Please comment on how patients with high-risk AML would be monitored and whether it might be at a different rate to those patients represented in the NCCN guidelines?

Company Response:

Treatment with intensive chemotherapy is associated with significant, potentially life threatening toxicities regardless of AML subtype or age. Daunorubicin and cytarabine intensive chemotherapy is used regardless of age or AML subtype (excluding APL) and therefore it is anticipated that high-risk AML will be monitored at the same rate as other intensively treated AML subtypes.

(iii) Please confirm the source of the estimates for monitoring patients in the CPX-351 group, as this does not appear to be extracted from the NCCN guidelines as for the 3+7 group. What is the rationale for the decreased need for blood count and chemistry panels during induction and consolidation compared with the 3+7 group?

Company Response:

The monitoring schedule for CPX-351 is reduced compared to the 3+7 group based on the reduced administration count.

(iv) Please clarify whether a complete blood count includes a platelet count, and if so, why platelet count is included as an additional monitoring test.

Company Response:

Separate platelet count was included as an additional test because the testing schedule differs from that of the complete blood count. However, there is some over counting of the number of platelet counts where they overlap with complete blood counts and the number of platelet tests may be reduced in the model correspondingly. This overlap, however, is very low in cost and balanced across arms, and thus has nominal impact on cost-effectiveness.

(v) The model only includes transfusions in the "best supportive care" health state. However, blood products are also recommended for use during intensive chemotherapy for AML. Please clarify why these were not included in the induction and consolidation health states.

Company Response:

Inclusion of transfusions in the induction and consolidation procedures is a reasonable addition to the model. These costs are balanced across treatment arms, however, and thus have nominal impact on cost-effectiveness.

B14. Please provide additional summary statistics for body surface area from the 301 trial (the standard error in addition to the mean value), in each treatment group separately and combined.

Company Response:

Please see BSA listed in Table 11*.*

Table 11. BSA (ITT population)

B15. Please incorporate any observed dose reductions requested in question A8 into the calculations on the vial usage of CPX-351 and 3+7, if appropriate.

Company Response:

Only minimal reductions in dose were observed in the clinical trial for either arm and these reductions were similar between the arms (

[redacted]*). As such, no reductions in vial usage associated with dose reductions were incorporated in the model.*

Quality of life

B16. In the analysis, CPX-351 is assumed to be associated with a lower treatment-related disutility than 3+7, which persists through induction and consolidation (table 33 of the company submission).

(i) Please describe the basis on which this assumption was made.

Company Response:

While all chemotherapy regimens are associated with risks and side effects, various regimens differ in terms of the patient experience. In this case, the types of differences (e.g., degree of fatigue and hair loss) were identified and described based on reports from clinicians who had personal experience treating patients with the more established regimens (e.g., 3+7, 2+5, HiDAC) and CPX-351.

(ii) The vignettes in the utility elicitation study (appendix N of the company submission) suggest a difference in levels of fatigue, risk of infection and hair loss, with less severe symptoms associated with CPX-351. Please provide supporting clinical evidence for these assumptions (e.g. rates of these events that suggest a meaningful difference between treatments).

Company Response:

In the clinician interviews conducted to inform health state development, physicians and nurses consistently and confidently reported that CPX-351 was associated with milder side effects than the standard of care. For example, one oncology nurse said "With 3+7, they are bald. With CPX-351, some of them would lose some hair, and some of them would lose only minor hair. I don't remember any of them being totally bald." She summarised adverse events of the standard of care (e.g., 3+2, 2+5) as "definite hair loss… major fatigue" compared with "minimal hair loss…mild fatigue" for CPX-351. Another oncology nurse said "I have seen a little bit less hair loss with [CPX-351]…less fatigued with [CPX-351]." Based on this clinician input, the health states included two of these differences (less fatigue, less hair loss).

(iii) The vignettes also appear to suggest that consolidation therapy has a similar safety profile to induction therapy, for both 3+7 and CPX-351. Please comment on whether this is a clinically plausible assumption, and provide any supporting evidence from the CLTR0310-301 trial (e.g. rates of these events during induction therapy and during consolidation therapy).

Company Response:

Health states (including descriptions of induction and consolidation) were drafted based on discussions with the clinicians who had direct experience treating patients with conventional chemotherapy and CPX-351. After the health states were drafted, they were presented to the clinicians so that they could be edited for clarity and accuracy. All clinicians agreed that the descriptions of induction and consolidation were plausible and reasonable representations of the typical patient experience. For example, the clinicians described the typical patient experience following consolidation, with one noting "they're still feeling less fatigued during the consolidation [with CPX-351] than the other treatment. Less fatigued with [CPX-351]."

Section C: Textual clarifications and additional points

C1. Table 16 states that 3 patients in the 3+7 group (2%) had bacteraemia, but the text states [redacted] Please clarify which figure is correct.

Company Response

Both values are correct. Table 16 lists grade 3-5 bacteraemia (3 patients, 2%) whereas the main text covers grade 1-5 bacteraemia ([redacted]).

C2. Please clarify whether the numbers of records identified, reported in the PRISMA diagram (figure 1, appendix D, page 15), are correct. They differ from the number of records reported from the searches of each database for MEDLINE, EMBASE and CENTRAL in tables 1, 2 and 3 in Appendix D.

Company Response:

The search results presented in Appendix D Tables 1, 2, and 3 provide the correct search strings that were run. The number of records identified in the table is however not correct because this table was not updated on the day the search was actually run in OVID and exported to Endnote. The correct number of records identified is reported in the PRISMA diagram (i.e., 953 for Medline, 1823 for Embase, and 1076 for Cochrane). All PRISMA figures represent the actual number of records exported from OVID to Endnote and then screened, excluded and included for review.

C3. The coefficient for consolidation therapy in post-transplant OS in the executable model (presented in the "Clinical Data" sheet cell E383) does not match the value in appendix M (table 77, page 306). Please clarify the correct estimate for this parameter.

Company Response:

The correct coefficients are denoted in the executable model, including the value of [redacted] in cell E383 of the "Clinical Data" sheet. Please see the updated table (Appendix M, Table 77) below, with the typos corrected in red text.

Table 77. Log-normal Survival Model for Post-transplant OS in Weeks among Patients in Pathways A, C, and E, 3+7 Final Model

[redacted]	[redacted]	[redacted]	[redacted]
[redacted]	[redacted]	[redacted]	[redacted]
[redacted]	[redacted]		
[redacted]		[redacted]	[redacted]
[redacted]		[redacted]	[redacted]
[redacted]	[redacted]		
[redacted]		[redacted]	[redacted]
[redacted]		[redacted]	[redacted]
[redacted]	[redacted]		
[redacted]		[redacted]	[redacted]
[redacted]		[redacted]	[redacted]

AML, acute myeloid leukaemia; CL, confidence limit; CMMoL, chronic myelomonocytic leukaemia; HMA, hypomethylating agent; MDS, myelodysplastic syndrome; t-r, treatment related

Bibliography

Burnett AK, R. N. (2017). A comparison of clofarabine with ara-C, each in combination with daunorubicin as induction treatment in older patients with acute myeloid leukaemia. *Leukemia, 31*(2), 310-317.

Hulegårdh E, N. C. (2015). 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, 90*(3), 208-14.

Martin PJ, C. G. (2010). Life expectancy in patients surviving more than 5 years after hematopoietic cell transplantation. *J Clin Oncol, 28*(6), 1011-6.

Milligan DW, G. D. (2006). Guidelines on the management of acute myeloid leukaemia in adults. *Br J Haematol, 135*(4), 450-74.

Østgård LS, N. J. (2015). Comorbidity and performance status in acute myeloid leukemia patients: a nation-wide population-based cohort study. *29*(3), 548-55.

Shimoni A, L. M. (2016). 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. *J Hematol Oncol., 9*(1), 118.

Wingard JR, M. N. (2011). Long-term survival and late deaths after allogeneic hematopoietic cell transplantation. *J Clin Oncol, 29*(16), 2230-9.

NICE National Institute for Health and Care Excellence

Patient organisation submission

Liposomal cytarabine and daunorubicin for untreated acute myeloid leukaemia [1225]

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, founded in 1969. We are dedicated to ensuring that anyone affected by blood cancer receives the right information, advice and support. Approximately 85-90% of our income comes from fundraising activities – such as legacies, community events, marathons etc. Leukaemia Care also received funding from a wide range of pharmaceutical companies, but in total those funds are less than 15% of our annual income. Leukaemia Care has undertaken a voluntary commitment to adhere to specific policies that regulate our involvement with the pharmaceutical industry set out at: http://www.leukaemiacare.org.uk/wp-content/uploads/2018/02/CODE-OF-PRACTICE.pdf
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?	This submission is informed by a patient experience survey of 373 adults diagnosed with acute myeloid leukaemia (AML), carried out by Leukaemia CARE in 2016. This was part of a wider survey of 2,019 leukaemia patients entitled 'Living with Leukaemia'. The results of this survey were published in September 2017 and are available online at: www.leukaemiacare.org.uk/living-with-leukaemia. AML specific breakdowns of the data have been used to inform this submission. We also gather information through our support services (helpline, support groups, conferences, communications with our membership) and one to one discussion with patients.

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) is a rapidly progressing form of leukaemia. In 2014, there were 2,590 cases diagnosed in England and 201 diagnosed in Wales. In 2014, there were 2,127 deaths in England and 130 in Wales. Patients with secondary AML - therapy related and myelodysplasia-related changes (e.g. prior MDS/CMML or de novo with MDS karyotype) – are high risk groups. These patients have poor survival outcomes, a clear end of life setting. The most common symptoms encountered by AML patients since their diagnosis are fatigue (73%), feeling weak or breathless (51%), memory loss or loss of concentration (38%), bleeding and bruising (37%), itchy skin (35%), nausea or vomiting (35%), sleeping problems (34%), infections (32%), bone or joint pain (31%), weight loss (28%) and muscle pain (23%). AML can have a huge emotional impact, prompting patients (and their families) to experience feelings of disbelief, denial, anger, fear, blame, guilt, isolation and depression. In our survey, 51% of AML patients reported that they have felt depressed or anxious more often since their diagnosis. The emotional impact does not affect the patient in isolation and is often also felt by carers and family members. This can place huge emotional strain on families and friends, many of whom may be affected by the diagnosis. As such, improvements in a patients' treatment and prognosis will also have a wider impact on the lives of their family and friends. AML also has a wider practical impact, with 52% of patients experiencing pain as a direct result of their condition (31% occasionally, 17% regularly and 4% constantly). Additionally, 51% of patients have difficulty moving around (sometimes 27%, often 15% and always 9%) and 69% of AML patients have difficulty performing some of their daily routines, such as cooking or cleaning. Another 38% reported that they have problems taking care of themselves. Of those in work or education before their diagnosis, 77% have been impacted (32% reduced hours, 45% no longer able to work or continue education).

	Consequently, 53% of AML patients reported a negative financial impact as a result of having cancer (increased costs or reduced income).
Current treatment of the condition in the NHS	
7. What do patients or carers think of current treatments and care available on the NHS?	For patients who are fit for intensive chemotherapy, treatment for newly diagnosed AML patients is induction with daunorubicin and cytarabine (3+10), followed by high-dose cytarabine in the consolidation phase. The most common side effects reported by AML patients were fatigue (76%), hair loss (54%), neutropenia (44%), diarrhoea (41%), sore mouth (40%), nausea or vomiting (39%), muscle or joint pain (34%), loss of concentration or memory (33%), constipation (29%), bone and joint pain (28%), sleeping problems (28%), anaemia (26%), weight loss (25%), fever (25%), bruising (22%), breathing difficulties (20%) and dizziness (20%).
8. Is there an unmet need for patients with this condition?	AML patients have an extremely poor prognosis, with AML accounting for over half of all leukaemia deaths. There has been limited progress in the treatment of AML since the 1990s. There is an urgent need for access to new treatment options. When asked what they considered to be important features of a new treatment, AML patients listed: improved or longer survival (86%), improved quality of life (70%), a remission or response (61%), tolerable side effects (56%), improved blood counts or test results (50%), a reduced impact on carers or family members (42%) and certainty of available treatment data (31%).

Advantages of the technology	
9. What do patients or carers think are the advantages of the technology?	• Bridge to transplant – higher transplant rate and better transplant outcomes. In our survey 89% of AML patients reported that they would positively welcome a treatment that enabled/bridged to stem cell transplant. • More tolerable – in particular, reduced alopecia – a key quality of life issue, particularly for women • Reduced early mortality (AML progression reduced) – 30 day and 60 day • CR and CRi rates – higher and quicker – could it reduce time to transplant in the future? This is a key issue, given the rapid early mortality in AML. • Preferable treatment administration schedule – infusion over 90 minutes
Disadvantages of the technology	
10. What do patients or carers think are the disadvantages of the technology?	Quality of Life was not measured in the clinical trial. However, some side effects were reported at a higher rate including infections and hypertension. Delayed count recovery
Patient population	
11. Are there any groups of patients who might benefit more or less from the technology than others? If so,	Secondary AML (as per the FDA licensed indication – EMA indication unknown) is a high-risk sub-group. The population is too small to further subgroup accurately.

please describe them and explain why.	
Equality	
12. Are there any potential equality issues that should be taken into account when considering this condition and the technology?	The Phase 3 trial only included patients aged 60-75. However, it is likely that it would be used to treat adults more broadly, as per the FDA indication (EMA indication unknown). Secondary AML patients of all ages are likely to be chemo-resistant, so the outcomes would be broadly consistent. It is anticipated that this regimen would be used to bridge patients to transplant. Looking at a breakdown of our AML patient survey data by age, the percentage of patients who have had a stem cell transplant is: 70% (35-44), 56% (45-54); 559% (55-64), 41% (65-74) and 5% (75-84). As such, it is likely to be extrapolated towards younger patients, rather than towards older patients. However, a further key question would be: does this regimen enable more older patients to proceed to transplant? Our survey highlighted that older patients were significantly less likely to receive a stem cell transplant. A full analysis by age group is available at: http://www.leukaemiacare.org.uk/get-involved/our-campaigns/leukaemia-i-wasnt-born-yesterday/ Quoting from the report: “Older patients are more likely to experience other issues such as comorbidities, social isolation and difficulties getting to and from hospital. However, the needs of active older people in otherwise good health are very different from those living with frailty and other conditions. Whilst stem cell transplants are associated with significant risks, there is evidence that stem cell transplantation is an option for patients over 70. As such, Macmillan’s recommendations remain applicable in this situation. Age alone should never be a barrier to treatment. Treatment decisions should be individual, based on their ‘ability to tolerate treatment, quality of life or personal preferences’ rather than chronological age.” (Page 15, Leukaemia: I wasn’t born yesterday).

Other issues	
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:

- AML is a rapidly progressing form of leukaemia, with significant mortality. There has been limited progress in the treatment of AML since the 1990s and there is an urgent need for access to improve treatment options.
- Patients with secondary AML - therapy related and myelodysplasia-related changes (e.g. prior MDS/CMML or de novo with MDS karyotype) – are high risk groups. These patients have extremely poor survival outcomes, a clear end of life setting.
- AML also has a significant symptom burden (fatigue, feeling weak or breathless, memory loss or loss of concentration, bleeding and bruising, itchy skin, nausea or vomiting, sleeping problems, infections, bone or joint pain, weight loss and muscle pain), as well as a practical, financial and emotional impact.
- Liposomal cytarabine and daunorubicin offers a number of benefits – in particular, improved transplant rates and outcomes. This is a key benefit, as SCT represents the only curative option for these patients.

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

Liposomal cytarabine and daunorubicin for untreated acute myeloid leukaemia [1225]

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? ☐ 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): Joint organisational response
5a. Brief description of the organisation (including who funds it).	NCRI AML Working Group – The Association of Cancer Physicians – Royal College of Physicans
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 treatment is curative therapy. Firstly by achieving remission and then by giving further chemotherapy / and allogeneic transplantation to prevent relapse.

or prevent progression or disability.)	
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 (CR) and overall survival. Proportion of patients 'bridged' to transplant
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. Secondary AML is associated with a particularly poor survival and is a devastating and perhaps unrecognised complication of chemo-radiotherapy for solid tumours.
What is the expected place of the technology in current practice?	
9. How is the condition currently treated in the NHS?	Patients with secondary AML considered fit enough for intensive therapy are entered into NCRI AML trials. Currently these are AML 19 for younger patients and AML 18 for patients >60 years.. If patients achieve a CR then they are candidates for allo transplant as a curative intervention. Older patients (>75 years) and those with significant co-morbidities that preclude intensive chemotherapy are likely to receive non-intensive therapies including azacytidine or BSC

• Are any clinical guidelines used in the treatment of the condition, and if so, which?	AML18 and 19 are used as guidelines. The ELN AML guidelines published in 2017 were written before the approval of CPX-351 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
• What impact would the technology have on the current pathway of care?	None although it is likely that more patients could receive chemotherapy in the ambulatory setting.
10. Will the technology be used (or is it already used) in the same way as current care in NHS clinical practice?	Yes, with proviso as in section 9 concerning ambulatory care
• How does healthcare resource use differ	CPX would replace standard DA3+10 chemotherapy

between the technology and current care?	
• 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.)	Our experts believe very little is required. Haematologists already have significant experience of using CPX from NCRI AML clinical trials
11. Do you expect the technology to provide clinically meaningful benefits compared with current care?	A clinical trial in patients with secondary AML (patients aged 60 to 75 years with a history of prior cytotoxic treatment, antecedent myelodysplastic syndrome (MDS) or chronic myelomonocytic leukemia, or AML with World Health Organization–defined MDS-related cytogenetic abnormalities or prior myelodysplastic syndromes) compared CPX-351 to standard daunorubicin/cytarabine showed a superior survival in those randomized to the novel agent (median, 9.56 vs 5.95 months; hazard ratio, 0.69; P = .005).. Furthermore a higher proportion of patients proceeded to BMT and the outcome after transplant was superior to those who had received standard therapy. Although the study was in patients aged 60-75 years CPX was recently approved for all patients with secondary AML by the FDA irrespective of age
• Do you expect the technology to increase length of life more than current care?	Yes the trial quoted about improved OS and allowed more patients to proceed to allograft
• Do you expect the technology to increase	

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?	The subgroup of patients with AML in the trial were those with a history of prior cytotoxic treatment for other tumours those with an antecedent myelodysplastic syndrome (MDS) or chronic myelomonocytic leukemia, or those with AML with World Health Organization–defined MDS-related cytogenetic abnormalities or prior myelodysplastic syndromes). These groups of patients can easily be identified in the clinic
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	CPX will be easier to use than DA chemotherapy. Treatment (particularly for consolidation cycles) can be ambulatory and although the regimen is myelotoxic there is evidence that there is less non-haematological toxicity

or ease of use or additional 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	CPX-351 is a lamellar-encapsulated vehicle that delivers daunorubicin and cytarabine in a fixed molar ratio with enhanced incorporation into the bone marrow and AML cells. While not strictly a 'targeted' therapy, in that there is no specific molecular target or determinant of response, this drug is a formulation of chemotherapy that allows more effective delivery to the malignant cells – so, in a way, it really is a targeted

benefits and how might it improve the way that current need is met?	agent. Clinical experience has shown that it is well tolerated with a favourable toxicity profile compared to standard DA 3+10 chemotherapy
• Is the technology a 'step-change' in the management of the condition?	Yes. There has been no improvement in the outcome of this group of patients for many years. Many patients are resistant to standard chemotherapy and do not make it to a BMT
• Does the use of the technology address any particular unmet need of the patient population?	Yes. By increasing CR rate, reducing relapse risk and delivering more patients to the possibility of allo transplant and cure
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?	No
Sources of evidence	
18. Do the clinical trials on the technology reflect current UK clinical practice?	Yes. Both the NCRI AMNL18 and AMI19 trials incorporate CPX into the treatment of high risk AML

• If not, how could the results be extrapolated to the UK setting?	
• What, in your view, are the most important outcomes, and were they measured in the trials?	CR rate, Overall survival. Numbers bridged to transplant. Outcome post BMT
• If surrogate outcome measures were used, do they adequately predict long-term clinical outcomes?	Surrogates were not used. The trial showed a survival benefit
• 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

21. How do data on real-world experience compare with the trial data?	There is a lack of 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.

- Secondary AML has a poor outcome with standard chemotherapy
- The only curative treatment is intensive chemotherapy followed by allogeneic BMT
- CPX impoves CR rates, increases the number of patients going to transplant and improves overall survival
- UK haematologists and pharmacists are familiar with the drug
- CPX is well tolerated and appears to have reduced non-haematological toxicity

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

Liposomal cytarabine and daunorubicin for untreated acute myeloid leukaemia [1225]

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]
2. Name of organisation	Royal College of Pathologists and British Society for Haematology
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? ☒ 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).	Royal College of Pathologists and British Society of Haematology: **healthcare professional organisation that represents clinicians**
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 disability.)	The treatment is a potentially curative therapy. Firstly by achieving remission, then further consolidation chemotherapy and/or allogeneic transplantation are then utilised to prevent relapse of the disease.

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 a complete response/remission- which will improve quality of life. Increased numbers of patients sufficiently responding, to then be successfully bridged to allogeneic transplant. Ultimately improvement in overall survival.
8. In your view, is there an unmet need for patients and healthcare professionals in this condition?	Yes. The standard chemotherapy employed has been unchanged for many years. Improvement in outcomes for this group of patients has largely been through the appropriate utilisation of stem cell transplantation and improvements in supportive care. More effective/better tolerated treatments are urgently required.
What is the expected place of the technology in current practice?	
9. How is the condition currently treated in the NHS?	Most patients are considered fit enough for intensive chemotherapy, generally induced with a combination of Daunorubicin and Cytarabine (DA). If they are at high risk of relapse they undergo allogeneic stem cell transplant. If their relapse risk is low then they have consolidation chemotherapy- most commonly with high dose Cytarabine therapy. A substantial proportion of these patients are entered into the NCRN AML 18 and 19 trials- which evealuate the role of dose escalation, new agents, MRD directed therapy and stem cell transplantation. Frailer/older patients with comorbidity receive non-intensive therapy such as Azacitidine/Low Dose Cytarabine or supportive care.
• Are any clinical guidelines used in the treatment of the	The current UK based AML treatment guidelines have not been updated in time to evaluate this technology. The European Leukaemia Network (ELN) guidelines were updated in 2017- prior to FDA/EMA approval of CPX-351.

condition, and if so, which?	
• 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.)	Designated Secondary and Tertiary centres for the management of AML are well defined within NHS England. Treatment application is uniform.
• What impact would the technology have on the current pathway of care?	This would be replacement of the current standard care (DA) in specifically defined settings- essentially the pathway would be unaltered.
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?	Replacement of DA- resource utilisation would likely be equivalent.
• In what clinical setting should the technology be	Secondary/Tertiary Haematology units with experience in the management of AML (well defined).

used? (For example, primary or secondary care, specialist clinics.)	
• What investment is needed to introduce the technology? (For example, for facilities, equipment, or training.)	Little- replacement technology with similar requirements to current standard. Many sites already experienced with CPX-351 as a consequence of the NCRN AML trails.
11. Do you expect the technology to provide clinically meaningful benefits compared with current care?	CPX-351 is designed as a liposomal formulation of the DA combination in a 5:1 ratio of cytarabine and daunorubicin, which was proved to be an optimal combination, with the highest level of synergy and the lowest level of antagonism (Mayer etal Mol Cancer Ther 2006). Two phase II randomized studies in 127 (Elderly untreated AML) and 125 patients (AML first relapse) both confirmed a higher rate of CR (66.7 vs 51.2%, and 49.3 vs 40.9%, respectively) for patients treated with CPX-315 compared with those receiving 7 + 3 regimen, although no difference in EFS or OS has been found in both phase II trials (Lancet etal Blood 2014, Cortes etal Cancer 2015). There was a trend for improved survival in secondary AML (p 0.01) which presumably inspired the design of the phase 3 study. In the phase 3 study for untreated high risk (Secondary AML) defined as:- ◦Therapy related AML: t-AML must have a documented history of prior cytotoxic therapy or ionizing radiotherapy for an unrelated disease ◦AML with a history of myelodysplasia: MDSAML must have bone marrow documentation of prior MDS ◦AML with a history of CMML: Must have bone marrow documentation of prior CMML ◦De novo AML with karyotypic abnormalities characteristic of MDS: de novoAML must have cytogenetics with abnormalities per WHO. A total of 309 patients were randomized (153 to CPX-351 + 156 to 7+3) and were well balanced for sex, race, age, performance status, AML-subtype, MDS-related cytogenetics and prior HMA therapy. After

	minimum follow-up of 13.7 months final analysis began.CPX-351 treatment resulted in superior overall survival (HR=0.69; P=0.005; median OS 9.56 vs. 5.95 months), EFS (HR=0.74; P=0.021), and CR+CRi response (47.7% vs. 33.3%; P=0.016). 60-day mortality favored CPX-351 (13.7% vs. 21.2%). Grade 3-5 AEs were equal (92% vs. 91%) and were similar in frequency and severity in both arms
• Do you expect the technology to increase length of life more than current care?	Yes- as above.
• Do you expect the technology to increase health-related quality of life more than current care?	The expectation is that achievement of CR correlates with an improvement in QOL- as such patients have less complications and supportive care requirements. Perhaps the data is lacking to directly support this with CPX-351 currently..
12. Are there any groups of people for whom the technology would be more or less effective (or appropriate) than the general population?	More appropriate for the defined groups of:- Therapy related AML: t-AML must have a documented history of prior cytotoxic therapy or ionizing radiotherapy for an unrelated disease ◦AML with a history of myelodysplasia: MDSAML must have bone marrow documentation of prior MDS ◦AML with a history of CMML: Must have bone marrow documentation of prior CMML ◦De novo AML with karyotypic abnormalities characteristic of MDS: de novoAML must have cytogenetics with abnormalities per WHO.
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 tests or monitoring needed.)	Comparable
14. Will any rules (informal or formal) be used to start or stop treatment with the technology? Do these include any additional testing?	Comparable to current standard therapy
15. Do you consider that the use of the technology will result in any substantial health-	No

related benefits that are unlikely to be included in the quality-adjusted life year (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?	This is the first example of utilisation of 'combiplex' technology in AML- which proposes a more targeted/synergistic delivery of conventional chemotherapeutic agents. The published (in abstract) phase 3 data in the pre-specified sub-populations of AML indicates improved response with lower toxicity.
• Is the technology a 'step-change' in the management of the condition?	Yes
• Does the use of the technology address any particular unmet need of the patient population?	The pre-specified sub-populations of AML are those with a particularly poor outcome. This is due to a reduced remission and increased relapse rate compared to primary/standard AML. As such CPX-351 appears to improve the response rate- such that more patients are able to undergo curative stem cell

	transplantation (34% vs 25 %) after CPX-351 compared to standard therapy. Additionally the transplant related complications were lower in the CPX-351 treated patients.
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?	Essentially comparable with standard therapy- Neutropenia, thrombocytopenia and rashes are more in evidence with CPX-351 therapy treated patients but seems very manageable.
Sources of evidence	
18. Do the clinical trials on the technology reflect current UK clinical practice?	CPX-351 is incorporated in the current NCRN AML clinical trials (18, 19). However the standard therapy arm in the phase 3 study has variation in that rather than the '7+3'being (Cytarabine 100 mg/m2/day x 7 days- continuous infusion , Daunorubicin 60 mg/m2 days 1, 2, 3) the UK standard is DA 3+10 (Daunorubicin 60 mg/m2 daily by IV infusion on days 1, 3 and 5 (3 doses) Cytarabine 100 mg/m2 12-hourly by i.v. push on days 1 – 10 inclusive (20 doses). The UK approach therefore has an alternative Daunorubicin schedule and significantly more cytarabine administered in an alternative approach intra-venously.
• If not, how could the results be extrapolated to the UK setting?	Unclear what impact this might have on response endpoints.

• What, in your view, are the most important outcomes, and were they measured in the trials?	Overall survival, complete response rate, treatment related death rate and proportion bridged to transplant. All assessed within the trial.
• 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
21. How do data on real-world experience compare with the trial data?	None available to my knowledge

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.	N/A
Key messages	
24. In up to 5 bullet points, please summarise the key messages of your submission. • Secondary AML has a poor outcome • Potentially curative therapy can be achieved through remission induction chemotherapy and allogeneic stem cell transplant • CPX-351 improves remission rates compared to '7+3', such that more patients are bridged to transplant • CPX-351 has a lower 60 day and transplant associated mortality • CPX-351 improves survival in the phase 3 study.	

Thank you for your time.

Please log in to your NICE Docs account to upload your completed submission.

NHS England submission on liposomal daunorubicin and cytarabine in newly diagnosed high risk acute myeloid leukaemia (AML)

1. Liposomal daunorubicin and cytarabine would potentially replace standard doses and scheduling of daunorubicin and cytarabine ie would be in the same place in the treatment pathway for those patients with high risk AML who are fit enough for intensive chemotherapy.
2. NHS England notes that though the clinical trial (study 301) was performed in patients restricted to the ages of 60-75 years, the likely marketing authorisation will be in adults without any such stipulation. There is no biologically plausible reason why benefit in the 60-75 year group would not translate into patients aged outside of this range.
3. NHS England notes the different doses and scheduling of liposomal daunorubicin and cytarabine between the 1st induction cycle and the 2nd induction cycle and then in the consolidation cycles. The body surface area used in the economic model (1.8m²) is less than that in the clinical trial (median 2.0m²) and NHS England notes that 60% of the trial participants were male (male patients are taller and tend to have higher body surface areas). There needs to be a clear justification for the use of 1.8m² in the economic model and NHS England supports the ERG in using a range of body surface areas. Wastage of drug also needs to be accounted for.
4. Overall survival data are immature as the median duration of follow-up is only about 20-21 months. There are few patients at risk beyond 21 months. This immaturity is curious as the company's submission is based on a data cut that was performed in December 2015 (over 2½ years ago). NHS England is surprised that the company does not have more mature follow up data.
5. NHS England notes that the stem cell transplant rate of 34% with liposomal daunorubicin and cytarabine was higher than that with standard treatment (24%) although this difference was not statistically significant.
6. NHS England notes that serious adverse events were similar in both arms.
7. There was no quality of life data collection incorporated into the design of the clinical trial. NHS England regrets this very significant omission as we are in 2018 and also because the company has had to resort to using non-trial utility data that is based on a very different age range and has not captured sufficiently the toxicity of treatment. The utilities of treated patients thus look too high.
8. NHS England notes that the company has chosen a very optimistic survival model with a [redacted] chance of long term survival. Given the immaturity of the survival data, this is almost certainly unrealistic.
9. The company has assumed that the mortality rates of patients who are cured after stem cell transplantation are those of the general population. This assumption is

incorrect as it has long been known of the much higher mortality rates post stem cell transplantation.

10. NHS England observes that there was no difference in inpatient stay between the 2 arms in study 301. There may be a small difference in inpatient stay as regards the administration of treatment but most inpatient stays relate to the consequences of the disease and the toxicities of treatment.
11. NHS England sees the advantages to patients of liposomal daunorubicin and cytarabine as it is better than standard therapy in high risk AML and may improve rates of stem cell transplantation. The survival data is immature and thus there is considerable uncertainty as to the degree of long term benefit. Without further follow-up data analyses, the cost of liposomal daunorubicin and cytarabine will have to be used to offset this considerable uncertainty as to the degree of benefit.

[redacted]

[redacted]

July 2018

NICE National Institute for Health and Care Excellence

Clinical expert statement

Liposomal cytarabine and daunorubicin for untreated acute myeloid leukaemia [ID1225]

Thank you for agreeing to give us your 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 expert statement

- 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	**Royal College of Pathologists**

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? ☒ 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):
5. Do you wish to agree with your nominating organisation's submission? (We would encourage you to complete this form even if you agree with your nominating organisation's submission)	☒ yes, I agree with it ☐ no, I disagree with it ☐ I agree with some of it, but disagree with some of it ☐ other (they didn't submit one, I don't know if they submitted one etc.)
6. If you wrote the organisation submission and/ or do not have anything to add, tick here. (If you tick this box, the rest of this form will be deleted after submission.)	☐ yes

Topic-specific questions	
24. Is azacitidine (excluded from company submission) considered to be established clinical practice in the NHS for the same people who would be eligible to receive liposomal cytarabine and daunorubicin?	I think these are, in the main, different populations. CPX is intensive therapy. Azacytidine is non-intensive therapy given in patients considered not fit for intensive therapy based on age, frailty and co-morbidity.
25. The key trial included people aged 60 to 75 years. How appropriate would it be to extrapolate the results of the trial to adults of all ages?	Yes this is fair as there is no evidence that secondary AML arising in younger patients is biologically different from the 60-75 year age group.

Thank you for your time.

Please log in to your NICE Docs account to upload your completed statement, declaration of interest form and consent form.

NICE National Institute for Health and Care Excellence

Clinical expert statement

Liposomal cytarabine and daunorubicin for untreated acute myeloid leukaemia [ID1225]

Thank you for agreeing to give us your 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 expert statement

- 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	**Dr Priyanka Mehta**
2. Name of organisation	**University Hospitals Bristol NHS Trust**

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? √☐ 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):
5. Do you wish to agree with your nominating organisation's submission? (We would encourage you to complete this form even if you agree with your nominating organisation's submission)	√☐ yes, I agree with it ☐ no, I disagree with it ☐ I agree with some of it, but disagree with some of it ☐ other (they didn't submit one, I don't know if they submitted one etc.)
6. If you wrote the organisation submission and/ or do not have anything to add, tick here. (If you tick this box, the rest of this form will be deleted after submission.)	☐ yes

The aim of treatment for this condition	
7. What is the main aim of treatment? (For example, to stop progression, to improve mobility, to cure the condition, or prevent progression or disability.)	The main aim of intensive treatment of AML is to cure. This is achieved by improving response to treatment and reducing the risk of relapse.
8. 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 response/ remission, allowing better chance of receiving a successful transplant without any increase in toxicity. Improved overall survival
9. In your view, is there an unmet need for patients and healthcare professionals in this condition?	Yes, there is an unmet need in both intensively and nonintensively treated AML. For the purpose of this technology appraisal, the unmet need in intensively treated AMLs is the CR, DFS and OS rates which needs to be improved. For the high risk AML's, the best chance of cure is an allogeneic stem cell transplantation and outside of the unmet needs in the area of transplantation (donor availability etc), there is a need to get higher proportion of patients to transplantation and reducing the risk of relapse and toxicity post transplantation.
What is the expected place of the technology in current practice?	

10. How is the condition currently treated in the NHS?	Conventional intensive chemotherapy is Daunorubicin + Cytarabine (3+10). Majority of the patients receive this treatment in the context of national clinical trials, where it has been combined with mylotarg in recent trials. Risk stratification post course 1 induction therapy helps with decision regarding consolidation with chemotherapy versus allogeneic stem transplantation, in patients less than 60 years of age. For patients over 60 years of age, allogeneic transplantation has shown to be superior to chemotherapy as consolidation and is offered to all, depending on factors such as disease status, performance status, comorbidities and suitable donor availability
• Are any clinical guidelines used in the treatment of the condition, and if so, which?	BCSH AML guidelines Diagnosis and management of AML in adults: 2017 ELN recommendations from an international expert panel Hartmut Döhner, Elihu Estey, David Grimwade, Sergio Amadori, Frederick R. Appelbaum, Thomas Büchner, Hervé Dombret, Benjamin L. Ebert, Pierre Fenaux, Richard A. Larson, Ross L. Levine, Francesco Lo-Coco, Tomoki Naoe, Dietger Niederwieser, Gert J. Ossenkoppele, Miguel Sanz, Jorge Sierra, Martin S. Tallman, Hwei-Fang Tien, Andrew H. Wei, Bob Löwenberg and Clara D. Bloomfield Blood 2017 129:424-447; doi: https://doi.org/10.1182/blood-2016-08-733196
• 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.)	The pathway of care in the UK is well defined. Disease assessment at diagnosis and risk stratification has evolved significantly based on identification of molecular and cytogenetic abnormalities in AML. Care pathway has predominantly been treating patients in the NCRI AML trials for all age groups and this includes risk stratification and additional/ targeted therapies in the treatment algorithms. The mainstay of induction treatment remains Daunorubicin and Cytarabine. FLAG-Ida is currently being compared for younger patients in the induction setting, and some clinicians prefer to use this regime over DA for younger patients with high risk MDS For patients over 60 yrs, there can be some variability amongst centres in the assessment for suitability for intensive treatments and/or transplantation. Assessments are mainly based on performance score and comorbidities (HCT-CI) but there is a need for a more comprehensive geriatric assessment which has been shown to influence post transplant outcomes.

• What impact would the technology have on the current pathway of care?	On the basis of the 301 study outcomes, CPX would influence the current care pathway by allowing a higher proportion of patients to get to transplant and having better survival post transplant. This would be a step change in improving outcomes for the high risk patients. There is considerable difference between CPX 351 and conventional chemo in the delivery i.e for induction, 3 day case attendances versus 10 -12 day inpatient stay in hospital respectively. Resource utilisation and patient experience on these treatments is significantly different.
11. Will the technology be used (or is it already used) in the same way as current care in NHS clinical practice?	Currently it is being used in NCRI AML 19 for induction and consolidation in high risk AML's and relapsed AML's. It is scheduled to be available in the AML 18 trial as part of an upfront randomisation.
• How does healthcare resource use differ between the technology and current care?	Standard chemotherapy requires inpatient stay for minimum10 days (most units 4 weeks until count recovery) and standard supportive care i.e transfusions, management of neutropenic sepsis, organ toxicity CPX 351 is delivered as 3 outpatient/day case 90 minute infusions on days 1,3,5. Supportive care is similar to the standard care, albeit potentially higher transfusion requirement for the delay in count recovery. This is not observed in clinical practice. Patients accept and tolerate CPX351 better as they spend much less time in hospital and this causes lesser disruption at home/ for carers
• In what clinical setting should the technology be used? (For example, primary or secondary care, specialist clinics.)	Secondary/ tertiary care units meeting Haemonc IOG for delivering intensive chemotherapies
• What investment is needed to introduce the technology? (For	No additional facilities are required. Most centres treating AML would have staff experienced and trained to use CPX351 in trials.

example, for facilities, equipment, or training.)	
12. Do you expect the technology to provide clinically meaningful benefits compared with current care?	Yes. A higher chance of remission and successful transplant as compared to current treatments would be beneficial. Patient experience on CPX 351 is much better and it is better tolerated than current care
• Do you expect the technology to increase length of life more than current care?	There is an expectation that better remission rates and survival, particularly post transplant should increase the length of life. Longer trial follow up data is awaited.
• Do you expect the technology to increase health-related quality of life more than current care?	Quality of life data is not available in the 301 trial and this should be available from the current trials. Clinical experience with managing patients on CPX shows significantly better tolerability and less time spent in the hospital; this may translate into an increase in health-related quality of life compared to current care
13. Are there any groups of people for whom the technology would be more or less effective (or appropriate) than the general population?	The findings of the 301 trial may be relevant to younger high risk AML patients, as biologically the same disease.

The use of the technology	
14. 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 tests or monitoring needed.)	CPX 351 is significantly easier to deliver than conventional chemotherapy from healthcare resource utilisation perspective. It is delivered as outpatients/ day case as a 90 minute infusion on days 1,3 and 5 as opposed to minimum 10 day inpatient stay, which has a significant favourable impact on inpatient bed requirements, latter being be a huge pressure in most NHS haematology units. Nursing time, monitoring and allied resource requirements is much lesser with CPX. CPX has short stability and hence most haematology units are unable to provide pharmacy to constitute it over weekends. However, as there is an increasing rationale for starting treatment after the diagnostic test results are available i.e cytogenetics and molecular, which gives extra time for planning the start of induction therapy to avoid weekends. Outpatient monitoring, ambulatory care and follow up is similar to conventional therapy Patients acceptability of CPX 351 is significantly better as day case treatment seems less daunting than a 10-28 day admission. Nursing staff caring for patients on CPX, with experience of treating patients on standard chemo, report better tolerability and side effect profile
15. Will any rules (informal or formal) be used to start or stop treatment with the technology?	Not any different to the standard care

Do these include any additional testing?	
16. 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?	Health related Quality of Life is likely to be superior with CPX due to improved remission rates, survival and reduction in hospital days. Clinical experience suggests CPX is better tolerated than conventional chemo. These factors are likely to impact QALY calculations favourably for CPX
17. 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. CPX 351 is the first example of utilisation of 'combiplex' technology in AML- which delivers a more targeted/synergistic combination of chemotherapeutic agents, resulting in superior PK/PD than achieved by the same drugs as free agents There is evidence of sparing non haematopoetic tissue and preferential localisation in the BM, in animal models. This achieves higher effective drug levels in the BM and thereby exposure and internalising in leukaemic cells. Bypassing the efflux mechanisms in the cell membranes may explain better efficacy in high risk AML's which are more likely to be chemoresistant. The published (in abstract) phase 3 data in the pre-specified sub-populations of AML indicates improved response with lower toxicity

• Is the technology a 'step-change' in the management of the condition?	Yes. This is the first development in the treatment of AML showing improvement in overall survival without excess toxicity, over the conventional DA3+10 which has been used for decades. Easier/ outpatient delivery of the drug is also a welcome change from resource utilisation, to the prolonged hospital stay with standard treatment.
• Does the use of the technology address any particular unmet need of the patient population?	Yes, the pre-specified sub-populations of AML are those with a particularly poor outcome. This is due to a reduced remission and increased relapse rate compared to primary/standard AML. As such CPX-351 appears to improve the response rate- such that more patients are able to undergo curative stem cell Professional organisation submission Liposomal cytarabine and daunorubicin for untreated acute myeloid leukaemia [1225] 9 of 11 transplantation (34% vs 25 %) after CPX-351 compared to standard therapy.
18. How do any side effects or adverse effects of the technology affect the management of the condition and the patient's quality of life?	Trial data shows comparable adverse events. QoL data was not available in the trial. Skin rashes and cytopenias are more with CPX, but have been manageable without requiring additional resources
Sources of evidence	
19. Do the clinical trials on the technology reflect current UK clinical practice?	No. The comparator DA3+7 is not the standard practice in the UK, DA3+10 is routinely used as first line.

• If not, how could the results be extrapolated to the UK setting?	There is no direct data comparing 3+10 with 3+7. Differences in treatment regimes in different countries is often a difficulty observed by the clinical community, whilst trying to interpret/ use data from non UK clinical trials. Consensus tends towards accepting DA3+7 as comparable to DA3+10
• What, in your view, are the most important outcomes, and were they measured in the trials?	CR, OS, SAE's and rate of transplantation: all observed in the trial EFS post transplantation, HRQoL and longer follow up data would be useful
• 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 subsequently?	None that I'm aware of in my clinical practice or in pharmacovigilance of the drug in the AML19 trial so far
20. Are you aware of any relevant evidence that might not be found by a systematic review of the trial evidence?	No

21. Are you aware of any new evidence for the comparator treatment(s) since the publication of NICE technology appraisal guidance [TAXXX]?	No
22. How do data on real-world experience compare with the trial data?	Not aware of any real world data
Equality	
23a. Are there any potential equality issues that should be taken into account when considering this treatment?	No
23b. Consider whether these issues are different from issues with current care and why.	NA
Topic-specific questions	

24. [To be added by technical team if required, after receiving the company submission. For example, if the company has deviated from the scope (particularly with respect to comparators) – check whether this is appropriate. Ask specific, targeted questions such as “Is comparator X [excluded from company submission] considered to be established clinical practice in the NHS for treating [condition Y]?”] **if not delete highlighted rows and renumber below**	

Key messages
25. In up to 5 bullet points, please summarise the key messages of your statement. • High risk AML is difficult to treat with poor outcomes • Best chance of cure is with remission inducing chemo followed by an allogeneic stem cell transplant • CPX 351 has used combiplex technology and significantly improved on the PK/PD over free drugs, enhancing synergy , localising in BM and bypassing resistance mechanisms • CPX 351 requires fewer days in hospital as inpatients; favourably impacting on busy wards and patient experience • Improves overall survival in high risk AML without increasing toxicity

Thank you for your time.

Please log in to your NICE Docs account to upload your completed statement, declaration of interest form and consent form.

……

Your privacy

The information that you provide on this form will be used to contact you about the topic above.

☐ **Please tick this box** if you would like to receive information about other NICE topics.

For more information about how we process your personal data please see our privacy notice.

……

NICE National Institute for Health and Care Excellence

Patient expert statement

Liposomal cytarabine and daunorubicin for untreated acute myeloid leukaemia [ID1225]

Thank you for agreeing to give us your 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 expert statement

- 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	***Mark Sandford***
2. Are you (please tick all that apply):	☐ a patient with the condition? ✓ ☐ a carer of a patient with the condition? ☐ a patient organisation employee or volunteer?

	☐ other (please specify):
3. Name of your nominating organisation	***Leukaemia Care***
4. Did your nominating organisation submit a submission?	☐ yes, they did ☐ no, they didn't ☐ I don't know ✓
5. Do you wish to agree with your nominating organisation's submission? (We would encourage you to complete this form even if you agree with your nominating organisation's submission)	☐ yes, I agree with it ☐ no, I disagree with it ☐ I agree with some of it, but disagree with some of it ☐ other (they didn't submit one, I don't know if they submitted one etc.) ✓

6. If you wrote the organisation submission and/ or do not have anything to add, tick here. (If you tick this box, the rest of this form will be deleted after submission.)	☐ yes
7. How did you gather the information included in your statement? (please tick all that apply)	☐ I have personal experience of the condition ✓ ☐ I have personal experience of the technology being appraised ☐ I have other relevant personal experience. Please specify what other experience: ☐ I am drawing on others' experiences. Please specify how this information was gathered:
Living with the condition	
8. What is it like to live with the condition? What do carers experience when caring for someone with the condition?	***Living with AML is worrying, stressful, irritating, a bit crap really. The biggest thing is to mentally get your head around everything and deal with the boredom. Physically it is only debilitating as it makes you feel so weak.***

Current treatment of the condition in the NHS	
9. What do patients or carers think of current treatments and care available on the NHS?	***Considering I am filling out this form and not six feet under I would say it is excellent. All the staff are fantastic and can't do enough for you. Some people really take the system and treatment for granted.***
10. Is there an unmet need for patients with this condition?	***I think it is fairly well managed but the need would possibly be for prevention of getting other illnesses and increasing the success rate of treatment.***
Advantages of the technology	
11. What do patients or carers think are the advantages of the technology?	***You don't feel as ill (not throwing up as much), don't lose your hair and you generally feel better in yourself.***
Disadvantages of the technology	
12. What do patients or carers think are the disadvantages of the technology?	***You feel ill with the chemo but this is a small price to pay. I had bowel problems either one way or the other, never a happy medium.***
Patient population	
13. Are there any groups of patients who might benefit	***Women would definitely benefit due to the fact I didn't lose my hair and this can be a greater concern to women than men.***

more or less from the technology than others? If so, please describe them and explain why.	
Equality	
14. Are there any potential equality issues that should be taken into account when considering this condition and the technology?	***No.***
Other issues	
15. Are there any other issues that you would like the committee to consider?	
Topic-specific questions	
16. [To be added by technical team if required, after receiving the company submission. For	

example, if the company has deviated from the scope (particularly with respect to comparators) – check whether this is appropriate. Ask specific, targeted questions such as "Is comparator X [excluded from company submission] considered to be established clinical practice in the NHS for treating [condition Y]?"] **if not delete highlighted rows and renumber below**	

Key messages

17. In up to 5 bullet points, please summarise the key messages of your statement:

- Successful treatment
- Positive attitude
- Lack of post chemo sickness

- Bowel issues
- Boredom

Thank you for your time.

Please log in to your NICE Docs account to upload your completed statement, declaration of interest form and consent form.

……

Your privacy

The information that you provide on this form will be used to contact you about the topic above.

☐ **Please tick this box** if you would like to receive information about other NICE topics.

For more information about how we process your personal data please see our privacy notice.

……

CONFIDENTIAL UNTIL PUBLISHED

Evidence Review Group's Report
Liposomal cytarabine and daunorubicin for untreated acute myeloid leukaemia

Produced by	CRD and CHE Technology Assessment Group, University of York, Heslington, York YO10 5DD
Authors	Lindsay Claxton, Research Fellow in Health Economics, CRD Alexis Llewellyn, Research Fellow, CRD Mousumi Biswas, Research Fellow in Health Economics, CRD Robert Hodgson, Lead Health Economist, CRD Lesley Stewart, Director, CRD Melissa Harden, Information Specialist, CRD Ros Wade, Research Fellow, CRD
Correspondence to	Ros Wade, Research Fellow, CRD, University of York, York YO10 5DD
Date completed	03/07/2018

Source of funding

This report was commissioned by the NIHR HTA Programme as project number 17/109/08.

Declared competing interests of the authors

None.

Acknowledgements

We would like to thank Dr Gail Jones, Consultant Haematologist at the Newcastle upon Tyne Hospitals NHS Foundation Trust 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:
Claxton L, Llewellyn A, Biswas M, Hodgson R, Stewart L, Harden M, Wade R. Liposomal cytarabine and daunorubicin for untreated acute myeloid leukaemia: A Single Technology Appraisal. CRD and CHE Technology Assessment Group, 2018.

