NICE National Institute for Health and Care Excellence

Single Technology Appraisal

Arsenic trioxide for treating acute promyelocytic leukaemia [ID466]

Committee Papers

NATIONAL INSTITUTE FOR HEALTH AND CARE EXCELLENCE

SINGLE TECHNOLOGY APPRAISAL

Arsenic trioxide for treating acute promyelocytic leukaemia [ID466]

Contents:

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

Pre-meeting briefing

Arsenic trioxide for treating acute promyelocytic leukaemia

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

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

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

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

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

1

Abbreviation	In full
AATO	ATRA + ATO / All-trans retinoic acid + arsenic trioxide
AIDA	ATRA + idarubicin
APL	Acute promyelocytic leukaemia
ATO	Arsenic trioxide
ATRA	All-trans retinoic acid
CI	Confidence intervals
CHR	Complete haematological remission
CMR	Complete molecular remission
CR	Complete remission
HSCT	Haematopoietic stem cell transplantation
ICER	Incremental cost-effectiveness ratio
NMB	Net monetary benefit
OS	Overall survival
PCR	Polymerase chain reaction
PFS	Progression-free survival
PSA	Probabilistic sensitivity analysis
QALY	Quality-adjusted life year
tMDS/AML	Treatment-related myelodysplastic syndrome or acute myeloid leukaemia [2]

Key issues – clinical effectiveness

- Are the results of the APL0406 trial generalisable to UK practice?
- Is arsenic trioxide with ATRA clinically effective in newly diagnosed APL?
- Is arsenic trioxide with ATRA clinically effective in relapsed or refractory APL?
  - Should data from studies other than randomised controlled trials be explored?
- Is arsenic trioxide innovative?
- Are there any equality issues?

3

Key issues – cost effectiveness

- Is the company's model appropriate for decision making?
- Are the model inputs used plausible (transition probabilities, extrapolation of treatment effectiveness)?
- Is arsenic trioxide with ATRA cost-effective in newly diagnosed APL?
- Should best supportive care and stem cell transplant be included as comparators for relapsed or refractory APL?
- Is arsenic trioxide alone used in NHS clinical practice?
- Is arsenic trioxide with ATRA cost-effective in relapsed or refractory APL?