Contributions of authors
Ros Wade and Alexis Llewellyn wrote the clinical effectiveness sections of the report. Lindsay Claxton and Mousumi Biswas wrote the cost-effectiveness sections and conducted the economic analyses. Robert Hodgson provided methodological and technical support for the cost-effectiveness sections. Melissa Harden wrote the sections on the search strategies. Lesley Stewart provided advice, commented on drafts of the report and took overall responsibility for the clinical and 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

AE	Adverse event
AHD-AML	Antecedent haematologic disorders
AIC	Akaike information criterion
AICC	Corrected Akaike information criterion
AML	Acute myeloid leukaemia
AML-MRC	AML with myelodysplasia-related changes
BCSH	British Committee for Standards in Haematology
BIC	Bayesian information criterion
BNF	British National Formulary
BSA	Body surface area
BSC	Best supportive care
CEA	Cost-effectiveness analysis
CEAC	Cost-effectiveness acceptability curve
CHMP	Committee for Medicinal Products for Human Use
CI	Confidence interval
CMMoL	Chronic myelomonocytic leukaemia
$_{CMMoL}$AML	AML with antecedent chronic myelomonocytic leukaemia
CR	Complete remission
CRi	Complete remission with incomplete platelet or neutrophil recovery
CCR	Conventional care regimen
CS	Company submission
CSR	Clinical study report
DA	Daunorubicin and cytarabine
DSA	Deterministic sensitivity analysis
EFS	Event-free survival
ELN	European LeukaemiaNet
EMA	European Medicines Agency
eMIT	Electronic market information tool
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
GCSF	Granulocyte colony stimulating factor
GVHD	Graft versus host disease
HMA	Hypomethylating agents

HMRN	Haematological Malignancies Research Network
HR	Hazard ratio
HRG	Healthcare resource group
HRQL	Health-related quality of life
HSCT	Haematopoietic stem cell transplantation
ICER	Incremental cost-effectiveness ratio
ITT	Intention-to-treat
IV	Intravenous
KM	Kaplan Meier
LDAC	Low-dose cytarabine
LE	Life expectancy
MDS	Myelodysplastic syndrome
$_{MDS}$AML	AML with antecedent myelodysplastic syndrome
mg	milligram
MPN	Myeloproliferative neoplasm
MRD	Minimal residual disease
NCCN	National Comprehensive Cancer Network
NHS	National Health Service
NICE	National Institute for Health and Care Excellence
OR	Odds ratio
OS	Overall survival
PAS	Patient Access Scheme
PRISMA	Preferred Reporting Items for Systematic Review and Meta-analyses
PSA	Probabilistic sensitivity analysis
PSS	Personal Social Services
QALY	Quality-adjusted life year
RCT	Randomised controlled trial
SAE	Serious adverse event
$_{S}$AML	Secondary AML
SD	Standard deviation
SE	Standard error
SLR	Systematic literature review
SMR	Standardised mortality ratio
SOC	Standard of care
STA	Single Technology Appraisal
t-AML	Therapy-related AML

TTO	Time trade-off
WHO	World Health Organisation
WTP	Willingness to pay

1 Summary

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

The patient population addressed in the company submission (CS) is people with newly diagnosed, high-risk (secondary) AML who are considered to be eligible for intensive therapy. High-risk (secondary) AML is defined by therapy-related AML (t-AML) and AML with myelodysplasia related changes (AML-MRC). Diagnosis of AML-MRC requires genotyping, with results usually available within 7-10 days, although it may take longer for FLT3, NPM1 and CEBPA molecular tests. In practice clinicians may begin the first cycle of treatment prior to receiving genetic test results, then review treatment after receiving the results.

Liposomal daunorubicin and cytarabine (CPX-351) is awaiting European Medicines Agency (EMA) marketing authorisation. The anticipated marketing authorisation for CPX-351 is for the treatment of adults with newly diagnosed therapy-related AML (t-AML) or AML with myelodysplasia-related changes (AML-MRC). On 28th June 2018 the Committee for Medicinal Products for Human Use (CHMP) adopted a positive opinion, recommending the granting of a marketing authorisation for CPX-351.

The comparator addressed in the CS is standard intensive chemotherapy (daunorubicin and cytarabine [DA]) for patients who are considered to be fit for intensive chemotherapy, which is narrower than that specified in the NICE scope. The company's rationale for excluding three other comparators specified in the NICE scope (azacitidine, midostaurin and gemtuzumab ozogamicin) appears appropriate. However, in clinical practice FLAG-Ida may be considered as an alternative to DA, particularly in younger patients with high-risk AML; this was not considered in the CS.

The outcome measures specified in the NICE scope were reported in the CS; although health-related quality of life (HRQL) outcomes were not presented in the CS as this was not assessed in the trial on which the CS focusses (Study 301).

1.2 Summary of clinical effectiveness evidence submitted by the company

The company conducted a systematic review to identify evidence on the efficacy, safety and tolerability of pharmacological interventions for the treatment of patients aged 60-75 years with untreated high-risk (secondary) AML. The systematic review excluded studies of patients younger than 60 or older than 75 in order to include only studies of patients aligned to the patient population in Study 301 (which is narrower than in the anticipated licence). Therefore, the systematic review eligibility criteria were designed to specifically select Study 301 and studies with a comparable patient population, rather than to undertake a more comprehensive systematic review of studies of

pharmacological interventions for high-risk AML in patients who may be eligible for CPX-351 in clinical practice.

Three randomised controlled trials (RCTs) met the eligibility criteria for the systematic review. However, two of the trials (Study 204 investigating CPX-351 and a study investigating two differing daunorubicin doses in the DA 3+7 regimen) were not described in the CS. No justification was given for excluding them. However, the ERG considers that it was reasonable for the CS to focus on Study 301.

Study 301 suggests that compared with DA 3+7, CPX-351 is associated with a significant improvement in overall survival (OS) (median OS: 9.56 months [95% CI: 6.60, 11.86] vs 5.95 months [95% CI: 4.99, 7.75]; hazard ratio (HR)=0.69 [95% CI: 0.52, 0.90], p=0.005) in patients with high-risk AML. Although results from the subgroup analyses should be interpreted with caution, there was some evidence to suggest that CPX-351 had a less beneficial impact on OS in the subgroup of patients with myelodysplastic syndrome (MDS) who had received prior treatment with hypomethylating agents (HMA)

[redacted]These patients constituted around a third of patients in the trial and a similar proportion of those who would be eligible for CPX-351 in clinical practice.

Overall CPX-351 may be a more effective bridge to haematopoietic stem cell transplant (HSCT) compared with 3+7 in patients with high-risk AML aged 60-75 years. Compared with 3+7, the proportion of patients undergoing HSCT was higher in the CPX-351 group (34.0% [52/153] vs 25.0% [39/156]) although the difference was not statistically significant (OR=1.54 [95% CI: 0.92, 2.56]). OS in patients who underwent HSCT was significantly greater with CPX-351: at the point of data cut the median OS was not reached in the CPX-351 group, and the median OS in the 3+7 group was 10.25 months (95% CI: 6.21, 16.69) (HR=0.46 [95% CI: 0.24, 0.89], p=0.0046). A substantial number of patients were censored in the CPX-351 arm and there were small numbers of patients in the tail of the survival curves. The company clarified that a number of deaths were known to have occurred since the 2015 data cut used for the primary analyses ([redacted] patients in the CPX-351 arm, and [redacted] patients in the 3+7 arm).

Overall the safety profiles of CPX-351 and 3+7 appear broadly comparable. The overall incidence of observed Grade 3-5 adverse events (AEs) was similar across groups. Although potentially concerning, the higher incidence in observed treatment-related serious AEs in CPX-351-treated patients may be largely explained by the higher number of cycles and longer duration of treatment in the CPX-351 arm compared with 3+7.

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

Study 301 was a phase 3 multi-centre trial that randomised 309 patients with high-risk (secondary) AML and used appropriate endpoints. The trial used 1:1 randomisation stratified by age and AML subtype. However, the ERG has some concerns about the validity of the clinical evidence. There was limited information on the selection and characteristics of patients not receiving HSCT and decisions on whether to transplant were made with the knowledge of which treatment a patient had received, meaning that the risk of patient selection bias at point of transplant cannot be excluded. The ERG considers the median follow up of 20.5 months in the CPX-351 group and 21.2 months in the 3+7 group to be inadequate for estimating long-term OS. The analyses used a December 2015 data cut which includes substantial censoring. The company were unable to provide the ERG with OS analyses based on more recent data. OS results post-HSCT should be interpreted with caution given the small number of patients, limited follow-up duration, extensive censoring and lack of randomisation at the point of transplant. Data on relapse post-HSCT was very limited and may not be reliable. Study 301 did not collect HRQL or utility data.

Study 301 included high-risk AML patients aged 60-75 years, which is a subpopulation of the patient population described in the NICE scope, in which no age restriction was applied, and the anticipated marketing authorisation. The clinical advisor to the ERG suggested that around 20-25% of high-risk AML patients seen in clinical practice are below the age of 60 years and that patients older than 75 years would be less likely to withstand intensive chemotherapy. Therefore, whilst the trial population is likely to be reflective of the majority of patients eligible for intensive chemotherapy for high-risk AML in clinical practice, the results of the trial may not be generalisable to patients under the age of 60 (who would be eligible for CPX-351 in practice).

The clinical advisor to the ERG advised that patients with *de novo* AML with MDS associated karyotypic changes (25% of patients in the trial) are difficult to confidently define until genetic test results are available.

[redacted] In this respect Study 301 is not reflective of clinical practice, where treatment may commence prior to cytogenetic test results becoming available.

The comparator used in the trial was DA using a 3+7 regimen, delivered as a continuous IV infusion. Standard induction therapy in the UK is DA using a 3+10 regimen, delivered as a twice daily IV push every 12 hours. The British Committee for Standards in Haematology (BCSH) guidelines for the management of AML state that there is no evidence that a 3+10 regimen is superior to a 3+7 regimen and the clinical advisor to the ERG agreed that it is reasonable to argue equivalence between the two

regimens. Therefore, the ERG considers that the comparator used in the trial and addressed in the CS is appropriate for older patients with high-risk AML. However, the ERG's clinical advisor stated that in clinical practice FLAG-Ida may be considered as an alternative to DA, particularly in younger patients with high-risk AML; the CS considered only DA as a comparator to CPX-351 and did not describe the use of FLAG-Ida in younger patients.

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 CPX-351 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 CPX-351.

The cost-effectiveness of CPX-351 compared with 3+7 was informed by an economic evaluation conducted by the company. The population included in the company's decision problem and economic model comprised adult patients with untreated, high-risk (secondary) AML. The company's model used a hybrid modelling approach. The initial part of the model was a decision tree, and determined whether patients achieved remission after induction therapy, and whether those achieving remission after induction received stem cell transplant ("post-transplant remission") or did not receive transplant ("post-consolidation remission"). Patients who did not achieve remission could also receive transplant. Following the decision tree, each group of patients transitioned through health states modelled with a partitioned survival approach. The model used the time-to-event data from Study 301 to determine the distribution of patients between the health states. The modelled health states included (i) newly diagnosed disease; (ii) remission (defined as CR or CRi, comprising post-consolidation and post-transplant remission); (iii) disease progression (comprising relapse after remission and progression in treatment non-responders); and (iv) death. The efficacy data, proportions of patients receiving induction and consolidation therapy, proportions of patients receiving transplant, adverse event rates and patient characteristics (age, AML types) used in the economic model were sourced from Study 301, with the remaining inputs informed by studies identified in the cost-effectiveness review and other sources.

The company found CPX-351 to be more costly (cost difference of [redacted]), but also more effective (gains of [redacted] QALYs). The estimated deterministic ICER for CPX-351 compared with 3+7 was [redacted] per QALY. The results for the base-case after applying a PAS, lowered the total costs for CPX-351 by [redacted], resulting in an ICER of £46,631 for CPX-351 versus 3+7. Without the PAS applied, the predicted probability that CPX-351 was cost-effective compared with 3+7 was [redacted] at a threshold of £50,000 per QALY, while the probability was [redacted] at both a threshold of £20,000 and

£30,000 per QALY. The company reported that the most influential parameters in the one-way sensitivity analysis included the post-transplant OS for CPX-351. In a series of scenario analyses, the results were notably most sensitive to variations in the time horizon, highlighting that the majority of the high-cost events occurred at the start of the model (induction and consolidation, hospitalisation and transplantation), while the benefits were accrued in the long-term. Across the scenarios explored, the ICER varied between a [redacted] decrease from the base-case ICER (all patients treated with CPX-351 receive second induction and consolidation therapy as outpatients, with an ICER of [redacted]), to an increase of [redacted] (adjustment of general population mortality for HSCT patients, with an ICER of [redacted]).

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

The ERG considers that the state-transition modelling approach taken by the company introduced unnecessary complexity and required that the limited clinical data available be subdivided, such that parameter estimates were often based on small numbers of patients. A simpler partitioned survival analysis approach may have been preferable, particularly given the minimal gains in flexibility or accuracy offered by the approach taken by the company.

The ERG considers that there is significant uncertainty surrounding the long-term survival of patients after transplant. The immaturity of the survival data analysed in Study 301 means that the life expectancy post-transplant predicted by the model was sensitive to the choice of OS survival models. The approach taken by the company also resulted in an unreliable estimation of EFS. Together, these result in significant uncertainty in the cost-effectiveness estimates for CPX-351.

The plateau in the CPX-351 post-transplant OS curve suggests that around [redacted] of patients are cured, and experience mortality rates similar to that of the general population. This is significantly higher than for 3+7 patients and what is suggested by external sources of data for other types less severe forms of AML. The OS benefit for CPX-351 is inconsistent with the post-transplant EFS survival analysis, which showed no benefit between treatment arms and predicted that the majority [redacted] of patients would experience an event within two years. The consequences of the post-transplant EFS in the model implies that patients alive at two years are in the relapsed health state. Given the poor prognosis of patients who relapse after transplant, it was thought to be unlikely that they would experience a mortality rate similar to that of the general population.

The ERG identified uncertainty in some of the parameter values used in the model. The quality of life for patients in post-transplant remission was considered to be too high, as it was greater than that in the general population. The utility values for the on-treatment health states were based on vignettes that included a description of side-effects for each treatment arm which suggested a higher quality of

life for CPX-351 patients that were not substantiated by clinical evidence. The ERG considered that the cost of transplant may have been over-estimated by the company, as it included the cost of provision of an unrelated stem cell donor; the ERG was unclear on whether this cost would be relevant to UK practice. The ERG was concerned that the length of hospitalisation was overestimated in the model, when compared with hospitalisation in Study 301. The CS assumed that during an outpatient consolidation course, patients would be hospitalised for [redacted] days in the CPX-351 arm and 30 days in the 3+7 arm, which is internally inconsistent with the health state vignettes.

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 sized phase 3 multi-centre RCT, which included CPX-351 at the recommended dose.

The company's economic submission met the majority of the requirements of the NICE reference case. The economic model accommodated a number of key clinical elements of the treatment of AML and incorporated a range of scenario analyses which allowed the impact of alternative assumptions to be explored. The company provided additional evidence and analyses in response to the ERG's points for clarification.

1.6.2 Weaknesses and areas of uncertainty

Study 301 included only patients aged 60-75 years. The anticipated marketing authorisation is for treatment of adults of any age. Although the population included in the trial is likely to be reflective of the majority of patients eligible for CPX-351 in clinical practice, the results may not be generalisable to patients under the age of 60. The ERG received clinical advice that approximately 20-25% of patients in this population are under the age of 60. The company provided a scenario analysis that extrapolated the model to these younger patients, but there was remaining uncertainty of the impact of treatment and survival and consequently the cost-effectiveness in the broader age group.

Patients with *de novo* AML with MDS associated karyotypic changes (25% of patients in the trial) are difficult to confidently define until genetic test results are available. In clinical practice treatment may commence prior to cytogenetic test results becoming available. The ERG also noted some differences in the magnitude of the treatment effect across the different subtypes of high-risk AML patients, which could not be explored in the model due to limitations in the company's analysis of trial data.

The length of follow-up in Study 301 was insufficient for measuring longer term post-transplant OS. In addition, there is lack of clarity on whether data on relapse after transplant were collected systematically in the trial.

The model structure implemented by the company analysed the survival of three sub-cohorts (non-responder patients, responder patients who received transplant, and responder patients who did not receive transplant). Each analysis was based on small patient numbers from Study 301. The survival analysis of non-responders included transplant patients as well as non-transplant patients, and the model did not capture relapses after transplant, and may overestimate QALYs in this group of patients.

There are significant areas of uncertainty in the cost-effectiveness analysis, relating to overall and event-free survival after transplant. The ERG raised concerns that the trial data used to predict long-term survival was not sufficiently mature to support the sustained benefit for CPX-351 in OS beyond the trial period. The EFS analysis in transplant patients was based on very small patient numbers as many patients [redacted]) were censored for event-free survival (because they received a transplant after their last examination in the study). The ERG also questioned whether the long-term survival projections for CPX-351 in the company's base case are clinically plausible, as they are substantially higher than 3+7 and survival observed in other studies of AML patients receiving transplant. Therefore, the results may not reliably reflect the true treatment effect.

Study 301 did not report quality of life outcomes. Instead the company conducted a time trade-off (TTO) study to estimate quality of life. The utility values generated from the TTO analysis were subject to a significant degree of uncertainty suggesting that these may not be fully generalizable to the decision problem population.

1.7 Summary of exploratory and sensitivity analyses undertaken by the ERG

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 impact in patients with different subtypes of high-risk AML, due to constraints relating to the available data and the model structure. The scenarios associated with the greatest impact on cost-effectiveness outcomes related to the alternative survival curves for post-transplant OS in CPX-351, post-transplant EFS, and inclusion of 30% patients under the age of 60. The ERG also presented an alternative base-case based on a combination of a number of these scenario analyses. The ERG alternative base-case analysis includes the following changes to the company base-case analysis:

- Post-transplant outcomes based on OS only,
- Post-transplant OS based on survival analysis weighted by goodness-of-fit (based on AICC weights),
- Long-term mortality in post-transplant remission patients adjusted for excess mortality,

- Utility estimate for patients in the post-transplant remission health state, further adjusted for age,
- Equivalent quality of life for CPX-351 and 3+7 patients while on induction and consolidation treatment,
- Vial usage reflecting the distribution of body surface area, and the mean body surface area re-weighted to reflect the gender distribution in Study 301,
- Reduced number of hospital days during the consolidation period,
- Provision of unrelated donor stem cells excluded from the costs of transplant.

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.

Table 1 Summary of ERG exploratory analyses

Scenarios	Treatments	Total costs	Total QALYs	Incremental costs	Incremental QALYs	ICER (incremental £/QALY)	Change in ICER (%)
Company base-case (including ERG corrections)	3+7	[redacted]	[redacted]	-	-	-	-
	CPX-351	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
10% of patients under 60 years	3+7	[redacted]	[redacted]	-	-	[redacted]	[redacted]
	CPX-351	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
30% of patients under 60 years	3+7	[redacted]	[redacted]	-	-	[redacted]	[redacted]
	CPX-351	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
Post-transplant OS in CPX-351 arm: Weibull	3+7	[redacted]	[redacted]	-	-	[redacted]	[redacted]
	CPX-351	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
Post-transplant OS in CPX-351 arm: Log-logistic	3+7	[redacted]	[redacted]	-	-	[redacted]	[redacted]
	CPX-351	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
Post-transplant OS in CPX-351 arm: Log-normal	3+7	[redacted]	[redacted]	-	-	[redacted]	[redacted]
	CPX-351	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
Post-transplant OS in CPX-351 arm: Exponential	3+7	[redacted]	[redacted]	-	-	[redacted]	[redacted]
	CPX-351	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
Post-transplant OS in CPX-351 arm: Generalised gamma	3+7	[redacted]	[redacted]	-	-	[redacted]	
	CPX-351	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
	3+7	[redacted]	[redacted]	-	-	-	

Scenarios	Treatments	Total costs	Total QALYs	Incremental costs	Incremental QALYs	ICER (incremental £/QALY)	Change in ICER (%)
Post-transplant OS in CPX-351 arm: Combining treatments (Gompertz)	CPX-351	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
Weighted OS curve using AICC weights	3+7	[redacted]	[redacted]	-	-	[redacted]	[redacted]
	CPX-351	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
Weighted OS curve using BIC weights	3+7	[redacted]	[redacted]	-	-	[redacted]	[redacted]
	CPX-351	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
Post-transplant outcomes are based on OS only	3+7	[redacted]	[redacted]	-	-	[redacted]	[redacted]
	CPX-351	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
Utility value of 0.75 for durable remission	3+7	[redacted]	[redacted]	-	-	[redacted]	[redacted]
	CPX-351	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
General population utility value (0.79) for durable remission	3+7	[redacted]	[redacted]	-	-	[redacted]	[redacted]
	CPX-351	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
Post-transplant remission utility, adjusted for aging	3+7	[redacted]	[redacted]	-	-	[redacted]	[redacted]
	CPX-351	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
Mean on-treatment utility during induction and consolidation treatment	3+7	[redacted]	[redacted]	-	-	[redacted]	[redacted]
	CPX-351	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
Vial usage reflecting the distribution of body surface area	3+7	[redacted]	[redacted]	-	-	[redacted]	[redacted]
	CPX-351	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
Reduced number of hospital days during consolidation period	3+7	[redacted]	[redacted]	-	-	[redacted]	[redacted]
	CPX-351	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
Alternative cost of transplant	3+7	[redacted]	[redacted]	-	-	[redacted]	[redacted]
	CPX-351	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
ERG alternative	3+7	[redacted]	[redacted]	-	-	[redacted]	[redacted]
	CPX-351	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]

Scenarios	Treatments	Total costs	Total QALYs	Incremental costs	Incremental QALYs	ICER (incremental £/QALY)	Change in ICER (%)
base-case analysis							
AICC, Akaike information criterion corrected; BIC, Bayesian information criterion; ERG, Evidence Review Group; ICER, incremental cost-effectiveness ratio; QALYs, quality-adjusted life years							

The ERG base-case analysis estimated CPX-351 to be more costly (cost difference [redacted]) and more effective ([redacted] QALY gain) compared with 3+7, and suggests that the ICER for CPX-351 compared with 3+7 is [redacted] per QALY.

The ERG also carried out a further series of exploratory subgroup analyses to explore the impact of alternative post-transplant OS and inclusion of patients under the age of 60. The ICER of CPX-351 vs. 3+7 varied between [redacted] (most favourable post-transplant OS for CPX-351) and [redacted] (least favourable post-transplant OS for CPX-351). The ICER decreased to [redacted] when 10% of patients were under the age of 60 and to [redacted] when 30% of patients were under the age of 60.

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.

Acute myeloid leukaemia (AML) is a rare, heterogeneous haematological malignancy characterised by the overproduction of immature myeloid cells, known as blasts. The result is reduced numbers of all three major, formed elements of blood (red blood cells, white blood cells and platelets) leading to an increased risk of bleeding, infection and mortality.[1, 2] The incidence of AML in the UK is around 3100 cases per year.[3]

High-risk AML is particularly aggressive and accounts for around 25% of all AML diagnoses.[4, 5] High-risk AML includes therapy-related AML (t-AML) and AML with myelodysplasia-related changes (AML-MRC). These diagnostic groups are often collectively referred to as 'secondary AML', as they have arisen from either prior exposure to chemotherapy or radiation therapy (t-AML) or previous clonal disorder of haematopoiesis (an antecedent haematologic disorder; AHD-AML). The terms 'high-risk AML' and 'secondary AML' are used interchangeably in the CS. High risk/secondary AML typically presents in older AML patients and is more resistant to chemotherapy treatment than *de novo* AML.

High-risk AML has a very poor prognosis, regardless of age at diagnosis. Data from a large Swedish registry reported median survival for patients with t-AML as 14 months in patients aged less than 55 years, 9 months for patients aged 55-74 and 8 months for patients aged 75 or above. Median survival for patients with AHD-AML was 7 months in the two younger age groups and 6 months for patients aged 75 or above.[4]

2.2 Critique of company's overview of current service provision

The company's overview of current service provision is generally appropriate and relevant to the decision problem under consideration.

Initial assessment evaluates whether a patient is a candidate for intensive induction chemotherapy or not, based on patient and disease-related factors and patient choice. For high-risk AML patients who are considered fit for intensive chemotherapy, the standard first induction regimen in the UK is daunorubicin (60 mg/m^2) plus cytarabine (100 mg/m^2) (DA) delivered over a 3 day and 10 day course, respectively (3+10). This is followed by the same drugs on a 3 day and 8 day schedule (3+8) for induction cycle two. These doses are delivered as a twice-daily intravenous (IV) push every 12 hours. In clinical practice FLAG-Ida may be considered as an alternative to DA, particularly in younger

patients with high-risk AML; the CS only considered DA as a comparator to CPX-351 and did not describe the use of FLAG-Ida in younger patients.

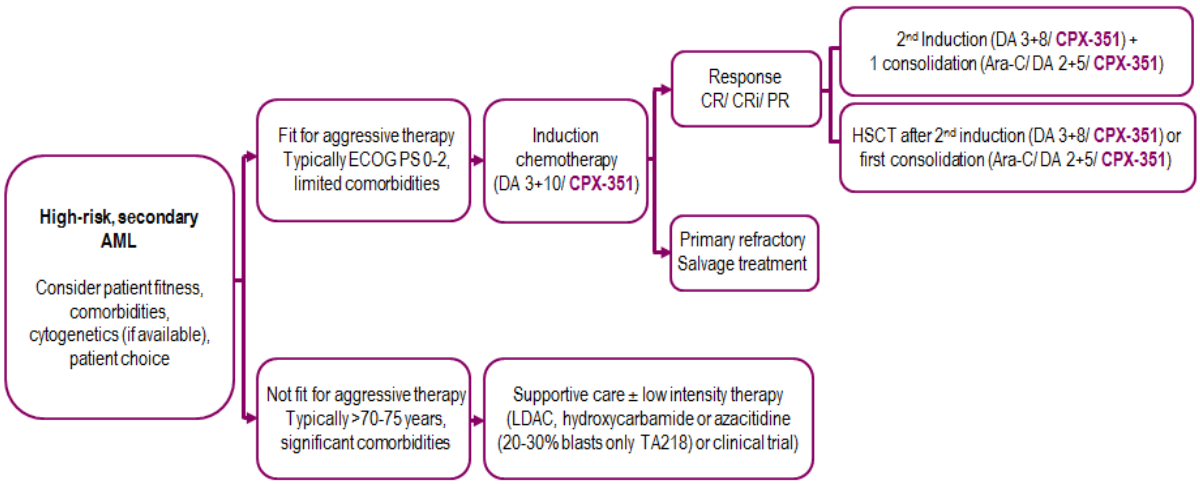
Patients who respond to induction therapy may receive a post-remission therapy to help prevent relapse. This can include consolidation chemotherapy with or without allogeneic haematopoietic stem cell transplantation (HSCT).[6] Standard consolidation chemotherapies include 2+5 (daunorubicin 50 mg/m^2 on days 1 and 3 and cytarabine 100 mg/m^2 12-hourly by IV push on days 1-5 inclusive) or intermediate/high dose cytarabine (1 to 3 g/m^2 daily for 5 days). The CS states that in total patients usually receive a maximum of two courses of induction therapy and one course of consolidation therapy. However, the clinical advisor to the ERG suggested that for younger patients, who are less likely to have comorbidities and are more able to tolerate intensive chemotherapy (those aged up to around 60 years), the standard of care is two courses of induction therapy and two courses of consolidation therapy (high dose cytarabine). In addition, those who have FLAG-Ida induction therapy would also receive FLAG-Ida as consolidation therapy. Patients with high-risk AML have a high risk of relapse, due to underlying resistance to chemotherapy, therefore, life expectancy remains poor in patients who do not receive HSCT.

High-risk AML patients who are not considered fit for intensive chemotherapy, or who do not wish to receive intensive chemotherapy, may be given dose reductions or non-intensive chemotherapy. This includes low dose cytarabine (LDAC) or, for patients with 20-30% bone marrow blasts and multi-lineage dysplasia, the hypomethylating agent (HMA) azacitidine.[7, 8] Other options for patients not fit for intensive chemotherapy include best supportive care, hydroxycarbamide to help control the white blood cell count, or clinical studies with investigational drugs.[7]

The CS states that CPX-351 is a novel, advanced dual-drug liposomal formulation designed to deliver daunorubicin and cytarabine in a synergistic, fixed 1:5 molar ratio to leukaemia cells for a prolonged period of time. CPX-351 is awaiting European Medicines Agency (EMA) marketing authorisation, the anticipated marketing authorisation for CPX-351 is for the treatment of adults with newly diagnosed therapy-related AML (t-AML) or AML with myelodysplasia-related changes (AML-MRC).

The current treatment pathway for high-risk AML patients, with the proposed positioning of CPX-351, is presented in Figure 1 (Figure 2 of the CS). This treatment pathway appears generally appropriate for older patients with high-risk AML. However in clinical practice FLAG-Ida may be considered as an alternative to DA, particularly in younger patients; this was not described as part of the current treatment pathway presented in the CS.

Figure 1 Current treatment pathway for high-risk AML patients with proposed positioning of CPX-351 (CS Figure 2, p25)



3+10, 3+8 or 2+5 refers to the number of days of drug delivery for daunorubicin and cytarabine, respectively; Ara-C, cytarabine; CR, complete remission; CRi, complete remission with incomplete blood count recovery; PR, partial response; ECOG PS, Eastern Cooperative Oncology Group performance status; LDAC, low dose cytarabine

3 Critique of company's definition of decision problem

3.1 Population

The patient population addressed in the CS is people with newly diagnosed, high-risk (secondary) AML who are considered to be eligible for intensive therapy. High-risk (secondary) AML is defined by therapy-related AML (t-AML) and AML with myelodysplasia related changes (AML-MRC); t-AML and AML-MRC are diagnostic subgroups in the 2016 Worth Health Organisation (WHO) classification.[9]

CPX-351 is awaiting European Medicines Agency (EMA) marketing authorisation, the anticipated marketing authorisation for CPX-351 is for the treatment of adults with newly diagnosed therapy-related AML (t-AML) or AML with myelodysplasia-related changes (AML-MRC).

The clinical advisor to the ERG advised that the diagnosis of AML-MRC requires genotyping, which generally takes about 7-10 days to complete, although it may take longer for FLT3, NPM1 and CEBPA molecular tests. In view of the very short timeframe between diagnosis and treatment in patients with high-risk AML, particularly younger patients, clinicians may begin the first cycle of treatment prior to receiving genetic test results, then review treatment after receiving the results.

The clinical effectiveness evidence presented is from a single randomised controlled trial (RCT); Study 301. The trial included 309 high-risk AML patients aged 60-75 years, with the following high-risk AML subtypes:

- t-AML (20.4% patients)
- myelodysplastic syndrome (MDS) AML with prior treatment with HMA (34% patients)
- MDS AML without prior treatment with HMA (12.9% patients)
- *de novo* AML with MDS associated karyotypic changes (25.2% patients)
- AML with antecedent chronic myelomonocytic leukaemia ($_{CMMOL}$AML) (7.4% patients)

The clinical advisor to the ERG confirmed that the proportion of patients within the different AML subtypes in the trial is broadly similar to the proportion with each subtype seen in clinical practice. The clinical advisor also stated that it is difficult to define '*de novo* AML with MDS associated karyotypic changes' subtype (25% of trial participants) until genetic test results are available. Prior treatment and history supports a diagnosis of high-risk AML for the other subtypes. The clinical study report (CSR) for Study 301 states that

[redacted] This is not reflective of clinical practice, where treatment may commence prior to cytogenetic test results becoming available. The clinical advisor to the ERG

stated that in clinical practice it would be preferable to commence treatment with DA 3+10 for patients in this subgroup, while waiting for genetic test results, as this is the current standard of care for *de novo* AML that is not high-risk.

The ERG requested clarification regarding whether additional subtypes of high-risk AML may be eligible for CPX-351 in clinical practice, but were not included in Study 301. The company clarified that

[redacted]

The clinical advisor to the ERG confirmed that this appears appropriate as standard WHO criteria were used.

Study 301 only included patients with high-risk AML who were aged 60-75 years. This is a sub-population of the patient population described in the NICE scope, in which no age restriction was applied, and the anticipated marketing authorisation, which only specifies 'adults'. The clinical advisor to the ERG stated that whilst incidence of high-risk AML goes up with age, older patients would be less likely to withstand intensive chemotherapy treatment, so the population of the trial is likely to be reflective of the majority of patients eligible for intensive chemotherapy for high-risk AML in clinical practice. Around 20-25% of high-risk AML patients seen in clinical practice are below the age of 60; the results of the trial may not be generalisable to this group of younger patients.

3.2 Intervention

The intervention specified in the NICE scope is liposomal daunorubicin and cytarabine. CPX-351 is awaiting EMA marketing authorisation. On 28th June 2018 the Committee for Medicinal Products for Human Use (CHMP) adopted a positive opinion, recommending the granting of a marketing authorisation for CPX-351.

CPX-351 is administered as an intravenous infusion. The recommended dose is daunorubicin 44 mg/m^2 and cytarabine 100 mg/m^2, infused over 90 minutes on days 1, 3 and 5 for the first course of induction therapy and on days 1 and 3 for subsequent courses of induction therapy, if needed. A subsequent course of induction therapy may be administered in patients who do not show disease progression or unacceptable toxicity. The first cycle of consolidation therapy can be administered 5 to 8 weeks after the start of the last induction, at a recommended dose of daunorubicin 29 mg/m^2 and cytarabine 65 mg/m^2, infused over 90 minutes on days 1 and 3. A subsequent course of consolidation

therapy may be administered in patients who do not show disease progression or unacceptable toxicity. Study 301 used the recommended dose of CPX-351.

In Section B.2.12.3 of the CS, the company states that results from a recent UK KOL advisory board reported an increasing trend for CPX-351 consolidation therapy to be administered in an outpatient setting (Data on file). It also states that US physicians have reported that CPX-351 is currently being used in the outpatient setting (either for full or part of the treatment) beginning as early as the induction phase in some cases (Data on file). The ERG requested further clarification regarding the US data. The company responded that the data were from US post-launch quantitative market research by Naxion Research Consulting, performed in February to March 2018. The research indicated that [redacted] of patients received the Day 1 dose of first induction CPX-351 therapy in the hospital's outpatient facility, [redacted] for the Day 3 dose and [redacted] for the Day 5 dose of first induction therapy (n=34 patients). In second induction [redacted] patients received therapy in an outpatient setting and 100% CPX-351 patients received consolidation therapy in an outpatient setting (number of patients not reported). In view of the low patient numbers and lack of reporting of participant characteristics, it is unclear how representative these data are to patients who would be eligible for CPX-351 in UK clinical practice.

3.3 Comparators

The comparators specified in the NICE scope are: standard intensive induction and consolidation therapy; azacitidine (for people who are not eligible for HSCT and have AML with 20-30% blasts and multilineage dysplasia); midostaurin (for people with FLT3-mutation-positive AML) (subject to ongoing NICE appraisal); and gemtuzumab ozogamicin (subject to ongoing NICE appraisal). The comparator addressed in the CS is standard intensive chemotherapy (daunorubicin and cytarabine) for patients who are considered to be fit for intensive chemotherapy, which was the comparator used in Study 301.

The company's rationale for excluding three comparators specified in the NICE scope was that azacitidine is not recommended for patients eligible for HSCT and is typically used in older, unfit AML patients as a palliative therapy and that both midostaurin and gemtuzumab ozogamicin are indicated as add-on therapies to standard intensive chemotherapy. The ERG's clinical advisor confirmed that patients eligible for azacitidine are different to those who would be eligible for CPX-351. In addition to midostaurin and gemtuzumab ozogamicin only being indicated as add-on therapies, gemtuzumab ozogamicin is also only indicated for patients with *de novo* AML. Therefore, the exclusion of midostaurin and gemtuzumab ozogamicin also appears to be appropriate.

As described in Section 2.2, standard first induction therapy in the UK is DA using a 3+10 regimen, delivered as a twice daily IV push every 12 hours. Study 301 used a 3+7 regimen delivered as a

continuous IV infusion. The CS states that there is no published head-to-head study of the UK 3+10 schedule versus US 3+7 schedule and that UK clinical advisory board feedback supports equivalence of the two schedules. The British Committee for Standards in Haematology (BCSH) guidelines on the management of AML in adults recommend either a 3+10 or 3+7 regimen and state that there is no evidence that a 3+10 regimen is superior to a 3+7 regimen.[6] The European LeukaemiaNet (ELN) guidelines for adult patients with AML recommend a 3+7 regimen.[7, 10] The ERG's clinical advisor considered it reasonable to argue equivalence between a 3+7 regimen and a 3+10 regimen. Therefore, the ERG considers that the comparator used in the trial and addressed in the CS is appropriate for older patients with high-risk AML. However, in clinical practice FLAG-Ida may be considered as an alternative to DA, particularly in younger patients with high-risk AML; this was not considered in the CS.

3.4 Outcomes

The outcome measures specified in the NICE scope were reported in the CS; overall survival (OS), event free survival (EFS), remission, health-related quality of life (HRQL) and adverse effects of treatment. The CS also included response rate, rate of achieving morphologic leukaemia-free state (MLFS) and the proportion and overall survival of patients receiving HSCT. The primary outcome of Study 301 was OS. Secondary outcomes were response rate, EFS, remission duration, rate of MLFS, rate of transfer to HSCT and time to recovery from neutropenia and thrombocytopenia. Study 301 did not collect HRQL or utility data.

3.5 Other relevant factors

The CS states that the use of CPX-351 is unlikely to raise any equality issues.

[redacted]

3.6 Summary

The clinical effectiveness evidence was based on a single trial, with a population of high-risk AML patients aged 60-75 years. This is a subpopulation of the unrestricted age group described in the NICE scope and the anticipated marketing authorisation. Around 20-25% of high-risk AML patients seen in clinical practice are younger than 60 years and patients older than 75 years would be less likely to withstand intensive chemotherapy, therefore the trial population is likely to be reflective of the majority of patients eligible for intensive chemotherapy for high-risk AML in UK clinical practice. However, the results of the trial may not be generalisable to patients with high-risk AML under the age of 60, who would be eligible for CPX-351 in practice.

The clinical advisor to the ERG advised that it is difficult to confidently define patients with *de novo* AML with MDS associated karyotypic changes (25% of patients in the trial) until genetic test results

are available.

[redacted] This differs from clinical practice, where treatment may commence prior to cytogenetic test results becoming available.

Three of the comparators specified in the NICE scope were not addressed in the CS; azacitidine, midostaurin and gemtuzumab ozogamicin. However, the company's rationale for excluding these comparators appears appropriate.

The comparator used in the trial was DA using a 3+7 regimen. Although standard induction therapy in the UK is DA using a 3+10 regimen, the ERG considers this reasonable. However, in clinical practice FLAG-Ida may be considered as an alternative to DA, particularly in younger patients with high-risk AML; this was not considered in the CS.

CPX-351 is administered as a 90 minute infusion on alternate days, which requires less patient contact time than the twice daily administration of DA for 10 days during induction cycle one, 8 days during induction cycle two and 5 days during consolidation therapy. The company states that data on file supports the proposal that CPX-351 could be administered in an outpatient setting. However, it is unclear how representative these data are to patients who would be eligible for CPX-351 in UK clinical practice.

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 CPX-351 for treating high-risk AML are presented at the end of this section.

4.1 Critique of the review methods

The company conducted a systematic review of RCTs assessing the efficacy, safety and tolerability of pharmacological interventions for the treatment of patients aged 60-75 years with untreated high-risk (secondary) AML. Details of the systematic review methods are presented in Appendix D of the CS.

4.1.1 Search strategy

The searches undertaken by the company to identify relevant clinical data on the use of pharmacological interventions for the treatment of patients with AML are provided in Section D.1.1.5.

The electronic databases MEDLINE, MEDLINE In Process, EMBASE and the Cochrane Central Register of Controlled Trials (CENTRAL) were searched on 4th October 2017. The database searches were restricted to publications in English. Searches of the abstracts from the following conferences were undertaken for the years 2014-2017: American Society of Hematology (ASH) and the European Haematology Association (EHA).

Parts of the searching carried out by the company were appropriate, however some weaknesses with their approach were identified by the ERG and this may have affected the comprehensiveness of the searches undertaken.

Appropriate electronic databases were searched to identify relevant published literature. However in their search for unpublished literature the company did not search any trial registers to identify relevant reports of unpublished trials (ongoing and completed) of drug treatments for AML. It is therefore possible that unpublished trials, particularly of comparator studies, could have been missed by the searches presented in the company submission.

The search strategies were appropriately structured, comprising of a set of terms for AML combined (using the AND operator) with a set of terms for drugs used to treat AML. A study design search filter was applied to the search strategies in MEDLINE and EMBASE to limit retrieval to randomised controlled trials, however the source of the search filter is not reported. The specific drugs included in the strategy were: CPX-351, cytarabine, daunorubicin, idarubicin, gemtuzumab ozogamicin, mitoxantrone, azacitidine, cladribine, decitabine, fludarabine, etoposide, midostaurin and filgrastim. In the description of the searches on page 7 of Appendix D the inclusion of search terms for outcomes

is mentioned, however outcomes were not included in the database search strategies presented in Tables 1, 2 and 3, which is appropriate.

Several weaknesses were noted with the search strategies presented in Tables 1, 2 and 3. Firstly, truncation and wildcards were missing from the search terms for AML. Therefore records containing the UK spelling of leukaemia could have been missed as well as records where the plural forms of these words are used (leukaemias or leukemias). Searching for acute myeloid leuk?emia$ would have ensured that all records containing the UK as well as US spelling of leukaemia were retrieved, along with records that contain the plural versions of these words. Secondly, in the search strategy for MEDLINE lines 17, 18, 19 (searches for the drugs cytarabine, daunorubicin and idarubicin) are missing from line 36. Line 36 brings together all of the search lines relating to the drug terms. Therefore the MEDLINE strategy is likely to have missed records about trials of cytarabine, daunorubicin, or idarubicin for the treatment of AML. Thirdly, the brand names for some of the drugs is missing from the search strategies, in particular the brand name Vyxeos™ is missing. The omission of brand names may have increased the likelihood of missing studies.

A difference in the reporting of the number of records identified in the Preferred Reporting Items for Systematic Review and Meta-analyses (PRISMA) diagram on page 15 of the CS appendices and the records identified from the database searches in Tables 1, 2 and 3 was found by the ERG and raised in their points for clarification. The company responded that the figures in the PRISMA diagram were correct: 953 for MEDLINE, 1823 for EMBASE and 1076 for Cochrane. The figures in the search tables were incorrect due to the search tables not being updated on the day that the searches were run. This reporting error could have been avoided by copying and pasting the search strategies from each database at the time of running the search and presenting these strategies without editing in the report. This is recommended in CRD's guidance for undertaking reviews in health care and helps increase transparency of the searches.

4.1.2 Inclusion criteria

The inclusion criteria for the systematic review are presented in Table 4, Section D.1.1.6 of the CS appendices. Eligible studies were RCTs of a range of pharmacological interventions (see Table 4 for full details) for the treatment of patients aged 60-75 years with untreated high-risk (secondary) AML. Outcomes of interest are also specified and appear to be appropriate. Only studies published in English were included.

The systematic review excluded studies of patients aged younger than 60 years or older than 75 years. When asked by the ERG for a justification for restricting the systematic review inclusion criteria to patients aged 60-75 years (which is a narrower population than in the anticipated licence), the company responded that this population was chosen as it is aligned to the patient population in Study

301, the results of which formed the basis of the effectiveness evidence for CPX-351 in this STA submission. The company stated that including studies of patients outside this patient population would have introduced biased comparisons with CPX-351. The ERG considers this to be a limitation of the review, since the company state that the population addressed in the CS is people with newly diagnosed, high-risk (secondary) AML who are considered eligible for intensive therapy, rather than the narrower population that were included in Study 301. The systematic review eligibility criteria were designed to specifically select Study 301 and studies with a comparable patient population, rather than to undertake a more comprehensive systematic review of studies of pharmacological interventions for high-risk AML in patients who may be eligible for CPX-351 in clinical practice.

The methods used to screen and select relevant studies was to a good standard, with two researchers independently screening titles and abstracts and full texts using pre-defined eligibility criteria, with disagreements resolved by a third researcher. A complete list of studies excluded at the full paper stage is provided in Section D.1.1.11 of the CS appendices.

4.1.3 Data extraction

Data were extracted into Microsoft Excel® data extraction forms by a single researcher and verified by a second researcher, reducing the risk of error and bias in data extraction. Adequate data for Study 301 were presented in the CS with a detailed summary of trial methods presented in Table 6.

4.1.4 Quality assessment

Study 301 was assessed for quality using appropriate criteria for RCTs; the trial was reasonably good quality (see Section 4.2.2 for further details). The quality assessment results are presented in Table 10 of the CS, with further details presented in Table 6 of the CS appendices, which was checked by the ERG.

4.1.5 Evidence synthesis

Eight publications met the eligibility criteria for the review; seven of which described two clinical trials that examined the safety and efficacy of CPX-351: Study 204[11-13] and Study 301.[14-17] The remaining publication investigated two versions of the DA 3+7 regimen, with differing daunorubicin doses (45 mg/m^2 vs 90 mg/m^2).[18]

The CS only describes Study 301. It is not stated why the other two trials, that met eligibility criteria for the systematic review, were not presented in the CS. However, the ERG considers that it was appropriate to exclude the trial comparing different doses of daunorubicin, since there was no CPX-351 comparator arm in this trial.[18] Study 204 was a phase II RCT of CPX-351 versus DA 3+7 in newly diagnosed older patients (aged 60-75 years) with AML. The aim of the trial is stated as being to determine efficacy and identify patient subgroups that may benefit from CPX-351 treatment.[11] A pre-

planned subgroup analysis of patients with secondary AML (n=52) demonstrated an improved response rate and prolonged EFS and OS, which provided the rationale for Study 301. Therefore, in view of the relatively small subgroup of secondary AML patients in Study 204, the ERG considers that it was acceptable for the CS to focus on Study 301.

Section B.2.8 of the CS describes the company's assessment of the feasibility of conducting an indirect comparison to determine the relative treatment effect between CPX-351, DA 3+7 and DA 3+10 in AML patients. The company concluded that the identified evidence did not allow a robust comparison of OS with 3+7 versus 3+10 in this population. The ERG considers this reasonable - as discussed in Section 3.3, BCSH guidelines state that there is no evidence that a 3+10 regimen is superior to a 3+7 regimen.[6]

4.1.6 Conclusions from the critique of systematic review methods

The company conducted a systematic review of RCTs assessing the efficacy, safety and tolerability of pharmacological interventions for the treatment of patients aged 60-75 years with untreated high-risk (secondary) AML. There were some weaknesses in the search strategy used to identify relevant studies, which may have resulted in relevant studies being missed. Inclusion and exclusion criteria for the review are clearly stated. The systematic review only included studies of patients aged 60-75 years, which is a narrower population than in the anticipated marketing authorisation for CPX-351; therefore, results may not be generalisable to the broader high-risk AML population who may be eligible for CPX-351 in clinical practice. Data extraction was undertaken by one researcher 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 Study 301 were presented, along with a table of quality assessment results; the trial was reasonably good quality. Two studies that met eligibility criteria for the systematic review (Study 204 investigating CPX-351 and a study investigating two differing daunorubicin doses in the DA 3+7 regimen) were not described in the CS, with no justification for why there were excluded. However, the ERG considers that it was acceptable for the CS to focus on Study 301.

4.1.7 Ongoing studies

Two ongoing trials of CPX-351 as first-line treatment for AML patients were listed in Table 21 of the CS: AML18 and AML19. AML18 (which has been extended to include CPX-351) is designed to assess the effects of adding one of two new treatment agents to commonly used chemotherapy combinations in older patients with AML and high risk myelodysplastic syndrome. AML19 is designed to assess 'risk-adapted' therapy in younger patients with AML and high risk myelodysplastic syndrome. The AML18 and AML19 trials include multiple interventions, including DA, gemtuzumab ozogamicin, CPX-351, FLAG-Ida, Cladribine, AC220 and Ganetispib. The estimated primary completion date of these studies is October 2019 and January 2021, respectively.

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

4.2.1 Trials included in the systematic review

Three RCTs were included in the systematic review: Study 301 (CLTR0310-301 NCT01696084),[14-17] Study 204[11-13] and Löwenberg (2009).[18] Study 301 (phase III) and 204 (phase II) evaluated the safety and efficacy of CPX-351 compared to DA 3+7 in patients with untreated AML. Löwenberg (2009) investigated the impact of differing daunorubicin doses in the 3+7 regimen.

Study 301 is the primary focus of the clinical effectiveness evidence presented in the CS. Although the company did not provide justification for this, the ERG believes this decision is acceptable as Study 301 was the only study of CPX-351 exclusively in high risk (t-AML or AML-MRC) AML patients (as per anticipated EMA marketing authorisation).

4.2.1.1 Study 301

4.2.1.2 Study design

Study 301 is a randomised, parallel-arm, open-label, standard therapy-controlled, phase III trial. The study was conducted in 39 sites, including 35 in the US and 4 in Canada.

The trial compared the safety and efficacy of CPX-351 versus DA 3+7 intensive chemotherapy as first-line therapy in older patients (60 to 75 years) with high-risk (secondary) AML. The study was open-label due to differences between the colour and administration of treatment regimens which the company stated made blinding impossible.

Figure 2 presents the study design, which illustrates treatment and follow-up phases. Participants were eligible to receive up to two inductions and up to two consolidations with either CPX-351 or DA 3+7. Table 6 of the CS provides details on doses used for induction and consolidation therapy. The number of inductions and consolidations depended upon attainment of response (complete remission [CR] or complete remission with incomplete platelet or neutrophil recovery [CRi]), confirmed by bone marrow assessment. The follow-up phase was planned to start from 30 days after the completion of the last induction or consolidation course until death or 5 years post-randomisation. However the extent to which patients were followed-up is partly unclear (see Section 4.2.3). Post-remission therapy with stem cell transplant was allowed either instead of or after post-remission chemotherapy.

Figure 2 Study 301: Study design

In response to the ERG's request for clarification on the criteria used to assess eligibility for HSCT, the company stated:

[redacted]

The trial used 1:1 randomisation stratified by age (60-69 years, 70-75 years) and AML subtype (t-AML, $_{MDS}$AML with prior HMA, $_{MDS}$AML without prior HMA, $_{de\ novo}$AML with MDS karyotype, $_{CMMoL}$AML). Following the ERG's request for clarification the company confirmed that no crossover was permitted.

Median duration of follow-up was approximately 20 months for each arm of the study.

Endpoints

Study endpoints and definitions were provided in CS Table 6. The primary outcome was overall survival (OS), measured from the date of randomisation to death. Secondary efficacy endpoints included response (achieving CR or CRi), event-free survival (EFS), remission duration, morphologic leukaemia-free state (MLFS), transfer to haematopoietic stem cell transplant (HSCT) after induction treatment, and time to recovery from neutropenia and thrombocytopenia. Health-related quality of life was not recorded. Safety endpoints included adverse events, adverse drug reactions and serious adverse events.

The ERG considers the study endpoints to be appropriate, although the long-term post-HSCT OS presented was subject to significant censoring and may not be reliable. No long-term post-HSCT EFS were presented. Treatment response was assessed independently and blinded to treatment allocation, but assessment of all other outcomes (including the decision to transplant) was not blinded.

Analysis

The primary efficacy analyses were conducted on an intention-to-treat (ITT) basis. Stratified log-rank tests were used to compare the treatment arms, and the distribution of overall survival was estimated in each treatment arm using Kaplan-Meier methodology.

In addition to the ITT population, three populations were defined for analysis: the safety population, the morphologic leukaemia-free state (MLFS) population and the per protocol population. The safety population comprised all participants who received at least one dose of either CPX-351 or 3+7 (304 patients). Adverse events were recorded from the start of the infusion on Day 1 to the last day of the treatment period, with the exception of serious adverse events which were also collected up to 30 days after treatment completion. The per protocol population included all participants who met the inclusion/exclusion criteria, received at least one dose of study medication, and had AML diagnosis and type confirmed. The MLFS population included all participants from the per protocol population who had at least one bone marrow assessment performed on or after Day 14 of the last induction.

To account for the effect of transplant as a confounding factor, a sensitivity analysis that censored patients at the start of HSCT was performed on OS and EFS.

The CS describes a range of means of imputing missing data. However, following the ERG's request for clarification, the company clarified that there were no missing dates of death and no imputation was required or done.

The target sample size [redacted] With an assumed median OS of 6.3 months with standard therapy, the study had 93.7% power to detect a hazard ratio of 0.635 in OS. [redacted]

4.2.1.3 Trial population

A summary of the eligibility criteria for Study 301 is presented in CS Table 7. Patients aged 60 to 75 years with no prior treatment for AML with confirmation of one of the following high-risk AML subtypes (per World Health Organisation [WHO] 2008 criteria) were enrolled:

- therapy-related AML (t-AML) arising from prior cytotoxic or radiation therapy for an unrelated disease
- AML with antecedent myelodysplastic syndrome (MDS) ($_{MDS}$AML) including:
    - $_{MDS}$AML with prior treatment with hypomethylating agents (HMA) azacitidine or decitabine
    - $_{MDS}$AML without prior HMA
- *de novo* AML with MDS associated karyotypic changes
- AML with antecedent chronic myelomonocytic leukaemia (CMMoL) ($_{CMMoL}$AML)

The ERG asked the company whether any other subtypes of high-risk AML not included in Study 301 would be eligible for CPX-351 in clinical practice under the anticipated marketing authorisation. [redacted] This appears to be appropriate and consistent with standard WHO criteria.[9]

All patients treated in Study 301 were fit for intensive chemotherapy. Following the ERG's request for clarification, the company stated [redacted] The decision to transplant was not blinded to treatment allocation.

The CPX-351 arm included 153 participants, and the 3+7 arm 156 patients. Baseline characteristics of participants are presented below in Table 2 and do not show any significant imbalance between treatment groups.

All included participants had high-risk AML, with generally poor prognosis. The most common subtype of AML was MDS with prior HMA (approximately one third of the trial population), followed by *de novo* AML with MDS associated karyotypic changes (25%), and therapy related-AML

(20%). Patients with *de novo* AML had to have in situ hybridization (FISH) or cytogenetic changes linked to MDS as per WHO criteria.[9] The clinical advisor to the ERG confirmed that the distribution of AML subtypes broadly represented that seen in UK practice.

Table 2 Baseline characteristics of patients in Study 301, ITT population (CS Table 8, page 43)

Characteristic	CPX-351	3+7
Number of patients (n)	153	156
Mean age (SD), years	[redacted]	[redacted]
Age, n (%)		
60-69 years	96 (62.7)	102 (65.4)
70-75 years	57 (37.3)	54 (34.6)
Male sex, n (%)	94 (61.4)	96 (61.5)
Race, n (%)		
White	128 (83.6)	139 (89.1)
Black or African American	7 (4.6)	6 (3.8)
Asian	6 (3.9)	2 (1.3)
American Indian or Alaska native	[redacted]	[redacted]
Other	11 (7.2)	9 (5.8)
Median weight (range), kg	82.0 (49.0, 134.0)	82.7 (46.0, 136.0)
Median height (range), cm	170.2 (149.0, 198.0)	170.2 (149.0, 189.0)
Median BSA (range), m^2	2.0 (1.0, 3.0)	2.0 (1.0, 3.0)
ECOG performance group, n (%)		
PS = 0	37 (24.2)	45 (28.8)
PS = 1	101 (66.0)	89 (57.1)
PS = 2	15 (9.8)	22 (14.1)
PS ≥ 3	0	0
Cytogenetic risk, n (%)[a]		
Favourable	7 (4.6)	5 (3.2)
Intermediate	64 (41.8)	58 (37.8)
Adverse	72 (47.8)	83 (53.2)
Unknown	10 (6.5)	10 (6.4)
White blood cell count, n (%)		
< 20 x 10^9/L	131 (85.6)	131 (84.0)
≥ 20 x 10^9/L	22 (14.4)	24 (15.4)
Unknown	0	1 (0.6)
Platelet count, n (%)		
≤ 50 x 10^9/L	95 (62.1)	91 (58.3)
> 50 x 10^9/L	58 (37.9)	63 (40.4)
Unknown	0	2 (1.3)
Haemoglobin, n (%)		
≤ 9 g/dL	[redacted]	[redacted]

Characteristic	CPX-351	3+7
> 9 g/dL	[redacted]	[redacted]
Unknown	[redacted]	[redacted]
Mean bone marrow blast (SD)		
Aspirate[b]	[redacted]	[redacted]
Biopsy[c]	[redacted]	[redacted]
Extra medullary disease, n (%)	[redacted]	[redacted]
AML subtype, n (%)		
Therapy related	30 (19.6)	33 (21.2)
MDS with prior HMA	50 (32.7)	55 (35.3)
MDS without prior HMA	21 (13.7)	19 (12.2)
CMMoL	11 (7.2)	12 (7.7)
de novo with MDS karyotype	41 (26.8)	37 (23.7)
Genetic mutations		
FLT3 mutated	22 (15.9)	21 (14.9)
NPM1 mutated	[redacted]	[redacted]
CEBPA mutated	[redacted]	[redacted]
Prior and concomitant medication, n (%)[d]	153 (100.0)	151 (100.0)
Prior anthracycline exposure	6 (3.9)	4 (2.6)

[a] Cytogenetic risk status was based on National Comprehensive Cancer Network (NCCN) guidelines for AML. [b] Mean bone marrow blast (aspirate) values based on n=141 patients in CPX-351 group and n=141 patients in 3+7 group. [c] Mean bone marrow blast (biopsy) values based on n=64 patients in CPX-351 group and n=60 patients in 3+7 group. [d] Prior and concomitant medication was assessed in the safety analysis population i.e. n=153 in the CPX-351 group and n=151 in the 3+7 group. AML, acute myeloid leukaemia; BSA, Body surface area; CMMoL, chronic myelomonocytic leukaemia; ECOG, Eastern Cooperative Oncology Group; HMA, hypomethylating agent; ITT, intent-to-treat; MDS, myelodysplastic syndrome; PS, performance status; SD, standard deviation

Study 301 was limited to high-risk AML patients between 60 and 75 years old, who may have poorer prognosis and may be harder to treat than the population of patients under 60 years old. The standard of care in UK practice for older patients (DA 3+10, 3+8, 2+5) is similar to the DA regimen used in the trial, but may be different in younger patients, who are less likely to have comorbidities and are more able to tolerate intensive chemotherapy.

Whilst the ERG does not have significant concerns about the generalisability of the trial population to the population of patients with high-risk AML aged between 60 and 75 in NHS practice, the generalisability of the trial results to the population of patients under 60 years is uncertain. As highlighted in Section 3.1 above, *de novo* AML with MDS associated karyotypic changes may not be diagnosed as 'high-risk' prior to receiving cytogenetic test results.

4.2.2 Summary of the quality of Study 301

Results of the quality assessment for Study 301 are presented in CS Table 10, with more detailed rationale for decisions in CS Appendix (Table 6).

Randomisation was performed using an interactive telephone or internet-based randomisation system. Patients were randomly assigned to treatment using a dynamic balancing randomisation algorithm to ensure a balanced distribution of the stratification variables (age and AML subtype) between the two treatment arms. There was no evidence of significant imbalance between the treatment arms at baseline.

Due to the open-label design of the trial, care providers and participants were not blind to treatment allocation. Treatment response was assessed independently and blinded to treatment allocation, but assessment of all other outcomes was not blinded. Lack of blinding is unlikely to have significantly affected objective outcomes such as OS. However the risk of bias due to lack of blinding cannot be ruled out for more subjective outcomes including the decision to transplant, which may ultimately affect post-HSCT efficacy and safety outcomes.

The ERG found no evidence that any additional outcomes were measured and not reported. Reasons for withdrawal were documented (CS Appendix Figure 2 and clarification response p7) and there was no evidence of a significant between-group difference in numbers of discontinuations. The efficacy analyses were conducted on an intention-to-treat (ITT) basis, and all randomised participants were evaluated for efficacy. Only five participants were excluded from the safety analyses (all were 3+7 treatment arm patients who were randomised but withdrew before receiving treatment).

An important concern is that the analyses presented in the CS and used in the economic model are based on a data cut from December 2015 (three years after the first patient was randomised). The trial follow-up phase was stated as being 5 years post-randomisation. Following the ERG's request for clarification, the company confirmed that follow-up is continuing for 5 years post-randomisation, but stated that no more recent data cut was available for OS or EFS. The clarification also stated that data on relapse following HSCT was not collected. The ERG finds this difficult to understand in the context of a trial designed to collect five-year follow up data. Elsewhere in their clarification response, and in apparent contradiction with this statement, the company, reported some additional data on EFS post-HSCT, although they note that this data is subject to significant limitations (see section 4.2.3.1).

The lack of any more recent analysis is disappointing given the extent of censoring in the current analyses. Updated analyses utilising the additional follow up that will have accrued since December 2015 would be informative and could reduce uncertainty.

The ERG considers Study 301 to be generally well conducted, although the lack of blinding means that the risk of bias for more subjective outcomes (notably transfer to HSCT and patient-reported adverse events) cannot be excluded. The lack of follow-up data beyond December 2015 and notably

the uncertainty about post-HSCT outcomes are significant limitations and mean that the longer-term relative efficacy and safety of CPX-351 compared with 3+7 is uncertain.

4.2.3 Summary of the results of Study 301

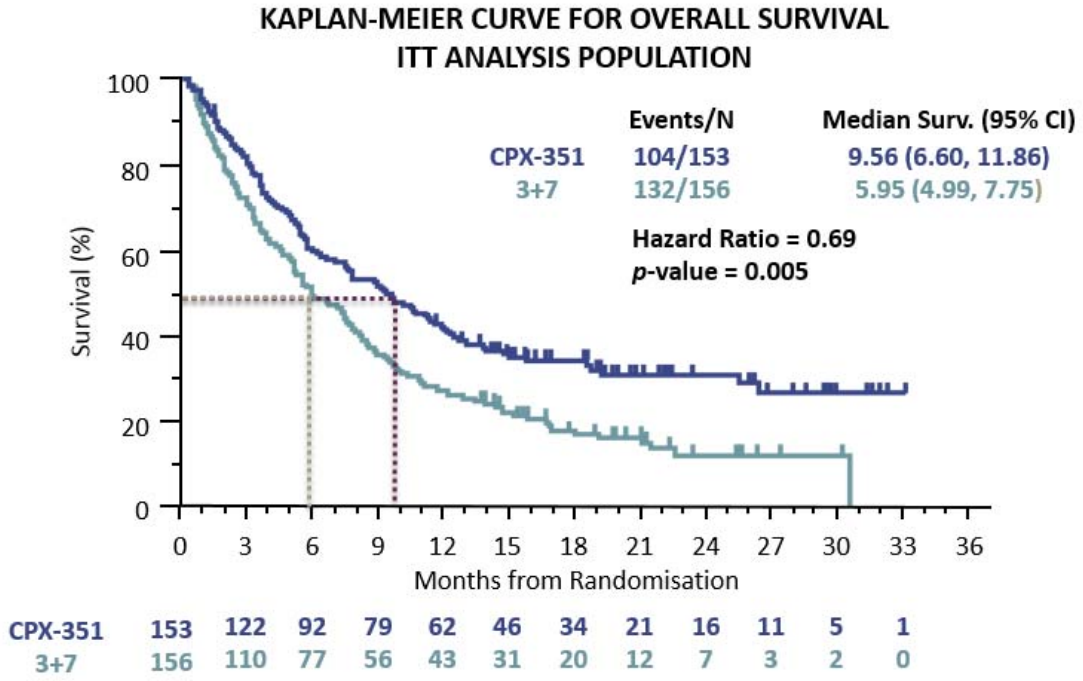
4.2.3.1 Efficacy results

Overall survival

The median length of follow-up was similar between treatment arms (20.5 months in CPX-351 group vs 21.2 months in 3+7 group). Figure 3 presents a Kaplan-Meier curve for OS in the ITT population. Median OS was significantly greater in the CPX-351 treatment group compared with 3+7 therapy (median OS: 9.56 months [95% CI: 6.60, 11.86] vs 5.95 months [95% CI: 4.99, 7.75]; HR=0.69 [95% CI: 0.52, 0.90], p=0.005). However, these results should be interpreted with caution due to substantial censoring and lack of mature OS data.

In the Safety population, 106 (69.3%) patients in the CPX-351 treatment group and 128 (84.8%) participants in the 3+7 group died during the treatment and follow-up phases of the study. The leading cause of death was progressive AML in both arms [redacted]. Following the ERG's request for more up to date data, the company stated that adverse event reporting indicates that to date there have been [redacted] known deaths on CPX-351 and [redacted] on 3+7 in the safety population.

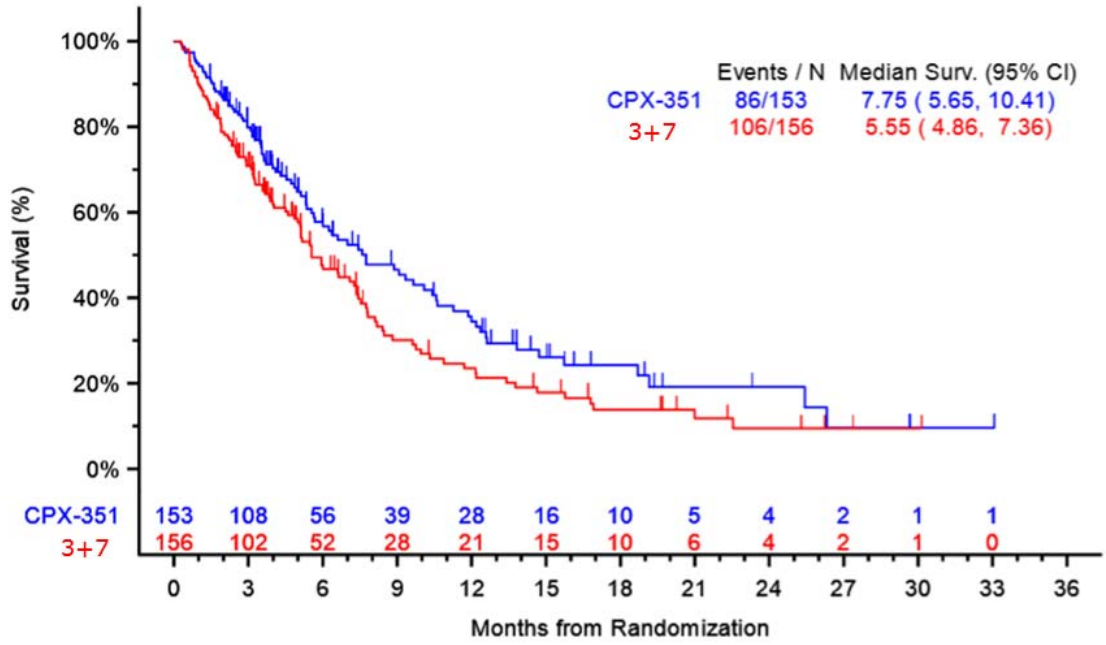
Figure 3 Primary endpoint of overall survival with CPX-351 and 3+7, ITT population (CS Figure 4, p49)



CI, confidence interval; ITT, intent-to-treat; OS, overall survival

Figure 4 presents the results of the sensitivity analysis that censored OS data at HSCT in the ITT population, which found that patients in the CPX-351 treatment group had a greater median OS compared with those in the control arm (7.75 months vs 5.55 months) although the difference was not statistically significant (HR=0.81 [95% CI: 0.60, 1.09], p=0.165).

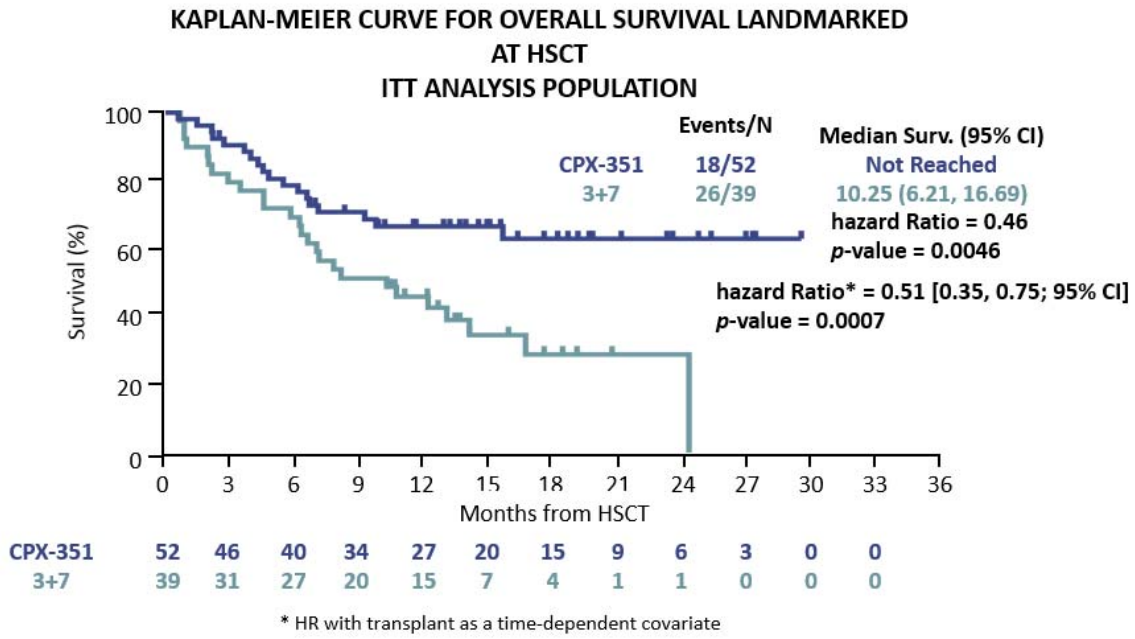
Figure 4 Sensitivity analysis (censored at HSCT): primary endpoint of overall survival with CPX-351 and 3+7, ITT population (CS Figure 5, p50)



CI, confidence interval; ITT, intent-to-treat; OS, overall survival

Results of a post-hoc analysis of the 91 patients (52 in the CPX-351 group and 39 in the 3+7 arm) who received HSCT landmarked from time of HSCT in the ITT population are presented in Figure 5 below. The survival curve shows a significant difference in OS post-HSCT favouring CPX-351 compared with 3+7, (curves diverging from approximately 6 months post-HSCT). These results should be interpreted with caution owing to small numbers, limited follow-up and associated censoring. Patients were selected for transplant meaning that groupings are no longer solely the product of randomisation. At the point of data cut the median OS was not reached in the CPX-351 group, and the median OS in the 3+7 group was 10.25 months (95% CI: 6.21, 16.69) (HR=0.46 [95% CI: 0.24, 0.89], p=0.0046). Adjustment for time of transplant as a time-dependent covariate made no substantial difference to the results (HR=0.51 [95% CI: 0.35, 0.75], p=0.0007).

Figure 5 Overall survival landmarked from time of HSCT (CS Figure 7, p54)

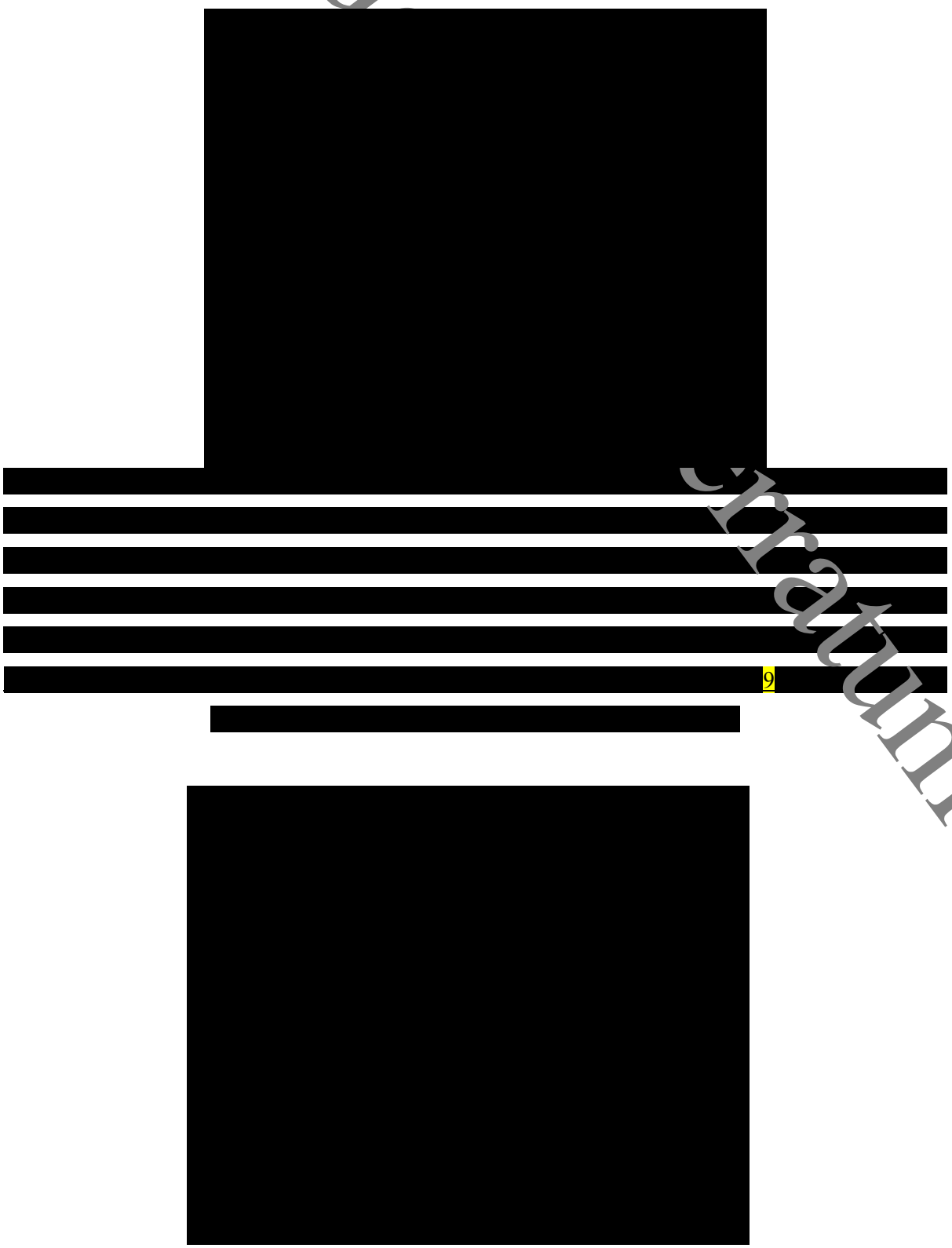


Following the ERG's request for clarification, the company provided further OS data for patients who received transplant by remission status (CR/CRi and no CR/CRi). Results showed improved OS post-HSCT in the CPX-351 arm vs. 3+7 regardless of remission status, although results should be interpreted with caution due to the concerns raised above.

[redacted]6[redacted]

[redacted]

[redacted]



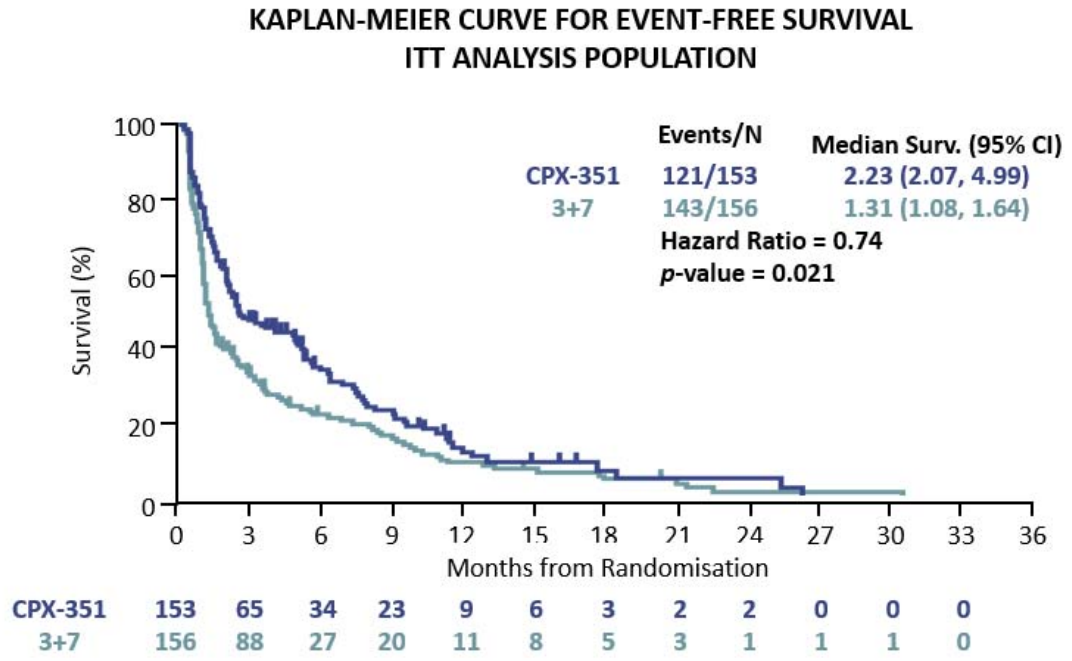
[redacted]

Event-free survival

Figure 10 (Figure 6 of CS) presents a Kaplan-Meier curve for EFS in the ITT population. Median EFS was significantly greater in the CPX-351 treatment group compared with 3+7 therapy (median EFS:

2.53 months [95% CI: 2.07, 4.99] vs 1.31 months [95% CI: 1.08, 1.64]; HR=0.74 [95% CI: 0.58, 0.96], p=0.021).

Figure 10 Event-free survival with CPX-351 and 3+7, ITT population (CS Figure 6, p52)



CI, confidence interval; ITT, intent-to-treat. The median survival value of 2.23 months for CPX-351 appears incorrect (2.53 months in the CSR)

[redacted]

In their response to the ERG request for clarification the company stated that

[redacted]

[redacted] For the sake of completion and as these results informed the economic eva uatio ney are presented below.

[redacted] [11]

[redacted] [11]

[redacted]

[redacted] [12]

[redacted]

12

[redacted]

13

[redacted]

[redacted]

[redacted]

[redacted]

14

[redacted]

Response

Table 3 shows that a significantly greater proportion of patients in the CPX-351 arm achieved an induction response. OR adjusted for age and AML subtype were reported for CR+CRi and CR, but not for CRi alone or non-response. The ERG calculated unadjusted OR using the Mantel-Haenszel test and found no statistically significant difference in CRi between the two arms, although these results should be interpreted with caution due to the relatively small number of events (unadjusted OR 1.40 [95% CI 0.64, 3.07]). Consistently with the CR+CRi results, the difference in proportion of non-responders was statistically significant (unadjusted OR 0.52 [95% CI [illegible].32, [illegible].82]).

Table 3 Proportion of patients with an induction response, ITT population (CS Table 11)

Endpoint, n (%)	CPX-351 (n=153)	3+7 (n=156)	Odds ratio (95% CI)[a]
CR+CRi	73 (47.7)	52 (33.3)	1.77 (1.11, 2.81), p=0.016
CR	57 (37.3)	40 (25.6)	1.69 (1.03, 2.78), p=0.040
CRi	16 (10.5)	12 (7.7)	NR
No response	80 (52.3)	104 (66.7)	NR

[a] Odds ratios were calculated using the 3+7 group as the reference. The resultant p-value is from a comparison of rates between treatment and is based on the Mantel-Haenszel test stratified by age and AML type groups
CI, confidence interval; CR, complete remission; CRi, complete remission with incomplete platelet or neutrophil recovery; NR, not reported

A greater proportion of patients receiving CPX-351 achieved remission after 1 induction cycle compared with 3+7 (CR: 47/105 [45%] with CPX-351 vs 28/100 [28%] with 3+7; CR+CRi: 58/105 [55%] with CPX-351 vs 34/100 [34%] with 3+7, respectively. Remission rates after 2 induction cycles were similar between treatment arms (CR: 10/48 [21%] vs 12/51 [24%]; CR+CRi: 15/48 [31%] vs 18/51 [35%], respectively).

Morphologic leukaemia-free state

In the MLFS population, a significantly greater proportion of patients in the CPX-351 group achieved a MLFS compared with the 3+7 group (69.0% [87/126] vs 55.5%, [66/119], OR=1.78 [95% CI: 1.05, 3.03], 2-sided p=0.034).

Remission duration

In the ITT population, there was no significant difference in remission duration between patients in the CPX-351 group compared with those in the 3+7 group. Remission duration was 6.93 months (95% CI: 4.60, 9.23) in the CPX-351 arm vs. 6.11 months (95% CI: 3.45, 8.71) in the 3+7 group (HR=0.77 [95% CI: 0.47, 1.26], 2-sided p=0.294).

Health-related quality of life

HRQL was not assessed in Study 301.

Administration setting

The proportion of patients receiving treatment in the outpatient setting was higher in the CPX-351 arm (51% in Cycle 1 and 61% in Cycle 2) compared with the 3+7 arm (6% in Cycle 1 and 0% in Cycle 2). It is unknown how many patients treated were hospitalised due to an adverse event related to treatment in an outpatient setting (see Section 4.2.3.2 for further details).

Stem cell transplant

In the ITT population, the proportion of patients undergoing HSCT was higher in the CPX-351 group (34.0% [52/153] vs 25.0% [39/156]) although the difference was not statistically significant (OR=1.54 [95% CI: 0.92, 2.56]).

Characteristics of patients who underwent HSCT are summarised in CS Table 12 and presented below (Table 4). The table shows that the CPX-351 arm had a higher proportion of participants who were 70 to 75 years (31% in CPX-351 arm vs. 15% in the 3+7 group) and a higher rate of participants with CR/CRi at time of transplant (77% in the CPX-351 group vs. 62% in the 3+7 arm). However, following the ERG's request for clarification the company reported that there were no statistically significant between-group differences in any of the characteristics of patients undergoing HSCT ($p<0.05$ significance level, Mantel-Haenszel Chi-Square tests and non-parametric Kruskal-Wallis tests), although the results of these tests should be interpreted with caution due to the relatively

small number of patients in each arm and evident lack of randomisation. As the decision to transplant was not blinded to treatment assigned, there is potential for bias for this outcome.

Table 4 Characteristics of patients undergoing HSCT (CS Table 12, p53)

n (%)	Characteristic	CPX-351 (n=52)	3+7 (n=39)
Age	60–69 70–75	36 (69) 16 (31)	33 (85) 6 (15)
ECOG PS	0–1 2	48 (92) 4 (8)	37 (95) 2 (5)
Karyotype	Intermediate Poor Unknown	27 (52) 21 (40) 4 (8)	18 (46) 19 (49) 2 (5)
Strata	t-AML MDS with prior HMA MDS without prior HMA CMMoL *de novo* with MDS karyotype	11 (21) 14 (27) 7 (14) 3 (6) 17 (33)	9 (23) 14 (36) 5 (13) 0 (0.0) 11 (28)
Induction response	CR CRi <CR	30 (58) 10 (19) 12 (23)	19 (49) 5 (13) 15 (38)
Disease status at time of transplant	CR/ CRi After salvage therapy Non-remission	39 (75) 5 (10) 8 (15)	24 (62) 12 (31) 3 (8)
Time from first dose	Median days (min, max)	114.5 (57.00, 418.00)	113.0 (50.00, 311.00)

AML, acute myeloid leukaemia; MDS, myelodysplastic syndrome; CMMoL, chronic myelomonocytic leukaemia; ECOG PS, Eastern Cooperative Oncology Group performance status; HMA, hypomethylating agent; CR, complete remission; CRi, complete remission with incomplete count recovery

Subgroup and exploratory analyses

Results of pre-planned subgroup analyses (age and AML subtypes) for OS in the ITT population are presented in Table 5. For convenience, the ERG summarised these results as a forest plot in [redacted] 15.

[redacted]

[redacted]The results of these pre-planned subgroup analyses should be subject to caution notably due to the relatively small number of patients.

Table 5 Overall survival by randomisation strata, ITT population (CS Appendix Table 7, p44)

Strata	Median OS, months (95%CI)		Hazard ratio (95%CI)
	CPX-351	3+7	
Age, years			
60-69	[redacted]	[redacted]	[redacted]
70-75	[redacted]	[redacted]	[redacted]
AML subtype			
CMMoL	[redacted]	[redacted]	[redacted]
de novo with MDS karyotype	[redacted]	[redacted]	[redacted]
MDS with prior HMA	[redacted]	[redacted]	[redacted]
MDS without prior HMA	[redacted]	[redacted]	[redacted]
t-AML	[redacted]	[redacted]	[redacted]

AML, acute myeloid leukaemia; CI, confidence interval; CMMoL, chronic myelomonocytic leukaemia; HMA, hypomethylating agent; ITT, intent-to-treat; MDS, myelodysplastic syndrome; NE, not estimable; OS, overall survival; t-AML, therapy-related acute myeloid leukaemia

[redacted] 15 [redacted]

AML, acute myeloid leukaemia; CI, confidence interval; CMMoL, chronic myelomonocytic leukaemia; HMA, hypomethylating agent; ITT, intent-to-treat; MDS, myelodysplastic syndrome; t-AML, therapy-related AML

Further post-hoc subgroup analyses of OS by genetic mutation were conducted and are reported in CS Appendix Table 8. The analyses found that in patients with FLT3, NPM1 or CEBPA genetic mutation, median OS was approximately twice as long with CPX-351 compared with 3+7, but these results should be interpreted with caution due to the small patient numbers.

Further subgroup analyses for all secondary endpoints in the ITT population by pre-specified strata (age and AML subtype) were reported in CS Appendix E.

4.2.3.2 Safety results

As stated above (Section 4.2.1.1), adverse events were recorded from the start of the infusion on Day 1 to the last day of the treatment period, and serious adverse events were also collected up to 30 days after treatment completion. CS Table 15 shows a summary of the number and rates of adverse events per arm in the safety population. [redacted]

[redacted]

Grade 5 adverse events (deaths)

Rates of adverse events leading to death were [redacted]

Most frequently reported severe adverse events (Grades 3 to 5)

[redacted] (

Table 6 Number of patients with treatment-emergent adverse events, incidence >10% by system organ class, Grades 3 to 5, Safety population (from CSR, Table 14)[redacted]

Table 6 Number of patients with treatment-emergent adverse events, incidence >10% by system organ class, Grades 3 to 5, Safety population (from CSR, Table 14)

[redacted]	[redacted]	[redacted]	[redacted]
[redacted]	[redacted]	[redacted]	[redacted]
[redacted]	[redacted]	[redacted]	[redacted]
[redacted]	[redacted]	[redacted]	[redacted]
[redacted]	[redacted]	[redacted]	[redacted]
[redacted]	[redacted]	[redacted]	[redacted]
[redacted]	[redacted]	[redacted]	[redacted]
[redacted]	[redacted]	[redacted]	[redacted]
[redacted]	[redacted]	[redacted]	[redacted]

Table 7

Table 7 Number of patients with treatment-emergent adverse events, incidence ≥5% by preferred term, Grades 3 to 5, Safety population (source: Study 301 CSR, Table 15)

Serious adverse events

Serious adverse events included any adverse events that either resulted in death (Grade 5), were life threatening, required inpatient hospitalisation or prolongation of existing hospitalisation, resulted in

persistent or significant disability or incapacity, or was a congenital anomaly. Serious adverse events were recorded from treatment start to 30 days after treatment completion.

Table 8 presents the rates of the most frequently reported serious adverse events with an occurrence of >3% by preferred term. The rate of serious adverse events was [redacted]

Table 8 Serious adverse events (>3% frequency) summary (CS Tables 15 & 18)

	CPX-351 (n=153)	3+7 (n=151)	All (n=304)
Preferred term	**n (%)**	**n (%)**	**n (%)**
Any serious adverse events	90 (58.8)	65 (43.0)	155 (51.0)
Febrile neutropenia	12 (7.8)	8 (5.3)	20 (6.6)
Respiratory failure	11 (7.2)	8 (5.3)	19 (6.3)
Ejection fraction decreased	9 (5.9)	9 (6.0)	18 (5.9)
Sepsis	12 (7.8)	5 (3.3)	17 (5.6)
Pneumonia	10 (6.5)	6 (4.0)	16 (5.3)

The CS stated that the higher observed rate of serious adverse events may be partly due to the greater proportion of patients in the CPX-351 arm who received consolidation in the outpatient setting (51% in Cycle 1 and 61% in Cycle 2) compared with the 3+7 arm (6% in Cycle 1 and 0% in Cycle 2), as a move to the hospital setting is one of the criteria for classifying an adverse event as serious. However, any prolongation of existing hospitalisation in either arm would also qualify as a serious adverse event, therefore the ERG are unconvinced that the difference in observed serious AEs may be explained by the higher rate of CPX-351 patients receiving treatment in the outpatient setting.

The company stated that the difference in observed serious adverse events may also be explained by the higher cumulative exposure to treatment in the CPX-351 group compared with the 3+7 arm. In the consolidation phase, a greater proportion of patients in the CPX-351 group received both an initial and second consolidation compared with the 3+7 treatment arm (32% vs. 21% for the first consolidation and 15% vs. 7.9% for the second consolidation). The median length of the treatment exposure and length of treatment phase [redacted] Following the ERG's request for clarification, the company provided further results from a more conventional Poisson distribution

adjusting for exposure duration which yielded comparable mean estimates between CPX-351 and 3+7 [redacted] The company stated the 95% confidence interval around the difference in the estimates included zero and concluded that there is no real difference in the SAE rates between the two study groups.

Based on these additional analyses the ERG's view is that although potentially concerning, the difference in observed treatment-related serious AEs between the study arms may be largely explained by the higher number of cycles and longer duration of treatment in the CPX-351 arm compared with 3+7.

Adverse events of interest

Infection-related adverse events

Table 9 presents the incidence of the most frequently reported infection-related adverse events (all grades) with an occurrence of ≥5%. Nearly all participants had at least one infection-related adverse event (92.8% in CPX-351 group vs 92.7% in 3+7 group). [redacted]

Table 9 Number of patients with Grade 1-5 infectious adverse events, ≥5% safety population (source: clarification response Table 2, p6)

	CPX-351 (n=153) n, %	**3+7 (n=151) n, %**
Any AE of infection	142 (92.8)	140 (92.7)
Febrile neutropenia	[redacted]	[redacted]
Chills	[redacted]	[redacted]
Pneumonia	[redacted]	[redacted]
Pyrexia	[redacted]	[redacted]
Sepsis	[redacted]	[redacted]
Cellulitis	[redacted]	[redacted]
Bacteraemia	[redacted]	[redacted]

Table 10 Number of patients with Grade 3-5 infectious adverse events, ≥5% safety population (source: clarification response Table 3, p6)

	CPX-351 (n=153) n (%)	**3+7 (n=151) n (%)**
Febrile neutropenia	[redacted]	[redacted]
Pneumonia	[redacted]	[redacted]
Sepsis	[redacted]	[redacted]
Bacteraemia	[redacted]	[redacted]

Bleeding-related adverse events

Table 11 presents the incidence of the most frequently reported bleeding-related adverse events (all grades) with an occurrence of ≥5%. More participants in the CPX-351 group had at least one bleeding-related adverse event (74.5% in CPX-351 group vs 59.6% in 3+7 group). [redacted]

Table 11 Number of patients with Grade 1-5 bleeding adverse events, ≥5% safety population (source: clarification response Table 4, p6)

	CPX-351 (n=153) n, %	3+7 (n=151) n, %
Any bleeding-related AE	[redacted]	[redacted]
Epistaxis	[redacted]	[redacted]
Petechiae	[redacted]	[redacted]
Mouth haemorrhage	[redacted]	[redacted]
Ecchymosis	[redacted]	[redacted]
Contusion	[redacted]	[redacted]
Haematuria	[redacted]	[redacted]
Blood blister	[redacted]	[redacted]
Gingival bleeding	[redacted]	[redacted]
Haemoptysis	[redacted]	[redacted]
Conjunctival haemorrhage	[redacted]	[redacted]

In response to the ERG's request for clarification the company reported that no Grade 3-5 bleeding-related adverse events occurred at a frequency of ≥5% in the safety population.

Cardiac adverse events

[redacted]

Alopecia

[redacted]

Fatigue

[redacted]

Time to recovery from neutropenia and thrombocytopenia

CS Table 17 summarises the time to recovery from neutropenia (Absolute Neutrophil Count [ANC] value ≥ 1000/μL) and thrombocytopenia (platelets ≥ 100,000/μL) for the ITT population. The median time from start of induction to recovery from neutropenia and thrombocytopenia was longer in the CPX-351 group than in the 3+7 group (44 days vs 35 days for ANC and 49 days vs 44 days for platelets). Median time to recovery for patients who achieved CR or CR/CRi was also longer in patients receiving CPX-351 compared with 3+7. The median length of recovery from neutropenia and thrombocytopenia from the start of induction in patients with two inductions was not significantly different from those with a single induction.

4.2.4 Supporting data from non-RCTs

No supporting data from non-RCTs was presented.

4.3 Conclusions of the clinical effectiveness section

Study 301 was a phase 3 multi-centre trial that randomised 309 patients with high-risk AML and used appropriate endpoints. The trial used 1:1 randomisation stratified by age and AML subtype. However, the ERG has some concerns about the validity of the clinical evidence. There was limited information on selection and characteristics of patients not receiving HSCT and decisions on whether to transplant were made with the knowledge of which treatment a patient had received, meaning that the risk of patient selection bias at point of transplant cannot be excluded. Analyses are based on a data cut from December 2015, which includes substantial censoring. The company was unable to provide the ERG with OS analyses based on more recent data.

Study 301 suggests that compared with 3+7, CPX-351 is associated with a significant improvement in OS (median OS: 9.56 months [95% CI: 6.60, 11.86] vs 5.95 months [95% CI: 4.99, 7.75]; HR=0.69 [95% CI: 0.52, 0.90], p=0.005) in patients with high-risk AML. Although results from the subgroup analyses should be interpreted with caution, there was some evidence to suggest that CPX-351 had a less beneficial impact on OS in patients with MDS with prior HMA [redacted] These patients constituted around a third of patients in the trial and a similar proportion of those who would be eligible for CPX-351 in clinical practice.

Overall CPX-351 may be a more effective bridge to stem cell transplant compared with 3+7 in patients with high-risk AML aged 60-75 years. Compared with 3+7, the proportion of patients undergoing HSCT was higher in the CPX-351 group (34.0% [52/153] vs 25.0% [39/156]) although the difference was not statistically significant (OR=1.54 [95% CI: 0.92, 2.56]) and the decision to transplant was not blinded to treatment allocation. OS in patients who underwent HSCT was significantly greater with CPX-351: at the point of data cut the median OS was not reached in the CPX-351 group, and the median OS in the 3+7 group was 10.25 months (95% CI: 6.21, 16.69)

(HR=0.46 [95% CI: 0.24, 0.89], p=0.0046). However, the OS results post-HSCT should be subject to caution given the small number of patients, limited follow-up duration, extensive censoring and lack of randomisation.

Overall the safety profiles of CPX-351 and 3+7 appear broadly comparable. The overall incidence of observed Grade 3-5 adverse events was similar across groups. Although potentially concerning, the higher incidence in observed treatment-related serious AEs in CPX-351-treated patients may be largely explained by the higher number of cycles and longer duration of treatment in the CPX-351 arm compared with 3+7.

The relative impact of CPX-351 vs. 3+7 on HRQL is unknown as HRQL and utility data were not collected. Due to the lack of up-to-date follow-up data and the associated substantial censoring of patients there is significant uncertainty about the longer-term efficacy and safety of CPX-351 including after stem cell transplant. The lack of evidence in patients under 60 years with high-risk AML means that the applicability of the trial results to this patient group is uncertain.

5 Cost-Effectiveness

This section focuses on the economic evidence, submitted by the company, and the additional information provided in response to the ERG's 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 areas of uncertainty. Section 6 presents additional analyses and scenarios, independently undertaken by the ERG, to further explore these uncertainties.

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

The company conducted a systematic literature review to identify relevant economic evidence associated with adults with AML. Searches for cost-effectiveness, costs and healthcare resource studies were carried out simultaneously. The ERG's critique of the systematic review presented by the company is given below.

5.1.1 Searches

The CS described the search strategies used to identify 1) relevant cost-effectiveness studies of adult patients with AML and 2) relevant studies of healthcare resource use or costs in adult patients with AML. The searches were described in Section B.3.1 of the submission appendix and full search strategies were presented in Appendix G. The strategies used and databases searched were considered appropriate.

The appropriate databases used for the cost-effectiveness systematic literature review were searched. Additional searches of conference websites were conducted to identify information. These are reported in Section G.1.2 of the CS. The search strategies used in MEDLINE, Embase, EconLIT and the NHS EED databases are fully reproduced in Tables 15 to 18 of the CS appendix and the number of records identified is given.

5.1.2 Inclusion/exclusion criteria used for study selection

The inclusion/exclusion criteria are reported in Appendix G (CS appendices, Table 19). Studies that assessed economic evidence in adult patients with AML were included in the review. The ERG considers that the inclusion/exclusion criteria used were reasonable.

5.1.3 Studies included and excluded in the cost-effectiveness review

Three of the records identified in the cost-effectiveness review were evaluations of treatments for AML from a UK payer perspective. Details of two additional studies presented by the company were previous submissions to NICE. A brief overview of those articles are presented in Table 12.

Table 12 Overview of UK economic evaluations (CS Table 23, p82)

Study	Patient population	Treatment and comparator	Model description	Estimated ICER
NICE TA10124 – 2018 [19 20]	Newly diagnosed FTL3-mutation positive AML	Midostaurin versus standard of care	Partitioned survival model with lifetime horizon	NR
NICE TA218 -2011 [21]	AML patients with 20-30% marrow blasts, not eligible for HSCT	Azacitidine versus CCR, BSC, low-dose chemotherapy	Partitioned survival model with lifetime horizon	Against BSC: £63,177 Against low-dose chemotherapy: £49,030 Against CCR: £59,954
Tremblay *et al.*, 2017 [22]	Newly diagnosed FTL3-mutation positive AML	Midostaurin (in combination with SOC) versus SOC	Partitioned survival model with lifetime horizon	£34,327
Tikhonova *et al.*, 2017 [23]	AML patients 65+ with greater than 30% marrow blasts, not eligible for HSCT	Azacitidine versus intensive chemotherapy with anthracycline in combination with cytarabine; non-intensive chemotherapy with low-dose cytarabine; and versus BSC only	Partitioned survival model with lifetime horizon	£273,308
Wang et al., 2014 [24]	Newly diagnosed AML	Induction chemotherapy (ADE - cytarabine, daunorubicin, etoposide; DA - cytarabine, daunorubicin) versus no induction	Probabilistic decision model with lifetime horizon	NR
BSC, best supportive care; CCR, conventional care regimen, HSCT, haematopoietic stem cell transplantation; ICER, incremental cost-effectiveness ratio; NR, not reported; QALYs, quality-adjusted life years; SOC, Standard of care				

5.1.4 Conclusions of the cost-effectiveness review

Five cost-effectiveness studies were identified and considered relevant for the cost-effectiveness review. However, none of these evaluated the cost-effectiveness of CPX-351. 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 13.

Table 13 Summary of the company's economic evaluation

	Approach	Source / Justification	Signpost (location in company submission)
Model	Cost-effectiveness (cost-utility) analysis using a hybrid model. Model consisting of an initial decision tree phase allocating patients to different cohorts according to whether or not patients achieved remission and reached transplant, followed by a partitioned survival approach, with progression and mortality based on survival curves derived from Phase III trial CLTR0310-301 (Study 301). A one-week model cycle is used in the first 104 weeks (2 years) of the model time horizon. In subsequent years of the model time horizon, an eight-week model cycle is used.	The partitioned survival approach is commonly used to model the cost-effectiveness of oncology treatments and has been present in previous NICE submissions for AML.	Section B.3.2.3 pg. 84-89
States and events	The model contains the following health states: newly diagnosed disease, remission (comprising remission and post-consolidation remission), transplant (comprising transplant and post-transplant remission), progression and death.	Designed to reflect the current clinical pathway and guidelines for people with high-risk (secondary) AML.	Section B.3.2.3 pg.85
Comparators	The CPX-351 regimen was compared to the 3+7 regimen (comprising daunorubicin and cytarabine)	3+7 was considered to be the most appropriate comparator to CPX-351, 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 this population.	Section B.3.2.4 pg. 89-91
Treatment effectiveness	Clinical outcomes included were: response rate, probability of transplant in responders, number of rounds of induction and consolidation therapy; time to post-induction response, time from post-induction response to last consolidation treatment, and time from last consolidation treatment to transplant; time to event analyses for post-consolidation OS and EFS in those who had response but did not receive transplant, post-transplant OS and EFS in those who had response, and OS, 'time to transplant or death' and 'time to progression or transplant or death' in those who do not achieve remission. Multivariate analyses were conducted to adjust all of the clinical outcomes with the stratification variables (age and high-risk AML subtype), and, where it was appropriate, with the rounds of induction and consolidation therapies. Parametric models were fitted to time to events analyses to extrapolate beyond the end of trial follow-up.	Data for the treatment effectiveness analyses were taken from Study 301 and post-hoc analyses conducted to estimate the values.	Section B.3.3 pg.92 and 100.