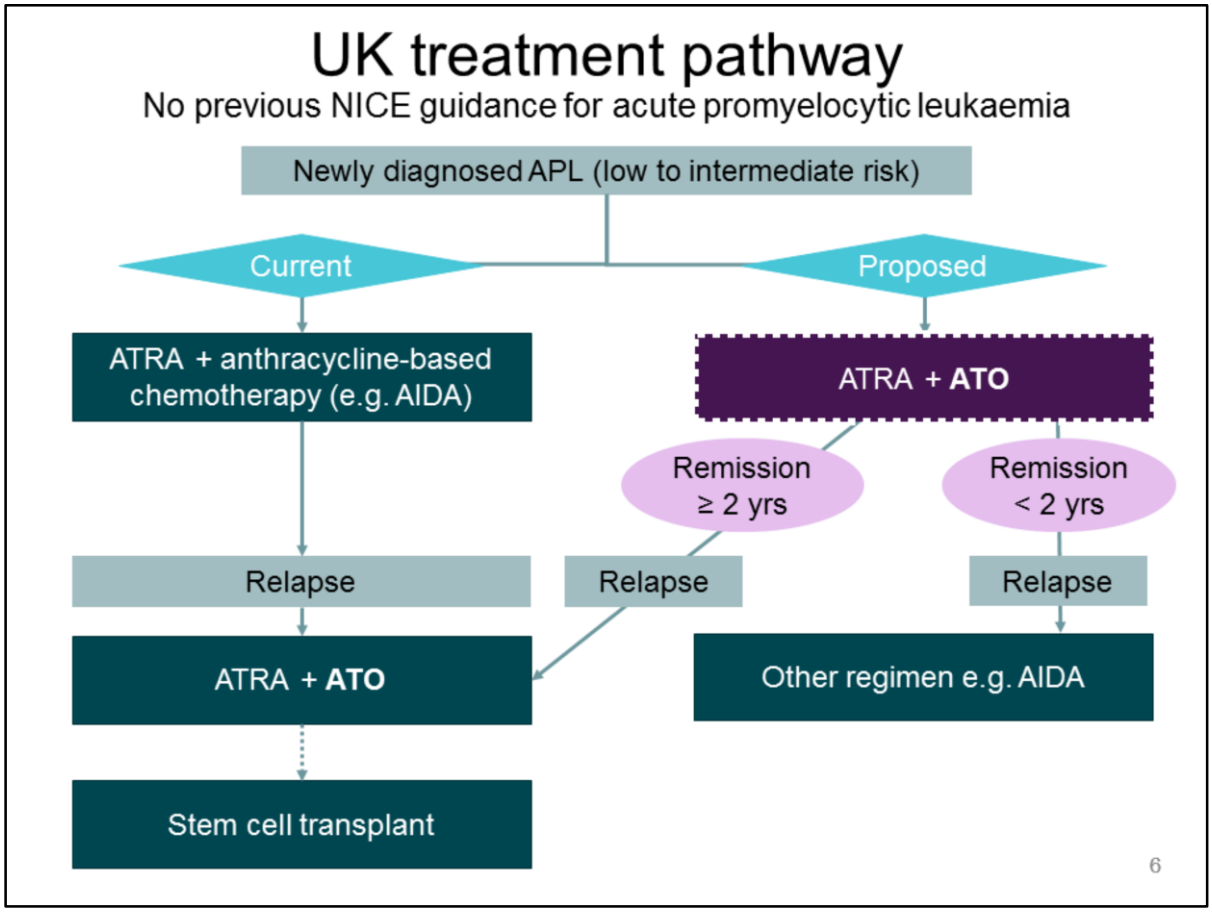
4

Disease background

- Acute promyelocytic leukaemia (APL) is a subtype of acute myeloid leukaemia, associated with a genetic abnormality
- Median age at diagnosis is about 47
- APL can progress rapidly and have a poor survival prognosis
- Assessment of relapse risk, primarily based on white blood cell count, is important in choosing the most appropriate treatment options
- Incidence in Europe is estimated to be 0.11-0.14 per 100,000 people

5

Source: company submission section 1.3.1



Source: Company submission B1.3.2.2 and figure 1.1; professional group submissions from RCPath and BSH, NCRI-ACP-RCP

Notes

- Company states arsenic trioxide (ATO) use in relapsed or refractory APL is well-established
- Company estimates that if ATO-based treatment offered for untreated disease, number of patients to be treated for relapsed or refractory disease would be approximately 10-16 in England.
- Company stated that maintenance treatment is rarely used in the UK, and is not included in the marketing authorisation for ATO.

Comments from patient and professional groups

- APL is a rapidly progressing condition: 62% of people experience symptoms for less than a month before visiting a GP
- Symptoms include bruising or bleeding, fatigue, feeling weak or breathless, bone/joint pain, and sleeping problems
- Little time to take in information before starting treatment – emotional impact for patient and families
- There is a current unmet need for an alternative to chemotherapy-based treatment that can cure APL with less toxicity
- ATO would remove the requirement to treat standard risk APL patients with chemotherapy and protracted molecular monitoring
- First line therapy with ATO is associated with a very low risk of relapse in APL, unlike current chemotherapy
  - Therefore if arsenic trioxide is used as a first-line treatment, the use of second-line treatment would decrease
- ATO has been routinely commissioned for relapsed/refractory disease for 10 years

7

Source: Submissions from Leukaemia CARE; NCRI-ACP-RCP; Royal College of Pathologists with the British Society for Haematology; Clinical expert

Arsenic trioxide (Trisenox, Teva)

UK marketing authorisation	Indicated for induction of remission, and consolidation in adult patients with: • Newly diagnosed low-to-intermediate risk acute promyelocytic leukaemia (white blood cell count, ≤ 10 x 103 /µl) in combination with all-trans-retinoic acid (ATRA) • Relapsed/refractory acute promyelocytic leukaemia (previous treatment should have included a retinoid and chemotherapy) characterised by the presence of the t(15;17) translocation and/or the presence of the Pro-Myelocytic Leukaemia/ Retinoic-Acid-Receptor-alpha (PML/RAR-alpha) gene.
Mechanism of action	Believed to have multiple mechanisms of action including inducing cell death by damaging or degrading the PML/RARα fusion protein in acute promyelocytic leukaemia
Administration and dosage	Administered intravenously at 0.15 mg/kg/day (duration of treatment varies for newly diagnosed/relapsed or refractory disease, and for induction and consolidation therapy)
List price	£2,920 for 10 ampoules of 10mg/10ml concentrate for solution for infusion (BNF)

8

Sources: Summary of product characteristics, British National Formulary

Notes

- Marketing authorisation granted by EMA for treating relapsed/refractory disease in 2002
- Marketing authorisation granted in November 2016 for treating newly diagnosed low- to intermediate-risk APL.

Decision problem [1]

	Final scope issued by NICE	Company submission	Rationale for difference
Population	Adults with: • untreated low-to-intermediate risk acute promyelocytic leukaemia • relapsed/refractory acute promyelocytic leukaemia (APL)	…characterised by the presence of the t(15;17) translocation and/or the presence of the promyelocytic leukaemia/retinoic-acid-receptor-alpha (PML/RAR-alpha) gene.	N/A
Intervention	Arsenic trioxide (ATO) (with or without ATRA)	ATO + ATRA	ATO alone rarely used in the relapsed/ refractory setting.

9

Source: company submission table 1.1

Decision problem [2]

	Final scope issued by NICE	Company submission	Rationale for difference
Comparators	• AIDA regimen • Haematopoietic stem cell transplantation (HSCT) (relapsed or refractory APL) • best supportive care (relapsed or refractory APL)	Single model evaluating ATO+ATRA vs AIDA as first-line treatment, with second-line treatments included	After relapse, choice of therapy depends on prior treatments - difficult to separate first- and second-line ATO. Use of ATRA+ATO usually precedes HSCT. Best supportive care used where disease is refractory to ATO in second-line.
Outcomes	Overall survival (OS) Progression-free survival (PFS) Response rates (bone marrow remission) Adverse effects of treatment Health-related quality of life	Additionally: • Event-free survival • Complete remission rates • Cumulative incidence of relapse • Disease-free survival or relapse-free survival	PFS not measured in trials – event-free survival presented instead

10

Source: company submission table 1.1

Clinical trials

Trials included in company's submission:

Newly diagnosed APL

- APL0406
- AML17

Relapsed/refractory APL

- Raffoux et al.
  - Compared ATRA+ATO with ATO
  - Used for supporting information

ERG comments on trials not included:

- No trials that compared ATO with haematopoietic stem cell transplant or best supportive care, as specified in scope
- No trials of ATO alone – company states ATO rarely used alone in UK
- Non-randomised clinical trials could have been included for relapsed/refractory APL as well as untreated APL as no directly relevant RCT evidence presented
  - Company states use of ATO in relapsed/refractory APL is so well-established it is difficult to provide novel information

Source: ERG report 4.2.1

Notes

- Results from Raffoux et al are described in the company submission B2.7.2

Summary of included trials

Newly diagnosed APL

	APL0406 (n=266, final cohort)	AML17 (n=235)
Design	Phase 3, randomised, open-label, non-inferiority trial	Phase 3, randomised, open-label trial
Population	• No UK patients • Low and intermediate risk disease only	• Based in UK, Denmark and New Zealand • Included people with high risk disease
Intervention	• Compared ATRA+ATO with ATRA+idarubicin (AIDA) • Dosing of ATO in line with marketing authorisation	• Compared ATRA with AIDA • Dosing of ATO different to marketing authorisation • 93% in high risk group, and 7 people in other risk groups received gemtuzumab ozogamicin
Primary outcome	Event-free survival at 2 years after diagnosis	Quality of life

12

Source: company submission tables 2.3 and 2.4

Notes

- APL0406:
  - Evidence presented from published papers only – company was not involved in the trial
  - 162 people were recruited initially. However compliance with quality of life questionnaires was lower than expected. Therefore enrolment to the trial was reopened to maximise quality of life information obtained.
- AML17
  - Evidence presented from published papers only – company was not involved in the trial
  - Dose of ATO was higher and less frequent than in the marketing authorisation

Baseline characteristics

Newly diagnosed APL

Study population	APL0406 final cohort		AML17	
Treatment arm	ATRA +ATO (n=129)	AIDA (n=137)	ATRA +ATO (n=116)	AIDA (n=119)
Median age (years)	46.6	46.6	47	47
Male gender; n (%)	60.0 (46.5)	70.0 (51.1)	60 (52)	60 (50)
White blood cell count, $\times 10^9$/L; median	1.4	1.5	3.0	2.2
Low risk, n (%)	57 (45.2)	55 (41.3)	86 (74)	92 (77)
Intermediate risk, n (%)	69 (54.7)	78 (58.6)	Not reported	Not reported
High risk, n (%)	N/A	N/A	30 (26)	27 (23)

ERG on APL0406: groups are similar and appear to reflect UK patients, based on comparing with patients in AML17