	Approach	Source / Justification	Signpost (location in company submission)
Mortality	Survival was estimated short-term and long-term using parametric extrapolation of Study 301 data. Additionally, age- and sex-specific mortality rates based on general UK population implemented, in any period and for any group of patients where modelled OS suggested lower mortality than the general population.	Data on short-term and long-term mortality was sourced from Study 301. Age- and sex-specific mortality rates were sourced from ONS 2017, National life tables: England. In a scenario analysis, these were adjusted by a HR estimated from Martin (2010).[25]	Section B.3.3.3 pg.96-99 and Section B.3.3.5 pg.100
Adverse events	Grade 3-5 AEs with at least 5% frequency of occurrence in at least one treatment arm in the study were included in the model. Included AEs were bacteraemia, diarrhoea, ejection fraction decreased, fatigue, febrile, neutropenia, hypertension, hypotension, hypoxia, pneumonia, respiratory failure and sepsis.	Adverse event rates were taken from Study 301. AEs were assumed to occur in Year 1 due to limitations in data availability.	Section B.3.3.4 pg.99-100.
Health-related quality of life	HRQL data was not collected in Study 301 and health state utilities were sourced from a utility elicitation study. [26] A SLR undertaken by the company provided an alternate set of utility values that were used in a scenario analysis.	Health state utilities were sourced from a utility elicitation study.[26] Data for scenario analysis were sourced from a vignette-based utility elicitation study conducted in the UK by Hensen et al., 2017.[27]	Section 3.4.5 pg.101-105
Resource utilisation and costs	These comprised: drug acquisition (first line and second line therapy), drug administration (delivery of chemotherapy and hospital stay), disease monitoring, transplant and treatment of adverse effects.	Drug acquisition costs for CPX-351 were based on list price with a confidential PAS applied. Unit costs for drug acquisition costs (generic compounds) were sourced from eMIT and BNF.[28] Unit costs for administration, monitoring, and adverse events were taken from NHS reference costs (2017) [29]. Post-consolidation costs were sourced from Mahmoud (2012) [30]. Costs associated with transplant were estimated from a combination of sources, including an NHS Blood and Transplant analysis (2014)[31] and NHS Reference costs (2017).[29]	Section B.3.5 pg. 107-126
Time horizon	30 years (assumed as lifetime).	Assumed to be a lifetime time horizon for this population (at least 60 years of age).	Section B.3.2.3 pg. 89
Discount rates	Beyond one year, the costs and benefits were discounted at 3.5% per annum.	In accordance with the NICE reference case.	Section B.3.2.3 pg. 89
Sensitivity analysis	Probabilistic sensitivity analysis was performed. Deterministic analysis was performed on a series of model parameters. A series of scenario analyses was also performed.	In accordance with the NICE reference case.	Section B.3.8 pg. 134-145
Subgroups	No subgroup analysis was conducted.	These were not considered applicable in this appraisal.	Section B.3.9 pg. 145

	Approach	Source / Justification	Signpost (location in company submission)
Note: N/A, not-applicable; NHS, National Health Service; NICE, National institute of health and care excellence; PSS, personal social services; QALY, quality-adjusted life years; SoC, standard of care			

5.2.1 Model structure

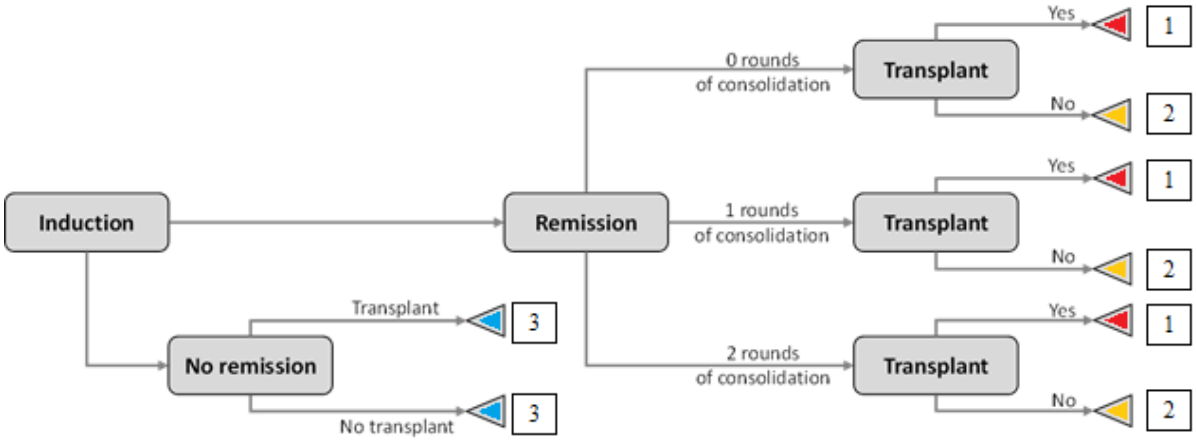
The CS presented a *de novo* cohort cost-effectiveness model to estimate the cost-effectiveness of CPX-351 compared with 3+7 in a population of adult AML patients with high-risk (secondary) AML.

Cost-effectiveness was assessed over a lifetime time horizon of 30 years. The cycle length used in the model was one week in the first two years of the model and eight weeks thereafter, which was considered to be sufficiently granular to accurately capture model costs and outcomes throughout the treatment pathway. A half-cycle correction was applied to costs and QALYs.

The modelled health states included newly diagnosed disease; remission (defined as CR or CRi, comprising post-consolidation and post-transplant remission); disease progression (comprising relapse after remission and progression in treatment non-responders); and death.

The analysis used a hybrid modelling approach. The initial part of the model was a decision tree (Figure 16), and determined whether patients achieved remission after induction therapy, and whether those achieving remission after induction received transplant ("post-transplant remission") or did not receive transplant ("post-consolidation remission"). Patients entered the model in the newly diagnosed health state, where they received either one or two rounds of induction therapy. If patients achieved remission, then they could receive up to two rounds of consolidation therapy. Probabilities of receiving transplant after remission, of requiring second induction, of receiving none or one or two rounds of consolidation therapy, and of achieving remission after induction were estimated from a regression analysis of data from Study 301, and are described in Section 5.2.6.1.

Figure 16 Model decision tree (presented in the company model)



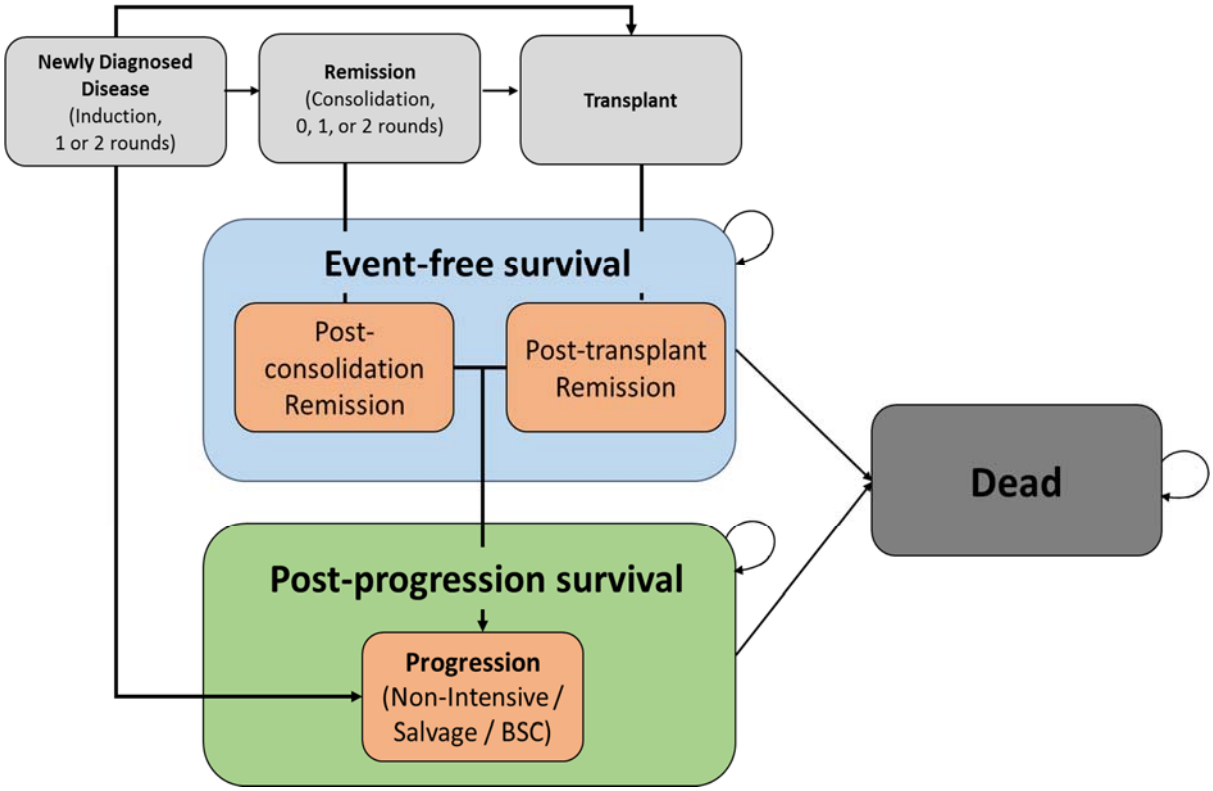
Following the decision tree, a partitioned survival approach was used. Overall survival and relapse-free survival outcomes for the three groups of patients, those in post-transplant remission (pathway 1 in Figure 17 Model schematic (CS Figure 11, pg. 86)), in post-consolidation remission (pathway 2 in Figure 17), and in non-responders (pathway 3 in Figure 17), were modelled separately. These were based on survival analyses of patient-level data from Study 301.

Survival models determined the mortality and rates of relapse of patients in the post-transplant and post-consolidation remission health states. OS and EFS were tracked from the point at which initial treatment was completed (i.e. a time-shift was applied from the start of the model), and no deaths or relapses occurred in these patients until the time-shift. For example, OS and EFS in the post-transplant state were applied starting from the time of transplant. A regression analysis of Study 301 data in responder patients provided the mean time to post-induction response, the time from post-induction response to last consolidation treatment, and the time from last consolidation treatment to transplant (Section 5.2.6.2).

In non-responder patients, mortality was determined by the overall survival, and progression-free survival was estimated using the difference between two curves: the time to progression, transplant or death and the time to transplant or death (Section 5.2.6.3). Survival models were estimated and applied from the start of the model i.e. unlike the responder patients, a time-shift was not performed in this analysis. Patients who did not achieve remission could also receive transplant, and again unlike the responder patients, the survival outcomes of these patients was not modelled separately but rather captured within the overall survival and progression-free survival models for the non-responder patients. The proportion of patients receiving transplant was determined from the time to transplant analysis, estimated as the difference between OS and the time to transplant or death analysis.

In the progressed health state, patients could receive non-intensive therapy, salvage therapy or best-supportive care (Section 5.2.4.2), and experienced a lower quality of life (Section 5.2.7). It was assumed that remission after second-lines of therapy could not be achieved, and no further relapses were captured.

Figure 17 Model schematic (CS Figure 11, pg. 86)



Many of the model parameters were based on regression analyses of patient-level data from Study 301, which included patient characteristics (ages as category 60-69 years or 70-75 years, AML subtype, and treatment arm in certain analyses) as covariates. These characteristics determined the probability of remission, the probability of transplant in responder patients, the number of induction and consolidation courses, the time to event-based probabilities described in Section 5.2.6.2, and the survival analyses described in Section 5.2.6.3. The proportion of patients in the age categories, gender and AML categories are presented in Section 5.2.3.

ERG comment

Modelling the treatment pathway in AML has historically taken a more sophisticated approach than for survival models in solid cancers, to allow for the impact of transplantation and the possibility of long-term cure [20, 32]. This is reflected in the complex structure of the company model. A strength of the model is that it allowed for explicit links between events that occurred at the start of the treatment

pathway, such as the number of induction therapy courses received and the rate of remission and the rate of transplantation. However, this approach requires more data to parameterise the model. Further, segregating the model into the three sub-cohorts requires that the trial data were further dissected for the survival analysis. These were then based on smaller patient numbers, and the analyses may not be sufficiently powered, possibly compromising their robustness.

The benefit of a parameterised model based on patient baseline characteristics is that is can be easily manipulated to explore the impact in individual subgroups. Importantly, this approach also allows the model to extrapolate parameter values outside the sample of patients recruited to the trial. This is potentially important given the restrictive inclusion criteria applied in Study 301 which included only patients between 60 and 75 years of age. However, the implementation of this by the company provided an additional layer of complexity in the model structure, and made the results of the model less transparent.

The ERG is generally satisfied with the approach taken to model patients in remission, that is, the use of separate models for patients following transplant or after consolidation therapy provided a framework whereby patient outcomes could be appropriately captured, and that the trial data was appropriately analysed to model these patients. A minor limitation of the model is that subsequent relapses were not explicitly modelled. However, the impact is expected to be minimal: data presented by the company at the clarification stage suggested that very few patients in this population receive third-line therapy. Furthermore, the survival of any patients with second relapses would be captured within the OS analysis.

However, the ERG is unclear on why this approach was not taken to model non-responder patients. A partitioned survival model is appropriate for use in advanced disease, where patients progress through a series of more and more severe health states, before death [33]. However, in reality, patients may not progress through a linear pathway: they could receive transplant before progression or after progression, and progression could occur before or after transplant. Given the model structure and the data used to parameterise the model, the model did not capture any relapses that occurred after transplant. The number of patients who progressed was based on the time to transplant or progression or death analysis (Section 5.2.6.3), which is censored at the first event that occurred. Therefore, if progression occurred after transplant, then it was not captured in the model. The exclusion of relapse events after transplant would result in an overestimation of QALYs and underestimation of costs. While there were only a small number of patients who received transplant in non-remission, these patients are associated with higher rates of survival and generate more QALYs than non-transplant patients. With a greater proportion of 3+7 patients with no response, this bias would result in a more conservative estimate of the ICER.

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

Table 14 summarises the economic submission and the ERG's assessment of whether the company's economic evaluation meets NICE's reference case and other methodological recommendations.

Table 14 Features of *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 NICE final scope lists the following comparators: • Standard intensive induction and consolidation therapy • Azacitidine (for people who are not eligible for haematopoietic stem cell transplantation and have AML with 20-30% blasts and multilineage dysplasia) • Midostaurin (for people with FLT3-mutation-positive AML) (subject to ongoing NICE appraisal) • Gemtuzumab ozogamicin (subject to ongoing NICE appraisal)	Yes	The comparator was restricted to daunorubicin and cytarabine (3+7) regimen for induction and consolidation. 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 for older patients. However, a comparator not listed in the NICE scope, FLAG-Ida, may be considered a relevant comparator for younger patients.
Type of economic evaluation	Cost-effectiveness analysis.	Yes	Cost-utility analysis (CUA) with the direct health effects expressed in terms of QALYs.
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	30 years of time horizon used in the economic model which is assumed to be equivalent to a life-time horizon for this population (at least 60 years of age).
Synthesis of evidence on outcomes	Systematic review.	NA	The evaluation used clinical evidence from Study 301. No meta-analysis was conducted, as no other relevant trials of CPX-351 compared with 3+7 were identified in the systematic review. Study 204 (to determine efficacy and identify patient subgroups that may benefit from CPX-351 treatment) included a subset of secondary AML patients, but no data was extracted for use in this appraisal.
Measure of health effects	QALYs.	Yes	Utility values were used from a utility elicitation study. The elicitation study aimed to estimate health state utilities associated with AML treatment options and was conducted to value the health states considered in the model. Vignette-based time trade-off (TTO) interviews with a one-year time horizon were conducted in a UK general population (Edinburgh and London).
Source of data for measurement of HRQL	Reported directly by patients and/or caregivers.	No	
Source of preference data for valuation of changes in HRQL	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	Probabilistic sensitivity analysis was undertaken.

5.2.3 Population

The population included in the company's decision problem and economic model comprises adult patients with untreated, high-risk (secondary) AML. This is defined by therapy-related AML (t-AML) or AML with myelodysplasia-related changes (AML-MRC).

The population in the economic analysis was based on the intention-to-treat population of Study 301. The inclusion criteria of the trial were patients aged 60 to 75, and thus the population represented a narrower population than the anticipated marketing authorisation, which includes all adult patients. Table 15 presents the characteristics of the patients in Study 301. These were included as covariates in a number of analyses of trial data that were then used to determine the value of key clinical parameters, including rate of response and survival, in the model (Section 5.2.6).

Table 15 Patient baseline characteristics (adapted from Table 24, CS, pg. 88)

Characteristic	Distribution
Gender	
Male	61.0%
Female	39.0%
Age group	
60–69 years	64.0%
70–75 years	36.0%
AML type	
Treatment-related AML	20.0%
MDS with prior HMA exposure	34.0%
MDS without prior HMA exposure	13.0%
CMMoL	7.4%
de novo AML with karyotype characteristic of MDS	25.6%
CMMoL, chronic myelomonocytic leukaemia; HMA, hypomethylating agents; MDS, myelodysplastic syndrome	

To account for the inclusion criteria restricting eligibility to patients aged between 60 and 75, the company included a scenario analysis that estimated the cost-effectiveness in a population that included a proportion of patients under the age of 60 years, one where there were *(i)* 10% of patients and *(ii)* 30% under 60. The ERG notes that these scenarios were both associated with a lower ICER (Table 34). The company modelled the impact of these patients by applying an odds ratio (OR) of 1.9 to the rate of remission for patients who were between the ages of 60 and 69 (Appendix O in the CS, pg. 571). The OR was estimated from Dumas (2017), a study of chemotherapy in older patients with

secondary or therapy-related AML. [34] The study found that age and type of treatment were significantly associated with response, where the probability of response decreases with increasing age.

ERG comment

The population discussed in the decision problem section (Section 3.1) highlighted a number of issues.

Firstly, the ERG notes that there are likely to be a number of patients who would be eligible for treatment with CPX-351 in this population who are under the age of 60, but who are excluded from the trial. The clinical advisor to the ERG stated that these patients potentially account for up to 25% of the eligible patient population unrestricted by age.

Secondly, the ERG had concerns regarding the approach taken by the company to extend the model to include these patients. The Dumas (2017) study that the company used to estimate the relative response rate in these younger patients did not include any patients under the age of 60, and the ERG was unclear on why the company used this source. Given the limited applicability of the Dumas (2017) study, the ERG considered that a similar analysis could have been performed using Study 301 data, which would have provided a relative response rate in the population pertinent to the decision problem.

Further, it is notable that the results of the Dumas (2017) study contrasted with the outcome of the regression analyses performed by the company, where age was not associated with a statistically significant change in response rates. The estimated odds ratios between these two age groups in Study 301 was also lower than the odds ratio estimated using the Dumas (2017) study ([redacted] vs 1.9, CS main submission, Table 27, pg. 93), suggesting that age may be a less important factor in determining response rates in this population. The ERG explored using a lower odds ratio in a scenario analysis presented in Section 6.

Thirdly, as highlighted in Section 3.1, the patients enrolled in Study 301 included a heterogeneous mix of patients with five different subtypes of AML. As described in Table 5, the magnitude of the treatment effect that was observed in each of these subtypes varied greatly. While these results were based on small patient numbers and provided indicative results only, they suggest that the effectiveness, and by extension, cost-effectiveness, of CPX-351 in these different populations may differ. Of particular note is the MDS with prior HMA subgroup, which comprises the largest patient group in the trial. The HR for OS observed in this group suggests a minimal or zero treatment effect (HR for OS: [redacted], 95% CI: [redacted]). Exploration of subgroups in the model is very difficult due to the model structure and the limitations of the clinical data it is based upon. However, given a HR of

[redacted], it is highly unlikely that CPX-351 would represent a cost-effective use of resources due to the minimal clinical benefits and substantially higher drug acquisition costs.

Finally, as also discussed in Section 3.1, there is an issue relating to the identification of patients with *de novo* AML with MDS-associated karyotypic changes, as doing so requires undertaking genetic tests. The implication of this is that it is often the case in clinical practice that patients are treated with the standard 3+7 regimen until cytogenetic test results are received. As such, the ERG questions how often these patients would receive the first cycle of induction therapy in clinical practice.

5.2.4 Interventions and comparators

5.2.4.1 First-line therapy

The company's economic model evaluated CPX-351 compared to 3+7 intensive chemotherapy. The dosing and administration schedule of the intervention and comparator regimens reflected that of the pivotal study and the decision problem. The dosing of CPX-351 also reflects the anticipated marketing authorisation license. Table 16 provides a summary of the dosing schedules applied in the model.

Treatment consisted of an induction phase and a consolidation phase. Patients who do not respond to the first round of induction therapy may receive a second round of induction therapy. In the second round of induction therapy, the number of administrations is reduced for both treatments (for the comparator treatment, this dosing schedule is also referred to as 2+5). Patients in remission after induction therapy may receive up to two courses of consolidation therapy in the model. CPX-351 is provided at a lower dose during consolidation, and provided at the same, lower, frequency as in the second induction.

Table 16 Dosing regimens of first-line treatments (CS Table 26, pg. 90)

Intervention: CPX-351 regimen	Comparator: 3+7 regimen
First induction: Daunorubicin 44 mg/m^2 and cytarabine 100 units/m^2 IV infusion over 90 minutes on Days 1, 3, and 5.	First induction: 100 mg/m^2/day of cytarabine administered by continuous infusion over seven days 60 mg/m^2/day of daunorubicin given over 15 minutes on Days 1, 2, and 3.
Second induction: Daunorubicin 44 mg/m^2 and cytarabine 100 units/m^2 IV infusion over 90 minutes on Days 1 and 3.	Second induction: 100 mg/m^2/day of cytarabine administered by continuous infusion over five days 60 mg/m^2/day of daunorubicin given over 15 minutes on Days 1 and 2.

Consolidation: Daunorubicin 29 mg/m^2 and cytarabine 65 mg/m^2 IV infusion over 90 minutes on Days 1 and 3.	Consolidation: 100 mg/m^2/day of cytarabine administered by continuous infusion over five days 60 mg/m^2/day of daunorubicin given over 15 minutes on Days 1 and 2.
IV, intravenous; m, metres; mg, milligrams	

ERG comment

As previously highlighted in the decision problem section (Section 3.2), there are potential differences in the dosing used in UK clinical practice for the cytarabine and daunorubicin (3+7) chemotherapy regimen. The model was based on the US 3+7 regimen as used in the clinical trial, rather than the typical 3+10 regimen used in the UK. However, the ERG considered that these two regimens could be considered equivalent in terms of efficacy based on the guidance provided from the British Committee for Standards in Haematology [6] and clinical advice received by the ERG, and so it was likely to be sufficiently representative of the current standard of care in the UK.

A number of comparators included in the NICE scope were not included in the company's evaluation (Section 3.3). The ERG was satisfied with the rationale for their exclusion. However, patients under the age of 60, who were younger than those enrolled in Study 301, may receive FLAG-Ida as standard care. The company did not submit any cost-effectiveness evidence for CPX-351 compared with FLAG-Ida.

Further to the above, the ERG also notes that the draft marketing authorisation for CPX-351 permits patients to receive up to four courses of consolidation, and that treatment can be continued as long as the patient benefits or until disease progression. This is inconsistent with the pivotal trial in which patients were restricted to a maximum of two rounds of consolidation therapy. Given this inconsistency, the ERG felt that it was uncertain how many patients would receive third and fourth consolidation courses in practice. Evidence from Study 301 suggests that few patients would continue on to third and fourth round consolidation therapy as only [redacted] of patients received a second round of consolidation therapy (Table 13, CSR) [35]. The clinical advisor to the ERG also confirmed that the number of courses of consolidation would not be likely to be above 2 or 3, since clinicians would aim to bridge patients to transplant as soon as possible. The number of patients who receive further courses of CPX-351 may therefore be small in practice. It is unclear what the impact of patients receiving additional lines of consolidation therapy would be on cost-effectiveness as while it would increase costs it would potentially also improve effectiveness.

5.2.4.2 Second-line therapy

Patients who do not respond to induction therapy and experience disease progression, or who experience a relapse following remission, could receive further treatment. This group was assumed to receive either salvage therapy, non-intensive therapy or best-supportive care (which consisted of monitoring only), or received no treatment (or monitoring) at all. The proportion of patients receiving each type of therapy was estimated from a published report (CancerMPact, 2015) [36], which provided details on treatment patterns for elderly AML patients in western Europe going through relapse. These values were applied to relapsed/progressed patients, regardless of response or transplant status.

Non-intensive therapy and salvage therapy consisted of a number of options of active therapy (Table 17). The relative proportions were extracted from the CancerMPact 2015 report [36]. Combined with the unit costs of the treatments (Section 5.2.8.2), these parameters were used to estimate a weighted cost of non-intensive and salvage therapy. The company applied standard treatment schedules for second-line therapies as recommended in ELN guidelines. [7]

Table 17 Treatments received as second-line therapy (CS Table 29, Table 37 and Table 38, pg. 95, 114)

Treatment	Proportion of patients treated
Non-intensive therapy	[redacted]
Azacitidine	[redacted]
Low-dose cytarabine	[redacted]
Salvage therapy	[redacted]
Intermediate-dose cytarabine + daunorubicin	[redacted]
Intermediate-dose cytarabine + idarubicin	[redacted]
Intermediate-dose cytarabine + mitoxantrone	[redacted]
Mitoxantrone + etoposide + cytarabine	[redacted]
Fludarabine + cytarabine + GCSF + idarubicin	[redacted]
Best supportive care only	[redacted]
No treatment	[redacted]
Source: CancerMPact (2016) [36]	

ERG comment

The ERG is satisfied that the selection of second-line therapies in the model is generalisable to UK practice.

However, the ERG has some concerns regarding the proportion of patients receiving second-line systemic therapy in the model.

Firstly, the ERG is unclear on how the company estimated resource use from the CancerMPact report (Table 24 in source) and it is not clear that the modelled proportion of patients receiving second-line therapy reflects what is reported in the CancerMPact report. The CancerMPact report states that only one-fifth of patients who undergo induction therapy undergo therapy at first relapse. This contrasted with the modelled assumption, where just under 50% of patients receive systemic therapy, either non-intensive therapy or salvage therapy, after relapse.

To determine whether second-line usage in the model aligned with that in the clinical trial, the ERG requested data from the company on the usage of second-line therapy in the trial. The company provided a list of chemotherapies that were used in the follow-up phase of the trial, i.e. at any time after induction and consolidation, and the proportion of patients who were treated with that therapy. Chemotherapy usage in the follow-up phase of the trial was higher than applied in the model, with [redacted] of CPX-351 patients and [redacted] 3+7 patients receiving chemotherapy. There were some limitations with this data which meant it could not be used to determine whether second-line therapy would vary depending on whether it was received after a failure to achieve response after first-line therapy or relapse following consolidation. Further, the data provided reported on individual chemotherapy items rather than regimens, and could not be used to comment on the generalisability of the regimens used in the model.

As such, it was difficult to determine whether the second-line therapy in the model and in the trial were consistent, and it was difficult to determine the generalisability of the trial and the model to the UK setting. Given these limitations with the trial data, the ERG did not consider it feasible to incorporate a scenario analysis that reflects second-line therapy in the trial, so this has not been explored further. The ERG, however, considers that it is unlikely that second-line therapies will have significant impact on cost-effectiveness, since the prognosis after relapse in this group of patients is very poor [37]. As such, it is unlikely that any salvage regimens would have a clinically meaningful survival benefit, and, therefore, any differences in salvage therapy options would likely relate to cost. Based on the differences in EFS observed between arms (Figure 10), the model does not appear to be sensitive to this parameter, but it is unclear on what the direction of bias would be.

5.2.5 Perspective, time horizon and discounting

The economic model adopted a National Health Service (NHS) and personal and social services (PSS) perspective in accordance with the NICE reference case.

The time horizon used in the economic model was 1,576 weeks or approximately 30 years. The NICE reference case indicates that the time horizon used for estimating clinical and cost-effectiveness should be sufficiently long to reflect any differences in costs and benefits between the technologies being compared. The ERG considers that 30 years was an appropriate time horizon, as it is very unlikely that patients would remain alive beyond 30 years.

The costs and benefits in the model were discounted at an annual rate of 3.5%, as per the NICE reference case.

5.2.6 Treatment effectiveness and extrapolation

The CS provided a description of the clinical data used in the model and analyses conducted to populate the model (CS appendix M pg. 215-470). All analyses conducted were based on the ITT populations, and post-hoc statistical analyses used individual patient-level data from Study 301 (CLTR0310-301, NCT01696084). [15-17, 38]

The analyses conducted to populate the model estimated the following outcomes:

- Treatment pathway probabilities;
    - Probability of receiving a second round of induction therapy,
    - Probability of remission (CR + CRi),
    - Probability of receiving 0 or 1 or 2 rounds of consolidation therapy,
    - Probability of receiving transplant after remission.
- Mean time elapsed after induction therapy (time-shift);
    - Time to post-induction response,
    - Time between post-induction response and end of consolidation,
    - Time between end of consolidation and transplantation.
- Survival analysis;
    - Overall survival,
    - Event free survival,
    - Time to HSCT or death in those who did not achieve remission.
- Rates of adverse events in each arm.

All analyses except adverse event rates were adjusted for the sampling stratification variables, including age (60–69 vs. 70–75 years old) and high-risk AML subtypes. All analyses included one or

more co-variate adjustment selecting from treatment arms, rounds of induction therapy and rounds of consolidation therapy, but the inclusion of co-variates for adjustment were not consistent in all analyses. As described in Section 5.2.1, the mean estimates predicted by the regression analyses were not used in the model directly; probabilities were estimated from the regression analyses undertaken by the company, and adjusted for each of the stratification variables.

In the sections below (5.2.6.1 to 5.2.6.5), the ERG provides descriptions of results and assumptions made in the economic model. Additionally, a brief description of the technical details and results of the multivariate analyses and a comprehensive assessment by the ERG is provided in Section 10 Appendix of this report.

5.2.6.1 Treatment pathway probabilities

As described in Section 5.2.1, the initial phase of the model allocated patients to different treatment pathways to determine long-term survival. Multivariate logistic regression analyses were performed to estimate the percentage of the following clinical pathway outcomes:

- Probability of receiving a second round of induction therapy,
- Probability of remission (CR + CRi),
- Probability of receiving 0, 1 or 2 rounds of consolidation therapy,

Probability of receiving transplant after remission. Table 18 reports the estimated mean probability of each event occurring in each treatment arm, and the probability of each event reported in the trial for comparison. The model appeared to predict the trial values well in each analysis.

Table 18 Modelled probabilities of events

Parameter	**CPX-351**		**3+7**	
	Mean probability in model	**Observed trial probability**	**Mean probability in model**	**Observed trial probability**
Probability of receiving a second round of induction therapy	[redacted]	[redacted]	[redacted]	[redacted]
Probability of remission, post-induction	[redacted]	[redacted]	[redacted]	[redacted]
Probability of receiving two rounds of consolidation therapy after post-induction remission	[redacted]	[redacted]	[redacted]	[redacted]
Probability of receiving one round of consolidation therapy after post-induction remission	[redacted]	[redacted]	[redacted]	[redacted]
Probability of receiving no rounds of consolidation therapy after post-induction remission	[redacted]	[redacted]	[redacted]	[redacted]

Probability of receiving transplant, after remission	[redacted]	[redacted]	[redacted]	[redacted]

These analyses used age and high-risk AML subtype as covariates, although these were not significant predictors in any of the analyses (except for age in the patients who received a second round of induction therapy). For the probability of receiving consolidation therapy and receiving transplant, AML subtype was grouped into two categories (one category consisting of MDS AML with or without prior HMA, and the remaining three in the other category).

The probability of remission also included a treatment effect because the higher rate of remission in CPX-351 patients than 3+7 patients observed in Study 351 was statistically significant. In the other analyses, the same probability was applied regardless of treatment arm as the trial found no statistically significant differences between arms. However, there were some small observed differences between treatment arms in the trial. For example, the observed probability of receiving transplant was somewhat higher in CPX-351 patients ([redacted] than in 3+7 patients ([redacted], and a lower proportion of 3+7 patients received two rounds of consolidation compared to CPX-351.

The regression analyses for all probabilities also included the number of induction courses (except the probability of receiving second round of induction therapy) as covariates.

ERG comment

The ERG is largely satisfied with the approach taken to estimate the probabilities, but has concerns about the combined analysis for treatment arms and inconsistent parameterisation in the multivariate analyses.

None of the analyses (except remission after induction) included treatment arms as a prognostic factor and therefore the variation by treatment arms may not be fully reflected in the analysis. Particularly, in the analysis for the patients who achieved remission and received transplant, the observed percentage was higher in CPX-351 patients ([redacted]) than in 3+7 patients ([redacted]); however, the treatment arms were combined in the regression analysis giving combined 51% probability for the patients who achieved remission and received transplant. The combined estimate slightly underestimates the probability of transplant in CPX-351 arm and overestimates in the 3+7 arm. Similarly, in the analysis of patients who receive consolidation, the observed second consolidation was lower in the 3+7 arm ([redacted]) compared to the CPX-351 arm ([redacted]). While the company's approach may provide less accurate predictions, the ERG considers that it may be clinically plausible that each group is subject to a similar rate of transplantation.

The categories of high-risk AML subtypes used as covariates in the regression analysis were not consistent. For example, the analysis of remission post-induction uses five categories of AML, but the analysis for rounds of consolidation in patients who had a remission uses two categories of AML. The ERG requested justification for this inconsistency at the clarification stage. The company clarified that the regrouping was due to small sample sizes, and adjusting for AML type using five levels led to either convergence issues when fitting regression models or resulted in unstable models. The ERG considers this an acceptable response, and does not believe that this will introduce any significant bias, given that results over all five AML subtypes are pooled in the base-case analysis. However, it prevents any further exploration of the impact of CPX-351 within specific AML subtypes, and the ERG believes that it is likely that the probability of each of these events would vary across the different AML subtypes, which would lead to biased results should subgroups be explored individually.

Despite the inconsistencies, the model predicted percentages were similar to the observed percentages, and the ERG did not explore this issue further.

5.2.6.2 Mean time elapsed after induction therapy (time-shift)

In the economic model, for patients who achieve response after induction therapy, OS and EFS are not tracked until after a certain amount of time had elapsed (as described in Section 5.2.1).

The CS conducted multivariate linear regression analyses to estimate the duration of these time-shifts for each subgroup in the model. The CS applied these time shifts only to patients who achieved remission after the induction phase. No time-shift was applied in relation to the survival of patients who did not respond to treatment in the induction phase; therefore, the parametric curves assessing the successive composite endpoints for this group of patients are applied from the beginning of the model.

The following three time periods were estimated using multivariate linear regression analyses:

- Time elapsed between randomisation and post-induction response among patients who achieve remission,
- Time elapsed between post-induction response and end date of consolidation among patients who achieve remission,
- Time elapsed between end date of consolidation and start date of transplantation among patients who received transplant.

Table 19 presents the estimated mean time elapsed in the model and mean time elapsed in the trial. The estimated mean times to post-induction response from the regression analysis were comparable with the observed mean times in the trial. However, the analyses predicted the trial less accurately in

the other two analyses. The mean time from last consolidation treatment to transplant predicted by the regression analysis was slightly higher than the observed mean time in the trial, and the regression analysis appeared to underestimate the mean time from post-induction response to last consolidation treatment.

Table 19 Modelled time-shift parameters

Parameter	**CPX-351**		**3+7**	
	Mean probability in model	**Observed trial probability**	**Mean probability in model**	**Observed trial probability**
Time to post-induction response	[redacted]	[redacted]	[redacted]	[redacted]
Time from post-induction response to last consolidation treatment	[redacted]	[redacted]	[redacted]	[redacted]
Time from last consolidation treatment to transplant	[redacted]	[redacted]	[redacted]	[redacted]

These analyses also used age and high-risk AML subtype as covariates, although these were not significant predictors in any of the analyses. For "time from last consolidation treatment to transplant", AML subtype was grouped into two categories (similar to the probability of transplant and receiving rounds of consolidation therapy). The "time to post-induction response" included a treatment effect as it was found to be significant, whereas in the other analyses, the same time-shift was applied regardless of treatment arm. Other covariates that were found to be statistically significant and included in the analysis included: the number of rounds of induction for "time to post-induction response", and the number of rounds of consolidation for "time from post-induction response to last consolidation treatment" and "time from last consolidation treatment to transplant".

ERG comment

The ERG is satisfied with the approach taken by the company in estimating these parameters. However, they, note two points. Firstly, similar to the analyses described in Section 5.2.6.1, the categorisation of AML subtypes is not consistent. Two AML categories are used for the "time from last consolidation treatment to transplant" analysis, but the rest of the analyses included five categories. Secondly, time elapsed analyses were conducted using a multivariate linear regression assuming a normal distribution, which does not reflect the typical skewness of the time variable. Assuming a linear relationship between covariates and time results in the regression poorly predicting the observed data. This may explain the observed differences between the predictions made by the regression analysis and the observed times in Study 301.

5.2.6.3 Survival analysis and extrapolation beyond the trial period

The company fitted parametric survival curves to the patient-level data from Study 301 to extrapolate over the model time horizon. These were stratified by response status, and further by transplant status in those who achieved a response. Parameterised survival curves are presented in Section 10 Appendix, and the corresponding Kaplan-Meier curves are described in Section 4.2.3.1.

The parametric curves included:

- *Overall survival and event-free survival* for patients who responded but did not receive transplant, where EFS and OS were calculated from the time of last consolidation therapy. This group is henceforth referred to as *post-consolidation patients*.

- *Overall survival and event-free survival* for patients who responded and received transplant, where EFS and OS were calculated from the time of transplant. This group is henceforth referred to as *post-transplant patients*.

- *Overall survival, time to transplant or death, and time to transplant or progression or death,* for patients who did not achieve response, where the survival models were calculated from the beginning of the model. This group is henceforth referred to as *non-responder patients*.

To extrapolate each of these survival curves, the company explored a range of conventional parametric models. Curves were adjusted for sampling stratification variables age (60–69 vs. 70–75 years old), AML subtype, treatment arm, number of rounds of induction therapy, and number of rounds of consolidation, where appropriate. AML subtype was grouped into two categories (one category consisting of MDS AML with or without prior HMA, and the remaining three in the other category) in the post-transplant analyses of OS and EFS. The Akaike information criterion (AIC) and Bayesian information criterion (BIC) were used to choose among the different parametric curves. Projected survival curves and median estimates were also examined to assess the clinical plausibility of the distributions. Table 20 presents a summary of the fitted parametric distributions and predictions from each of the time-to-event analyses used in the base-case economic model.

Table 20 Summary of fitted parametric distributions and predictions used in base-case economic model

Parameters	CPX-351 (N = 153)				3+7 (N = 156)			
	Number of patients	Parametric curves	Mean LE (weeks)	Median	Number of patients	Parametric curves	Mean LE (weeks)	Median
Responder - models have been fitted separately for each treatment								
Post-consolidation OS	[redacted]	Loglogistic	[redacted]	[redacted]	[redacted]	Loglogistic	[redacted]	[redacted]
Post-consolidation EFS	[redacted]	Weibull	[redacted]	[redacted]	[redacted]	Weibull	[redacted]	[redacted]
Post-transplant OS	[redacted]	Gompertz	[redacted]	[redacted]	[redacted]	Gompertz	[redacted]	[redacted]
Post-transplant EFS	[redacted]	Lognormal	[redacted]	[redacted]	[redacted]	Lognormal	[redacted]	[redacted]
Non-responder – models have been fitted without treatment as a covariate (i.e. merging both treatments)								
Non-responder OS	[redacted]	Lognormal	[redacted]	[redacted]	n/a	n/a	n/a	n/a
Non-responder time to transplant or death	[redacted]	Lognormal	[redacted]	[redacted]	n/a	n/a	n/a	n/a
Non-responder time to progression or transplant or death	[redacted]	Generalised gamma	[redacted]	[redacted]	n/a	n/a	n/a	n/a

$, [redacted] people ([redacted] in CPX arm and [redacted] in 3+7 arm) were censored; ^, non-responder who did not have transplant and who had transplant were combined together; CI, ; LE, life expectancy; N, number of patients randomised; n/a, not applicable;

Post-consolidation OS and EFS

Both analyses provided mean estimates that were comparable with the observed Kaplan Meier (KM) estimates. However, these analyses were based on small numbers of patients in the tail of the KM curve. At 84 weeks, [redacted] patients in CPX-351 group and [redacted] patients in 3+7 group remained in the survival analysis.

EFS was shown to be similar in each treatment arm, and each parametric model provided similar long-term projections. For OS, the choice of parametric curves in the CPX-351 arm had relatively little impact on long-term predictions (median OS varies between [redacted] to [redacted] weeks). In contrast, the choice of parametric model in the 3+7 arm had a significant impact on estimated rates of long-term survival (median OS varies between [redacted] to [redacted] weeks). Two of the models provided by the company for 3+7 provided more favourable survival estimates than for CPX-351 (Gompertz and Weibull).

Post-transplant OS and EFS

These analyses were based on a small number of patients from the trial, and a very small number of patients remained at the end of the KM curves. At 84 weeks, [redacted] patients in the CPX-351 group and only [redacted] patient in the 3+7 group remained in the KM curve for OS.

For patients receiving CPX-351, the predicted median time to relapse or death from the log-normal fit was [redacted] weeks, which is slightly lower than the results observed with the KM survival curve [redacted] weeks. For patients receiving 3+7, the predicted median time to relapse or death from the log-normal fit was [redacted], which was much lower than the results observed with the KM survival curve ([redacted]). The CS noted that differences in the median values were due to the small sample size involved in this analysis (as only [redacted] patients received 3+7 treatment and had a positive value for the time to relapse or death post-consolidation).

EFS was demonstrated as being similar in each treatment arm. However, there was a large difference between arms for OS.

While the different models for 3+7 provided similar long-term predictions for OS, there appeared to be heterogeneity across outcomes in the CPX-351 analysis.

Table 21 presents the predicted medians and model fit statistics for the parametric distributions for OS. For both treatments, the model fit statistics differ very little for various parametric distributions, but the predicted difference in median life expectancies between CPX-351 and 3+7 varied greatly (a difference of [redacted] weeks for exponential, to [redacted]weeks for generalized gamma). Predicted

median post-transplant OS from the Gompertz model did not reach the 50% survival rate at any time point, as it plateaued at [redacted] after approximately 2 years. In contrast, the predicted median post-transplant OS from the Gompertz model for 3+7 arm was [redacted] weeks.

Table 21 Post-transplant OS - Predicted Life Expectancy and Model Fit Statistics (CS Appendix Table 74 and Table 75, pg. 300)

Distribution	**CPX-351**				**3+7**				**Difference in predicted LE CPX-351 vs. 3+7**	**Difference in predicted medians CPX-351 vs. 3+7**
	Mean LE (weeks)	**Median LE (weeks) (95% CI)**	**AICC**	**BIC**	**Mean LE (weeks)**	**Median LE (weeks) (95% CI)**	**AICC**	**BIC**		
Weibull	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
Log-normal	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
Log-logistic	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
Exponential	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
Generalized gamma	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
Gompertz	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]

*: Incorrect value reported; AICC, Akaike information criteria corrected; BIC, Bayesian information criteria; LE: Life expectancy;

Non-responder OS, time to transplant or death, time to transplant or progression or death

In the analyses among patients who do not achieve remission, the KM curves for both CPX-351 and standard care were combined, justified by the lack of a detected treatment effect. Within each analysis, the survival models displayed little variation regarding their long-term predictions. Predicted median times from each analysis were comparable with the observed KM data. For time to transplant or death, a very small number of patients were included in the analysis (a total of [redacted], [redacted] in CPX-351 and [redacted] in 3+7).

ERG comment

The ERG has significant concerns related to survival analyses and extrapolation beyond the trial period, which are discussed in turn below.

Post-transplant OS for CPX-351

Immature data and high censoring

There are a substantial number of patients ([redacted]) who were censored in the CPX-351 arm after 1 year in the post-transplant OS analysis (CS response to clarification, Figure 3) and very small numbers of patients remained after 84 weeks ([redacted] patients in the CPX-351 group and only [redacted] patient in the 3+7 group). The ERG, therefore, has concerns that the available data were too immature to robustly estimate the survival benefit for post-transplant patients.

In response to the ERG's request for clarification, the company provided additional data on deaths captured through adverse event monitoring that have occurred since the 2015 data cut used for the survival analyses ([redacted] people in the CPX-351 arm, and [redacted] people in the 3+7 arm). While the transplant status of these patients was not known, the ERG considered that it was plausible that the additional patients who died after the data cut-off time point could have been in the post-transplant group. These additional deaths were not included in the current analysis. The higher number of deaths in the CPX-351 arm might suggest a degree of convergence in the OS curves.

Variability of long-term predictions

While the model is relatively robust to the choice of parametric curve for most outcomes, the parametric curve used to extrapolate post-transplant OS has a very significant impact on predicted OS gains and cost-effectiveness, with mean OS ranging from [redacted] years with the exponential curve to ~[redacted] years with the log-logistic model. Examination of the model fit statistics AIC and BIC, however, showed very little difference in the goodness of fit for the different survival models and as such, there is no statistical reason to prefer one curve over another. The ERG did note that the impact

of treatment was statistically significant for OS after transplant in the overall population (i.e. including non-responder transplant patients), but was highly concerned whether the magnitude of this benefit was sustained in the long-term, as suggested by the company.

Face validity

The ERG question whether it is clinically plausible for CPX-351 patients to gain an additional survival benefit after transplant compared to 3+7 patients, and even if it were plausible, whether a benefit of this magnitude would be realised.

The curve selected by the company for its base-case analysis was the Gompertz curve, which was selected on the grounds that it had best fit to the trial data and was able to capture the plateau in survival that is observed in long-term transplant survivors. However, the same plateau was not modelled for 3+7 patients. The ERG also consider that other survival models, such as mixture cure models and flexible spline models, are commonly used in the modelling of AML patients for this purpose, and would be able to capture any observed plateau more accurately, although they acknowledge that typically more mature data are required to be able to robustly estimate the cure point.

The plateau displayed by Gompertz suggested that, at two years, [redacted] of these patients can be considered "cured", and experience mortality at a similar rate to that of the general population. The ERG, however, question this choice, as the survival predictions of the Gompertz model appear to lack face validity. The survival for CPX-351 was substantially higher than is suggested by external sources of data for other types of AML that are less severe and, therefore, would be expected to have a higher cure rate than patients in this decision problem. For example, ten-year survival after transplant for AML patients over the age of 50 as reported by Shimoni (2016) [39] (which included patients with more favourable risk factors than Study 301), was approximately 35% compared with [redacted] for CPX-351 and [redacted] for 3+7. For patients in CR, 3-year survival after transplant was estimated as 49% [40].

The model assumption that the survival of CPX-351 is substantially better than for 3+7 patients post-transplant is a key driver of cost-effectiveness in the model. While this is consistent with the trial, the ERG have substantive concerns as to whether such apparent benefits would be observed in practice. The clinical advisor to the ERG suggested that a number of factors drive post-transplant OS including depth of remission and age of recipient, the ERG, however, could not find any clinical data reported in Study 301, such as minimal residual disease (MRD) status, to support this potential rationale for a benefit, and note that the patients in the CPX-351 arm of Study 301 who received transplant are older (on average) than the patients who received transplant in the 3+7 arm (Table 5).

Given these uncertainties regarding the extrapolation of post-transplant OS, the ERG explored the impact of using alternative parametric curves.

Post-transplant EFS

Immature data and high censoring

The company noted that the post-transplant EFS analysis was based on a small number of patients, a subset of [redacted] patients of those who received transplant after induction response ([redacted]). In total, [redacted] patients were excluded because they received transplant after their last examination. These data were collected close to the end-of-study follow-up, which suggests that the data is not sufficiently mature to predict outcomes in these patients. Among the [redacted] patients included in analysis, only [redacted] progressed (all of which subsequently died), [redacted] died (including those who progressed before dying), and [redacted] were censored.

Additionally, the company provided information about the collection of EFS data that appeared contradictory: in the clarification response, the company stated that "*information about relapse after HSCT was not collected*" (CS response to clarification, A2 priority question), while in the same document, data on relapse after transplant was presented. Therefore, it appears plausible that the data after transplant was not adequately collected.

Face validity

The ERG considered that EFS appears to lack face validity when considered in the context of OS in these patients. Unlike the post-transplant OS, there was little difference in post-transplant EFS between the two treatment arms, and the ERG questioned the clinical plausibility of a benefit in OS but not EFS. This is particularly important, as a substantial proportion of the benefit of CPX-351 derives from improvements in post-transplant OS. At the clarification stage the ERG requested that the company explain this inconsistency. In their response the company suggested that this is because the results from the post-transplant EFS analysis are an unreliable indication of the true treatment effect, and that no clinical inferences should be drawn from estimates of post-transplant EFS.

The consequences of the use of the post-transplant EFS analysis in the model implies that patients alive at two years are in the relapsed health state. Given the poor prognosis of these patients [37], it seems unlikely that they would experience a mortality rate similar to that of the general population (as suggested by the OS data at that point in the model), and the ERG considers that they would only spend a short amount of time in the relapsed health state before death.

While the ERG recognises that issues with how progression events were captured leads to some uncertainty the interpretation of the EFS data, the ERG can see no plausible reason why the EFS data

would be unrepresentative of patient outcomes and does not consider the company's explanation of this inconsistency to be satisfactory. Alternative interpretations of these data are explored by the ERG in Section Impact on the ICER of additional clinical and economic analyses undertaken by the ERG6.

Inconsistent approach to covariate adjustment

The ERG noted some inconsistencies in the inclusion of stratification variables, and some inconsistencies in the justification for their inclusion/exclusion. Covariates were often excluded due to a lack of statistical significance.

For example, in the post-transplant OS analysis for CPX-351, the number of induction rounds was statistically significant and included in the model. In contrast, for 3+7, the number of induction rounds was not included on the ground that it is not statistically significant. The approach to include the variable was inconsistent to the company's previous covariate inclusion for post-consolidation OS, where the number of consolidation rounds was not statistically significant for CPX-351, but was significant for 3+7.

The ERG generally questions the merits of the approach to adjust parameters by stratification variables: given the sparsity of data in various subgroups, the ERG was concerned that it has created additional computational challenges without having a benefit to the economic model. It added a source of uncertainty into the analysis, but it was difficult to determine how these inconsistencies may affect the results of the regression and subsequently bias the model.

Similarly to the regression analyses presented in Sections 5.2.6.1 and 5.2.6.2, the covariate grouping for AML type were not consistent. The ERG notes that regrouping AML types prevents any further exploration of the survival impact of CPX-351 within specific AML subtypes.

5.2.6.4 All-cause mortality

To ensure that the mortality rate in the model never fell below that of the general population, the mortality rate applied in the model was based on the highest rate predicted by either age- and sex-specific general population mortality rates (sourced from the ONS [41]) or the mortality rates predicted by the parametric curves used to extrapolate the trial survival data.

To explore the uncertainty in this assumption, the CS also conducted a scenario analysis to assess the impact of post-transplant mortality on the cost-effectiveness results considering a higher mortality than the general population mortality, based on the findings of Martin *et al.*, 2010 [25]. In this scenario, the general mortality was adjusted by means of a standardised mortality ratio (SMR) to reflect that patients would not fully return to normal life expectancy after transplant. This was calculated by estimating that that life expectancy would be [redacted] lower compared with that in the general population.

This resulted in a SMR of 2.34 being applied (2.25 for male and 2.47 for female), using the AML gender distribution in the UK [3]. This assumption implies that patients will experience significantly shorter life expectancy following transplant and results in a significant increase in the ICER, as it reduces the QALY benefits of achieving long-term cure.

ERG comment

The ERG considers the application of the SMR to be the more plausible assumption, as previous studies suggested patients would not fully return to normal life expectancy after transplant [25 42 43 44]. However, the impact of this scenario varies according to the choice of parametric curve used for the long-term survival prediction and are maximised when the Gompertz model is used for post-transplant OS as it was associated with lower long-term mortality. The ERG is satisfied with the HR adjustment in the model and considers it appropriate for inclusion in the ERG's base case (Section 6), although it is unclear on the method used to estimate the SMR.

5.2.6.5 Adverse events

In the economic model, the CS included Grade 3-5 adverse events (AEs) with at least 5% frequency of occurrence in at least one treatment arm in Study 301. AEs were associated with a one-off treatment cost (Section 5.2.8.5), applied in year one of the model. Table 22 presents the percentage of Grade 3-5 AE rates as reported in Study 301 and used in the model.

Table 22 The percentage of patients experiencing Grade 3-5 adverse events (CS Table 30, pg. 99)

Adverse Event	CPX-351	3+7
Bacteraemia	9.8%	2.0%
Diarrhoea	[redacted]	[redacted]
Ejection Fraction Decreased	5.2%	5.3%
Fatigue	7.2%	6.0%
Febrile Neutropenia	68.0%	70.6%
Hypertension	10.5%	5.3%
Hypotension	[redacted]	[redacted]
Hypoxia	13.1%	15.2%
Pneumonia	19.6%	14.6%
Respiratory Failure	7.2%	6.6%
Sepsis	9.2%	7.3%
Source: Medeiros et al., 2017 [45], Clinical Study Report for Study 301 [35]		

ERG comment

The ERG is satisfied with the implementation and selection of the adverse events during treatment.

5.2.7 Health related quality of life

Since HRQL data was not collected in Study 301, the company conducted a systematic literature review on utility values in AML and undertook a utility elicitation study [26]. The values from the elicitation study were applied in the company base-case analysis, while a study identified in the SLR that provided utility values was used in a scenario analysis.

Elicitation study

The elicitation study aimed to estimate health state utilities associated with AML treatment options and was conducted to value the health states considered in the model. Vignette-based time trade-off (TTO) interviews with a one-year time horizon were conducted in a UK general population (Edinburgh and London). Participants were asked to value twelve differing pre-specified health states, and four of these health states pertained to patients on induction and consolidation therapy (summarised in this section in Table 23). A total of 200 valid utility interviews were conducted, providing 193 complete valuations for the health states. The vignette descriptions were developed from a number of discussions with clinical experts with experience of treating AML patients. Appendix N of the CS provides further details on the utility elicitation study and the methods that were used to estimate utility values.

The TTO study provided utility values for the health states of AML (induction), progression, post-consolidation remission, and post-transplant remission. The company applied disutilities relating to induction therapy and consolidation therapy with CPX-351 and with 3+7 which captured the respective safety profiles of each treatment regimen and disutilities associated with transplant. The disutility associated with each course of treatment was multiplied by the number of cycles of treatment given, to provide the total disutility. Disutilities for each state were calculated as follows:

- Induction therapy: utility for remission between treatment minus the utility for induction therapy;
- Consolidation therapy: utility for remission between treatment minus the utility for consolidation therapy;
- Transplant disutility: utility for durable remission minus the transplant health state utility.

Table 23 provides a summary of the key assumptions in the vignettes for the on-treatment health states.

Table 23 Description of on-treatment health state vignettes (adapted from Tables in CS appendix N.5, pg. 487)

Vignette	Symptoms	Treatment process
Induction with 3+7	• Complete hair loss and risk of infection. • Feel very fatigued and tired.	• In hospital for about 4 weeks in total. • First treatment is administered with a continuous infusion pump for 7 days. You are connected to this pump 24 hours per day for all 7 days but can walk around the hospital with the pump • A second treatment is administered in a 15-minute infusion on days 1, 2 and 3. • After treatment, remain in the hospital for 3 more weeks to recover, receiving weekly blood transfusions to reduce symptoms and aid recovery.
Induction with CPX	• Thinning hair and risk of infection. • Feel a little fatigued and tired.	• In hospital for about 5 weeks in total. • Receive treatment by intravenous (IV) infusions treatment is administered with an infusion pump in a 90 minute infusion on days 1, 3, and 5. • Other than these three 90-minute periods, you are not connected to the infusion pump, and can move around freely • After treatment, remain in the hospital for 4 more weeks to recover, receiving weekly blood transfusions to reduce symptoms and aid recovery.
Consolidation with 2+5	• Complete hair loss and risk of infection. • Feel very fatigued and tired.	• In hospital for about 1 week in total. • Treatment is administered with a continuous infusion pump for five days. You are connected to this pump 24 hours per day for all five days but can walk around the hospital with the pump • Second treatment is administered in a 15-minute infusion on days 1 and 2 • After the week in hospital, you go home. For 2 weeks, you attend outpatient follow-up appointments twice per week for monitoring, receiving weekly blood transfusions to reduce symptoms and aid recovery.
Consolidation with CPX (inpatient)	• Thinning hair and risk of infection. • Feel a little fatigued and tired.	• In hospital for about 1 week in total. • Treatment is administered with an IV infusion pump in a 90-minute infusion on days 1 and 3. • Other than these two 90-minute periods, you are not connected to the infusion pump, and can move around freely. • For 3 weeks after, you attend outpatient follow-up appointments twice per week for monitoring, receiving weekly blood transfusions to reduce symptoms and aid recovery.
Consolidation with CPX (outpatient)	• Thinning hair and risk of infection. • Feel a little fatigued and tired.	• Receive this treatment at an outpatient clinic. • Treatment is administered with an IV infusion pump in a 90-minute infusion on days 1 and 3. • Other than these two 90-minute periods, you are not connected to the infusion pump, and can move around freely • For 3 weeks after, you attend outpatient follow-up appointments twice per week for monitoring, receiving weekly blood transfusions to reduce symptoms and aid recovery.

Disutilities for second-line therapy were captured within the utility value for the progression health state. The symptoms relating to this health state included anaemia, bleeding risk, infection risk, fatigue and shortness of breath, and treatment included weekly blood transfusions, and treatment with antibiotics. A disutility relating to treatment with HMA (azacitidine) was estimated by the company which was associated with a lower utility value than for the progressed health state, but was not used in the economic model.

Health state utility values and their application in the model are summarised in Table 24.

Table 24 Summary of health state utility values (CS Table 32, pg. 102)

Health state	Utility value: mean (SE)
Health state utility values	
AML (induction) [1]	0.550 (0.023)
Remission (post-induction/consolidation) [2]	0.656 (0.021)
[redacted] [3]	[redacted]
[redacted] [4]	[redacted]
Utility decrements	
3+7: Disutility of an induction cycle	[redacted]
3+7: Disutility of a consolidation cycle (inpatient)	[redacted]
3+7: Disutility of a consolidation cycle (outpatient)	[redacted]
CPX-351: Disutility of an induction cycle	[redacted]
CPX-351: Disutility of a consolidation cycle (inpatient)	[redacted]
CPX-351: Disutility of a consolidation cycle (outpatient)	[redacted]
Disutility of a transplant	[redacted]

SE, standard error; AML, acute myeloid leukaemia; BSC, best supportive care.
[1] Applied to those non-responders prior to progression or transplant. [2] Applied to patients in remission after consolidation therapy (who do not receive transplant). [3] Applied to those achieving remission after transplant (for responders and non-responders). [4] Applied to those experience relapse after transplant, after consolidation, or progression after a non-response.

Scenario analysis

The systematic literature review (SLR) conducted by the company did not identify any studies that were considered suitable for inclusion in the economic model as they did not provide a consistent set of estimates for stages in the AML treatment pathway. Of those studies that provided utility values estimated by EQ-5D, the preferred instrument according to the NICE reference case, none were considered by the company to be generalizable to the population of the decision problem.

A vignette-based TTO study conducted in members of the UK general population was considered in a scenario analysis [27]. In this scenario, utility decrements were equal for 3+7 and CPX-351. The study provided similar utility values for a number of health states, including progression and transplant, although the treatment-related and durable remission utilities were lower. The utility values are presented in Table 34 and 35 of the CS main submission pg. 105. The inclusion of these utility values results in fewer QALYs in both arms, but this has a relatively small impact on the results (Section 5.2.9.2).

ERG comment

The ERG was concerned about the generalisability of the utility values used in company model, specifically the use of TTO studies in both the base-case and scenario analyses. The NICE reference case [46] indicates that utilities should be directly elicited from patients, with a preference for the use of the EQ-5D generic instrument. The study used by the company in their base-case analysis and in their scenario analysis enrolled patients from the general population. In the base-case analysis, the sample had a mean age of 45.5 years which was lower than that of the typical AML patient in this decision problem, and no participants reported having AML. Analyses undertaken by the company explored the effect of age on utility, with older patients associated with higher utility scores for all health states except "Durable Remission", "Induction, CPX-351", and "Transplant" (CS appendix, Section N.3). This suggests that the utility values from the TTO study may not be fully generalizable to the decision problem population, which comprises patients older than those enrolled in the TTO study. However, a comparison of the utility values in the TTO study to those in other AML models submitted to NICE showed little variation across the majority of health states, with the exception of durable remission.

A key concern with the utility values generated by the TTO analysis performed by the company is the utility value applied to the post-transplant remission health state, which appears to be implausibly high. The utility value for these patients is similar, if not higher, than that of the general population. In one study undertaken in the UK on population norms for EQ-5D, the utility value was 0.79 for patients aged 65 to 74 (compared with [redacted] in the company model) [47]. Further, the value used in the company model is inconsistent with values reported in the literature. Leunis et al, which investigated quality of life in AML patients who had survived two years following transplant, reported a utility value of 0.81 [48]. This study also suggested that patients who survive following transplant experience lower quality of life than the general population. Given these uncertainties in the long-term utilities applied in the model the ERG implements scenario analysis in Section 6 exploring alternative utility values.

A further issue with the long-term utility values applied in the model relates to the fact that health state utility values were not adjusted for aging. While survival is generally low for the majority of patients, there are some (particularly in the CPX-351 arm) who may see a large OS benefit after

transplant. This OS benefit is a key driver of the model, and the high value for post-HSCT remission utility value coupled with the lack of an aging adjustment potentially overestimates HRQL in this health state. Given the greater proportion of CPX-351 patients who receive transplant after treatment response, any overestimation of utility in this health state would bias the model in favour of CPX-351. As such, this issue was explored further by the ERG in Section 6.

The ERG had some concerns that the treatment-related disutilities estimated from the company's TTO study did not capture the full impact of the safety profile associated with each treatment. Firstly, the vignettes for the health states for CPX-351 and 3+7 describe a side effect profile including fatigue, risk of infection and hair loss, the basis of which were a number of interviews with clinicians who had personal experience treating patients with the more established regimens, rather than any clinical evidence from the trial. It was felt that the impact of inpatient versus outpatient administration was reasonable to include, but that any differences relating to side effects were not fully substantiated. There did not appear to be any differences in the likelihood of Grade ≥3 fatigue in the CS (Table 8) or Grade 1-5 fatigue in the CSR; however CPX-351 was described as having less severe fatigue symptoms. While the rate of alopecia reported in the CSR was lower in the CPX-351 arm than in the 3+7 arm ([redacted] vs [redacted]), the ERG did not feel that this substantiated statements suggesting that CPX-351 patients experienced hair thinning while 3+7 patients experienced full hair loss. Furthermore, the clinical advisor to the ERG noted that there were some events associated with CPX-351 that may be associated with a poorer quality of life than those on 3+7, including the longer duration of count recovery. Quality of life during the on-treatment period was explored further by the ERG in Section 6. Secondly, the disutilities did not include the impact of the more serious adverse events included in the model, described in Section 5.2.6.5. Given the relatively low incident rate of AEs and similar profile between arms, the exclusion of these were felt to have minimal impact.

The ERG also noted that the results of the model lacked face validity in relation to the number of QALYs gained, suggesting an error in the implementation of utility values. The 3+7 arm was associated with a mean of [redacted] life years and [redacted] total QALYs, which suggests that the mean utility value for these patients was around [redacted]. However, none of the health states, including those that related to the treatment period, had an associated utility value of less than 0.5 (Table 31 in CS). The error was identified as relating to the application of the treatment-related disutility, which was estimated relative to the 'between-treatment' utility. However, the accrual of QALYs based on the health state was tracked from the first cycle of the model, and did not take into account the time-shift. For example, responder patients post-transplant gained QALYs in the durable remission health state (with an associated value of [redacted]) from the start of the model, which implies that CPX-351 patients on induction treatment would have a utility value of [redacted],

compared with [redacted], as estimated by the company. This error was amended by the ERG, described in Section 6.

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, monitoring, and costs associated with adverse events.

The company conducted a systematic review to identify published evidence regarding the resource use and costs associated with the management of patients with AML in the UK. The company found six studies relating to the UK setting. One study, Mahmoud 2012 [30], estimated total treatment costs of AML patients throughout the treatment pathway, and was used to provide the cost of post-consolidation monitoring for the model. Of the remaining studies, three were conducted prior to 2002, and were considered outdated [49 50 51], one was not considered to provide costs in the appropriate format [52] and one provided limited information on resource use (provision of antifungals only [53]). In addition to the Mahmoud study identified in the SLR, resource use estimates in the company's model health states were supplemented by a number of other sources [31 24].

5.2.8.1 Costs of first-line therapy

Drug acquisition

The drug acquisition cost for CPX-351 was based on the confidential list price. The cost per 100 unit vial of CPX-351 was [redacted]. CPX-351 has an associated confidential PAS, consisting of a simple [redacted] discount. Unit costs of the components of the comparator treatment, cytarabine and daunorubicin, were sourced from eMIT and BNF, respectively. The company based the calculations on a 1000mg vial for cytarabine and a 50mg vial for daunorubicin.

Dosing and vial usage

The dosing of first-line (induction and consolidation) therapies was based on Study 301, and reflects the expected marketing authorisation for CPX-351. As described in Section 5.2.4, the dosing for 3+7 reflects the standard US schedule. In the UK, 3+10 is more commonly used [6]. The company conducted a scenario analysis where the comparator was costed as the 3+10 regimen (Table 62 in CS). This resulted in a lower ICER, due to the increased costs in the comparator arm.

The dose of CPX-351, cytarabine and daunorubicin are based on body surface area (BSA). The model assumed a mean BSA of 1.79 m^2, estimated from a UK multicentre retrospective study of adult cancer patients in the UK [54]. In estimating the mean vial usage per dose, the company assumed that a whole

number of vials would be used for each patient, i.e. that some of the vial would be wasted. Table 25 summarises the unit costs and vial usage per course of treatment.

Table 25 Drug acquisition costs (adapted from Table 36, CS, pg. 109)

Treatment	Dosing regimen	Units per vial	Cost per vial	Number of vials required per cycle	Cost per cycle
First induction					
CPX-351	100 units/m^2 over 90 minutes on Days 1, 3, and 5	100 units	[redacted]	[redacted]	[redacted]
3+7 regimen	Cytarabine: 100 mg/m^2 on Days 1–7	1000 mg	£6.13	7	£43
	Daunorubicin: 60 mg/m^2 over 15 minutes on Days 1, 2, and 3	50 mg	£250	9	£2,250
Second induction					
CPX-351	100 units/m^2 over 90 minutes on Days 1 and 3	100 units	[redacted]	[redacted]	[redacted]
2+5 Regimen	Cytarabine: 100 mg/m^2 on Days 1–5	1000 mg	£6.13	5	£31
	Daunorubicin: 60 mg/m^2 over 15 minutes on Days 1 and 2	50 mg	£250	6	£1,500
Consolidation					
CPX-351	65 units/m^2 over 90 minutes on Days 1 and 3	100 units	[redacted]	[redacted]	[redacted]
2+5 Regimen	Cytarabine: 100 mg/m^2 on Days 1–5	1000 mg	£6.13	5	£31
	Daunorubicin: 60 mg/m^2 over 15 minutes on Days 1 and 2	50 mg	£250	6	£1,500

Administration

Administration costs comprised hospitalisations and chemotherapy delivery costs. CPX-351 and daunorubicin are administered as IV infusion, and cytarabine is administered as continuous infusion. Unit costs of drug administration were extracted from NHS Reference Costs (Table 40 in CS) [29], and were estimated as a weighted average of day case, outpatient and "other" activity. CPX-351 was assumed to be administered as a "Simple Parenteral Chemotherapy" procedure, cytarabine as a "Complex Chemotherapy, including Prolonged Infusion Treatment", given its requirement for continuous infusion, and daunorubicin as a "Subsequent Elements of a Chemotherapy Cycle".

The daily cost of hospitalisation was calculated using a weighted average of activity and length of stay across elective and non-elective AML patients, reflecting that patients would be non-elective at first induction and elective from this point onwards.

The CS stated that the length of hospitalisation for each treatment was derived from the CSR for Study 301, although it was not clear to the ERG how these values were derived, as the CSR did not provide hospitalisation by treatment phase. It was assumed that 50% of patients on consolidation therapy with CPX-351 would be administered as an outpatient, as observed in Study 301. No hospitalisation stay is required for consolidation therapy administered in an outpatient setting.

Table 26 Drug administration costs (adapted from Table 41, CS, pg. 119)

Treatment		Number of hospital days	Total hospitalisation cost	Number of administrations	Total delivery cost	Total admin cost
First induction therapy						
CPX-351	CPX-351	[redacted]	[redacted]	[redacted]	[redacted] [1]	[redacted]
3+7	Daunorubicin	28	£18,066	3	£913 [2]	£21,486
	Cytarabine			7	£2,492 [3]	
Second induction therapy						
CPX-351	CPX-351	[redacted]	[redacted]	[redacted]	[redacted] [1]	[redacted]
2+5	Daunorubicin	28	£18,066	2	£609 [2]	£20,453
	Cytarabine			5	£1,778 [3]	
Consolidation therapy						
CPX-351 (inpatient)	CPX-351	[redacted]	[redacted]	[redacted]	[redacted] [1]	[redacted]
CPX-351 (outpatient)	CPX-351	0	£0	[redacted]	[redacted] [1]	[redacted]
2+5 (inpatient)	Daunorubicin	30	£19,536	2	£609 [2]	£21,743
	Cytarabine			5	£1,778 [3]	
2+5 (outpatient)	Daunorubicin	0	£0	2	£609 [2]	£2,387*
	Cytarabine			5	£1,778 [3]	

[1] Deliver Simple Parenteral Chemotherapy at First Attendance
[2] Deliver Subsequent Elements of a Chemotherapy Cycle
[3] Deliver Complex Chemotherapy, including Prolonged Infusion Treatment, at First Attendance
* Error in CS, table presented corrected value

ERG comment

The ERG considers that the method used by the company to estimate vial usage does not provide an accurate estimate of the true level of vial usage. The calculations are based on the mean body surface area (BSA), and does not account for variation in BSA across the population. For example, based on a mean BSA of 1.79m^2, CPX-351 treatment would require two vials per dose; however, there will be a proportion of patients with a BSA higher than 2.0m^2 that would require three vials per dose. The ERG

preferred that the calculations reflected the distribution of BSA in the population, which can be used to estimate vial use through a "method of moments" [55]. The method assumes that weight is normally distributed, and estimates the proportion of patients requiring each dose. Given that drug acquisition costs are a key driver of the model, the ERG felt that it was important that they were represented accurately, and so this method was implemented in the model by the ERG, in Section 6.

The ERG was also concerned that BSA may have been underestimated in the model. The study that provided the estimate of BSA included patients with solid tumours, with a higher proportion of female patients than Study 301 (59.3% female in Sacco et al, compared with 39% in Study 301). Re-weighting the BSA data from Sacco *et al* using the gender proportions in Study 301 resulted in a mean BSA of 1.83m^2 [54]. The impact of using this BSA in the vial usage calculations is explored in Section 6.

The ERG was concerned that the length of hospitalisation was overestimated in the model. It was noted that hospital stay was higher for 3+7 patients in the consolidation period than in the induction period, despite the number of administrations being reduced. The CSR for Study 301 provided information on hospitalisation over the whole study period, reporting that, overall, there was little difference in hospital length of stay between arms (mean days [redacted] vs [redacted] for CPX-351 and 3+7 respectively, Table 14.4.1.4). Based on the model assumptions, the ERG estimated that CPX-351 was associated with a mean [redacted] hospital days during the induction and consolidation period, and 3+7 had mean [redacted] days, both of which are higher than that observed in Study 301. There is also an internal inconsistency with the health state vignettes for quality of life regarding the hospitalisation assumption (Section 5.2.7). The vignettes for consolidation therapy were based on the assumption that patients, regardless of treatment arm, would spend one week in hospital, with the remaining follow-up care taking place as an outpatient. The clinical advisor to the ERG also noted that the hospitalisation assumption for consolidation therapy was overestimated. The ERG explored the impact of a lower number of hospital days during the consolidation phase in Section 6.

The infusion costs could have been estimated more accurately by applying a day case and the outpatient infusion cost to those particular patients, rather than using a weighted cost overall throughout the treatment periods. The ERG concluded that since these were a relatively minor cost component, the issue was unlikely to impact on cost-effectiveness and did not explore further.

5.2.8.2 Costs of second-line therapy

Subsequent lines of therapy in the model consisted of non-intensive therapy, salvage therapy, and best-supportive care (no active treatment) (Section 5.2.4).

Drug acquisition

Unit costs for generic compounds were sourced from eMIT, with the remainder of drug acquisition costs being sourced from the British National Formulary. [28] Table 36 in the CS summarises the dosing regimen and costs associated with each cycle of treatment.

The mean cost per dose and cost per cycle were estimated in the same manner as that of first line therapies, as described in the section above. It was assumed that each cycle would be repeated in cycles of 4 weeks. The mean number of cycles per treatment was obtained from the CancerMPact 2015 report. [36]

A summary of drug cost and admin cost per cycle is provided in Table 27.

Table 27 Summary of drug and admin costs of second-line therapy

	Number of cycles	**Total drug costs**	**Total administration costs**	**Proportion of patients**
Non-intensive therapy				
Azacitidine	5.7	25,616	14,135	[redacted]
Low-dose cytarabine	4.4	539	27,666	[redacted]
Total weighted drug acquisition costs				£12,316
Total weighted drug administration costs				£21,312
Salvage therapy				
Intermediate-dose cytarabine + daunorubicin	3.6	8,365	11,697	[redacted]
Intermediate-dose cytarabine + idarubicin	3.2	3,747	15,739	[redacted]
Intermediate-dose cytarabine + mitoxantrone	4.0	1,192	18,867	[redacted]
Mitoxantrone + etoposide + cytarabine	4.3	1,619	22,761	[redacted]
FLAG-Ida	2.8	13,532	18,504	[redacted]
Total weighted drug acquisition costs				£6,882
Total weighted drug administration costs				£18,327

ERG comment

The ERG is satisfied with the approach taken by the company to estimate costs associated with second-line therapy. A minor limitation was noted with the use of the CancerMPact report which was used to estimate the mean number of cycles of each treatment [36]. The report covers a broad population of AML and is not specific to secondary AML, which is more aggressive than other types of AML and is characterised by resistance to therapy. As such, it may not be representative of the population in this analysis, but the ERG considered that this was unlikely to have material impact given the relatively balanced levels of second-line therapy in the two treatment groups.

5.2.8.3 Monitoring costs

On-treatment monitoring

Monitoring costs for patients on first-line and second-line therapy were estimated from the unit cost of the monitoring resource and healthcare resource use. Unit costs of monitoring were extracted from NHS Reference Costs (Table 43 in the CS).

Given an absence of UK- or European-specific guidelines on monitoring in AML, the company identified a US-specific guideline, the NCCN Guidelines for Acute Myeloid Leukemia [56]. These guidelines were used to inform monitoring requirements for patients on the 3+7 and 2+5 regimens for induction therapy and consolidation therapy, and for best-supportive care. For non-intensive therapy and salvage therapy, resource use was assumed to be equal to that of the 3+7 regimen. The company estimated a reduced monitoring schedule for CPX-351, based on the reduced administration. Further details were requested at the clarification stage, but no justification for this assumption was provided.

The monitoring requirements are summarised in Table 44 of the CS. These reflect the number of units of each resource for treatment cycle. Monitoring resources comprised a number of items, including blood count, platelets, chemistry panels, liver function tests, coagulation panels, bone marrow biopsy and blood transfusion. Resource use was higher in the induction phase than the consolidation phase. For non-intensive therapy, salvage therapy and best-supportive care, a total monitoring cost per therapy type was estimated using the mean number of treatment cycles. The total weighted cost, estimated with the total cost per cycle of each treatment type and the relative proportion of patients receiving each type of treatment, was applied to patients as they entered the relapse or progressed health state.

Monitoring of patients in remission

Monitoring of patients who achieve EFS in the post-consolidation phase was also captured. The company derived these costs from Mahmoud et al., 2012, [30] a study that was identified in the SLR of published cost and resource use studies. The cost primarily consisted of supportive care and

laboratory tests conducted during six cycles of four weeks at an outpatient clinic, and was estimated as being £4,311 (inflated from £4,097 in 2012). The weekly cost of £179 was applied to patients while they remained in the post-consolidation remission and the post-transplant remission health states.

Post-transplantation monitoring

The model also included follow-up costs for transplant recipients. A cost of £26,742 was extracted from UK Stem Cell Strategic Forum (2009) [31], and inflated to the current cost of £30,097. The cost reflects 6 months of follow-up resources, and was adjusted to reflect costs of those remaining alive after the procedure (90% alive at 6 months, in the source report). The cost was applied in the model as a one-off cost to patients who received transplant.

ERG comment

There were some minor concerns with the on-treatment monitoring requirements, which were informed by non-UK guidelines, and may not be fully generalizable to this population in this analysis. At the clarification stage, the company provided details of the validation process with two UK AML experts, who confirmed that these were generally consistent with UK practice, with the potential exception of bone marrow examinations. The reduction in blood count and chemistry panels for CPX-351 patients does not appear to be fully substantiated, but a reduction in monitoring due to fewer administrations does not appear to be unreasonable, and is unlikely to drive the model. The ERG considered that there may be some double counting of platelet counts in the model, which were also included in the complete blood count. However, these are low-cost items, and the overlap is unlikely to impact on cost-effectiveness. The ERG also considered that there were some monitoring items that were not included, such as blood transfusions during treatment, but the use of these appeared relatively well-balanced between arms (page 690, CSR).

The ERG is also satisfied with the application and estimation of post-consolidation monitoring costs. It noted that they were slightly higher than those applied in previous AML models submitted to NICE [32] but considered this not unreasonable given that these were higher risk patients. It did not appear that any monitoring costs were applied to non-responder patients after induction therapy until either transplantation or relapse.

The ERG noted some limitations with the study used to inform the cost of follow-up after transplantation. These costs were obtained from a costing study conducted in the Netherlands between 1994 and 1999 [57].The transplant process has changed substantially in the intervening period, and that inflating these costs to 2017 may not accurately reflect the current costs of treating patients after transplant. Additionally, this study provided costs for up to two years following transplant, reflecting the risk of more long-term sequelae of transplantation, such as chronic graft versus host disease

(GVHD). The company applied costs only up to six months, and the ERG considers that the costs following this period should be also included.

5.2.8.4 Transplant costs

The cost of transplantation comprised the identification of a compatible unrelated donor, outpatient visits prior to transplant, stem cell harvest and stem cell transplant. The components of a bone marrow allogeneic transplant and frequency of outpatient visits prior to transplant were derived from Wang et al. [24], and the unit cost of these components were sourced from the NHS National Schedule of Reference Costs 2016/17 [29]. As transplants in AML tend to be allogeneic rather than autologous, the cost of a transplant was extracted from NHS reference costs. The cost of providing unrelated adult stem cells was extracted from an NHS Blood and Transplant report [31], and included the cost of donor recruitment, sampling, typing and testing, registration, maintenance, and a registry search. The total cost of providing a transplant was estimated as being £64,235.

Table 28 Transplant costs (CS Table 42, pg. 121)

Transplant component	Unit cost	Frequency	Source (HRG code)	Total cost
Cost of providing unrelated adult stem cells for transplantation	£34,894	1	UK Stem Cell Strategic Forum	£34,894
Outpatient consulting visits before transplant	£302	3	NHS Reference Costs (308 - Blood and marrow transplantation - Total outpatient attendances) Wang *et al.*, 2011	£905
Stem cell harvest	£2,202	1	NHS Reference Costs (SA18Z)	£2,202
Stem cell transplant	£26,233	1	NHS Reference Costs (SA20A, Bone Marrow Transplant, Allogeneic Graft (Sibling), 19 years and over)	£26,233
Total				**£64,235**
HRG: healthcare resource group				

ERG comment

The ERG agrees that the transplant received by AML patients are more likely to be provided by a related donor, which is supported by data provided in the Wang study [24]. However, the report, which was used to identify cost components of transplant provision, found that the transplants in the study were all provided by sibling donor, and as such did not include provision of unrelated adult stem cells. The clinical expert consulted by the ERG confirmed that stem cells from a matched sibling would be preferred; however, this was not explicitly stated as a requirement for transplantation in the ELN 2017

guidelines. In the absence of any definitive guidelines for the treatment of AML patients in the UK, the ERG felt there was remaining uncertainty on the type of transplant that is typically provided, and whether the cost of providing unrelated adult stem cells should be included in the transplant cost. The ERG was also unclear on whether the service cost of providing unrelated stem cells should be included in the total transplant cost, even if there are some transplants using unrelated donor stem cells. Previous AML models submitted to NICE have not included this cost. The ERG understands that two of these registries in the UK are charities, and that their running costs may not fall to the NHS. It was also unclear whether the costs estimated by NHS Blood and Transplant would actually be incurred in practice.

After drug acquisition costs, the cost of transplantation is the largest component of the total cost of the treatment pathway (Table 28), and the cost of providing unrelated adult stem cells is a substantial proportion of this cost. Given the uncertainty discussed in this section, the ERG explored the impact of reducing this cost in Section 6.

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 (Section 5.2.6.5). The weighted cost of AEs was similar between CPX-351 and 3+7, although slightly higher for CPX-351, with CPX-351 associated with a weighted cost of £[redacted] and 3+7 associated with a weighted cost of £[redacted].

This cost was applied as a one-off cost in the model. No AEs were associated with subsequent treatment costs. Unit costs were extracted from NHS Reference Costs, and were a weighted average of elective inpatient, non-elective long-stay and non-elective short-stay [29].

Table 29 Summary of costs associated with adverse events (CS Table 47, pg. 126)

Adverse Event	Cost	Source (HRG Code)
Event		
Bacteremia	£1,895	WJ06
Diarrhoea	£1,354	FD10
Ejection Fraction Decreased	£1,837	EB03
Fatigue	£875	WH17
Febrile Neutropenia	£1,727	SA35
Hypertension	£593	EB04Z
Hypotension	£1,730	EB14
Hypoxia	£1,847	DZ27
Pneumonia	£1,698	DZ11

Respiratory Failure	£1,847	DZ27
Sepsis	£1,895	WJ06
Weighted average cost		
CPX-351	[redacted]	
3+7	[redacted]	
Source: NHS reference costs [29] HRG: health resource group, 3+7: daunorubicin + cytarabine		

ERG comment

The ERG is generally satisfied with the approach to implementing the AE-related costs for first-line therapy.

The ERG did not consider that the exclusion of costs for treating any AEs relating to second-line therapy would introduce bias into the analysis, since the use of second-line therapy appeared relatively balanced across arms.

5.2.9 Cost-effectiveness results

In this section, the results of cost-effectiveness analyses are presented for the deterministic base-case analysis, probabilistic sensitivity analysis, deterministic sensitivity analyses and scenario analyses.

Both CPX-351 and azacitidine, a treatment used as second-line therapy, have a confidential patient access scheme (PAS), comprising a simple discount. For CPX-351, this is [redacted].

The results in this section reflect the outcomes of the analysis *i)* when neither PAS was applied and *ii)* when the PAS for CPX-351 was applied. The confidential appendix presented the results including both the CPX-351 and azacitidine PAS.

5.2.9.1 Base-case incremental cost-effectiveness analysis results

Table 30 presents the results of the company base-case analysis. Costs and QALYs, using a 3.5% discount rate, were estimated over a lifetime time horizon. The company found CPX-351 to be more costly (cost difference of [redacted]), but also more effective (gains of [redacted] QALYs). The estimated deterministic ICER for CPX-351 compared with 3+7 was [redacted] per QALY. The results for the base-case after applying a PAS, lower the total costs for CPX-351 by [redacted], resulting in an ICER of £46,631 for CPX-351 versus 3+7.

Table 30 Results of the company base-case analysis (CS Table 51 and 52, pg. 133)

Technologies	Total costs	Total QALYs	Incremental costs	Incremental QALYs	ICER (incremental £/QALY)
Without CPX-351 PAS					
3+7	[redacted]	[redacted]			
CPX-351	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
With CPX-351 PAS					
3+7	[redacted]	[redacted]	-	-	-
CPX-351	[redacted]	[redacted]	[redacted]	[redacted]	£46,631
ICER, incremental cost-effectiveness ratio; QALYs, quality-adjusted life years					

Total costs (presented without the CPX-351 PAS applied) and QALYs broken down by health state are presented in Table 31 and

Table 32, respectively. The majority of costs were incurred in the induction and consolidation phase, due to the large hospitalisation costs and the drug acquisition cost of CPX-351. The largest QALY gains for CPX-351 were observed in the post-progression health state. This appeared to be a consequence of the similar post-transplant EFS for each arm but longer post-transplant OS for CPX-351, resulting in the CPX-351 patients remaining in the post-transplant relapse health state for longer.

Table 31 Total costs, by health state

Health state	**3+7**	**CPX-351**
Induction and consolidation phase	[redacted]	[redacted]
Drug acquisition costs	[redacted]	[redacted]
Drug administration costs	[redacted]	[redacted]
Treatment monitoring costs	[redacted]	[redacted]
Adverse event costs	[redacted]	[redacted]
Transplant	[redacted]	[redacted]
Remission (post-consolidation and post-transplant)	[redacted]	[redacted]
Non-Intensive/Salvage/BSC	[redacted]	[redacted]
Total costs	[redacted]	[redacted]
BSC, best-supportive care		

Table 32 Total QALYs, by health state

Health state	3+7	CPX-351
Event-free QALYs	[redacted]	[redacted]
Accrued prior to transplant	[redacted]	[redacted]
Post-consolidation remission	[redacted]	[redacted]
Post-transplant remission	[redacted]	[redacted]
Post-progression QALYs	[redacted]	[redacted]
QALY decrements	[redacted]	[redacted]
Attributable to induction and consolidation	[redacted]	[redacted]
Attributable to transplant	[redacted]	[redacted]
Total QALYs	[redacted]	[redacted]
QALY, quality-adjusted life year.		

5.2.9.2 Sensitivity analysis

The results presented in this section refer to that without the CPX-351 PAS applied, and can be directly compared with the base-case ICER of [redacted]. Results with PAS applied are provided in the confidential appendix to the ERG report.

Probabilistic sensitivity analysis

The company undertook a probabilistic sensitivity analysis (PSA) to explore and quantify uncertainty in the outcomes of the analysis. Probabilistic results were estimated from 500 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 [redacted] per QALY. The probabilistic results were relatively similar to those estimated in the deterministic base-case.

Table 33 Probabilistic sensitivity analysis results (CS Table 54, pg. 138)

Treatment	Total mean costs (£)	Total mean QALYs	Mean incremental costs (£)	Mean incremental QALYs	ICER incremental (£/QALY)
3+7	[redacted]	[redacted]	-	-	-
CPX-351	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
ICER, incremental cost-effectiveness ratio; QALYs, quality-adjusted life years					

The probability that CPX-351 is cost-effective compared with 3+7 was [redacted] at a threshold of £50,000 per QALY, while the probability was [redacted] at both a threshold of £20,000 and £30,000 per QALY (Figure 18).

Figure 18 [redacted] (CS Figure 17, pg. 139)

[redacted]

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 analysis.

Parameters included in the DSA were those relating to: probability of achieving post-induction response, time to post-induction response, post-consolidation EFS, post-HSCT EFS, post-consolidation OS, post-HSCT OS, alternative fit for 3+7 post-HSCT EFS, alternative fit for CPX-351 post-HSCT OS, cost of transplant procedure, medical resource use of treatment monitoring, adverse event costs, and health state utilities and utility decrements.

The company presented a tornado diagram depicting the results of the DSA (CS, Figure 19, pg. 143). The ERG identified an error in the DSA, where the company model contained an incorrect cell reference when applying the 3+7 coefficients to CPX-351 for post-HSCT OS. Figure 19 presents the tornado diagram with the ERG's correction depicting the results of the DSA. The ICER was most sensitive to use of the 3+7 post-transplant OS applied to CPX-351. The alternative fit for post-transplant OS for CPX-351 was the second most impactful driver of the model.

Figure 19 [redacted] (Error corrected) (CS Figure 19, pg. 143)

Scenario analyses

The company presented a range of scenario analyses within their base-case analysis. The results of the scenarios explored are presented in Table 34.

The results were notably most sensitive to variations in the time horizon. The ICER increases by [redacted] when a five year time horizon was assumed (ICER [redacted]) and [redacted] when ten years was assumed (ICER [redacted]). This highlights the difference in timings of the impact of treatment, where high-cost events (transplant, induction and consolidation therapy) were typically within one year of the model, and the QALY benefits occurred over the long-term.

Across the other set of scenarios explored, the ICER varied between a [redacted] decrease from the base-case ICER (all patients treated with CPX-351 receive second induction and consolidation therapy as outpatients, with an ICER of [redacted]), to an increase of [redacted] (adjustment of general population mortality for HSCT patients, with an ICER of [redacted]).

Table 34: Scenario analysis results (CS Table 60-68, pg. 143-145)

Scenario Analysis	Base-case	Incremental costs (£)	Incremental QALYs	ICER incremental (£/QALY)	Difference from Base-case ICERs
Base-case	-	[redacted]	[redacted]	[redacted]	-
Extension of model to younger population					
30% of modelled cohort <60 years of age	No extension to younger population	[redacted]	[redacted]	[redacted]	[redacted]
10% of modelled cohort <60 years of age	No extension to younger population	[redacted]	[redacted]	[redacted]	[redacted]
3+7 vs. 3+10 regimen					
3+10 regimen based on UK	3+7 Comparator	[redacted]	[redacted]	[redacted]	[redacted]
Percentage of patients treated with CPX-351 who receive second induction and consolidation therapy as outpatients					
Alter percentage of patients treated with CPX-351 who receive second induction and consolidation therapy as outpatients: in second induction, 50% of CPX-351 patients receive therapy in an outpatient setting and 100% of CPX-351 patients receive consolidation in an outpatient setting.	50% patients treated with CPX-351 who receive second induction and consolidation therapy as outpatients	[redacted]	[redacted]	[redacted]	[redacted]
Time horizon					
Five-year time horizon	Life time horizon (~30 year)	[redacted]	[redacted]	[redacted]	[redacted]
Ten-year time horizon	Life time horizon (~30 year)	[redacted]	[redacted]	[redacted]	[redacted]
Alternative source of utility scores					
Utilities sourced from Hensen *et al.*, 2017	Utilities sources from AML utility study	[redacted]	[redacted]	[redacted]	[redacted]
Mortality adjustment					
Adjust general population mortality for HSCT patients 2.3-fold	No general population mortality for HSCT patients	[redacted]	[redacted]	[redacted]	[redacted]
Transplant who do not response					
No transplantation in non-responders	Transplantation in non-responders	[redacted]	[redacted]	[redacted]	[redacted]
ICER, incremental cost-effectiveness ratio; LYs, life years; QALYs, quality-adjusted life years					

5.2.9.3 Conclusion

The company analyses show that CPX-351, with a deterministic ICER of [redacted] per QALY with no PAS applied, is not cost-effective at the £50,000 willingness to pay (WTP) threshold, which applies when EOL criteria is met (Section 7). The probabilistic analysis found that CPX-351 therapy has a [redacted] chance of being the cost-effective treatment at the £50,000 WTP threshold.

Additionally, the probabilistic and deterministic sensitivity analyses results and pre-defined scenario testing, demonstrate that there is significant uncertainty relating to post-transplant OS and alternative assumption to fit post-transplant OS. The ICER was most sensitive to use of the same OS for both treatments for post-transplant OS, five-year time horizon and alternative fit for post-transplant OS for CPX-351.

5.2.10 Model validation and face validity check

A comparison of modelled and trial-based overall survival in the overall patient 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. However, after approximately 80 weeks, a plateau effect occurred in the KM curve for CPX-351 likely owing to the small number of patients and number of events, which was not reflected in the modelled survival.

Figure 20 [redacted]

The company also provided details of the technical validation provided by senior health economists. No details were provided on whether the model was externally validated by clinical experts.

5.3 Conclusions of the cost-effectiveness section

The ERG considered the company's economic submission to meet the majority of 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 taken by the company introduced unnecessary complexity and required that the limited clinical data available be subdivided, meaning that parameter estimates were often based on small numbers of patients A simpler partitioned survival analysis approach, may have been preferable, particularly given the minimal benefits in terms of increased flexibility or accuracy from the state-transition modelling approach taken by the company. 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 Study 301 results.

2. *The modelling of non-responder patients*

A number of structural assumptions were imposed in the modelling of non-responders to treatment which potentially lack face validity. These included combining patients who receive transplant with non-transplant patients. The model did not capture relapses after transplant, and so overestimates QALYs in this group of patients. The ERG considered that it may have been more accurate to split patients into transplant and non-transplant, since response is not a requirement for transplant for these high risk patients.

3. *Heterogeneity within the population of high-risk (secondary) AML*

The ERG noted the heterogeneity in the magnitude of the treatment effect that was observed across the different subtypes of high-risk AML patients. Subgroup analyses conducted by the company suggest that the effectiveness, and by extension, cost-effectiveness, of CPX-351 in these different populations may differ substantially. Of note is the MDS with prior HMA subgroup, which is the largest patient group in the trial, and for which the reported hazard ratio for OS suggests a negligible treatment effect. However, given the company's modelling approach, it was not possible to explore the cost-effectiveness of CPX-351 within specific AML subtypes.

4. *Age of the patient population*

Study 301 excluded patients under the age of 60, who would be eligible for treatment with CPX-351. The clinical advisor to the ERG stated that these patients were a non-negligible part of the population,

potentially accounting for up to 25% of the eligible patient population. Exclusion of these patients may have important implications on the cost-effectiveness of CPX-351, as evidence suggests that they are more likely to achieve remission, tend to receive a different standard of care to patients over 60 and are more likely to receive transplant due to being fitter. The benefits of curative therapy such as transplant will also be greater in a younger population due to their greater life expectancy.

5. *Immature data for post-transplant event-free and overall survival*

There are a significant number of patients (approximately [redacted]) who were censored in the CPX-351 arm after one year in the post-transplant OS analysis. This makes extrapolation of the survival curves highly uncertain. Additional evidence provided by the company on deaths that have occurred since the 2015 data cut used for the survival analyses ([redacted] people in the CPX-351 arm, and [redacted] people in the 3+7 arm) suggests a degree of convergence in the survival curves. It is, however, difficult to draw strong conclusions from this additional evidence because the transplant status of these patients is not known.

The immaturity of the survival data resulted in a large degree of variation in the predictions of life expectancy across the different post-transplant OS survival models. The company model applied the curve associated with the most optimistic projections, which included a plateau-effect after approximately two years. However the ERG considers that the trial data is too immature to confirm such a plateau in the survival data, and that the long-term survival projections are subject to a degree of uncertainty. Since the parametric curve used to extrapolate post-transplant OS has a very significant impact on predicted OS gains and cost-effectiveness, this results in significant uncertainty in the cost-effectiveness of CPX-351.

Related to the above, the ERG also noted an additional issue with the way EFS data was captured, which means that many patients [redacted] were censored for event-free survival, because they received a transplant after their last examination in the study before the data cut-off time point. As a result, the EFS analysis in transplant patients was based on a very small patient number. The EFS results used in the model are, therefore, an unreliable indication of the true treatment effect, and the company advised that no clinical inferences should be drawn from estimates of post-transplant EFS.

6. *Clinical plausibility of the projected post-transplant event-free and overall survival*

While a statistically significant difference in OS was demonstrated for the whole population, the ERG does not consider that the clinical data supports the predicted sustained benefits beyond the trial period (as discussed above). In addition, the ERG does not believe that the long-term survival projections for CPX-351 are clinically plausible. When validating the long-term survival projections, the company stated that between one and two years after transplant is when the majority of deaths

would occur. The plateau in the CPX-351 post-transplant OS curve suggests that around [redacted] of patients are cured. Firstly, this is significantly higher than for 3+7 patients and what is suggested by external sources of data for other types of AML that are less severe and therefore would be expected to have a higher cure rate than patients in this decision problem. Secondly, the OS benefit for CPX-351 is inconsistent with the post-transplant EFS survival analysis, which showed no benefit between treatment arms, and predicted that the majority [redacted] of patients would experience an event within two years. While the company advised that this data was subject to limitations and should not be over-interpreted, the consequences of its use in the model implies that patients alive at 2 years are in the relapsed health state. Given the poor prognosis of these patients, it seems unlikely that they would experience a mortality rate similar to that of the general population.

7. Quality of life during the treatment period and after transplant

Quality of life in the base case was estimated from a TTO study conducted by the company. The NICE reference case states a preference for utilities directly elicited from patients. The company analysis of the TTO utilities found that age had an influence on quality of life. This suggested that the utility values in the study may not be fully generalizable to the decision problem population, which comprises patients older than those enrolled in the TTO study, who had a mean age of 45.5 years.

Further, the TTO vignettes for the on-treatment phases included a description of the side effect profile of each treatment, which were based on discussions with clinical experts. These vignettes described less severe symptoms described for CPX-351 compared with 3+7 patients, which resulted in a higher utility values being elicited for CPX-351 patients than for 3+7 patients. The ERG, however, considered that the evidence from Study 301 on the comparative safety profile of CPX-351 and 3+7 did not substantiate the differences described in the vignettes.

The ERG also considered that the utility for remission after transplant was implausibly high, as it was similar, if not higher, than that of the general population utility.

8. Hospitalisation during the treatment period

The ERG was concerned that the length of hospitalisation was overestimated in the model. Over the whole study period, there was little difference in hospital length of stay between arms, as reported in the CSR (approximately [redacted]. In comparison, the company model predicted higher values for each arm ([redacted] for CPX-351, and [redacted] for 3+7). The CS assumed that during a consolidation course, patients would be hospitalised for 30 days, which is internally inconsistent with the health state vignettes for quality of life regarding the hospitalisation assumption, where it was assumed that patients in consolidation would spend one week in hospital.

9. Uncertainties surrounding the costing assumptions for transplant

The ERG identified a number of uncertainties surrounding the costing assumptions for transplant. The source of the follow-up cost may not be generalizable to current practice, and the company only applied the cost of the first six months of follow-up. The ERG considered that the cost of transplant may have been over-estimated by the company, as it included the cost of provision of an unrelated stem cell donor, which accounted for around half of the total cost of the transplant procedure. The ERG was unclear on whether this cost would be relevant to this population, where many transplants use stem cells from a sibling donor. Further, it was unclear whether it should be included even if there are some transplants using unrelated donor stem cells.

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 5.2. This section is organised in five parts. Section 6.2 details the impact of corrections implemented by the ERG to rectify calculation errors in 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:

- Inclusion of patients under the age of 60;
- Post-transplant OS for CPX-351;
- Post-transplant EFS;
- HRQL:
    - Alternative value for post-transplant remission,
    - Utility values adjusted for aging,
    - Equivalent quality of life for CPX-351 and 3+7 patients while on induction and consolidation treatment.
- Cost and resource utilisation:
    - Vial usage reflecting the distribution of body surface area,
    - Reduced number of hospital days during the consolidation period,
    - Alternative costs of transplant.

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 analyses are presented in Section 6.5, exploring the impact of a number of specific assumptions 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. The results in this section are presented without the confidential PAS for CPX-351 and for azacitidine. Results with the application of both PAS are provided in the confidential appendix that accompanies the ERG's report.

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. The amendments made by the ERG are as follows:

- Treatment-related disutilities were removed, and the health state utility values for patients during the initial treatment period were applied instead, corresponding to the time on induction therapy and on consolidation therapy. The values applied were those estimated by the company, and presented in Table 24. The utility value for patients who had completed consolidation but not had transplant was also amended to reflect the "between treatment remission" utility.
- Post-consolidation monitoring was applied to responder transplant patients from the point of transplant. The model was amended so that this cost was applied at the point when consolidation was completed.
- The DSA was amended so that the scenario for post-transplant OS was correctly implemented, as described in Section 5.2.9.2.

Table 35 presents the results of the ERG corrections to the company model: the ICER increased by approximately [redacted] from [redacted] to [redacted] per QALY.

Table 35 Results of the ERG-corrected company base-case model

Technologies	Total costs	Total QALYs	Incremental costs	Incremental QALYs	ICER (incremental £/QALY)	Change in ICER (%)
Company base-case						
3+7	[redacted]	[redacted]	-	-	-	-
CPX-351	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	-
Company base-case (including ERG corrections)						
3+7	[redacted]	[redacted]	-	-	-	-
CPX-351	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
ERG, Evidence Review Group; ICER, incremental cost-effectiveness ratio; QALYs, quality-adjusted life years						

6.3 Additional ERG analyses

6.3.1 Inclusion of patients under the age of 60

As discussed in Section 5.2.3, the ERG noted that there are likely to be a number of patients who would be eligible for treatment with CPX-351 in this population who are under the age of 60, who were excluded from Study 301. The ERG explored the impact on the ICER including *(i)* 10% and *(ii)* 30% patients under 60 years old.

In this scenario, the ERG implemented an alternative odds ratio to the value applied by the company in the scenario analysis in the CS. As discussed in Section 5.2.3, the ERG considered that there were some limitations with the company's approach, namely that the study from which the value was extracted did not enrol patients under the age of 60, and that the study suggested a higher age-related effect on the rate of response than might be estimated from Study 301 analyses. In this scenario, the ERG applied the lower odds ratio of [redacted] for post-induction remission to the patients aged 60 to 69 to estimate the rate of remission in patients under 60, which was estimated by the company from Study 301 (Table 27 in CS).

The ERG also made some further adjustments to this scenario, based on advice from their clinical advisor. The ERG were advised that patients under the age of 60 receive a different standard of care to patients over 60, and would be more likely to receive an intermediate-dose cytarabine in the consolidation phase. The ERG, therefore, adjusted the drug acquisition costs applied to reflect these differences.

Results

The results of these scenario analyses are presented in Table 36. In each scenario, there was a small increase in total costs and QALYs in both arms, leading to a decrease in the ICERs. The increase in costs appeared to be attributable to the additional patients receiving transplant, as a result of the increased rate of remission leading to an absolute higher rate of transplantation in these patients. Given the lower odds of response relative to older patients that was applied in the company's scenario analysis (Table 34), the ICERs are higher in the ERG's scenarios.

Table 36 Scenario with the inclusion of patients under 60 years old

Technologies	Total costs	Total QALYs	Incremental costs	Incremental QALYs	ICER (incremental £/QALY)	Change in ICER (%)
Company base-case (including ERG corrections)						
3+7	[redacted]	[redacted]	-	-	-	-
CPX-351	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	-
Scenario: 10% of patients under 60 years						
3+7	[redacted]	[redacted]	-	-	-	-
CPX-351	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
Scenario: 30% of patients under 60 years						
3+7	[redacted]	[redacted]	-	-	-	
CPX-351	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]

ERG, Evidence Review Group; ICER, incremental cost-effectiveness ratio; QALYs, quality-adjusted life years

6.3.2 Post-transplant OS for CPX-351

As discussed in Section 5.2.6.3, the choice of parametric curve to extrapolate post-transplant OS for CPX-351 had a significant impact on predicted OS gains and cost-effectiveness, with mean OS ranging from [redacted] years with the exponential model, to approximately [redacted] years with the log-logistic model. As mentioned, the model fit statistics AICC and BIC, however, showed very little difference in the goodness of fit for the different survival models. Therefore, the impact on ICERs were explored using a series of survival distributions, including the Weibull, loglogistic, lognormal, exponential, generalised gamma and Gompertz (used in the company base-case). In addition, a common treatment effect for both arms was explored combining the CPX-351 and 3+7 post-transplant OS.

Within cost-effectiveness studies, it is common practice to use a single survival distribution in the base-case analysis, which is chosen based on goodness of fit statistics (AIC and BIC), the fit of each distribution to the Kaplan Meier curves, and the clinical plausibility of subsequent model projections. However, it is unlikely that a single survival distribution can adequately characterise uncertainties over the longer-term extrapolation period. The robustness of the ICER estimates to alternative distributions can be considered within separate sensitivity analyses or scenarios. However, transparency concerns may exist regarding this approach if their weighting is not explicitly specified in subsequent policy decisions.

Therefore, the ERG explores the uncertainty surrounding the choice of survival distribution adopting a model averaging approach using the methods outlined in Jackson et al. [58] and Hettle et al. [59]. This technique involves the parameterisation of uncertainty surrounding the choice of distribution, through including all plausible survival functions as part of a weighted distribution, and sampling both the parametric uncertainty associated within each distribution and the uncertainty (or weights) surrounding the choice of preferred method.

Using the model averaging approach, each model was assigned a weight that represents the adequacy of that distribution in predicting the lifetime survival of the modelled cohort, in comparison to all other distributions considered in the model. The weighted distribution was then applied in the base case analysis to estimate the impact. Figure 21 shows the predicted post-transplant OS using the weighted distribution against the Gompertz and KM curve. Using AICC weights, the predicted 1 year, 2 years and 5 years survival were [redacted], [redacted] and [redacted], respectively. Using BIC weights, the predicted 1 year, 2 years and 5 years survival were [redacted], [redacted] and [redacted], respectively.

Figure 21 [redacted]

In addition, the ERG implemented a scenario where post-transplant OS was pooled for both treatment arms, i.e. CPX-351 was not associated with a survival benefit after transplant compared with 3+7. In this scenario, survival was based on the Gompertz distribution, which predicted approximately [redacted] survival at 5 years for both 3+7 and CPX-351. The ERG acknowledges the limitations of this approach, as highlighted by the company in their response to clarification questions, namely that the application of the common OS curve post-transplant yields predicted survival that deviates from the observed clinical data. Nonetheless, given the limitations associated with the long-term predictions of the post-transplant OS, the ERG considered it important to include this conservative scenario.

Results

Table 37 presents the implications on the cost-effectiveness results of using different distributions for post-transplant OS. The results show that the QALYs in the CPX-351 arm are highly influenced by the choice of survival distribution. The QALY gained are much lower compared with the company's base-case, which leads to much higher estimates of the ICER. When post-transplant OS was implemented for each treatment arm, separately, the resulting ICERs vary between [redacted] and

[redacted] per QALY. When post-transplant OS was pooled for both arms, the number of QALYs increased for 3+7 and the QALYs decreased for CPX-351, and the ICER increased to [redacted] per QALY. The scenario analysis using AICC weights to pool post-transplant OS for CPX-351 was associated with a lower ICER than the analysis using BIC weights, as it implied a greater weight on the Gompertz and lognormal survival models, which were associated with higher survival estimates.

Table 37 ERG exploratory analyses of alternative model fits for post-transplant OS

Technologies	Total costs	Total QALYs	Incremental costs	Incremental QALYs	ICER (incremental £/QALY)	Change in ICER (%)
Company base-case (including ERG corrections) - Gompertz						
3+7	[redacted]	[redacted]	-	-	-	-
CPX-351	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	-
Scenario: Weibull for CPX-351 post-transplant OS						
3+7	[redacted]	[redacted]	-	-	-	-
CPX-351	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
Scenario: Log-logistic for CPX-351 post-transplant OS						
3+7	[redacted]	[redacted]	-	-	-	-
CPX-351	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
Scenario: Log-normal for CPX-351 post-transplant OS						
3+7	[redacted]	[redacted]	-	-	-	-
CPX-351	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
Scenario: Exponential for CPX-351 post-transplant OS						
3+7	[redacted]	[redacted]	-	-	-	-
CPX-351	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
Scenario: Generalised gamma for CPX-351 post-transplant OS						
3+7	[redacted]	[redacted]	-	-	-	
CPX-351	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
Scenario: Combining treatment arms for post-transplant OS (using Gompertz distribution)						
3+7	[redacted]	[redacted]	-	-	-	
CPX-351	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
Scenario: Weighted CPX-351 post-transplant OS curve (weighted average of Gompertz, Weibull, log-normal, log-logistic and exponential using AICC weights)						
3+7	[redacted]	[redacted]	-	-	-	-
CPX-351	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
Scenario: Weighted CPX-351 post-transplant OS curve (weighted average of Gompertz, Weibull, log-normal, log-logistic and exponential using BIC weights)						
3+7	[redacted]	[redacted]	-	-	-	-
CPX-351	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]

AICC, Akaike information criterion corrected; BIC, Bayesian information criterion; ERG, Evidence Review Group; ICER, incremental cost-effectiveness ratio; QALYs, quality-adjusted life years

6.3.3 Post-transplant EFS

Given the concerns of the ERG regarding event-free survival in post-transplant patients, the ERG considered a scenario where this data was excluded from the model. For the transplant responder patients, the ERG explored the use of a two-state model, where patients are either in remission or are dead (informed only by the OS analysis). The ERG considers the OS analysis to provide more reliable predictions of patients after transplant (although it is associated with its own limitations, such as a lack of face validity in the long-term, which were explored in Section 6.3.2). This scenario is also consistent with the assumptions for patients who received transplant but did not achieve response (Section 5.2.1). Given that the prognosis of these high-risk patients who experience a relapse after receiving transplant is poor [37], the ERG considers that they would only spend a short amount of time in the relapsed health state before death, and so the bias that this scenario introduces is likely to be small and is associated with an overestimation in the number of QALYs. The ERG acknowledges that this is a simplifying assumption, but considers that the removal of the bias associated with the inclusion of the post-transplant EFS analysis outweighs these limitations.

In this scenario analysis, the ICER reduced to [redacted] (Table 38). This reduction in the ICER is because the changes mean that fewer patients are in the post-transplant relapse health state, and the number of QALYs increased due to the higher utility value of remission patients. While this impacts on both model arms the effect is larger in the CPX-351 arm, where overall survival is longer and there were more patients implicated in this analysis.

Table 38 ERG exploratory analysis: Post-transplant outcomes based on OS only

Technologies	Total costs	Total QALYs	Incremental costs	Incremental QALYs	ICER (incremental £/QALY)	Change in ICER (%)
Company base-case (including ERG corrections)						
3+7	[redacted]	[redacted]	-	-	-	-
CPX-351	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	-
Scenario: Post-transplant outcomes based on OS only						
3+7	[redacted]	[redacted]	-	-	-	-
CPX-351	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
ERG, Evidence Review Group; ICER, incremental cost-effectiveness ratio; QALYs, quality-adjusted life years						

6.3.4 HRQL

Post-transplant remission utility

The utility applied by the company to the post-transplant remission health state was higher than those typically reported for the general population. The ERG did not consider this plausible and, therefore, explored two scenarios where post-transplant remission patients would have *(i)* lower quality of life than that of the general population, and *(ii)* quality of life equal to the general population. In the first scenario, the ERG applied a utility value of 0.75, which was the value applied in the company's scenario analysis for patients in post-transplant remission. This value is also consistent with studies examining the HRQL of AML post-transplant [48], and in other NICE submissions for AML [32]. The general population quality of life was derived from Szende (2014) [47], who reported the mean utility value for people aged 65 to 74 in England as 0.79. This is also the mean utility value for a 70 year old, as estimated by Ara & Brazier [60].

Utility age adjustment

The potential benefits of CPX-351 include that it allows a greater proportion of patients to receive transplant and therefore achieve long-term cure. The benefits of CPX-351 are therefore extended over a long period. Given this, the ERG believes it is appropriate to apply age adjusted utilities, to account for the natural reductions in HRQL with increased age. A scenario analysis was implemented, adjusting the utility value for durable remission in post-transplant patients for aging using the values from Ara & Brazier [60].

Treatment-related disutility

As described in Section 5.2.7, the treatment-related disutilities applied in the model were based on vignettes which described more favourable safety profiles for CPX-351 than for 3+7, resulting in smaller disutility being applied for CPX-351 than 3+7. The ERG felt that these differences in the safety profile of CPX-351 were not fully substantiated by clinical evidence, and a scenario was presented in which the mean utility was estimated by the ERG for each treatment phase, and applied to both the CPX-351 arm and the 3+7 arm.

Results

Results of these analyses are presented in Table 39. Each scenario was associated with lower QALYs due to the lower utility values; however, within the context of the company base-case for assumptions regarding post-transplant OS and EFS, the scenarios have only a small impact on the resulting ICERs, since the majority of patients gained QALYs in the relapsed state. The ERG considered that in a scenario where the durable remission utility value was varied while the outcome of patients after

transplant was purely modelled on the basis of being either in remission or dead (as implemented in Section 6.3.3), the impact of reducing the utility value in durable remission would be greater, as all of the QALY gains after transplant would be based on this value.

Table 39 ERG exploratory analysis with alternative HRQL values

Technology	Total costs	Total QALYs	Incremental costs	Incremental QALYs	ICER (incremental £/QALY)	Change in ICER (%)
Company base-case (with ERG corrections)						
3+7	[redacted]	[redacted]	-	-	-	-
CPX-351	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	-
Scenario: Utility value of 0.75 for durable remission						
3+7	[redacted]	[redacted]	-	-	-	-
CPX-351	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
Scenario: General population utility value (0.79) for durable remission						
3+7	[redacted]	[redacted]	-	-	-	-
CPX-351	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
Scenario: Post-transplant remission utility, adjusted for aging						
3+7	[redacted]	[redacted]	-	-	-	-
CPX-351	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
Scenario: Mean on-treatment utility during induction and consolidation treatment						
3+7	[redacted]	[redacted]	-	-	-	-
CPX-351	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]

ERG, Evidence Review Group; ICER, incremental cost-effectiveness ratio; QALYs, quality-adjusted life year;

6.3.5 Cost and resource utilisation

Vial usage

The ERG implemented an alternative method to estimate vial usage (the "method of moments" [55]) that took into account the distribution of body surface area in the population. This provided a more accurate estimate of vial usage (Section 5.2.8.1). This method was not applied to estimate usage of second-line therapy, since the usage was mostly balanced between arms, and the ERG felt that the increased level of precision would not provide sufficient benefit in the model. Additionally, the ERG re-estimated the mean BSA, using gender-specific body surface area from Sacco (2010) [54], re-weighted to reflect the gender distribution in Study 301. This increased the mean BSA from 1.79 m^2

to 1.83 m^2. As a result, the number of required vials per dose increased for CPX-351 during induction and decreased during consolidation (Table 40).

Table 40 Vial usage per dose during the induction and consolidation period

Treatment	Company analysis	ERG analysis*
CPX-351 induction (dosage: 100 units/m^2)	[redacted]	[redacted]
CPX-351 consolidation (dosage: 65 units/m^2)	[redacted]	[redacted]
Cytarabine (dosage: 100 mg/m^2)	1	1.00
Daunorubicin (50ml vial) (dosage: 60 mg/m^2)	3	2.82
Daunorubicin (20ml vial) (dosage: 60 mg/m^2)	6	6.00
Vial estimated assuming a BSA standard deviation of 0.18 ERG, Evidence Review Group, BSA; body surface area		

Hospitalisation

A scenario analysis was conducted in which the number of hospital days in the consolidation period was reduced to 7 days (from [redacted] days for CPX-351, and from 30 days for 3+7). As discussed in 5.2.8.1, the ERG considered that hospitalisation was overestimated in the model when compared with hospitalisation in the trial, and noted that the assumption that patients were in hospital for [redacted] days during the consolidation period was inconsistent with the assumption made for quality of life, where patients were in hospital for 7 days when receiving consolidation therapy.

Stem cell transplant

In this scenario analysis, two adjustments to the cost associated with transplant were made. Firstly, the ERG removed the cost of providing unrelated adult stem cells from the total transplant cost. The ERG considered that the majority of transplants would be from matched sibling donors, as reported in Wang (2010) [24], which is consistent with the unit cost applied by the company for the procedure. This reduced the cost of transplant from £64,235 to £29,340, which is similar to the cost applied for transplant in other recent NICE submissions in AML [20, 32]. Secondly, the 6-month follow-up cost applied by the company was increased to reflect the two-year cost, as the ERG felt that it was important to capture the impact of all possible long-term sequelae of transplant. This increased the follow-up cost from £30,097 to £44,447.

Results

Results of these analyses are presented in Table 41. The largest impact on the ICER was observed in the vial usage scenario, which resulted in the ICER increasing from [redacted] to [redacted] per QALY.

This is due to an increase in the number of required vials of CPX-351 per dose, increasing incremental costs. The relatively minor impact of the scenarios on hospitalisation and transplant cost is due to the usage of these items being relatively balanced between arms.

Table 41 Results of the ERG exploratory analyses on cost and resource use

	Total costs	Total QALYs	Incremental costs	Incremental QALYs	ICER (incremental £/QALY)	Change in ICER (%)
Company base-case (with ERG corrections)						
3+7	[redacted]	[redacted]	-	-	-	-
CPX-351	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	-
Scenario: Vial usage reflecting the distribution of body surface area						
3+7	[redacted]	[redacted]	-	-	-	
CPX-351	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
Scenario: Reduced number of hospital days during consolidation period						
3+7	[redacted]	[redacted]	-	-	-	
CPX-351	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
Scenario: Alternative cost of transplant						
3+7	[redacted]	[redacted]	-	-	-	
CPX-351	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
ERG, Evidence Review Group; ICER, incremental cost-effectiveness ratio; QALYs, quality-adjusted life year;						

6.4 ERG alternative base-case

Table 42 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:

- Post-transplant outcomes based on OS only,
- Post-transplant OS based on survival analysis weighted by goodness-of-fit (based on AICC weights),
- Adjustment of general population mortality for transplant patients (using the method described in Section 5.2.6.4, and explored in the company scenario analyses, Table 34),
- Utility estimate for patients in the post-transplant remission health state, further adjusted for age,

- Equivalent quality of life for CPX-351 and 3+7 patients while on induction and consolidation treatment,
- Vial usage reflecting the distribution of body surface area, and the mean body surface area re-weighted to reflect the gender distribution in Study 301,
- Reduced number of hospital days during the consolidation period,
- Provision of unrelated donor stem cells excluded from the costs of transplant.

Under the ERG's alternative set of assumptions, the ICER for CPX-351 versus 3+7 is [redacted] per QALY.

Table 42 Results of the ERG alternative base-case analysis

	Total costs	Total QALYs	Incremental costs	Incremental QALYs	ICER (incremental £/QALY)	Change in ICER (%)
Company base-case (with ERG corrections)						
3+7	[redacted]	[redacted]	-	-	-	-
CPX-351	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	-
ERG alternative base-case						
3+7	[redacted]	[redacted]	-	-	-	
CPX-351	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
ERG, Evidence Review Group; ICER, incremental cost-effectiveness ratio; QALYs, quality-adjusted life year;						

6.5 Exploratory analysis on the ERG alternative base-case

6.5.1 Post-transplant survival

In Section 6.3.2, the ERG explored the impact of implementing a range of survival models for post-transplant OS in CPX-351 patients. Due to the immaturity of the data for patients after transplant, there was a large degree of variation in the long-term survival estimates projected by each survival model. As demonstrated by the scenario analyses conducted by the ERG on the company base-case, these had consequences on the estimated cost-effectiveness of CPX-351.

In this section, the ERG explores the impact of the most favourable post-transplant OS and the least favourable post-transplant OS within their alternative base-case. The most favourable model was Gompertz, which estimated that [redacted] of CPX-351 patients would be alive at 5 years. The least favourable model was exponential, which estimated that [redacted] of CPX-351 patients would be alive at 5 years. In comparison, the 5-year survival for 3+7 was [redacted] in both survival models explored by the company.

Table 43 summarises the range of ICERs across the different survival models for CPX-351 post-transplant OS. The ICER of CPX-351 versus 3+7 varied between [redacted] (most favourable post-transplant OS) and [redacted] (least favourable post-transplant OS).

Table 43 Scenario analysis on the ERG base-case: alternative post-transplant OS survival models

	Total costs	Total QALYs	Incremental costs	Incremental QALYs	ICER (incremental £/QALY)	Change in ICER (%)
ERG alternative base-case						
3+7	[redacted]	[redacted]	-	-	-	-
CPX-351	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	-
Scenario: Most favourable post-transplant OS for CPX-351 (Gompertz)						
3+7	[redacted]	[redacted]	-	-	-	-
CPX-351	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
Scenario: Least favourable post-transplant OS for CPX-351 (exponential)						
3+7	[redacted]	[redacted]	-	-	-	-
CPX-351	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]

ERG, Evidence Review Group; ICER, incremental cost-effectiveness ratio; OS, overall survival; QALYs, quality-adjusted life year;

6.5.2 Inclusion of patients under the age of 60

The ERG previously explored the impact of including a proportion of patients under the age of 60. The inclusion of these patients was associated with a lower ICER: this was attributable to the increased rate of response, which was proportionally higher for CPX-351.

The ERG explored this assumption within the context of the alternative base-case. The ERG did not include this assumption in their alternative base-case due to the remaining uncertainty in the proportion of patients in practice that would be under the age of 60, and the relative impact of treatment on response rate, rate of transplant, and survival in younger patients.

In this analysis, the ICER for CPX-351 versus 3+7 was decreased to [redacted] when 10% of patients were under the age of 60 and to [redacted] when 30% of patients were under the age of 60 (Table 44).

Table 44 Scenario analysis on the ERG base-case: inclusion of patients under the age of 60

	Total costs	Total QALYs	Incremental costs	Incremental QALYs	ICER (incremental £/QALY)	Change in ICER (%)
ERG alternative base-case						
3+7	[redacted]	[redacted]	-	-	-	-
CPX-351	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	-
Scenario: 10% patients under the age of 60						
3+7	[redacted]	[redacted]	-	-	-	-
CPX-351	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
Scenario: 30% patients under the age of 60						
3+7	[redacted]	[redacted]	-	-	-	-
CPX-351	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
ERG, Evidence Review Group; ICER, incremental cost-effectiveness ratio; QALYs, quality-adjusted life year;						

6.6 Conclusions from ERG analyses

The ERG has presented a number of additional analyses 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 from [redacted] to [redacted] per QALY.

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

- Inclusion of patients under the age of 60,
- Post-transplant EFS,
- Post-transplant OS for CPX-351 patients,
- The quality of life in post-transplant remission,
- The quality of life while on induction and consolidation treatment,
- The cost of transplantation,
- Resource use in the treatment phases (vial usage, and hospitalisation during the consolidation period).

The scenarios associated with the greatest impact on cost-effectiveness outcomes related to changes made by the ERG to the post-transplant OS, post-transplant EFS, and to the number of vials required for treatment. All scenarios exploring the impact of alternative survival models for post-transplant OS resulted in an increase to the ICER. The majority of scenarios on resource use and quality of life were associated with an increase to the ICER, but within the context of the company base-case these were

insubstantial. 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 CPX-351 to be more costly (cost difference [redacted]) and more effective ([redacted] QALY gain) compared with 3+7, and suggests that the ICER for CPX-351 compared with 3+7 is [redacted] per QALY.

The final part of this section carried a further series of exploratory analyses that explored the impact of alternative survival models for post-transplant OS in CPX-351 patients within the ERG alternative base-case. The results of this analysis show the ICER is very sensitive to OS in this group of patients. This is partly due to the immaturity of the OS data from Study 301, which leads to considerable uncertainty around the extrapolation. The ICER of CPX-351 compared with 3+7 varied between [redacted] (most favourable post-transplant OS) and [redacted] (least favourable post-transplant OS). The ERG also explored the impact within their alternative base-case of expanding the patient population to include a proportion of patients under the age of 60 years. In this scenario, the ICER decreased to [redacted] when 10% of patients were under the age of 60 and to [redacted] when 30% of patients were under the age of 60.

7 End of life

The life expectancy of adult patients with high-risk AML is short. The CS states that in the Haematological Malignancies Research Network (HMRN) Yorkshire registry, outcomes for patients with t-AML and AML-MRC were very poor; median survival was around 3-4 months, with a 5-year OS for AML-MRC and t-AML of 2.8% and 2.4% respectively.[61] Data from a large Swedish registry reported median survival for patients with t-AML as 14 months in patients aged less than 55 years, 9 months for patients aged 55-74 and 8 months for patients aged 75 or above. Median survival for patients with AHD-AML was 7 months in the two younger age groups and 6 months for patients aged 75 or above.[4] This is generally consistent with the overall survival results from Study 301, which included patients aged 60-75 years of age.

The median overall survival in Study 301 was 9.56 months (95% CI: 6.60 to 11.86) in the CPX-351 treatment group and 5.95 months (95% CI: 4.99 to 7.75) in the DA 3+7 treatment group. Therefore, the median difference between CPX-351 and current standard NHS treatment was longer than 3 months.

The number of patients indicated for CPX-351 treatment in the UK is small. The incidence of AML in the UK is around 3100 cases per year.[3] High-risk AML accounts for around 25% of all AML diagnoses.[4, 5] CPX-351 was granted orphan drug status by the European Medicines Agency in January 2012.

8 Overall conclusions

8.1 Clinical effectiveness

Evidence from one phase 3 multi-centre randomised trial suggests that CPX-351 is associated with a significant improvement in OS compared with 3+7 (median OS: 9.56 months [95% CI: 6.60, 11.86] vs 5.95 months [95% CI: 4.99, 7.75]; HR=0.69 [95% CI: 0.52, 0.90], p=0.005) in patients with high-risk AML. Although results from the subgroup analyses should be interpreted with caution, there was some evidence to suggest that CPX-351 had a less beneficial impact on OS in patients with MDS with prior HMA [redacted] These patients constituted around a third of patients in the trial and a similar proportion of those who would be eligible for CPX-351 in clinical practice.

Overall CPX-351 may be a more effective bridge to stem cell transplant compared with 3+7 in patients with high-risk AML aged 60-75 years. Compared with 3+7, the proportion of patients undergoing HSCT was higher in the CPX-351 group (34.0% [52/153] vs 25.0% [39/156]) although the difference was not statistically significant (OR=1.54 [95% CI: 0.92, 2.56]). OS in patients who underwent HSCT was significantly greater with CPX-351: at the point of data cut the median OS was not reached in the CPX-351 group, and the median OS in the 3+7 group was 10.25 months (95% CI: 6.21, 16.69) (HR=0.46 [95% CI: 0.24, 0.89], p=0.0046). Overall the safety profiles of CPX-351 and 3+7 appear broadly comparable.

Analyses of Study 301 data presented in the CS and used in the model are based on a data cut dated December 2015 with a median follow up of 20.5 months in the CPX-351 group and 21.2 months in the 3+7 group. The ERG considers the length of follow-up insufficient for measuring long term post-HSCT OS. A substantial number of patients were censored in the CPX-351 arm and there were small numbers of patients in the tail of the survival curves. The company clarified that a number of deaths were known to have occurred since the 2015 data cut ([redacted] people in the CPX-351 arm, and [redacted] people in the 3+7 arm), which might suggest that updated analyses will reveal a degree of convergence in the OS curves. Data on relapse after HSCT is very limited. Study 301 did not collect HRQL or utility data.

The anticipated marketing authorisation for CPX-351 is for the treatment of adults with newly diagnosed therapy-related AML (t-AML) or AML with myelodysplasia-related changes (AML-MRC). Therefore, the age range of patients in the trial (60 – 75 years) is narrower than the anticipated marketing authorisation. The majority of patients with high-risk AML in clinical practice are over the age of 60 years and patients older than 75 years would be less likely to withstand intensive chemotherapy, therefore, the population of the trial is likely to be reflective of the majority of patients

eligible for intensive chemotherapy for high-risk AML in clinical practice. However, results of the trial may not be generalisable to patients under the age of 60.

Patients with *de novo* AML with MDS associated karyotypic changes are difficult to confidently define until genetic test results are available. [redacted] Therefore, this is not reflective of clinical practice, where treatment may commence prior to cytogenetic test results becoming available.

8.2 Cost-effectiveness

The economic evidence presented by the company primarily consisted of a *de novo* model. The company's model used a cohort state-transition approach which directly used the time-to-event data from Study 301 to determine the patient transitions between the health states.
W[redacted]he company found CPX-351 to be more costly (cost difference of £[redacted]) and more effective ([redacted] QALY gain) compared with 3+7. The deterministic base-case ICER was [redacted] and the mean probabilistic ICER was [redacted] 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. The ERG's key concerns related to the long-term survival predictions after transplant. The ERG carried 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.
W[redacted]he ERG's alternative base-case analysis estimated CPX-351 to be more costly ([redacted]) and more effective ([redacted] QALY gain) compared with 3+7 alone, and suggests that the ICER for CPX-351 compared with 3+7 is [redacted] per QALY.

8.3 Implications for research

Follow-up of Study 301 is continuing for 5 years post-randomisation, therefore, longer term data are anticipated. In addition, there are two ongoing studies of CPX-351 as first-line treatment for AML patients: AML 18 (which has been extended to include CPX-351) and AML 19.

There is no evidence for the effectiveness and safety of CPX-351 in patients with high-risk AML aged less than 60 years or over 75 years. Data on relapse after HSCT is very limited and no data on HRQL are available.

9 References

1. Kantarjian H. Acute myeloid leukemia - major progress over four decades and glimpses into the future. *Am J Hematol* 2016;**91**:131-45.

2. Andreeff M, Quintas-Cardama A. *Pathobiology of Acute Myeloid Leukemia in Hematology: Basic Principles and Practice* 13th ed; 2013.

3. Cancer Research UK. *Cancer statistics: Acute myeloid leukaemia*. 2016. [accessed 2nd July 2018].

4. 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.

5. Ostgard LSG, Medeiros BC, Sengelov H, Norgaard M, Andersen MK, Dufva IH, et al. Epidemiology and clinical significance of secondary and therapy-related acute myeloid leukemia: a national population-based cohort study. *J Clin Oncol* 2015;**33**:3641-9.

6. Milligan DW, Grimwade D, Cullis JO, Bond L, Swirsky D, Craddock C, et al. British Committee for Standards in Haematology: guidelines on the management of acute myeloid leukaemia in adults. *Br J Haematol* 2006;**135**:450-74.

7. Döhner H, Estey E, Grimwade D, Amadori S, Appelbaum FR, Büchner T, et al. Diagnosis and management of AML in adults: 2017 ELN recommendations from an international expert panel. *Blood* 2017;**129**:424-47.

8. National Institute for Health and Care Excellence (NICE). *NICE Pathways: Azacitidine for people with acute myeloid leukaemia who are ineligible for stem cell transplants*. NICE; 2018 Available from: https://pathways.nice.org.uk/pathways/blood-and-bone-marrow-cancers/leukaemia#path=view%3A/pathways/blood-and-bone-marrow-cancers/myeloid-leukaemia.xml&content=view-node%3Anodes-azacitidine-for-people-with-acute-myeloid-leukaemia-who-are-ineligible-for-stem-cell-transplants [accessed 2nd July 2018].

9. Arber DA, Orazi A, Hasserjian R, Thiele J, Borowitz MJ, Le Beau MM, et al. The 2016 revision to the World Health Organization classification of myeloid neoplasms and acute leukemia. *Blood* 2016;**127**:2391-405.

10. 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.

11. Lancet JE, Cortes JE, Hogge DE, Tallman MS, Kovacsovics TJ, Damon LE, et al. Phase 2 trial of CPX-351, a fixed 5:1 molar ratio of cytarabine/ daunorubicin, vs cytarabine/daunorubicin in older adults with untreated AML. *Blood* 2014;**123**:3239-46.

12. Lancet J, Cortes J, Kovacsovics T, Hogge D, Kolitz J, Tallman M, et al. The impact of prior hypomethylating agent treatment among secondary AML patients treated with CPX-351 or 7+3 chemotherapy. *Haematologica* 2011;**96**:27.

13. Lancet JE, Cortes JE, Kovacsovics T, Hogge D, Kolitz JE, Tallman MS, et al. CPX-351 versus cytarabine (CYT) and daunorubicin (DNR) therapy in newly diagnosed AML patients age 60-75: safety and efficacy in secondary AML (sAML). *J Clin Oncol* 2011;**29**:6519.

14. Lancet JE, Uy GL, Cortes JE, Newell LF, Lin TL, Ritchie EK, et al. Final results of a phase III randomized trial of CPX-351 versus 7+3 in older patients with newly diagnosed high risk (secondary) AML. *J Clin Oncol* 2016;**34**:7000.

15. Lancet JE, Rizzieri D, Schiller GJ, Stuart RK, Kolitz JE, Solomon SR, et al. Overall survival (OS) with CPX-351 versus 7+3 in older adults with newly diagnosed, therapy-related acute myeloid leukemia (tAML): Subgroup analysis of a phase III study. *J Clin Oncol* 2017;**35**:7035.

16. Lancet JE, Rizzieri D, Schiller GJ, Stuart RK, Kolitz JE, Solomon SR, et al. Overall survival with CPX-351 versus 7+3 in older adults with newly diagnosed, therapy-related acute myeloid leukemia: subgroup analysis of a phase 3 study. *Haematologica* 2017;**102**:214-5.

17. Medeiros BC, Lancet JE, Cortes JE, Newell LF, Lin TL, Ritchie EK, et al. Analysis of efficacy by age for patients aged 60-75 with untreated secondary acute myeloid leukemia (AML) treated with CPX-351 liposome injection versus conventional cytarabine and daunorubicin in a phase III trial. *Blood* 2016;**128**:902.

18. Lowenberg B, Ossenkoppele GJ, van Putten W, Schouten HC, Graux C, Ferrant A, et al. High-dose daunorubicin in older patients with acute myeloid leukemia. *N Engl J Med* 2009;**361**:1235-48.

19. National Institute for Health and Care Excellence (NICE). *Midostaurin for untreated acute myeloid leukaemia. In development [GID-TA10124]*. 2018. Available from: https://www.nice.org.uk/guidance/indevelopment/gid-ta10124 [accessed 20th April 2018].

20. National Institute for Health and Care Excellence (NICE). *Midostaurin for untreated acute myeloid leukaemia [ID894]*. NICE; 2018. Available from: https://www.nice.org.uk/guidance/indevelopment/gid-ta10124 [accessed 20th April 2018].

21. National Institute for Health and Care Excellence (NICE). *Azacitidine for the treatment of myelodysplastic syndromes, chronic myelomonocytic leukaemia and acute myeloid leukaemia [TA218]*. NICE; 2011. Available from: https://www.nice.org.uk/guidance/ta218 [accessed 20th April 2018].

22. Tremblay G, Dolph M, Patel S, Brandt P, Forsythe A. Cost-effectiveness analysis of midostaurin (MIDO) with standard chemotherapy (SOC) for acute myeloid leukemia (AML) in the united kingdom (UK). *Value Health* 2017;**20**:A399.

23. Tikhonova IA, Hoyle MW, Snowsill TM, Cooper C, Varley-Campbell JL, Rudin CE, 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.

24. Wang HIA, E.; Howell, D.; Roman, E.; Patmore, R.; Jack, A.; Smith, A. Long-term medical costs and life expectancy of acute myeloid leukemia: a probabilistic decision model. *Value Health* 2014;**17**:205-14.

25. Martin PJ, Counts GW, Jr., Appelbaum FR, Lee SJ, Sanders JE, Deeg HJ, et al. Life expectancy in patients surviving more than 5 years after hematopoietic cell transplantation. *J Clin Oncol* 2010;**28**:1011-6.

26. Matza LS, Deger K, Howell T, Hillgruber NK, Yeager AM, Hogge D, et al. Health state utilities associated with treatment options for acute myeloid leukemia (AML). *Value Health* 2017;**20**:A448.

27. Hensen M, Joshi N, Xu W, Patel S, Bal V, Lasch K. Assessment of utility values for treatment-related health states of acute myeloid leukemia in the United Kingdom. *Value Health* 2017;**20**:A115.

28. *Electronic market information tool (eMIT). eMIT national database*. Department of Health and Social Care; 2018. Available from: https://www.gov.uk/government/publications/drugs-and-pharmaceutical-electronic-market-information-emit [accessed 20th April 2018].

29. NHS Improvement. *National schedule of reference costs 1*. NHS Improvement; 2017. Available from: https://improvement.nhs.uk/resources/reference-costs/ [accessed 20th April 2018].

30. Mahmoud D, Skikne BS, Kucmin-Bemelmans I, Alleman C, Hensen M. Overall Economic Burden of Total Treatment Costs in Acute Myeloid Leukemia throughout the Course of the Disease. *Blood* 2012;**120**:3614-.

31. NHS Blood and Transplant (NHSBT). *The future of unrelated donor stem cell transplantation in the UK. Part 2: aAnnexes. A report from the UK Stem Cell Strategic Forum.* Watford: NHSBT; 2010. Available from: www.nhsbt.nhs.uk

32. National Institute for Health and Care Excellence (NICE). *Single Technology Appraisal. Gemtuzumab ozogamicin for untreated acute myeloid leukaemia [ID982]. Committee Papers*. London: NICE; 2018. Available from: https://www.nice.org.uk/guidance/gid-ta10142/documents/committee-papers

33. Woods B, Sideris E, Palmer S, Latimer N, Soares M. *NICE DSU Technical Support Document 19. Partitioned survival analysis for decision modelling in health care: a critical review*. Sheffield: Decision Support Unit; 2017. Available from: http://nicedsu.org.uk/wp-content/uploads/2017/06/Partitioned-Survival-Analysis-final-report.pdf

34. Dumas PY, Bertoli S, Berard E, Mediavilla C, Yon E, Tavitian S, et al. Azacitidine or intensive chemotherapy for older patients with secondary or therapy-related acute myeloid leukemia. *Oncotarget* 2017;**8**:79126-36. Available from: https://www.ncbi.nlm.nih.gov/pubmed/29108292

35. Jazz Pharmaceuticals Inc. *Clinical Study Report: Study CLTR0310-301. Phase III, Multicenter, Randomized, Trial of CPX-351 (Cytarabine: Daunorubicin) Liposome Injection versus Cytarabine and Daunorubicin in Patients 60 -75 years of Age with Untreated High Risk (Secondary) AML*; 2017 2/24/2017. Report No.: V1.0.

36. Kantar Health. *Treatment Architecture: Western Europe Leukemia, Acute Myeloid. CancerMPact® Western Europe*; 2016. Available from: http://www.kantarhealth.com/docs/datasheets/Kantar_Health_CancerMPact_datasheet%20.pdf?sfvrsn=68

37. Bejanyan N, Weisdorf DJ, Logan BR, Wang H-L, Devine SM, de Lima M, et al. Survival of Patients with Acute Myeloid Leukemia Relapsing after Allogeneic Hematopoietic Cell Transplantation: A Center for International Blood and Marrow Transplant Research Study. *Biol Blood Marrow Transplant* 2015;**21**:454-9. Available from: http://www.sciencedirect.com/science/article/pii/S108387911400679X

https://ac.els-cdn.com/S108387911400679X/1-s2.0-S108387911400679X-main.pdf?_tid=41cd6a98-48ed-4ab7-a822-da0030042b93&acdnat=1530310891_a877e91598a6c782650c6c2d42522678

38. Lancet JE, Uy G, Cortes JE, Newell LF, Lin TL, al. e. Final results of a phase III randomized trial of VYXEOS™ (CPX-351) vs 7 + 3 in older patients with newly diagnosed high-risk (secondary) AML [abstract/oral presentation], ASCO Annual Meeting (June 2016). *J Clin Oncol* 2016;**34**:7000.

39. Shimoni A, Labopin M, Savani B, Volin L, Ehninger G, Kuball J, 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. Available from: https://www.ncbi.nlm.nih.gov/pubmed/27821187

40. Pohlen M, Groth C, Sauer T, Gorlich D, Mesters R, Schliemann C, et al. Outcome of allogeneic stem cell transplantation for AML and myelodysplastic syndrome in elderly patients (60 years). *Bone Marrow Transplant* 2016;**51**:1441-8.

41. Office of National Statistics (ONS). *National life tables, UK: 2014 to 2016*. 2017. Available from: https://www.ons.gov.uk/peoplepopulationandcommunity/birthsdeathsandmarriages/lifeexpectancies/bulletins/nationallifetablesunitedkingdom/2014to2016 [accessed

42. Battiwalla M, Hashmi S, Majhail N, Pavletic S, Savani BN, Shelburne N. National Institutes of Health hematopoietic cell transplantation late effects initiative: developing recommendations to improve survivorship and long-term outcomes. *Biol Blood Marrow Transplant* 2017;**23**:6-9.

43. Socié G, Stone JV, Wingard JR, Weisdorf D, Henslee-Downey PJ, Bredeson C, et al. Long-term survival and late deaths after allogeneic bone marrow transplantation. *N Engl J Med* 1999;**341**:14-21.

44. Wingard JR, Majhail NS, Brazauskas R, Wang Z, Sobocinski KA, Jacobsohn D, et al. Long-term survival and late deaths after allogeneic hematopoietic cell transplantation. *J Clin Oncol* 2011;**29**:2230-9.

45. Medeiros BC, Lancet JE, Cortes JE, Newell LF, Lin TL, Ritchie EK, et al. *Rates of adverse events per patient-year in a randomized, phase 3 study of CPX-351 versus 7+ 3 in older adults with newly diagnosed, treatment-related Acute Myeloid Leukemia (AML) or AML with myelodysplasia-related changes*. In: Am Soc Hematology; 2017.

46. National Institute for Health and Care Excellence (NICE). *Guide to the methods of technology appraisal*. NICE; 2013. [accessed

47. Szende A, Janssen B, Cabases J. *Self-Reported Population Health: An International Perspective based on EQ-5D*. London: Springer; 2014.

48. 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.

49. Kay HEP, R. L.; Lawler, S. D.; Clink, H. M. Cost of bone-marrow transplants in acute myeloid leukaemia. *Lancet* 1980;**1**:1067-9.

50. Lobo PJP, R. L.; Hanrahan, A.; Reynold, D. K. Acute myeloblastic leukaemia--a model for assessing value for money for new treatment programmes. *BMJ* 1991;**302**:323-6.

51. Standaert BG, J.; Lu, Z. J.; Erder, M. H.; Yin, J. L. Economic analysis of filgrastim use for patients with acute myeloid leukaemia in the UK: a comparison of collection methods of resource use data. *Pharmacoeconomics* 2002;**20**:665-74.

52. Smith BK-B, I.; Alleman, C.; Hensen, M.; Skikne, B.; Mahmoud, D. Estimates of burden of disease associated with management of acute myeloid leukemia in UK and US. *Haematologica* 2013;**98**:211.

53. Agrawal SG, Dexter TL. Cost of antifungal drugs in different haemato-oncology settings. *Mycoses* 2015;**58**:157-8.

54. Sacco JJ, Botten J, Macbeth F, Bagust A, Clark P. The Average Body Surface Area of Adult Cancer Patients in the UK: A Multicentre Retrospective Study. *PLOS ONE* 2010;**5**:e8933.

55. Hatswell AJ, Porter J, Lee D, Hertel N, Latimer NR. The cost of costing treatments incorrectly: Errors in the application of drug prices in economic evaluation due to failing to account for the distribution of patient weight. *Value Health* 2016;**19**:1055-8.

56. O'Donnell MR, Tallman MS, Abboud CN, Altman JK, Appelbaum FR, Arber DA, et al. Acute Myeloid Leukemia, Version 3.2017, NCCN Clinical Practice Guidelines in Oncology. *Journal of the National Comprehensive Cancer Network* 2017;**15**:926-57.

57. van Agthoven M, Groot MT, Verdonck LF, Lowenberg B, Schattenberg AV, 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.

58. Jackson CH, Thompson SG, Sharples LD. Accounting for uncertainty in health economic decision models by using model averaging. *J R Stat Soc Ser A Stat Soc* 2009;**172**:383-404.

59. Hettle R, Corbett M, Hinde S, Hodgson R, Jones-Diette J, Woolacott N, et al. The assessment and appraisal of regenerative medicines and cell therapy products: an exploration of methods for review, economic evaluation and appraisal. *Health Technol Assess. 2017 Feb;21(7):1-204. doi: 10.3310/hta21070.* 2017.

60. Ara R, Brazier JE. Using health state utility values from the general population to approximate baselines in decision analytic models when condition-specific data are not available. *Value Health* 2011;**14**:539-45.

61. Roman E, Smith A, Appleton S, Crouch S, Kelly R, Kinsey S, et al. Myeloid malignancies in the real-world: occurrence, progression and survival in the UK's population-based Haematological Malignancy Research Network 2004-15. *Cancer Epidemiol* 2016;**42**:186-98.

10 Appendices

10.1 Treatment effectiveness and extrapolation

The CS provided a description of the clinical data used in the model and analyses conducted to populate the model (CS appendix M pg. 215-470). All analyses conducted were based on the ITT populations, and post-hoc statistical analyses used individual patient-level data from Study 301 (CLTR0310-301, NCT01696084). [15-17, 38]

The analyses conducted to populate the model estimated the following outcomes:

- Treatment pathway probabilities;
    - Probability of receiving a second round of induction therapy,
    - Probability of remission (CR + CRi),
    - Probability of receiving 0 or 1 or 2 rounds of consolidation therapy,
    - Probability of receiving transplant after remission.
- Mean time elapsed after induction therapy (time-shift);
    - Time to post-induction response,
    - Time between post-induction response and end of consolidation,
    - Time between end of consolidation and transplantation.
- Survival analysis;
    - Overall survival,
    - Event free survival,
    - Time to HSCT or death in those who did not achieve remission.
- Rates of adverse events in each arm.

All analyses except adverse event rates were adjusted for the sampling stratification variables, including age (60–69 vs. 70–75 years old) and high-risk AML subtypes (referred as AML types in this section). All analyses included one or more co-variate adjustment selecting from treatment arms, rounds of induction therapy and rounds of consolidation therapy, but the inclusion of co-variates for adjustment were not consistent in all analyses.

10.1.1 Treatment pathway probabilities

As described in the model structure section (Section 5.2.1), the initial phase of the model allocated patients to different treatment pathways to determine long-term survival. Multivariate logistic regression analyses were performed to estimate the percentage of the following clinical pathway outcomes:

- Probability of receiving a second round of induction therapy,
- Probability of remission (CR + CRi),
- Probability of receiving 0, 1 or 2 rounds of consolidation therapy,
- Probability of receiving transplant after remission.

A brief description of the technical details are presented in the Table 45 and results of the multivariate analyses are presented in Table 46. Further details are provided in CS Appendix M pg. 215-470.

Table 45 Description of the technical details used to estimate probability of Induction, Remission, Consolidation, and Transplant receiving (adapted from Table-35 in CS Appendix M)

Description	Technical Details	Co-variates adjustment	Purpose
Percentage of patients who receive a second round of induction therapy	Multivariate logistic regression ITT population	Adjusting for age (60–69 vs. 70–75 years old) and five AML types	Required for simulating who have 1 or 2 rounds of induction therapy, cost and treatment disutility estimates
Percentage of patients who achieve remission post-induction	Multivariate logistic regression, Considering both CR and CR_i as remission Evaluating remission at "Date of Post-induction Response"	Adjusting for age (60–69 vs. 70–75 years old), five AML types, treatment, and number of inductions	Required for simulating who achieve remission, and determining the survival of patients who achieve remission post-induction
Percentage of patients who receive 0, 1, or 2 rounds of consolidation among those who achieve remission	Multivariate multinomial logistic regression	Adjusting for age (60–69 vs. 70–75 years old), AML types combined in two categories, and number of inductions	Required for simulating number of consolidation rounds, and cost and treatment disutility estimates
Percentage of patients who receive HSCT among those who achieve remission	Multivariate logistic regression	Adjusting for age (60–69 vs. 70–75 years old), AML types combined in two categories, and number of inductions	Required for simulating who receive HSCT among those who achieve remission and survival among those who receive HSCT among those who achieve remission
CR, complete response; CRi, complete response with incomplete platelet or neutrophil recovery; HSCT, hematopoietic stem cell transplant; ITT, intent-to-treat; OS, overall survival;			

Table 46 Treatment pathway: Odds ratios from of multivariate analyses (adapted from Tables in CS Appendix M)

Parameter	Second Round of Induction Therapy	Remission Post-Induction	1 (versus 0) round of consolidation therapy	2 (versus 0) rounds of consolidation therapy	HSCT, Given Remission
	Odds Ratio [95% CI]	Odds Ratio [95% CI]	Odds Ratio [95% CI]	Odds Ratio [95% CI]	Odds Ratio [95% CI]
Age					
60-69 years	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
70-75 years	(reference)	(reference)	(reference)	(reference)	(reference)
AML Type					
t-AML	[redacted]$	[redacted]	(reference)	(reference)	(reference)
MDS without prior HMA exposure	[redacted]$	[redacted]	[redacted]	[redacted]	[redacted]
MDS with prior HMA exposure	(reference)	(reference)			
CMMoL	[redacted]$	[redacted]	(combined with reference)	(combined with reference)	(combined with reference)
de novo AML with karyotype characteristic of MDS	[redacted]$	[redacted]	(combined with reference)	(combined with reference)	(combined with reference)
Induction therapy rounds (ref: 2 rounds)					
1 round		[redacted]	[redacted]$	[redacted]$	[redacted]$
2 rounds		(reference)	(reference)	(reference)	(reference)
Treatment					
CPX-351		[redacted]$			
3+7		(reference)			

$: p-value <0.05; AML, acute myeloid leukaemia; CMMoL, chronic myelomonocytic leukaemia; HMA, hypomethylating agents; HSCT; haematopoietic stem cell transplant; MDS, myelodysplastic syndrome

10.1.1.1 Patients who receive a second round of induction therapy

Overall, [redacted] of the ITT population ([redacted] received a second round of induction therapy. This percentage did not vary much by treatment or by age. However, some variations were observed among the AML types. A lower percentage of patients received a second round of induction in the "tAML", "MDSAML without HMA", and "CMMoLAML" types with [redacted] and [redacted] of the patients, respectively. The "MDSAML with HMA" type was similar to the overall population

([redacted] but the percentage of patients receiving a second round of induction was higher with the "denovoAML" type ([redacted] (CS appendix M, Table 45 pg. 264)

The results of the multivariate analysis showed that the age-group effect was not statistically significant (p=0.527), but the estimated odds ratio (OR) suggests that among younger patients, a smaller proportion received a second round of induction. Treatment effect was not included in this analysis as a covariate. The AML-type effect was statistically significant (p=0.030), but all 95% CI of OR included 1.0 and overlap with each other. As compared to patients with the "MDSAML with HMA" type, patients with the "denovoAML" type had higher odds of a second round of induction. On the other hand, the OR point estimates for "CMMoLAML", "MDSAML without HMA", and "tAML" types demonstrated lower odds of a second round of induction as compared to "MDSAML with HMA". (See Table 46)

10.1.1.2 Patients who receive post-induction remission

Overall, [redacted] patients ([redacted] achieved remission. This percentage was higher in CPX-351 patients ([redacted] than in 3+7 patients ([redacted] and it was higher in younger patients ([redacted] than in older ones ([redacted] For the five AML types, the percentage of remission varied from [redacted] with the "CMMoLAML" type to [redacted] with the "MDSAML without HMA" type. The percentage of patients achieving remission was higher in those who received only one round of induction ([redacted] than in those who received two rounds ([redacted] (CS appendix M, Table 46 pg. 265)

The multivariate analysis included treatment and the number of inductions in addition to the stratification variables age group and AML types. Even though the 95% CI of the OR included 1.0 and overlap with each other, the direction of the results suggests that the odds of remission were higher when the patient received no rounds or one round of induction as compared to two rounds. The CPX-351 patients were more likely to reach remission than 3+7 patients with an OR estimate of ([redacted]). (See Table 46)

10.1.1.3 Patients who receive consolidation

CS appendix M Table 47 (pg. 266) presents results on the percentages of patients who received 0, 1, or 2 rounds of consolidation among the 125 patients who achieved remission. Overall, the distribution of the number of rounds of consolidation received was relatively balanced with [redacted] and [redacted]% receiving 0, 1, and 2 rounds of consolidation. A similar distribution was observed among patients receiving CPX-351 treatment, but a lower percentage of 3+7 patients received two rounds of consolidation ([redacted] In younger patients, more patients received zero rounds of consolidation than received one or two rounds (decreasing trend with [redacted] receiving no rounds, [redacted] receiving one round, and [redacted] receiving two rounds). In the older age group, the opposite trend was observed, with [redacted] and [redacted] of patients receiving no, one, and two rounds of

consolidation, respectively. Finally, fewer patients required consolidation after two rounds of induction ([redacted] than after one round ([redacted] Similarly, fewer patients required two rounds of consolidation after two rounds of induction ([redacted] than after one round of induction ([redacted] (CS appendix M, Table 47 pg. 266)

The multivariate analysis included the number of inductions, and the stratification variables age group and AML type. The results comparing two rounds of consolidation to no rounds, the younger patients were less likely to receive two rounds of consolidation than were older ones. The AML types were combined to two categories due to fewer sample size per level of AML types. Therefore, the "CMMoLAML", "denovoAML", and "tAML" types were combined and compared to the combination of "MDSAML without HMA" and "MDSAML with HMA". Finally, patients who received one round of induction were more likely to receive two rounds of consolidation than were patients who received two rounds of induction. When comparing one round of consolidation to no rounds, similar patterns were observed for all explanatory variables. (See Table 46)

The CS noted that the treatment was not statistically significant suggesting no direct effect of treatment on the number of consolidation therapy rounds, therefore, it was not included in the analysis. However, the results with p-values were not presented from the multivariate logistic regression analyses including treatment as a co-variate.

10.1.1.4 Patients who receive a transplant

Overall, [redacted] of the 125 patients who achieved remission received a transplant. This percentage was slightly higher in CPX-351 patients ([redacted] than in 3+7 patients ([redacted] and it was higher in younger patients ([redacted] than in older patients ([redacted]%). A higher percentage of patients were transplanted after two rounds of inductions ([redacted] than after one round ([redacted] Regarding the number of rounds of consolidation, the lower the number of rounds received, the higher the chance to be transplanted, with [redacted] and [redacted] of patients transplanted after no, one, and two rounds of consolidation, respectively. (CS appendix M, Table 48 pg. 267)

The number of rounds of induction was included in the multivariate analysis in addition to the stratification variables age group and AML type. The AML types were combined in two categories – "CMMoLAML", "denovoAML", and "tAML" types versus the combination of "MDSAML without HMA" and "MDSAML with HMA") due to fewer sample size per level of AML types. The OR point estimate for age group suggests that younger patients were more likely to be transplanted than were older ones and patients who received only one round of induction as compared to two had less chances to receive a transplant. (See Table 46)

The CS stated that the treatment effect was not statistically significant, nor was the number of rounds of consolidation, therefore, CS multivariate analysis did not include these as covariates. However, the OR with p-values were not presented from multivariate logistic regression analyses including treatment as a covariate.

10.1.2 Mean time elapsed after induction therapy (time-shift)

In the economic model, for patients who achieve response after induction therapy, OS and EFS are not tracked until after a certain amount of time had elapsed (as described in Section 5.2.1).

The CS conducted multivariate linear regression analyses to estimate the duration of these time-shifts for each subgroup in the model. The CS applied these time shifts only to patients who achieved remission after the induction phase. No time-shift was applied in relation to the survival of patients who did not respond to treatment in the induction phase; therefore, the parametric curves assessing the successive composite endpoints for this group of patients are applied from the beginning of the model.

The following three time periods were estimated using multivariate linear regression analyses.

- Time elapsed between randomisation and post-induction response among patients who achieve remission,
- Time elapsed between post-induction response and end date of consolidation among patients who achieve remission,
- Time elapsed between end date of consolidation and start date of transplantation among patients who received transplant.

A brief description of the technical details are presented in the Table 47. Detailed information on the statistical analyses are presented in CS appendix M.

Table 47 Description of the technical details used to estimate time to shift (adapted from Table-35 in CS Appendix M)

Analysis	Description	Technical Details	Co-variates adjustment	Purpose
1	Mean time elapsed between "Date of Randomization" and "Date of Post-induction Response" among patients who achieve remission	Summary statistics obtained by fitting a linear regression model for the time elapsed between "Date of Randomization" and "Date of Post-induction Response"	Adjusting for age (60–69 vs. 70–75 years old), five AML types, treatment, and number of rounds of induction therapy	Required to estimate the shift of survival outcomes for patients who achieve remission
2	Mean time elapsed between "Date of Post-induction Response" and "End Date of Consolidation Last	Summary statistics obtained by fitting a linear regression model for the time elapsed between "Date of Post-induction Response" and "End Date of	Adjusting for age (60–69 vs. 70–75 years old), five AML types, and number of rounds	Required to estimate the shift of survival outcomes for patients who achieve remission

	Treatment" (last round) among patients who achieve remission had one or two rounds of consolidation therapy	Consolidation Last Treatment" (last round)	(0 or 1 vs. 2) of consolidation	
3	Mean time elapsed between "End Date of Consolidation Last Treatment" (last round) and "Start Date of Stem Cell Transplant" among patients who achieve remission and had transplant	Summary statistics obtained by fitting a linear regression model for the time elapsed between "End Date of Consolidation Last Treatment" (last round) and "Start Date of Stem Cell Transplant" For patients who achieve remission and had transplant , but did not receive consolidation therapy, "Date of Post-induction Response" will be used as the "End Date of Consolidation Last Treatment"	Adjusting for age (60–69 vs. 70–75 years old), five AML types combined in two categories and number of rounds (0 vs. 1 vs. 2) of consolidation	Required to estimate the shift of survival outcomes for patients who achieve remission

The results of multivariate linear regression analyses used to estimate three time periods are presented in Table 48.

Table 48 Results of multivariate linear regression analyses to estimate time to shift (adapted from Tables in CS Appendix M)

Time-shift	**Time to post-induction response**	**Time from post-induction response to last consolidation treatment**	**Time from last consolidation treatment to HSCT**
Parameter	**Estimate (95% CI)**	**Estimate (95% CI)**	**Estimate (95% CI)**
Intercept	[redacted]	[redacted]	[redacted]
Treatment			
CPX-351	[redacted]	-	-
3+7	(reference)	-	-
Age			
60-69 years old	[redacted]	[redacted]	[redacted]
70-75 years old	(reference)	(reference)	(reference)
AML TYpe			
"$_{CMMoL}$AML"	[redacted]	[redacted]	(combined with reference)
"$_{denovo}$AML"	[redacted]	[redacted]	(combined with reference)
"$_{MDS}$AML with HMA"	[redacted]	[redacted]	[redacted]
"$_{MDS}$AML without HMA"	[redacted]	[redacted]	
"t-r AML"	(reference)	(reference)	(reference)
Number of Inductions			
1	(reference)	-	-
2	[redacted]	-	-
Number of Consolidations			
0	-	(reference)	(reference)
1	-		[redacted]
2	-	[redacted]	[redacted]

$, p-value <0.05; ^: combined "MDSAML with and without HMA" together; AML, acute myeloid leukemia; CL, confidence limit; CMMoL, chronic myelomonocytic leukemia; HMA, hypomethylating agent; MDS, myelodysplastic syndrome; t-r, treatment related

10.1.2.1 Time-shift: Time to post-induction response

Overall, among the 125 patients who achieved remission, there were [redacted] between the date of randomisation and the date when the response achievement was confirmed. This duration was similar with regard to treatment received (CPX-351: [redacted] 3+7: [redacted] and between age groups (60–69: [redacted] 70-75: [redacted] The duration was longer when two rounds of induction were received ([redacted]) as compared to only one round ([redacted]). (CS appendix M Table 54 pg. 273)

Treatment and the number of rounds of induction were included in the final multivariate linear regression analysis as they were statistically significant. They were included along with the

stratification variables. The number of rounds of consolidation was not included in the final model. The regression coefficients were in line with the descriptive summary - patients receiving CPX-351 had a longer delay in response as compared to those receiving 3+7 (mean difference of [redacted] and patients receiving two rounds of induction had a longer delay than those receiving only one round (mean difference of [redacted]). (See Table 48)

10.1.2.2 Time-shift: Time from post-induction response to last consolidation treatment

Overall, 80 patients were included in this pathway who achieve remission had one or two rounds of consolidation therapy. On average, 5.0 weeks (SD 3.62) estimated between the date when the response achievement was confirmed and the end date of the last consolidation treatment. This duration did not vary by treatment (CPX-351: [redacted] Overall results by age group and number of inductions were very similar, but, patients receiving two rounds of consolidation experienced a longer time between the date of response and the end of their last consolidation treatment. Patients receiving one round of consolidation experienced an elapsed time of [redacted] as compared to [redacted] for patients receiving two rounds of consolidation. (CS appendix M Table 55 pg. 274)

The stratification variables were not statistically significant, but they were included in the final model. The treatment and the number of rounds of induction were not included in the final model justified as they were not statistically significant. Only the number of rounds of consolidation was included in the final model addition to stratification variables (mean difference: [redacted]). (See Table 48)

10.1.2.3 Time-shift: Time from last consolidation treatment to transplant

Overall, 63 patients were part of the combined pathways who achieve remission and had transplant. On average, [redacted] were recorded between the end of the last consolidation treatment (last induction treatment for the group who achieve remission and had transplant, but did not receive consolidation therapy) and the start date of the transplant. The time period were similar between two treatment groups (CPX-351: [redacted] 3+7: [redacted] and also between age groups (60–69: [redacted] 70–75: [redacted] The time period slightly varied between the two rounds of induction (1: [redacted], 2: [redacted] The time between the end of the consolidation last treatment and the transplant was longer when two rounds of consolidation were given ([redacted]) than when no rounds or one round of consolidation were given (0: [redacted] 1: 8.9 [[redacted]). (CS appendix M Table 56 pg. 275-276).

The stratification variables were included in the final model, but they were not statistically significant. Treatment and number of rounds of induction were not included in the final model and justified as

they were not statistically significant. Only the number of rounds of consolidation was added to the stratification variables. When two-round patients were compared to no-round patients, the mean difference in the modelled outcome was [redacted]. Patients receiving no rounds of consolidation were comparable to those receiving one round in terms of time elapsed between the end of consolidation last treatment and the transplant (mean difference: [redacted] (See Table 48).

10.1.3 Survival analysis and extrapolation beyond the trial period

The company fitted parametric survival curves to the patient-level data from Study 301 to extrapolate over the model time horizon. These were stratified by response status, and further by transplant status in those who achieved a response.

The parametric curves included:

- *Overall survival and event-free survival* for patients who responded but did not receive transplant, where EFS and OS were calculated from the time of last consolidation therapy. This group is henceforth referred to as *post-consolidation patients*.

- *Overall survival and event-free survival* for patients who responded and received transplant, where EFS and OS were calculated from the time of transplant. This group is henceforth referred to as *post-transplant patients*.

- *Overall survival, time to transplant or death, and time to transplant or progression or death,* for patients who did not achieve response, where the survival models were calculated from the beginning of the model. This group is henceforth referred to as *non-responder patients*.

To extrapolate each of these survival curves, the company explored a range of conventional parametric models. Curves were adjusted for sampling stratification variables age (60–69 vs. 70–75 years old), AML subtype, treatment arm, number of rounds of induction therapy, and number of rounds of consolidation, where appropriate. The Akaike information criterion (AIC) and Bayesian information criterion (BIC) were used to choose among the different parametric curves. Projected survival curves and median estimates were also examined to assess the clinical plausibility of the distributions.

A brief description of the technical details implemented in the base-case economic model are presented in Table 49. Detailed information on the statistical analyses are presented in CS appendix M.

Table 49 Description of the technical details: Time to event analyses implemented in the base-case economic model (adapted from Table-35 in CS Appendix M)

Analysis	Description	Technical Details	Adjustment	Purpose
1	Post-consolidation OS among patients who achieved remission but didn't receive transplant	Parametric survival model: Log-logistic curve fitted to the base-case economic model, and curves fitted separately for each treatment Outcome: Time to death Entry time: "End Date of Consolidation Last Treatment" (last round) Censoring time: "Date of Last Contact"	Adjusting for age (60–69 vs. 70–75 years old), five AML types, and number of consolidation rounds (0 vs. 1 vs. 2)	Required for a post-consolidation partitioning survival model among patients who achieved remission but didn't receive transplant
2	Post-consolidation EFS among patients who achieved remission but didn't receive transplant	Parametric survival model: Weibull curve fitted to the base-case economic model, and curves fitted separately for each treatment Outcome: Time to relapse or death Entry time: "End Date of Consolidation Last Treatment" (last round) Censoring time: "Date of Last Examination"	Adjusting for age (60–69 vs. 70–75 years old), AML types combined in two categories, and number of consolidation rounds (0 vs. 1 or 2)	Required for a post-consolidation partitioning survival model among patients who achieved remission but didn't receive transplant
3	Post-HSCT OS who achieved remission and received transplant	Parametric survival model: Gompertz curve fitted to the base-case economic model, and curves fitted separately for each treatment Outcome: Time to death Entry time: "Start Date of Stem Cell Transplant" Censoring time: "Date of Last Contact"	Adjusting for age (60–69 vs. 70–75 years old), AML types combined in two categories, and number of consolidation rounds (0 vs. 1 or 2) Additionally, CPX-351 was adjusted for number of induction rounds	Required for a post-HSCT partitioning survival model among patients who achieved remission and received transplant
4	Post-HSCT EFS among patients who achieved remission and received transplant	Parametric survival model: Log-normal curve fitted to the base-case economic model, and curves fitted separately for each treatment Outcome: Time to relapse or death Entry time: "Start Date of Stem Cell Transplant" Exit time: "Date of Last Examination"	Adjusting for age (60–69 vs. 70–75 years old), AML types combined in two categories, and number of consolidation rounds (0 vs. 1 or 2)	Required for a post-HSCT partitioning survival model among patients who achieved remission and received transplant
5	OS among those do not achieve remission	Parametric survival models: Log-normal curve fitted to the base-case economic model, and has been fitted without treatment as a covariate (i.e. merging both treatments)	Adjusting for age (60–69 vs. 70–75 years old), five AML types, and number of induction rounds (0 or1 vs. 2)	Required for a partitioning survival model among those who do not achieve remission

		Outcome: Time to death Entry time: "Date of Randomization" Censoring time: "Date of Last Contact"		
6	Time to transplant or death among those do not achieve remission	Parametric survival models: Log-normal curve fitted to the base-case economic model, and has been fitted without treatment as a covariate (i.e. merging both treatments) Outcome: Time to HSCT or death Entry time: "Date of Randomization" Censoring time: "Date of Last Contact"	Adjusting for age (60–69 vs. 70–75 years old), and five AML types	Required for a partitioning survival model among those who do not achieve remission
7	Time to progression or transplant or death among those do not achieve remission	Parametric survival models: Generalised gamma curve fitted to the base-case economic model, and has been fitted without treatment as a covariate (i.e. merging both treatments) Outcome: Time to progression or HSCT or death Entry time: "Date of Randomization" Censoring time: "Date of Last Examination"	Adjusting for age (60–69 vs. 70–75 years old), five AML types, and number of induction rounds (0 or1 vs. 2)	Required for a partitioning survival model among those who do not achieve remission
HSCT, hematopoietic stem cell transplant; ITT, intent-to-treat; OS, overall survival; EFS, event free survival				

10.1.3.1 Post-consolidation OS

When modelling post-consolidation OS among patients who achieved remission but didn't receive transplant, a number of distributions were fitted to survival data in both treatment groups separately (see [redacted] 22**Error! Reference source not found.**). The log-logistic curve selected by the company for their base-case analysis for both treatments. Predicted median OS from the log-logistic model was [redacted] weeks for CPX-351 and [redacted] weeks for 3+7, which were comparable against the observed KM estimates.

[redacted] 22 [redacted]

[redacted]

The log-logistic survival models for each treatment for the post-consolidation OS among those who achieved remission but didn't receive transplant are presented in Table 50.

Table 50 Log-logistic Survival Model for Post-consolidation OS in Weeks among Patients who achieved remission but didn't receive transplant, CPX-351 and 3+7 (adapted from Tables in CS Appendix M)

Parameter	**CPX-351**	**3+7**
	Estimate (95% CI)	**Estimate (95% CI)**
Intercept	[redacted]	[redacted]
Age		
60-69 years old	[redacted]	[redacted]
70-75 years old	(reference)	(reference)
AML Type		
"$_{CMMoL}$AML"	[redacted]	[redacted]
"$_{denovo}$AML"	[redacted]	[redacted]
"$_{MDS}$AML with HMA"	[redacted]	[redacted]
"$_{MDS}$AML without HMA"	[redacted]	[redacted]
"t-r AML"	(reference)	(reference)
Number of Consolidations		
0	[redacted]	[redacted]
1	[redacted]	[redacted]
2	(reference)	(reference)
$: p-value<0.05; AML, acute myeloid leukemia; CI, confidence interval; CMMoL, chronic myelomonocytic leukemia; HMA, hypomethylating agent; MDS, myelodysplastic syndrome; t-r, treatment related		

The number of rounds of consolidation therapy was not statistically significant for CPX-351, but it was significant for 3+7. Despite its lack of statistical significance for CPX-351, the model included

number of consolidation rounds for each treatment. Therefore, the final models included number of rounds of consolidation in addition to the stratification variables age group and AML type. For patients receiving 3+7, the negative sign of the regression coefficient for rounds of consolidation ([redacted], 95% CI: [redacted]) indicates that dying post-consolidation in happens earlier (i.e., higher risk) for patients receiving no consolidation than it does in patients receiving two rounds. Similarly, time to death after consolidation is shorter (i.e., higher risk) for patients receiving one round versus two rounds of consolidation [redacted], 95% CI: [redacted] The same pattern was observed in patients receiving CPX-351, but much smaller in magnitude.

In both cases, the number of rounds of induction was not retained in the final model since it was not statistically significant.

10.1.3.2 Post-consolidation EFS

When modelling post-consolidation EFS (time to relapse or death) among patients who achieved remission but didn't receive transplant, a number of distributions were fitted to survival data in both treatment groups separately (see [redacted]23). The Weibull curve selected by the company for their base-case economic analysis for both treatments. Predicted median time to relapse or death from the log-logistic model was [redacted] weeks for CPX-351 and [redacted] weeks for 3+7, which were comparable against the observed KM estimates.

[redacted]23

[redacted]

The Weibull survival models for each treatment for the post-consolidation EFS among patients who achieved remission but did not receive transplant is presented in Table 51.

Table 51 Weibull Survival Model for Post-consolidation EFS (Time to Relapse or Death) in Weeks among Patients in Pathways B, D, and F, CPX-351 and 3+7 (adapted from Tables in CS Appendix M)

Parameter	**CPX-351**	**3+7**
	Estimate (95% CI)	**Estimate (95% CI)**
Intercept	[redacted]	[redacted]
Age		
60-69 years old	[redacted]	[redacted]
70-75 years old	(reference)	(reference)
AML Type		
"$_{CMMoL}$AML"	[redacted]	[redacted]
"$_{denovo}$AML"	[redacted]	[redacted]
"$_{MDS}$AML with HMA"	[redacted]	[redacted]
"$_{MDS}$AML without HMA"	[redacted]	[redacted]
"t-r AML"	(reference)	(reference)
Number of Consolidations		
0	[redacted]	[redacted]
1	[redacted]	[redacted]
2	(reference)	(reference)

$: p-value<0.05; AML, acute myeloid leukemia; CI, confidence interval; CMMoL, chronic myelomonocytic leukemia; HMA, hypomethylating agent; MDS, myelodysplastic syndrome; t-r, treatment related

Similar to post consolidation OS analysis, the number of consolidation rounds was included in the final model despite its lack of statistical significance. The final models therefore included number of rounds of consolidation in addition to the stratification variables age group and AML type. For patients receiving 3+7, the estimate indicates that the time to relapse or death post-consolidation is shorter (i.e., higher risk) for patients receiving no consolidation than for patients receiving two rounds. Similarly, after consolidation, time to relapse or death is shorter (i.e., higher risk) for patients receiving one round versus two rounds of consolidation. For patients receiving CPX-351, the estimates indicate lower risk for patients receiving no consolidation than for patients receiving two rounds but higher risk for patients receiving one round versus two rounds of consolidation, however, the magnitude of effects are small.

In both cases, the number of rounds of induction was not retained in the final model since it was not statistically significant.

10.1.3.3 Post-transplant OS

For the post-transplant OS analyses among patients who achieved remission and received transplant, Gompertz distribution was chosen as the best fit to the trial data among the various distributions tested (see [redacted] 24). While the different models for 3+7 provided similar long-term predictions for OS, there appeared to be heterogeneity across outcomes in the CPX-351 analysis.

[redacted] 24

[redacted]

The Gompertz survival models for each treatment for the post-transplant OS among patients who achieved remission and received transplant are presented in Table 52.

Table 52 Gompertz Survival Model for Post-transplant OS in Weeks among Patients who achieved remission and received transplant, CPX-351 and 3+7 (adapted from Tables in CS Appendix M)

Parameter	CPX-351	3+7
	Estimate (95% CI)	**Estimate (95% CI)**
Intercept	[redacted]	[redacted]
Age		
60-69 years old	[redacted]	[redacted]
70-75 years old	(reference)	(reference)
AML Type, Combined		
"$_{CMMoL}$AML", "$_{denovo}$AML", or "t-r AML"	(reference)	(reference)
"$_{MDS}$AML without HMA" or "$_{MDS}$AML with HMA"	[redacted]	[redacted]
Number of Inductions		
1	[redacted]	Didn't include in the base-case economic model
2	(reference)	Didn't include in the base-case economic model
Number of Consolidations		
0	(reference)	(reference)
1, or 2	[redacted]	[redacted]
$: p-value<0.05; AML, acute myeloid leukemia; CI, confidence interval; CMMoL, chronic myelomonocytic leukemia; HMA, hypomethylating agent; MDS, myelodysplastic syndrome; t-r, treatment related		

For patients receiving CPX-351, the number of rounds of induction (one vs two) was statistically significant and the estimate indicates that the time of death in patients receiving only one round of induction was later (i.e., lower risk of dying) than in patients receiving two rounds. For patients receiving 3+7, the number of rounds of induction was not statistically significant and so was not included in the model. The number of rounds of consolidation was included in the final model despite its lack of statistical significance.

10.1.3.4 Post-transplant EFS

Like as other analyses, for the post-transplant EFS (time to relapse or death) among patients who achieved remission and received transplant, log-normal distribution was chosen as the best fit among the various distributions tested (see [redacted]**Error! Reference source not found.**). For patients receiving CPX-351, the predicted median time to relapse or death from the log-normal fit was [redacted] weeks, which is lower than the results observed with the KM survival curve [redacted] weeks.

For patients receiving 3+7, the predicted median time to relapse or death from the log-normal fit was [redacted], which was much lower than the results observed with the KM survival curve

([redacted]). The CS noted that difference in the median values were due to the small sample size involved in this analysis (as only [redacted] patients received 3+7 treatment and had a positive value for the time to relapse or death post-consolidation).

[redacted]

[redacted]

The log-normal survival model for the post-transplant time to relapse or death among patients who achieved remission and received transplant is presented in Table 53.

Table 53 Log-normal Survival Model for Post-transplant Time to Relapse or Death in Weeks among Patients in Pathways A, C, and E, CPX-351 and 3+7 (adapted from Tables in CS Appendix M)

Parameter	**CPX-351**	**3+7**
	Estimate (95% CI)	**Estimate (95% CI)**
Intercept	[redacted]	[redacted]
Age		
60-69 years old	[redacted]	[redacted]
70-75 years old	(reference)	(reference)
AML Type, Combined		
"$_{CMMoL}$AML", "$_{denovo}$AML", or "t-r AML"	[redacted]	[redacted]
"$_{MDS}$AML without HMA" or "$_{MDS}$AML with HMA"	(reference)	(reference)
Number of Consolidations		
0	[redacted]	[redacted]
1, or 2	(reference)	(reference)
AML, acute myeloid leukemia; CI, confidence interval; CMMoL, chronic myelomonocytic leukemia; HMA, hypomethylating agent; MDS, myelodysplastic syndrome; t-r, treatment related		

For patients receiving CPX-351, the number of rounds of consolidation (zero vs. one or two) included in the model addition to the stratification variables age and AML types combined in two categories. However, none of these variables were statistically significant in the multivariate analysis. Similarly, for patients receiving 3+7, the model included the age, the combined AML types, and the number of rounds of consolidation (zero vs. one or two), even though these were not statistically significant.

In both arm it was reported that the number of rounds of induction (one vs two) was not statistically significant. However, the results of the multivariate analyses including number of rounds of induction (one vs two) were not presented.

10.1.3.5 Non-responder OS

In the analyses for OS among those who do not achieve remission, the treatments were combined and the choice of the most appropriate parametric distribution was made based on this combined group of patients. The log-normal distribution was chosen as the best fit among the various distributions tested after considering several criteria (see [redacted]26**Error! Reference source not found.**). Predicted median OS from the log-normal model was [redacted] weeks, which was comparable with the KM survival curve.

[redacted]26

[redacted]

The log-normal survival model for OS among those who do not achieve remission is presented in Table 54.

Table 54 Log-normal Survival Model for OS in Weeks among those who do not achieve remission, Combined Treatments (adapted from Tables in CS Appendix M)

Parameter	Estimate (95% CI)
Intercept	[redacted]
Age	
60-69 years old	[redacted]
70-75 years old	(reference)
AML Type	
"$_{CMMoL}$AML"	[redacted]
"$_{denovo}$AML"	[redacted]
"$_{MDS}$AML with HMA"	[redacted]
"$_{MDS}$AML without HMA"	[redacted]
"t-r AML"	(reference)
Number of Inductions	
0, or 1	[redacted]
2	(reference)
$: p-value <0.05; AML, acute myeloid leukemia; CL, confidence interval; CMMoL, chronic myelomonocytic leukemia; HMA, hypomethylating agent; MDS, myelodysplastic syndrome; t-r, treatment related	

In addition to age group and AML type, the number of rounds of induction was included as it was identified as a statistically significant prognostic factor where zero and one round were combined and compared to two rounds of induction. The estimate indicates that higher risk of death for patients receiving no rounds or one round of induction as compared to two rounds.

10.1.3.6 Non-responder Time to Transplant or Death

In the analyses for Time to Transplant or Death among those who do not achieve remission, the treatments were combined and the choice of the most appropriate parametric distribution was made based on this combined group of patients. The log-normal distribution was chosen as the best fit among the various distributions tested after considering several criteria (see [redacted] 27**Error! Reference source not found.**). Predicted median Time to Transplant or Death from the log-normal model was [redacted] weeks, which was comparable with the KM survival curve.

[redacted]27[redacted]

[redacted]

The log-normal survival model for Time to Transplant or Death among those who do not achieve remission in Pathways G and H is presented in Table 55. Only stratification variables age group and AML type were included in the model.

Table 55 Log-normal Survival Model for Time to Transplant or Death in Weeks among those who do not achieve remission, Combined Treatments (adapted from Tables in CS Appendix M)

Parameter	Estimate (95% CI)
Intercept	[redacted]
Age	
60-69 years old	[redacted]
70-75 years old	(reference)
AML Type	
"$_{CMMoL}$AML"	[redacted]
"$_{denovo}$AML"	[redacted]
"$_{MDS}$AML with HMA"	[redacted]
"$_{MDS}$AML without HMA"	[redacted]
"t-r AML"	(reference)
AML, acute myeloid leukemia; CL, confidence interval; CMMoL, chronic myelomonocytic leukemia; HMA, hypomethylating agent; MDS, myelodysplastic syndrome; t-r, treatment related	

10.1.3.7 Non-responder Time to progression or Transplant or Death

In the analyses for Time to Progression or Transplant or Death among those who do not achieve remission, the treatments were combined and the choice of the most appropriate parametric distribution was made based on this combined group of patients. The generalised gamma distribution was chosen as the best fit among the various distributions tested after considering several criteria (see **Error! Reference source not found.**). Predicted median OS from the generalised gamma model was [redacted] weeks, which was comparable with the KM survival curve.

[redacted]28[redacted]

[redacted]

The generalized gamma survival model for the time to progression or transplant or death among those who do not achieve remission is presented in Table 56.

Table 56 Generalized Gamma Survival Model for the time to progression or transplant or death among those who do not achieve remission, Combined Treatments (adapted from Tables in CS Appendix M)

Parameter	Estimate (95% CI)
Intercept	[redacted]
Age	
60-69 years old	[redacted]
70-75 years old	(reference)
AML Type	
"$_{CMMoL}$AML"	[redacted]
"$_{denovo}$AML"	[redacted]
"$_{MDS}$AML with HMA"	[redacted]
"$_{MDS}$AML without HMA"	[redacted]
"t-r AML"	(reference)
Number of Inductions	
0, or 1	[redacted]
2	(reference)
$: p-value <0.05; AML, acute myeloid leukemia; CL, confidence interval; CMMoL, chronic myelomonocytic leukemia; HMA, hypomethylating agent; MDS, myelodysplastic syndrome; t-r, treatment related	

In addition to age group and AML type, the number of rounds of induction was included as it was statistically significant. No and one rounds of induction were combined and compared to two rounds of induction. The estimate indicates that the time to disease progression, receiving a transplant, or death (i.e., higher risk of disease progression, receiving a transplant, or death) is shorter for patients receiving no rounds or one round of induction as compared to two rounds.

National Institute for Health and Care Excellence
Centre for Health Technology Evaluation

Pro-forma Response

ERG report

Liposomal cytarabine and daunorubicin for untreated acute myeloid leukaemia [ID 1225]

You are asked to check the ERG report from the 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 **Friday 13 July 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.

ERG additional comment:

In addition, the ERG noted an omission in their description to the alternative base-case described in Section 6.4 of the ERG report. In the ERG alternative base case (page 131), we should have stated that the utility estimate for patients in the post-transplant health state should have been *based on the utility value of 0.75 for patients in durable remission*, further adjusted for age.

The text on page 22 and page 132 has been amended to reflect this change.

Issue 1 Post-transplant outcomes based on OS only

Description of problem	Description of proposed amendment	Justification for amendment	ERG response
The ERG implemented a scenario in which patient outcomes post-HSCT were based on the OS curve only (pg. 129 and pg. 133). This change interacts with the EFS post-transplant costs (cell G220 of the cost-inputs page in the model) and results in an inappropriate extension of short-term monitoring costs to persist for a patient lifetime for patients surviving post-transplant. This cost should be zeroed out for post-transplant patients in analyses using the "Post-transplant outcomes based on OS only" setting, including the ERG base case.	Cells H181:182 in the Parameters sheet in the model should be amended to include a switch to set these costs to zero if the "Post-transplant outcomes based on OS only" scenario is selected. On pg. 129, Table 38, this reduces the incremental costs in the scenario by [redacted], and correspondingly reduces the ICER by [redacted]. Tables 43 and 44 (pg. 135 and 136) will have their numerical results changed as well, as the scenario is included in the ERG base case.	The EFS post-transplant costs reflect a six month period of additional monitoring for patients in remission. Applying it over longer periods lacks face validity. The ERG change materially increases the ICER by over £10,000/QALY. The company proposed amendment is needed to accurately interpret cost-effectiveness in the ERG analyses.	The ERG notes that the company's statement on that monitoring costs should not continue indefinitely to be valid, but does not believe that these patients should have no associated monitoring cost. As such, the ERG proposes an amendment to their analysis, whereby monitoring costs are applied for six months after achieving remission. The results of these are included in the updated report on page 128, and the impact was to lower the ICER by approx [redacted] per QALY for the single change analysis and [redacted] per QALY for the all change analysis.

Issue 2 Post-transplant cost

Description of problem	Description of proposed amendment	Justification for amendment	ERG response
The ERG proposed an increase in the follow-up costs post-transplant to reflect two-year costs rather than six-month costs (pg. 132 and 134). This is inappropriate, as not all post-transplant patients survive for two years. The additional costs (i.e. difference between two-year and six-month costs) should only be applied to those patients who survive sufficiently long.	Post-transplant costs should be incorporated in the model over time. As an approximation, and to account for the fact that costs are not accrued continuously, the initial six month costs could be applied as a lump sum following transplant (as in the company base case) with the additional costs applied after six months to those patients that survive when the ERG scenario is being considered.	The incorporation of all costs through 2 years as a lump sum at the time of transplant overestimates costs post-transplant and creates a modest bias in ICERs. Inclusion of the company's proposed amendment along with the recommended amendment for issue 1, reduces the ERG base case ICER to [redacted]/QALY.	The ERG's estimation of post-transplant monitoring is weighted by the proportion of patients who are alive after transplant, using data estimated from the study on which the costs were estimated. The overall cost that is applied to each patient receiving transplant is, therefore, reflective of the mean total cost incurred by patients who experience a range of mortality. The ERG, however, acknowledges that the reporting of this method was lacking and has amended the report to provide details of the methods implemented on page 132. The ERG retains their original method in the analysis.

Issue 3 Application of disutility

Description of problem	Description of proposed amendment	Justification for amendment	ERG response
The ERG notes that patients accrue less QALYs than might be expected based on the scores from the submitted time trade-off	The description of the face validity of QALY gains in the 3+7 arm (full paragraph on pg. 101) should be removed.	The company's position is that the application of disutilities is consistent with the design of the TTO study and is not an error in the	The TTO study and the QALYs predicted by the model suggest an on-treatment utility that the ERG consider to be implausibly

Description of problem	Description of proposed amendment	Justification for amendment	ERG response
(TTO) study (pg. 101) and attributes this to an error in the calculation of utility in the model (pg. 101). The ERG proposes a change to the model programming to address this (pg. 101 and pg. 124). The application of disutilities, however, is consistent with the design of the TTO study and is not an error in the model programming. The vignettes in the TTO study used to assess the impact of induction and consolidation describe path states in which the state of the patient changes during the year being valued. Because of the changing state of the patient, path states values cannot be converted to utility scores directly. Instead, the differences between the states can be valued as QALY decrements by using the difference in the TTO values for those states. The QALY accumulation cited in the ERG report (pg. 101) is below the value of any of the individual states because there are separate vignettes for induction and consolidation, with no vignette in the TTO study that values both	The first bullet on pg. 124 describing the application of treatment-related disutilities as an error (and the associated model change) should be removed and results in table 35 amended. Analyses and results in Tables 36 through 41 should be amended to include the original application of treatment-related disutility.	model programming. The change made by the ERG is responsible for virtually all of the ICER difference between the ERG's correction of the company base case and the submitted company base case (roughly [redacted]/QALY). All results from Table 35 through Table 41 are based on that correction and would have the ICERs reported reduced similarly.	low. The company model predicted [redacted] and [redacted] QALYs accrued prior to transplant for CPX-351 and 3+7 respectively, and [redacted] and [redacted] QALY loss attributable to induction and consolidation therapy. This implies that patients would experience a quality of life worse than death (a negative utility value) during this time. As a result, the ERG believed that this was due to an error in how quality of life was implemented in the model Upon further consideration of this issue, the ERG believes that the implication of negative utility while on-treatment is due to limitations in the TTO study and how the health states were designed. The ERG acknowledges that there are some limitations with the suggested correction, but maintains that it provides an estimation of QALYs with greater face validity than those predicted by the company. Furthermore, the ERG's approach applies utility values

Description of problem	Description of proposed amendment	Justification for amendment	ERG response
induction and consolidation. As would be expected, a year in which a patient experiences both induction and consolidation is valued lower than either a year in which the patient experiences only induction or only consolidation.			that are more consistent with those used in other NICE appraisals of treatments for AML and in the relevant studies identified in the company's SLR.

Issue 4 Typographical error

Description of problem	Description of proposed amendment	Justification for amendment	ERG response
On pg. 82, there is a typographical error in the title of the study: "The probability of remission also included a treatment effect because the higher rate of remission in CPX-351 patients than 3+7 patients observed in Study 351 was statistically significant."	"Study 351" should be amended to "Study 301".	Correction of typographical error.	The text has been amended as per the company's suggestion.

Issue 5 Reported number of additional deaths

Description of problem	Description of proposed amendment	Justification for amendment	ERG response
On pg. 16, 91, 118 and 137, the ERG state that there have been [redacted]	Modify [redacted] throughout document, and clarify that this is from a safety analysis population. Example:	The incorrect calculation of deaths is misleading and may lead to the conclusion that the OS curves are	Changes were made to numbers of events according to company suggestion on pages.

Description of problem	Description of proposed amendment	Justification for amendment	ERG response
[redacted] since the 2015 data cut, and infer that this might suggest a degree of convergence in the OS curves. The numbers of deaths have been calculated incorrectly, by comparing two different populations. The additional data provided by the company were the number of known deaths ([redacted]) that have occurred since the 2015 data cut from a safety analysis population. The number of additional deaths should have been calculated against the safety analysis population (see CSR pg. X). At the primary analysis, there had been [redacted] in the safety analysis population. Therefore, the numbers of additional deaths should have been correctly calculated as [redacted]. The ERG calculated [redacted] by	[redacted] versus the safety analysis population	converging.	16 and 137. The reference to converging OS curves was removed p137. Additional detail was provided on page 91 regarding the relative number of deaths. The statement on page 91 on the curve converging was not removed, as this is a statement of a possible consequence of the omission of these deaths. This statement was modified in page 120 to reflect the original speculative tone taken by the ERG in their previous statement.

Description of problem	Description of proposed amendment	Justification for amendment	ERG response
inappropriately comparing the safety analysis population versus the ITT efficacy population. As this is a safety update rather than a formal survival analysis, these results only capture known deaths currently documented. It therefore may not reflect actual deaths on both arms nor provide information on whether these are transplant patients. Thus, these data should be interpreted with caution.			

Issue 6 Number of consolidations

Description of problem	Description of proposed amendment	Justification for amendment	ERG response
On pg. 77, the ERG states that the draft marketing authorisation for CPX-351 permits patients to receive up to four courses of consolidation and that it was uncertain how many patients would receive third and fourth consolidation courses in practice. Section 4.2 of the latest CPX-351 SmPC states: [redacted].	Remove paragraph	Treatment with up to 2 consolidation cycles is inconsistent with the updated SmPC.	This is not a factual inaccuracy, as the ERG did not have a copy of the updated CPX-351 SmPC at the time of drafting the ERG report. However, a note has been added to page 77 to clarify that the updated draft SmPC allows up to maximum of 2 consolidation courses.

Description of problem	Description of proposed amendment	Justification for amendment	ERG response
Treatment beyond 2 consolidation cycles would therefore be inconsistent with both the SmPC and the data from the pivotal trial.			

Issue 7 Prior-HMA patient population

Description of problem	Description of proposed amendment	Justification for amendment	ERG response
Pg. 16, 55, 64, 76 and 137 refer to the OS results for the stratified subgroup of MDS patients previously treated with HMAs [redacted] This analysis does not consider all patients previously treated with HMAs. In the 301 study, there were additional patients previously treated with HMAs but who were classified under different strata and therefore excluded from this analysis. In a new analysis, when all patients previously treated with HMAs are included [redacted]	Update text to include full details on efficacy: [redacted]	Focussing only on the subgroup of MDS patients previously treated with HMAs may potentially undervalue the benefit of CPX-351 in the wider population of patients who have been treated with HMAs ([redacted] and are known to do poorly.	The text on pages 16, 64 and 137 has been amended to clarify that these results relate to the stratified subgroup of '$_{MDS}$AML with prior treatment with HMA' patients, rather than a previously undefined subgroup which included additional patients previously treated with HMAs but who were classified under different strata. Text on pages 55 and 76 has not been amended, as it is already clear that the subgroup analysis results relate to the stratified subgroup of '$_{MDS}$AML with prior treatment with HMA' patients.

Description of problem	Description of proposed amendment	Justification for amendment	ERG response
[redacted].			
On pg. 16 the ERG propose that MDS patients previously treated with HMAs represent a third of potential CPX-351 patients in routine practice. The company believes that the ERG report may have overestimated the size of this population in routine practice Within the subgroup AML patients transformed from MDS, patients previously treated with HMA treatment accounts approximately 10-15% of patients in routine practice (Boddu et al. 2017; Subari et al. 2016).	Change “around a third of patients” to “10-15% of patients”	The company believes that the ERG report may have overestimated the size of this population in routine practice, and therefore the potential cost impact to the NHS.	See previous response, the text on pages 16, 64 and 137 has been amended to clarify the subgroup of ‘$_{MDS}$AML with prior treatment with HMA’ patients this refers to.

Issue 8 EFS limited data

Description of problem	Description of proposed amendment	Justification for amendment	ERG response
On pg. 92, the ERG states: [redacted]	It should be clarified that [redacted]	There is no contradiction because the data were not required to be collected. The reference to a contradiction undermines confidence in the execution of the study.	This is not a factual inaccuracy. Some data on post-HSCT relapse was presented in the company’s response to the ERG’s points for clarification, despite the company stating that this information was not collected. The company response was not entirely

Description of problem	Description of proposed amendment	Justification for amendment	ERG response
[redacted]			clear, hence the ERG's comment that the information about the collection of EFS data appeared contradictory.

Issue 9 ACIC marking

Description of problem	Description of proposed amendment	Justification for amendment	ERG response
The AIC marking in Table 2 on pg. 41 does not reflect that in the company submission.	Data corresponding to the following items should be marked AIC: • Mean age, years • Race: Black or African American, Asian, American Indian or Alaska native, Other • Haemoglobin • Mean bone marrow blast • Extra medullary disease • Genetic mutations: NPM1 mutated, CEBPA mutated	Amendment ensures consistency in AIC marking.	AIC markings have been added.

Issue 10 Figure labelling

Description of problem	Description of proposed amendment	Justification for amendment	ERG response
Table 8 on pg. 60 is incorrectly labelled “Table 8 Serious adverse events (>3% frequency) summary (CS Tables 15 & 18)”; an additional column has been added to the table, with data calculated by the ERG.	“Table 8 Serious adverse events (>3% frequency) summary (CS Tables 15 & 18)” should be amended to “Table 8 Serious adverse events (>3% frequency) summary (adapted from CS Tables 15 & 18)”	Amendment ensures consistency in citing.	Amended
Table 12 on pg. 66 is incorrectly labelled “Table 12 Overview of UK economic evaluations (CS Table 23, p82)”; information in the table has been reordered and some information added/reworded by the ERG.	“Table 12 Overview of UK economic evaluations (CS Table 23, p82)” should be amended to “Table 12 Overview of UK economic evaluations (adapted from CS Table 23, p82)”	Amendment ensures consistency in citing.	Amended
Table 24 on pg. 97 is incorrectly labelled “Table 24 Summary of health state utility values (CS Table 32, pg. 102)”; this table also includes information from CS Table 33, p4. 104.	“Table 24 Summary of health state utility values (CS Table 32, pg. 102)” should be amended to “Table 24 Summary of health state utility values (CS Table 32 and 33, pg. 102-104)”	Amendment ensures consistency in citing.	Amended
Table 29 on pg. 108 is incorrectly labelled “Table 29 Summary of costs associated with adverse events (CS Table 47, pg. 126)”; a column has been removed from	“Table 29 Summary of costs associated with adverse events (CS Table 47, pg. 126)” should be amended to “Table 29 Summary of costs associated with adverse events (adapted from CS Table 47, pg. 126)”	Amendment ensures consistency in citing.	Amended

the table by the ERG.			
Table 33 on pg. 111 is incorrectly labelled “Table 33 Probabilistic sensitivity analysis results (CS Table 54, pg. 138)”; columns have been removed from the table by the ERG.	“Table 33 Probabilistic sensitivity analysis results (CS Table 54, pg. 138)” should be amended to “Table 33 Probabilistic sensitivity analysis results (adapted from CS Table 54, pg. 138)”	Amendment ensures consistency in citing.	Amended
Table 34 on pg. 114 is incorrectly labelled “Table 34 Scenario analysis results (CS Table 60-68, pg. 143-145)”; the table has been reordered by the ERG.	“Table 34 Scenario analysis results (CS Table 60-68, pg. 143-145)” should be amended to “Table 34 Scenario analysis results (adapted from CS Table 60-68, pg. 143-145)”	Amendment ensures consistency in citing.	Amended
Figure 6 on pg. 48 is incorrectly labelled “Figure 6 OS in patients who achieve remission and receive HSCT”; this figure should be cited to the clarification response provide by the company.	“Figure 6 OS in patients who achieve remission and receive HSCT” should be amended to “Figure 6 OS in patients who achieve remission and receive HSCT (source: clarification response Figure 6, pg. 18)”	Amendment ensures consistency in citing.	Amended
Figure 7 on pg. 48 is incorrectly labelled “Figure 7 OS in patients who do not achieve remission and receive HSCT”; this figure should be cited to the clarification response provide by the company.	“Figure 7 OS in patients who do not achieve remission and receive HSCT” should be amended to “Figure 7 OS in patients who do not achieve remission and receive HSCT (source: clarification response Figure 10, pg. 20)”	Amendment ensures consistency in citing.	Amended
Figure 8 on pg. 49 is incorrectly labelled “Figure 8 OS in patients who achieve remission and do not receive HSCT”; this figure should be cited to the clarification response provide by the company.	“Figure 8 OS in patients who achieve remission and do not receive HSCT” should be amended to “Figure 8 OS in patients who achieve remission and do not receive HSCT (source: clarification response Figure 8, pg. 19)”	Amendment ensures consistency in citing.	Amended

Figure 9 on pg. 49 is incorrectly labelled “Figure 9 OS in patients who do not achieve remission or receive HSCT”; this figure should be cited to the clarification response provide by the company.	“Figure 9 OS in patients who do not achieve remission or receive HSCT” should be amended to “Figure 9 OS in patients who do not achieve remission or receive HSCT (source: clarification response Figure 12, pg. 21)”	Amendment ensures consistency in citing.	Amended
Figure 11 on pg. 51 is incorrectly labelled “Figure 11 EFS in patients who achieve remission and receive HSCT”; this figure should be cited to the clarification response provide by the company.	“Figure 11 EFS in patients who achieve remission and receive HSCT” should be amended to “Figure 11 EFS in patients who achieve remission and receive HSCT (source: clarification response Figure 7, pg. 18)”	Amendment ensures consistency in citing.	Amended
Figure 12 on pg. 52 is incorrectly labelled “Figure 12 EFS in patients who do not achieve remission and receive HSCT”; this figure should be cited to the clarification response provide by the company.	“Figure 12 EFS in patients who do not achieve remission and receive HSCT” should be amended to “Figure 12 EFS in patients who do not achieve remission and receive HSCT (source: clarification response Figure 11, pg. 20)”	Amendment ensures consistency in citing.	Amended
Figure 13 on pg. 52 is incorrectly labelled “Figure 13 EFS in patients who achieve remission and do not receive HSCT”; this figure should be cited to the clarification response provide by the company.	“Figure 13 EFS in patients who achieve remission and do not receive HSCT” should be amended to “Figure 13 EFS in patients who achieve remission and do not receive HSCT (source: clarification response Figure 9, pg. 19)”	Amendment ensures consistency in citing.	Amended
Figure 14 on pg. 53 is incorrectly labelled “Figure 14 EFS in patients who do not achieve remission or receive HSCT”; this figure should be cited to the clarification response provide by the company.	“Figure 14 EFS in patients who do not achieve remission or receive HSCT” should be amended to “Figure 14 EFS in patients who do not achieve remission or receive HSCT (source: clarification response Figure 13, pg. 22)”	Amendment ensures consistency in citing.	Amended

Figure 19 on pg. 113 is incorrectly labelled “[redacted] (Error corrected) (CS Figure 19, pg. 143)”; this figure has been substantially adapted from the original.	“[redacted] (Error corrected) (CS Figure 19, pg. 143)” should be amended to: “[redacted] adapted from CS Figure 19, pg. 143”	Amendment ensures consistency in citing.	Amended

pharmacological interventions for high-risk AML in patients who may be eligible for CPX-351 in clinical practice.

Three randomised controlled trials (RCTs) met the eligibility criteria for the systematic review. However, two of the trials (Study 204 investigating CPX-351 and a study investigating two differing daunorubicin doses in the DA 3+7 regimen) were not described in the CS. No justification was given for excluding them. However, the ERG considers that it was reasonable for the CS to focus on Study 301.

Study 301 suggests that compared with DA 3+7, CPX-351 is associated with a significant improvement in overall survival (OS) (median OS: 9.56 months [95% CI: 6.60, 11.86] vs 5.95 months [95% CI: 4.99, 7.75]; hazard ratio (HR)=0.69 [95% CI: 0.52, 0.90], p=0.005) in patients with high-risk AML. Although results from the subgroup analyses should be interpreted with caution, there was some evidence to suggest that CPX-351 had a less beneficial impact on OS in the stratified subgroup of patients with 'myelodysplastic syndrome (MDS) AML with prior treatment with hypomethylating agents (HMA)'

[redacted]

[redacted]These patients constituted around a third of patients in the trial and a similar proportion of those who would be eligible for CPX-351 in clinical practice.

Overall CPX-351 may be a more effective bridge to haematopoietic stem cell transplant (HSCT) compared with 3+7 in patients with high-risk AML aged 60-75 years. Compared with 3+7, the proportion of patients undergoing HSCT was higher in the CPX-351 group (34.0% [52/153] vs 25.0% [39/156]) although the difference was not statistically significant (OR=1.54 [95% CI: 0.92, 2.56]). OS in patients who underwent HSCT was significantly greater with CPX-351: at the point of data cut the median OS was not reached in the CPX-351 group, and the median OS in the 3+7 group was 10.25 months (95% CI: 6.21, 16.69) (HR=0.46 [95% CI: 0.24, 0.89], p=0.0046). A substantial number of patients were censored in the CPX-351 arm and there were small numbers of patients in the tail of the survival curves. The company clarified that safety data indicated a number of deaths were known to have occurred since the 2015 data cut used for the primary analyses ([redacted] patients in the CPX-351 arm, and [redacted] patients in the 3+7 arm), although they noted that as this is a safety update rather than a formal survival analysis, these results only capture known deaths currently documented and may not include all actual deaths.

Overall the safety profiles of CPX-351 and 3+7 appear broadly comparable. The overall incidence of observed Grade 3-5 adverse events (AEs) was similar across groups. Although potentially concerning, the higher incidence in observed treatment-related serious AEs in CPX-351-treated

patients may be largely explained by the higher number of cycles and longer duration of treatment in the CPX-351 arm compared with 3+7.

- Utility estimate for patients in the post-transplant remission health state based on the 0.75 value for patients in durable remission, further adjusted for age,
- Equivalent quality of life for CPX-351 and 3+7 patients while on induction and consolidation treatment,
- Vial usage reflecting the distribution of body surface area, and the mean body surface area re-weighted to reflect the gender distribution in Study 301,
- Reduced number of hospital days during the consolidation period,
- Provision of unrelated donor stem cells excluded from the costs of transplant.

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.

Table 1 Summary of ERG exploratory analyses

Scenarios	Treatments	Total costs	Total QALYs	Incremental costs	Incremental QALYs	ICER (incremental £/QALY)	Change in ICER (%)
Company base-case (including ERG corrections)	3+7	[redacted]	[redacted]	-	-	-	-
	CPX-351	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	-
10% of patients under 60 years	3+7	[redacted]	[redacted]	-	-	[redacted]	-
	CPX-351	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
30% of patients under 60 years	3+7	[redacted]	[redacted]	-	-	[redacted]	[redacted]
	CPX-351	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
Post-transplant OS in CPX-351 arm: Weibull	3+7	[redacted]	[redacted]	-	-	[redacted]	[redacted]
	CPX-351	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
Post-transplant OS in CPX-351 arm: Log-logistic	3+7	[redacted]	[redacted]	-	-	[redacted]	[redacted]
	CPX-351	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
Post-transplant OS in CPX-351 arm: Log-normal	3+7	[redacted]	[redacted]	-	-	[redacted]	[redacted]
	CPX-351	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
Post-transplant OS in CPX-351 arm: Exponential	3+7	[redacted]	[redacted]	-	-	[redacted]	[redacted]
	CPX-351	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
Post-transplant OS in CPX-351 arm: Generalised gamma	3+7	[redacted]	[redacted]	-	-	[redacted]	
	CPX-351	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
	3+7	[redacted]	[redacted]	-	-	-	

Scenarios	**Treatments**	**Total costs**	**Total QALYs**	**Incremental costs**	**Incremental QALYs**	**ICER (incremental £/QALY)**	**Change in ICER (%)**
Post-transplant OS in CPX-351 arm: Combining treatments (Gompertz)	CPX-351	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
Weighted OS curve using AICC weights	3+7	[redacted]	[redacted]	-	-	[redacted]	[redacted]
	CPX-351	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
Weighted OS curve using BIC weights	3+7	[redacted]	[redacted]	-	-	[redacted]	[redacted]
	CPX-351	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
Post-transplant outcomes are based on OS only	3+7	[redacted]	[redacted]	-	-	[redacted]	[redacted]
	CPX-351	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
Utility value of 0.75 for durable remission	3+7	[redacted]	[redacted]	-	-	[redacted]	[redacted]
	CPX-351	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
General population utility value (0.79) for durable remission	3+7	[redacted]	[redacted]	-	-	[redacted]	[redacted]
	CPX-351	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
Post-transplant remission utility, adjusted for aging	3+7	[redacted]	[redacted]	-	-	[redacted]	[redacted]
	CPX-351	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
Mean on-treatment utility during induction and consolidation treatment	3+7	[redacted]	[redacted]	-	-	[redacted]	[redacted]
	CPX-351	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
Vial usage reflecting the distribution of body surface area	3+7	[redacted]	[redacted]	-	-	[redacted]	[redacted]
	CPX-351	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
Reduced number of hospital days during consolidation period	3+7	[redacted]	[redacted]	-	-	[redacted]	[redacted]
	CPX-351	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
Alternative cost of transplant	3+7	[redacted]	[redacted]	-	-	[redacted]	[redacted]
	CPX-351	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
ERG alternative base-case analysis	3+7	[redacted]	[redacted]	-	-	[redacted]	[redacted]
	CPX-351	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]

AICC, Akaike information criterion corrected; BIC, Bayesian information criterion; ERG, Evidence Review Group; ICER, incremental cost-effectiveness ratio; QALYs, quality-adjusted life years

The ERG base-case analysis estimated CPX-351 to be more costly (cost difference [redacted]) and more effective ([redacted] QALY gain) compared with 3+7, and suggests that the ICER for CPX-351 compared with 3+7 is [redacted] per QALY.

The ERG also carried out a further series of exploratory subgroup analyses to explore the impact of alternative post-transplant OS and inclusion of patients under the age of 60. The ICER of CPX-351 vs. 3+7 varied between [redacted] (most favourable post-transplant OS) and [redacted] (least favourable post-transplant OS for CPX-351). The ICER decreased to [redacted] when 10% of patients were under the age of 60 and to [redacted] when 30% of patients were under the age of 60.

Pages 42-43

Table 2 Baseline characteristics of patients in Study 301, ITT population (CS Table 8, page 43)

Characteristic	CPX-351	3+7
Number of patients (n)	153	156
Mean age (SD), years	[redacted]	[redacted]
Age, n (%) 60-69 years 70-75 years	 96 (62.7) 57 (37.3)	 102 (65.4) 54 (34.6)
Male sex, n (%)	94 (61.4)	96 (61.5)
Race, n (%) White Black or African American Asian American Indian or Alaska native Other	[redacted]	[redacted]
Median weight (range), kg	82.0 (49.0, 134.0)	82.7 (46.0, 136.0)
Median height (range), cm	170.2 (149.0, 198.0)	170.2 (149.0, 189.0)
Median BSA (range), m^2	2.0 (1.0, 3.0)	2.0 (1.0, 3.0)
ECOG performance group, n (%) PS = 0 PS = 1 PS = 2 PS ≥ 3	 37 (24.2) 101 (66.0) 15 (9.8) 0	 45 (28.8) 89 (57.1) 22 (14.1) 0
Cytogenetic risk, n (%)[a] Favourable Intermediate Adverse Unknown	f 7 (4.6) 64 (41.8) 72 (47.8) 10 (6.5)	 5 (3.2) 58 (37.8) 83 (53.2) 10 (6.4)
White blood cell count, n (%) < 20 x 10^9/L ≥ 20 x 10^9/L Unknown	 131 (85.6) 22 (14.4) 0	 131 (84.0) 24 (15.4) 1 (0.6)
Platelet count, n (%) ≤ 50 x 10^9/L > 50 x 10^9/L Unknown	 95 (62.1) 58 (37.9) 0	 91 (58.3) 63 (40.4) 2 (1.3)
Haemoglobin, n (%) ≤ 9 g/dL > 9 g/dL Unknown	[redacted]	[redacted]
Mean bone marrow blast (SD) Aspirate[b] Biopsy[c]	[redacted]	[redacted]
Extra medullary disease, n (%)	[redacted]	[redacted]

Characteristic	CPX-351	3+7
AML subtype, n (%)		
Therapy related	30 (19.6)	33 (21.2)
MDS with prior HMA	50 (32.7)	55 (35.3)
MDS without prior HMA	21 (13.7)	19 (12.2)
CMMoL	11 (7.2)	12 (7.7)
de novo with MDS karyotype	41 (26.8)	37 (23.7)
Genetic mutations	[redacted]	[redacted]
FLT3 mutated		
NPM1 mutated		
CEBPA mutated		
Prior and concomitant medication, n (%)[d]	153 (100.0)	151 (100.0)
Prior anthracycline exposure	6 (3.9)	4 (2.6)

[a] Cytogenetic risk status was based on National Comprehensive Cancer Network (NCCN) guidelines for AML. [b] Mean bone marrow blast (aspirate) values based on n=141 patients in CPX-351 group and n=141 patients in 3+7 group. [c] Mean bone marrow blast (biopsy) values based on n=64 patients in CPX-351 group and n=60 patients in 3+7 group. [d] Prior and concomitant medication was assessed in the safety analysis population i.e. n=153 in the CPX-351 group and n=151 in the 3+7 group. AML, acute myeloid leukaemia; BSA, Body surface area; CMMoL, chronic myelomonocytic leukaemia; ECOG, Eastern Cooperative Oncology Group; HMA, hypomethylating agent; ITT, intent-to-treat; MDS, myelodysplastic syndrome; PS, performance status; SD, standard deviation

Study 301 was limited to high-risk AML patients between 60 and 75 years old, who may have poorer prognosis and may be harder to treat than the population of patients under 60 years old. The standard of care in UK practice for older patients (DA 3+10, 3+8, 2+5) is similar to the DA regimen used in the trial, but may be different in younger patients, who are less likely to have comorbidities and are more able to tolerate intensive chemotherapy.

Whilst the ERG does not have significant concerns about the generalisability of the trial population to the population of patients with high-risk AML aged between 60 and 75 in NHS practice, the generalisability of the trial results to the population of patients under 60 years is uncertain. As highlighted in Section 3.1 above, *de novo* AML with MDS associated karyotypic changes may not be diagnosed as 'high-risk' prior to receiving cytogenetic test results.

4.2.2 Summary of the quality of Study 301

Results of the quality assessment for Study 301 are presented in CS Table 10, with more detailed rationale for decisions in CS Appendix (Table 6).

Randomisation was performed using an interactive telephone or internet-based randomisation system. Patients were randomly assigned to treatment using a dynamic balancing randomisation algorithm to ensure a balanced distribution of the stratification variables (age and AML subtype) between the two

Page 48-49

[redacted]

[redacted]

[redacted]

[redacted]

[redacted]

[redacted]

[redacted]

[redacted]

[redacted]

[redacted]

[redacted]

[redacted]

[redacted]

[redacted]

[redacted]

[redacted]

[redacted]

[redacted]

[redacted]

[redacted]

[redacted]

[redacted]3[redacted]

Error! Reference source not found.[redacted]

[redacted]4[redacted]

Event-free survival

Pages 51-54

[redacted]

For the sake of completion and as these results informed the economic evaluation, they are presented below.

[redacted]

5

[redacted]

5

[redacted]

[redacted]

6

[redacted]

[redacted]

[redacted]

[redacted]

[redacted]

[redacted]

[redacted]

[redacted]

[redacted]

[redacted]

[redacted]

[redacted]

[redacted]

[redacted]

[redacted]

[redacted]

[redacted]

[redacted]

[redacted]8[redacted]

[redacted]

[redacted]

Response

Table 3 shows that a significantly greater proportion of patients in the CPX-351 arm achieved an induction response. OR adjusted for age and AML subtype were reported for CR+CRi and CR, but not for CRi alone or non-response. The ERG calculated unadjusted OR using the Mantel-Haenszel test and found no statistically significant difference in CRi between the two arms, although these results should be interpreted with caution due to the relatively small number of events (unadjusted OR 1.40 [95% CI 0.64, 3.07]). Consistently with the CR+CRi results, the difference in proportion of non-responders was statistically significant (unadjusted OR 0.52 [95% CI 0.32, 0.82]).

Table 3 Proportion of patients with an induction response, ITT population (CS Table 11)

Endpoint, n (%)	CPX-351 (n=153)	3+7 (n=156)	Odds ratio (95% CI)[a]
CR+CRi	73 (47.7)	52 (33.3)	1.77 (1.11, 2.81), p=0.016
CR	57 (37.3)	40 (25.6)	1.69 (1.03, 2.78), p=0.040
CRi	16 (10.5)	12 (7.7)	NR
No response	80 (52.3)	104 (66.7)	NR

[a] Odds ratios were calculated using the 3+7 group as the reference. The resultant p-value is from a comparison of rates between treatment and is based on the Mantel-Haenszel test stratified by age and AML type groups
CI, confidence interval; CR, complete remission; CRi, complete remission with incomplete platelet or neutrophil recovery; NR, not reported

[redacted]

Table 4 Serious adverse events (>3% frequency) summary (adapted from CS Tables 15 & 18)

	CPX-351 (n=153)	3+7 (n=151)	All (n=304)
Preferred term	**n (%)**	**n (%)**	**n (%)**
Any serious adverse events	90 (58.8)	65 (43.0)	155 (51.0)
Febrile neutropenia	12 (7.8)	8 (5.3)	20 (6.6)
Respiratory failure	11 (7.2)	8 (5.3)	19 (6.3)
Ejection fraction decreased	9 (5.9)	9 (6.0)	18 (5.9)
Sepsis	12 (7.8)	5 (3.3)	17 (5.6)
Pneumonia	10 (6.5)	6 (4.0)	16 (5.3)

The CS stated that the higher observed rate of serious adverse events may be partly due to the greater proportion of patients in the CPX-351 arm who received consolidation in the outpatient setting (51% in Cycle 1 and 61% in Cycle 2) compared with the 3+7 arm (6% in Cycle 1 and 0% in Cycle 2), as a move to the hospital setting is one of the criteria for classifying an adverse event as serious. However, any prolongation of existing hospitalisation in either arm would also qualify as a serious adverse event, therefore the ERG are unconvinced that the difference in observed serious AEs may be explained by the higher rate of CPX-351 patients receiving treatment in the outpatient setting.

The company stated that the difference in observed serious adverse events may also be explained by the higher cumulative exposure to treatment in the CPX-351 group compared with the 3+7 arm. In the consolidation phase, a greater proportion of patients in the CPX-351 group received both an initial and second consolidation compared with the 3+7 treatment arm (32% vs. 21% for the first consolidation and 15% vs. 7.9% for the second consolidation). The median length of the treatment exposure and length of treatment phase [redacted] Following the ERG's request for clarification, the company provided further results from a more conventional Poisson distribution adjusting for exposure duration which yielded comparable mean estimates between CPX-351 and 3+7 [redacted] The company stated the 95% confidence interval around the difference in the estimates included zero and concluded that there is no real difference in the SAE rates between the two study groups.

Based on these additional analyses the ERG's view is that although potentially concerning, the difference in observed treatment-related serious AEs between the study arms may be largely

[95% CI: 0.52, 0.90], p=0.005) in patients with high-risk AML. Although results from the subgroup analyses should be interpreted with caution, there was some evidence to suggest that CPX-351 had a less beneficial impact on OS in the stratified subgroup of patients with '$_{MDS}$AML with prior treatment with HMA' [redacted] These patients constituted around a third of patients in the trial and a similar proportion of those who would be eligible for CPX-351 in clinical practice.

Overall CPX-351 may be a more effective bridge to stem cell transplant compared with 3+7 in patients with high-risk AML aged 60-75 years. Compared with 3+7, the proportion of patients undergoing HSCT was higher in the CPX-351 group (34.0% [52/153] vs 25.0% [39/156]) although the difference was not statistically significant (OR=1.54 [95% CI: 0.92, 2.56]) and the decision to transplant was not blinded to treatment allocation. OS in patients who underwent HSCT was significantly greater with CPX-351: at the point of data cut the median OS was not reached in the CPX-351 group, and the median OS in the 3+7 group was 10.25 months (95% CI: 6.21, 16.69) (HR=0.46 [95% CI: 0.24, 0.89], p=0.0046). However, the OS results post-HSCT should be subject to caution given the small number of patients, limited follow-up duration, extensive censoring and lack of randomisation.

Overall the safety profiles of CPX-351 and 3+7 appear broadly comparable. The overall incidence of observed Grade 3-5 adverse events was similar across groups. Although potentially concerning, the higher incidence in observed treatment-related serious AEs in CPX-351-treated patients may be largely explained by the higher number of cycles and longer duration of treatment in the CPX-351 arm compared with 3+7.

The relative impact of CPX-351 vs. 3+7 on HRQL is unknown as HRQL and utility data were not collected. Due to the lack of up-to-date follow-up data and the associated substantial censoring of patients there is significant uncertainty about the longer-term efficacy and safety of CPX-351 including after stem cell transplant. The lack of evidence in patients under 60 years with high-risk AML means that the applicability of the trial results to this patient group is uncertain.

Table 5 Overview of UK economic evaluations (adapted from CS Table 23, p82)

Study	Patient population	Treatment and comparator	Model description	Estimated ICER
NICE TA10124 – 2018 [19 20]	Newly diagnosed FTL3-mutation positive AML	Midostaurin versus standard of care	Partitioned survival model with lifetime horizon	NR
NICE TA218 -2011 [21]	AML patients with 20-30% marrow blasts, not eligible for HSCT	Azacitidine versus CCR, BSC, low-dose chemotherapy	Partitioned survival model with lifetime horizon	Against BSC: £63,177 Against low-dose chemotherapy: £49,030 Against CCR: £59,954
Tremblay *et al.*, 2017 [22]	Newly diagnosed FTL3-mutation positive AML	Midostaurin (in combination with SOC) versus SOC	Partitioned survival model with lifetime horizon	£34,327
Tikhonova *et al.*, 2017 [23]	AML patients 65+ with greater than 30% marrow blasts, not eligible for HSCT	Azacitidine versus intensive chemotherapy with anthracycline in combination with cytarabine; non-intensive chemotherapy with low-dose cytarabine; and versus BSC only	Partitioned survival model with lifetime horizon	£273,308
Wang et al., 2014 [24]	Newly diagnosed AML	Induction chemotherapy (ADE - cytarabine, daunorubicin, etoposide; DA - cytarabine, daunorubicin) versus no induction	Probabilistic decision model with lifetime horizon	NR
BSC, best supportive care; CCR, conventional care regimen, HSCT, haematopoietic stem cell transplantation; ICER, incremental cost-effectiveness ratio; NR, not reported; QALYs, quality-adjusted life years; SOC, Standard of care				

5.1.4 Conclusions of the cost-effectiveness review

Five cost-effectiveness studies were identified and considered relevant for the cost-effectiveness review. However, none of these evaluated the cost-effectiveness of CPX-351. 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 **Error! Reference source not found.**.

Second induction: Daunorubicin 44 mg/m^2 and cytarabine 100 units/m^2 IV infusion over 90 minutes on Days 1 and 3.	Second induction: 100 mg/m^2/day of cytarabine administered by continuous infusion over five days 60 mg/m^2/day of daunorubicin given over 15 minutes on Days 1 and 2.
Consolidation: Daunorubicin 29 mg/m^2 and cytarabine 65 mg/m^2 IV infusion over 90 minutes on Days 1 and 3.	Consolidation: 100 mg/m^2/day of cytarabine administered by continuous infusion over five days 60 mg/m^2/day of daunorubicin given over 15 minutes on Days 1 and 2.
IV, intravenous; m, metres; mg, milligrams	

ERG comment

As previously highlighted in the decision problem section (Section **Error! Reference source not found.**), there are potential differences in the dosing used in UK clinical practice for the cytarabine and daunorubicin (3+7) chemotherapy regimen. The model was based on the US 3+7 regimen as used in the clinical trial, rather than the typical 3+10 regimen used in the UK. However, the ERG considered that these two regimens could be considered equivalent in terms of efficacy based on the guidance provided from the British Committee for Standards in Haematology [6] and clinical advice received by the ERG, and so it was likely to be sufficiently representative of the current standard of care in the UK.

A number of comparators included in the NICE scope were not included in the company's evaluation (Section **Error! Reference source not found.**). The ERG was satisfied with the rationale for their exclusion. However, patients under the age of 60, who were younger than those enrolled in Study 301, may receive FLAG-Ida as standard care. The company did not submit any cost-effectiveness evidence for CPX-351 compared with FLAG-Ida.

Further to the above, the ERG also notes that the draft marketing authorisation for CPX-351 permits patients to receive up to four courses of consolidation[1], and that treatment can be continued as long as the patient benefits or until disease progression. This is inconsistent with the pivotal trial in which patients were restricted to a maximum of two rounds of consolidation therapy. Given this inconsistency, the ERG felt that it was uncertain how many patients would receive third and fourth consolidation courses in practice. Evidence from Study 301 suggests that few patients would continue on to third and fourth round consolidation therapy as only [redacted] of patients received a second round of consolidation therapy (Table 13, CSR) [35]. The clinical advisor to the ERG also confirmed that the

[redacted]

number of courses of consolidation would not be likely to be above 2 or 3, since clinicians would aim to bridge patients to transplant as soon as possible. The number of patients who receive further

Probability of receiving transplant, after remission	[redacted]	[redacted]	[redacted]	[redacted]

These analyses used age and high-risk AML subtype as covariates, although these were not significant predictors in any of the analyses (except for age in the patients who received a second round of induction therapy). For the probability of receiving consolidation therapy and receiving transplant, AML subtype was grouped into two categories (one category consisting of MDS AML with or without prior HMA, and the remaining three in the other category).

The probability of remission also included a treatment effect because the higher rate of remission in CPX-351 patients than 3+7 patients observed in Study 301 was statistically significant. In the other analyses, the same probability was applied regardless of treatment arm as the trial found no statistically significant differences between arms. However, there were some small observed differences between treatment arms in the trial. For example, the observed probability of receiving transplant was somewhat higher in CPX-351 patients ([redacted] than in 3+7 patients ([redacted], and a lower proportion of 3+7 patients received two rounds of consolidation compared to CPX-351.

The regression analyses for all probabilities also included the number of induction courses (except the probability of receiving second round of induction therapy) as covariates.

ERG comment

The ERG is largely satisfied with the approach taken to estimate the probabilities, but has concerns about the combined analysis for treatment arms and inconsistent parameterisation in the multivariate analyses.

None of the analyses (except remission after induction) included treatment arms as a prognostic factor and therefore the variation by treatment arms may not be fully reflected in the analysis. Particularly, in the analysis for the patients who achieved remission and received transplant, the observed percentage was higher in CPX-351 patients ([redacted]) than in 3+7 patients ([redacted]); however, the treatment arms were combined in the regression analysis giving combined 51% probability for the patients who achieved remission and received transplant. The combined estimate slightly underestimates the probability of transplant in CPX-351 arm and overestimates in the 3+7 arm. Similarly, in the analysis of patients who receive consolidation, the observed second consolidation was lower in the 3+7 arm ([redacted]) compared to the CPX-351 arm ([redacted]). While the company's approach may provide less accurate predictions, the ERG considers that it may be clinically plausible that each group is subject to a similar rate of transplantation.

Non-responder OS, time to transplant or death, time to transplant or progression or death

In the analyses among patients who do not achieve remission, the KM curves for both CPX-351 and standard care were combined, justified by the lack of a detected treatment effect. Within each analysis, the survival models displayed little variation regarding their long-term predictions. Predicted median times from each analysis were comparable with the observed KM data. For time to transplant or death, a very small number of patients were included in the analysis (a total of [redacted], [redacted] in CPX-351 and [redacted] in 3+7).

ERG comment

The ERG has significant concerns related to survival analyses and extrapolation beyond the trial period, which are discussed in turn below.

Post-transplant OS for CPX-351

Immature data and high censoring

There are a substantial number of patients [redacted]) who were censored in the CPX-351 arm after 1 year in the post-transplant OS analysis (CS response to clarification, Figure 3) and very small numbers of patients remained after 84 weeks ([redacted] patients in the CPX-351 group and only [redacted] patient in the 3+7 group). The ERG, therefore, has concerns that the available data were too immature to robustly estimate the survival benefit for post-transplant patients.

In response to the ERG's request for clarification, the company provided additional data on deaths captured through adverse event monitoring that have occurred since the 2015 data cut used for the survival analyses ([redacted] deaths on CPX-351 and [redacted] on 3+7, compared with [redacted] and [redacted] deaths for CPX-351 and 3+7 respectively in the original safety analysis, and 104 and 132 for CPX-351 and 3+7 respectively in the ITT analysis on which the survival analyses in this section were based). While the transplant status of these patients was not known, the ERG considered that it could be plausible that the additional patients who died after the data cut-off time point could have been in the post-transplant group. These additional deaths were not included in the current analysis. The higher number of deaths in the CPX-351 arm might suggest a degree of convergence in the OS curves.

Variability of long-term predictions

While the model is relatively robust to the choice of parametric curve for most outcomes, the parametric curve used to extrapolate post-transplant OS has a very significant impact on predicted OS gains and cost-effectiveness, with mean OS ranging from [redacted] years with the exponential curve to

Page 97

Disutilities for second-line therapy were captured within the utility value for the progression health state. The symptoms relating to this health state included anaemia, bleeding risk, infection risk, fatigue and shortness of breath, and treatment included weekly blood transfusions, and treatment with antibiotics. A disutility relating to treatment with HMA (azacitidine) was estimated by the company which was associated with a lower utility value than for the progressed health state, but was not used in the economic model.

Health state utility values and their application in the model are summarised in Table 6.

Table 6 Summary of health state utility values (CS Table 32 and 33, pp 102-104)

Health state	Utility value: mean (SE)
Health state utility values	
AML (induction) [1]	0.550 (0.023)
Remission (post-induction/consolidation) [2]	0.656 (0.021)
[redacted] [3]	[redacted]
[redacted] [4]	[redacted]
Utility decrements	
3+7: Disutility of an induction cycle	[redacted]
3+7: Disutility of a consolidation cycle (inpatient)	[redacted]
3+7: Disutility of a consolidation cycle (outpatient)	[redacted]
CPX-351: Disutility of an induction cycle	[redacted]
CPX-351: Disutility of a consolidation cycle (inpatient)	[redacted]
CPX-351: Disutility of a consolidation cycle (outpatient)	[redacted]
Disutility of a transplant	[redacted]

SE, standard error; AML, acute myeloid leukaemia; BSC, best supportive care.

[1] Applied to those non-responders prior to progression or transplant. [2] Applied to patients in remission after consolidation therapy (who do not receive transplant). [3] Applied to those achieving remission after transplant (for responders and non-responders). [4] Applied to those experience relapse after transplant, after consolidation, or progression after a non-response.

Scenario analysis

The systematic literature review (SLR) conducted by the company did not identify any studies that were considered suitable for inclusion in the economic model as they did not provide a consistent set of estimates for stages in the AML treatment pathway. Of those studies that provided utility values estimated by EQ-5D, the preferred instrument according to the NICE reference case, none were considered by the company to be generalizable to the population of the decision problem.

guidelines. In the absence of any definitive guidelines for the treatment of AML patients in the UK, the ERG felt there was remaining uncertainty on the type of transplant that is typically provided, and whether the cost of providing unrelated adult stem cells should be included in the transplant cost. The ERG was also unclear on whether the service cost of providing unrelated stem cells should be included in the total transplant cost, even if there are some transplants using unrelated donor stem cells. Previous AML models submitted to NICE have not included this cost. The ERG understands that two of these registries in the UK are charities, and that their running costs may not fall to the NHS. It was also unclear whether the costs estimated by NHS Blood and Transplant would actually be incurred in practice.

After drug acquisition costs, the cost of transplantation is the largest component of the total cost of the treatment pathway (**Error! Reference source not found.**), and the cost of providing unrelated adult stem cells is a substantial proportion of this cost. Given the uncertainty discussed in this section, the ERG explored the impact of reducing this cost in Section 6.

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 (Section 5.2.6.5). The weighted cost of AEs was similar between CPX-351 and 3+7, although slightly higher for CPX-351, with CPX-351 associated with a weighted cost of £[redacted] and 3+7 associated with a weighted cost of £[redacted].

This cost was applied as a one-off cost in the model. No AEs were associated with subsequent treatment costs. Unit costs were extracted from NHS Reference Costs, and were a weighted average of elective inpatient, non-elective long-stay and non-elective short-stay [29].

Table 7 Summary of costs associated with adverse events (adapted from CS Table 47, pg. 126)

Adverse Event	Cost	Source (HRG Code)
Event		
Bacteremia	£1,895	WJ06
Diarrhoea	£1,354	FD10
Ejection Fraction Decreased	£1,837	EB03
Fatigue	£875	WH17
Febrile Neutropenia	£1,727	SA35
Hypertension	£593	EB04Z
Hypotension	£1,730	EB14
Hypoxia	£1,847	DZ27
Pneumonia	£1,698	DZ11

Table 8 Total QALYs, by health state

Health state	**3+7**	**CPX-351**
Event-free QALYs	[redacted]	[redacted]
Accrued prior to transplant	[redacted]	[redacted]
Post-consolidation remission	[redacted]	[redacted]
Post-transplant remission	[redacted]	[redacted]
Post-progression QALYs	[redacted]	[redacted]
QALY decrements	[redacted]	[redacted]
Attributable to induction and consolidation	[redacted]	[redacted]
Attributable to transplant	[redacted]	[redacted]
Total QALYs	[redacted]	[redacted]
QALY, quality-adjusted life year.		

5.2.9.2 Sensitivity analysis

The results presented in this section refer to that without the CPX-351 PAS applied, and can be directly compared with the base-case ICER of £[redacted]. Results with PAS applied are provided in the confidential appendix to the ERG report.

Probabilistic sensitivity analysis

The company undertook a probabilistic sensitivity analysis (PSA) to explore and quantify uncertainty in the outcomes of the analysis. Probabilistic results were estimated from 500 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 [redacted] per QALY. The probabilistic results were relatively similar to those estimated in the deterministic base-case.

Table 9 Probabilistic sensitivity analysis results (adapted from CS Table 54, pg. 138)

Treatment	**Total mean costs (£)**	**Total mean QALYs**	**Mean incremental costs (£)**	**Mean incremental QALYs**	**ICER incremental (£/QALY)**
3+7	[redacted]	[redacted]	-	-	-
CPX-351	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
ICER, incremental cost-effectiveness ratio; QALYs, quality-adjusted life years					

The probability that CPX-351 is cost-effective compared with 3+7 was [redacted] at a threshold of £50,000 per QALY, while the probability was [redacted] at both a threshold of £20,000 and £30,000 per QALY (**Error! Reference source not found.**).

Pages 114-115

Figure 9 [redacted] (Error corrected) (adapted from CS Figure 19, pg. 143)

Scenario analyses

The company presented a range of scenario analyses within their base-case analysis. The results of the scenarios explored are presented in Table 10.

The results were notably most sensitive to variations in the time horizon. The ICER increases by [redacted]% when a five year time horizon was assumed (ICER [redacted]) and [redacted]% when ten years was assumed (ICER [redacted]). This highlights the difference in timings of the impact of treatment, where high-cost events (transplant, induction and consolidation therapy) were typically within one year of the model, and the QALY benefits occurred over the long-term.

Across the other set of scenarios explored, the ICER varied between a [redacted]% decrease from the base-case ICER (all patients treated with CPX-351 receive second induction and consolidation therapy as outpatients, with an ICER of [redacted]), to an increase of [redacted]% (adjustment of general population mortality for HSCT patients, with an ICER of [redacted]).

Table 10 Scenario analysis results (adapted from CS Table 60-68, pg. 143-145)

Scenario Analysis	Base-case	Incremental costs (£)	Incremental QALYs	ICER incremental (£/QALY)	Difference from Base-case ICERs
Base-case	-	[redacted]	[redacted]	[redacted]	-
Extension of model to younger population					
30% of modelled cohort <60 years of age	No extension to younger population	[redacted]	[redacted]	[redacted]	[redacted]
10% of modelled cohort <60 years of age	No extension to younger population	[redacted]	[redacted]	[redacted]	[redacted]
3+7 vs. 3+10 regimen					
3+10 regimen based on UK	3+7 Comparator	[redacted]	[redacted]	[redacted]	[redacted]
Percentage of patients treated with CPX-351 who receive second induction and consolidation therapy as outpatients					
Alter percentage of patients treated with CPX-351 who receive second induction and consolidation therapy as outpatients: in second induction, 50% of CPX-351 patients receive therapy in an outpatient setting and 100% of CPX-351 patients receive consolidation in an outpatient setting.	50% patients treated with CPX-351 who receive second induction and consolidation therapy as outpatients	[redacted]	[redacted]	[redacted]	[redacted]
Time horizon					
Five-year time horizon	Life time horizon (~30 year)	[redacted]	[redacted]	[redacted]	[redacted]
Ten-year time horizon	Life time horizon (~30 year)	[redacted]	[redacted]	[redacted]	[redacted]
Alternative source of utility scores					
Utilities sourced from Hensen *et al.*, 2017	Utilities sources from AML utility study	[redacted]	[redacted]	[redacted]	[redacted]
Mortality adjustment					
Adjust general population mortality for HSCT patients 2.3-fold	No general population mortality for HSCT patients	[redacted]	[redacted]	[redacted]	[redacted]
Transplant who do not response					
No transplantation in non-responders	Transplantation in non-responders	[redacted]	[redacted]	[redacted]	[redacted]
ICER, incremental cost-effectiveness ratio; LYs, life years; QALYs, quality-adjusted life years					

potentially accounting for up to 25% of the eligible patient population. Exclusion of these patients may have important implications on the cost-effectiveness of CPX-351, as evidence suggests that they are more likely to achieve remission, tend to receive a different standard of care to patients over 60 and are more likely to receive transplant due to being fitter. The benefits of curative therapy such as transplant will also be greater in a younger population due to their greater life expectancy.

5. *Immature data for post-transplant event-free and overall survival*

There are a significant number of patients (approximately [redacted]) who were censored in the CPX-351 arm after one year in the post-transplant OS analysis. This makes extrapolation of the survival curves highly uncertain. Additional evidence provided by the company on deaths that have occurred since the 2015 data cut used for the survival analyses suggests there may be additional deaths in the CPX-351 arm, which could result in some degree of convergence in the survival curves over time. It is, however, difficult to draw strong conclusions from this additional evidence because the transplant status of these patients is not known.

The immaturity of the survival data resulted in a large degree of variation in the predictions of life expectancy across the different post-transplant OS survival models. The company model applied the curve associated with the most optimistic projections, which included a plateau-effect after approximately two years. However the ERG considers that the trial data is too immature to confirm such a plateau in the survival data, and that the long-term survival projections are subject to a degree of uncertainty. Since the parametric curve used to extrapolate post-transplant OS has a very significant impact on predicted OS gains and cost-effectiveness, this results in significant uncertainty in the cost-effectiveness of CPX-351.

Related to the above, the ERG also noted an additional issue with the way EFS data was captured, which means that many patients [redacted] were censored for event-free survival, because they received a transplant after their last examination in the study before the data cut-off time point. As a result, the EFS analysis in transplant patients was based on a very small patient number. The EFS results used in the model are, therefore, an unreliable indication of the true treatment effect, and the company advised that no clinical inferences should be drawn from estimates of post-transplant EFS.

6. *Clinical plausibility of the projected post-transplant event-free and overall survival*

While a statistically significant difference in OS was demonstrated for the whole population, the ERG does not consider that the clinical data supports the predicted sustained benefits beyond the trial period (as discussed above). In addition, the ERG does not believe that the long-term survival projections for CPX-351 are clinically plausible. When validating the long-term survival projections, the company stated that between one and two years after transplant is when the majority of deaths

6.3.3 Post-transplant EFS

Given the concerns of the ERG regarding event-free survival in post-transplant patients, the ERG considered a scenario where this data was excluded from the model. For the transplant responder patients, the ERG explored the use of a two-state model, where patients are either in remission or are dead (informed only by the OS analysis). The ERG considers the OS analysis to provide more reliable predictions of patients after transplant (although it is associated with its own limitations, such as a lack of face validity in the long-term, which were explored in Section **Error! Reference source not found.**). This scenario is also consistent with the assumptions for patients who received transplant but did not achieve response (Section **Error! Reference source not found.**). Given that the prognosis of these high-risk patients who experience a relapse after receiving transplant is poor [37], the ERG considers that they would only spend a short amount of time in the relapsed health state before death, and so the bias that this scenario introduces is likely to be small and is associated with an overestimation in the number of QALYs. The ERG acknowledges that this is a simplifying assumption, but considers that the removal of the bias associated with the inclusion of the post-transplant EFS analysis outweighs these limitations. In this scenario, the ERG limited the duration of time whereby patients would accrue monitoring costs while in remission to six months.

In this scenario analysis, the ICER reduced to [redacted] (**Table 11**). This reduction in the ICER is because the changes mean that fewer patients are in the post-transplant relapse health state, and the number of QALYs increased due to the higher utility value of remission patients. While this impacts on both model arms the effect is larger in the CPX-351 arm, where overall survival is longer and there were more patients implicated in this analysis.

Table 11 ERG exploratory analysis: Post-transplant outcomes based on OS only

Technologies	Total costs	Total QALYs	Incremental costs	Incremental QALYs	ICER (incremental £/QALY)	Change in ICER (%)
Company base-case (including ERG corrections)						
3+7	[redacted]	[redacted]	-	-	-	-
CPX-351	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	-
Scenario: Post-transplant outcomes based on OS only						
3+7	[redacted]	[redacted]	-	-	-	
CPX-351	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]

ERG, Evidence Review Group; ICER, incremental cost-effectiveness ratio; QALYs, quality-adjusted life years

to 1.83 m^2. As a result, the number of required vials per dose increased for CPX-351 during induction and decreased during consolidation (Table 12).

Table 12 Vial usage per dose during the induction and consolidation period

Treatment	Company analysis	ERG analysis*
CPX-351 induction (dosage: 100 units/m^2)	[redacted]	[redacted]
CPX-351 consolidation (dosage: 65 units/m^2)	[redacted]	[redacted]
Cytarabine (dosage: 100 mg/m^2)	1	1.00
Daunorubicin (50ml vial) (dosage: 60 mg/m^2)	3	2.82
Daunorubicin (20ml vial) (dosage: 60 mg/m^2)	6	6.00
Vial estimated assuming a BSA standard deviation of 0.18 ERG, Evidence Review Group, BSA; body surface area		

Hospitalisation

A scenario analysis was conducted in which the number of hospital days in the consolidation period was reduced to 7 days (from [redacted] days for CPX-351, and from 30 days for 3+7). As discussed in **Error! Reference source not found.**, the ERG considered that hospitalisation was overestimated in the model when compared with hospitalisation in the trial, and noted that the assumption that patients were in hospital for [redacted] days during the consolidation period was inconsistent with the assumption made for quality of life, where patients were in hospital for 7 days when receiving consolidation therapy.

Stem cell transplant

In this scenario analysis, two adjustments to the cost associated with transplant were made. Firstly, the ERG removed the cost of providing unrelated adult stem cells from the total transplant cost. The ERG considered that the majority of transplants would be from matched sibling donors, as reported in Wang (2010) [24], which is consistent with the unit cost applied by the company for the procedure. This reduced the cost of transplant from £64,235 to £29,340, which is similar to the cost applied for transplant in other recent NICE submissions in AML [20, 32]. Secondly, the 6-month follow-up cost applied by the company was increased to reflect the two-year cost, weighted by the number of patients predicted to be alive to incur these costs, as the ERG felt that it was important to capture the impact of all possible long-term sequelae of transplant. This increased the follow-up cost from £30,097 to £44,447.

Results

Page 133-136

- Utility estimate for patients in the post-transplant remission health state based on the 0.75 value for patients in durable remission, further adjusted for age,
- Equivalent quality of life for CPX-351 and 3+7 patients while on induction and consolidation treatment,
- Vial usage reflecting the distribution of body surface area, and the mean body surface area re-weighted to reflect the gender distribution in Study 301,
- Reduced number of hospital days during the consolidation period,
- Provision of unrelated donor stem cells excluded from the costs of transplant.

Under the ERG's alternative set of assumptions, the ICER for CPX-351 versus 3+7 is [redacted] per QALY.

Table 13 Results of the ERG alternative base-case analysis

	Total costs	Total QALYs	Incremental costs	Incremental QALYs	ICER (incremental £/QALY)	Change in ICER (%)
Company base-case (with ERG corrections)						
3+7	[redacted]	[redacted]	-	-	-	-
CPX-351	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	-
ERG alternative base-case						
3+7	[redacted]	[redacted]	-	-	-	
CPX-351	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
ERG, Evidence Review Group; ICER, incremental cost-effectiveness ratio; QALYs, quality-adjusted life year;						

6.5 Exploratory analysis on the ERG alternative base-case

6.5.1 Post-transplant survival

In Section **Error! Reference source not found.**, the ERG explored the impact of implementing a range of survival models for post-transplant OS in CPX-351 patients. Due to the immaturity of the data for patients after transplant, there was a large degree of variation in the long-term survival estimates projected by each survival model. As demonstrated by the scenario analyses conducted by the ERG on the company base-case, these had consequences on the estimated cost-effectiveness of CPX-351.

In this section, the ERG explores the impact of the most favourable post-transplant OS and the least favourable post-transplant OS within their alternative base-case. The most favourable model was Gompertz, which estimated that [redacted] of CPX-351 patients would be alive at 5 years. The least favourable model was exponential, which estimated that [redacted] of CPX-351 patients would be alive at 5

years. In comparison, the 5-year survival for 3+7 was [redacted] in both survival models explored by the company.

Table 14 summarises the range of ICERs across the different survival models for CPX-351 post-transplant OS. The ICER of CPX-351 versus 3+7 varied between [redacted] (most favourable post-transplant OS) and [redacted] (least favourable post-transplant OS).

Table 14 Scenario analysis on the ERG base-case: alternative post-transplant OS survival models

	Total costs	Total QALYs	Incremental costs	Incremental QALYs	ICER (incremental £/QALY)	Change in ICER (%)
ERG alternative base-case						
3+7	[redacted]	[redacted]	-	-	-	-
CPX-351	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	-
Scenario: Most favourable post-transplant OS for CPX-351 (Gompertz)						
3+7	[redacted]	[redacted]	-	-	-	
CPX-351	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
Scenario: Least favourable post-transplant OS for CPX-351 (exponential)						
3+7	[redacted]	[redacted]	-	-	-	
CPX-351	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]

ERG, Evidence Review Group; ICER, incremental cost-effectiveness ratio; OS, overall survival; QALYs, quality-adjusted life year;

6.5.2 Inclusion of patients under the age of 60

The ERG previously explored the impact of including a proportion of patients under the age of 60. The inclusion of these patients was associated with a lower ICER: this was attributable to the increased rate of response, which was proportionally higher for CPX-351.

The ERG explored this assumption within the context of the alternative base-case. The ERG did not include this assumption in their alternative base-case due to the remaining uncertainty in the proportion of patients in practice that would be under the age of 60, and the relative impact of treatment on response rate, rate of transplant, and survival in younger patients.

In this analysis, the ICER for CPX-351 versus 3+7 was decreased to [redacted] when 10% of patients were under the age of 60 and to [redacted] when 30% of patients were under the age of 60 (**Table 15**).

Table 15 Scenario analysis on the ERG base-case: inclusion of patients under the age of 60

	Total costs	Total QALYs	Incremental costs	Incremental QALYs	ICER (incremental £/QALY)	Change in ICER (%)
ERG alternative base-case						
3+7	[redacted]	[redacted]	-	-	-	-
CPX-351	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	-
Scenario: 10% patients under the age of 60						
3+7	[redacted]	[redacted]	-	-	-	
CPX-351	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
Scenario: 30% patients under the age of 60						
3+7	[redacted]	[redacted]	-	-	-	
CPX-351	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
ERG, Evidence Review Group; ICER, incremental cost-effectiveness ratio; QALYs, quality-adjusted life year;						

6.6 Conclusions from ERG analyses

The ERG has presented a number of additional analyses 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 from [redacted] to [redacted] per QALY.

Using the corrected and updated model, the ERG then presented a number of analyses considering a range of issues raised in Section 0. These scenario analyses addressed the following issues:

- Inclusion of patients under the age of 60,
- Post-transplant EFS,
- Post-transplant OS for CPX-351 patients,
- The quality of life in post-transplant remission,
- The quality of life while on induction and consolidation treatment,
- The cost of transplantation,
- Resource use in the treatment phases (vial usage, and hospitalisation during the consolidation period).

The scenarios associated with the greatest impact on cost-effectiveness outcomes related to changes made by the ERG to the post-transplant OS, post-transplant EFS, and to the number of vials required for treatment. All scenarios exploring the impact of alternative survival models for post-transplant OS resulted in an increase to the ICER. The majority of scenarios on resource use and quality of life were associated with an increase to the ICER, but within the context of the company base-case these were insubstantial. 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 CPX-351 to be more costly (cost difference [redacted]) and more effective ([redacted] QALY gain) compared with 3+7, and suggests that the ICER for CPX-351 compared with 3+7 is [redacted] per QALY.

The final part of this section carried a further series of exploratory analyses that explored the impact of alternative survival models for post-transplant OS in CPX-351 patients within the ERG alternative base-case. The results of this analysis show the ICER is very sensitive to OS in this group of patients. This is partly due to the immaturity of the OS data from Study 301, which leads to considerable uncertainty around the extrapolation. The ICER of CPX-351 compared with 3+7 varied between [redacted] (most favourable post-transplant OS) and [redacted] (least favourable post-transplant OS). The ERG also explored the impact within their alternative base-case of expanding the patient population to include a proportion of patients under the age of 60 years. In this scenario, the ICER decreased to [redacted] when 10% of patients were under the age of 60 and to [redacted] when 30% of patients were under the age of 60.

8 Overall conclusions

8.1 Clinical effectiveness

Evidence from one phase 3 multi-centre randomised trial suggests that CPX-351 is associated with a significant improvement in OS compared with 3+7 (median OS: 9.56 months [95% CI: 6.60, 11.86] vs 5.95 months [95% CI: 4.99, 7.75]; HR=0.69 [95% CI: 0.52, 0.90], p=0.005) in patients with high-risk AML. Although results from the subgroup analyses should be interpreted with caution, there was some evidence to suggest that CPX-351 had a less beneficial impact on OS in the stratified subgroup of patients with '$_{MDS}$AML with prior treatment with HMA'

[redacted]

[redacted]These patients constituted around a third of patients in the trial and a similar proportion of those who would be eligible for CPX-351 in clinical practice.

Overall CPX-351 may be a more effective bridge to stem cell transplant compared with 3+7 in patients with high-risk AML aged 60-75 years. Compared with 3+7, the proportion of patients undergoing HSCT was higher in the CPX-351 group (34.0% [52/153] vs 25.0% [39/156]) although the difference was not statistically significant (OR=1.54 [95% CI: 0.92, 2.56]). OS in patients who underwent HSCT was significantly greater with CPX-351: at the point of data cut the median OS was not reached in the CPX-351 group, and the median OS in the 3+7 group was 10.25 months (95% CI: 6.21, 16.69) (HR=0.46 [95% CI: 0.24, 0.89], p=0.0046). Overall the safety profiles of CPX-351 and 3+7 appear broadly comparable.

Analyses of Study 301 data presented in the CS and used in the model are based on a data cut dated December 2015 with a median follow up of 20.5 months in the CPX-351 group and 21.2 months in the 3+7 group. The ERG considers the length of follow-up insufficient for measuring long term post-HSCT OS. A substantial number of patients were censored in the CPX-351 arm and there were small numbers of patients in the tail of the survival curves. The company clarified that a number of deaths were known to have occurred in the safety population since the 2015 data cut ([redacted] people in the CPX-351 arm, and [redacted] people in the 3+7 arm). Data on relapse after HSCT is very limited. Study 301 did not collect HRQL or utility data.

The anticipated marketing authorisation for CPX-351 is for the treatment of adults with newly diagnosed therapy-related AML (t-AML) or AML with myelodysplasia-related changes (AML-MRC). Therefore, the age range of patients in the trial (60 – 75 years) is narrower than the anticipated marketing authorisation. The majority of patients with high-risk AML in clinical practice are over the age of 60 years and patients older than 75 years would be less likely to withstand intensive chemotherapy, therefore, the population of the trial is likely to be reflective of the majority of patients

eligible for intensive chemotherapy for high-risk AML in clinical practice. However, results of the trial may not be generalisable to patients under the age of 60.

Patients with *de novo* AML with MDS associated karyotypic changes are difficult to confidently define until genetic test results are available.

[redacted] Therefore, this is not reflective of clinical practice, where treatment may commence prior to cytogenetic test results becoming available

8.2 Cost-effectiveness

The economic evidence presented by the company primarily consisted of a *de novo* model. The company's model used a cohort state-transition approach which directly used the time-to-event data from Study 301 to determine the patient transitions between the health states. When the confidential PAS for CPX-351 was not applied, the company found CPX-351 to be more costly (cost difference of £[redacted]) and more effective ([redacted] QALY gain) compared with 3+7. The deterministic base-case ICER was [redacted] and the mean probabilistic ICER was [redacted] 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. The ERG's key concerns related to the long-term survival predictions after transplant. The ERG carried 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. When the confidential PAS for CPX-351 was not applied, the ERG's alternative base-case analysis estimated CPX-351 to be more costly ([redacted]) and more effective ([redacted] QALY gain) compared with 3+7 alone, and suggests that the ICER for CPX-351 compared with 3+7 is [redacted] per QALY.

8.3 Implications for research

Follow-up of Study 301 is continuing for 5 years post-randomisation, therefore, longer term data are anticipated. In addition, there are two ongoing studies of CPX-351 as first-line treatment for AML patients: AML 18 (which has been extended to include CPX-351) and AML 19.

There is no evidence for the effectiveness and safety of CPX-351 in patients with high-risk AML aged less than 60 years or over 75 years. Data on relapse after HSCT is very limited and no data on HRQL are available.