13

Source: company submission table 2.7, ERG report table 4.6

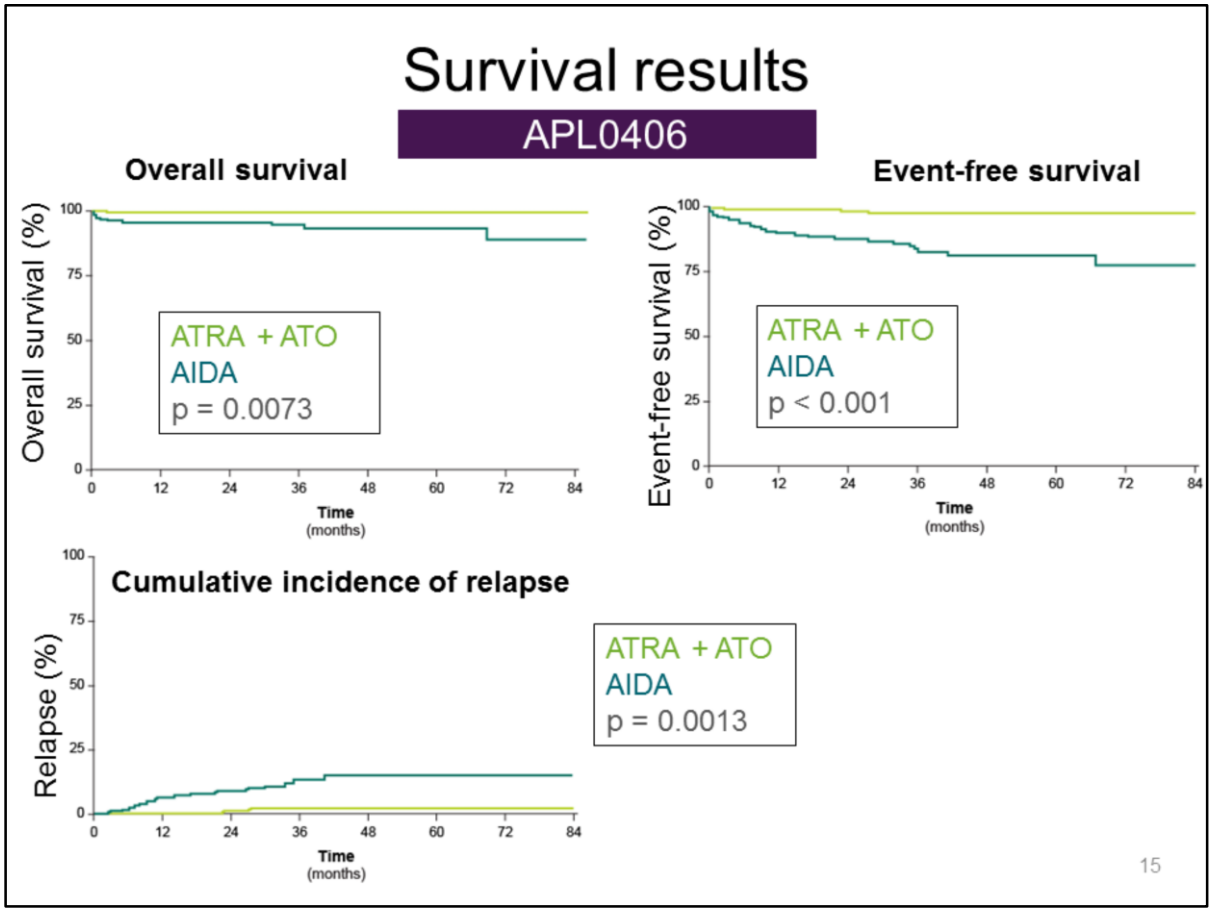
APL0406 results

Summary

Endpoint	Final cohort		
	ATRA+ATO (n = 129)	AIDA (n = 137)	P value
Event-free survival at 50 months, % (95% CI)	97.3 (94.3 to 100)	80.0 (72.9 to 88.0)	< 0.001
Overall survival at 50 months, % (95% CI)	99.2 (97.7 to 100)	92.6 (87.9 to 97.5)	0.007
Disease-free survival at 50 months, % (95% CI)	97.3 (94.3 to 100)	82.6 (75.6 to 90.3)	< 0.001
Haematological CR rate after induction; n (%)	127 (100)	132 (97.0)	0.120
Molecular CR rate after third consolidation cycle; n (%)	115 (100)	117 (98.3)	Not reported
Cumulative incidence of relapse at 50 months, % (95% CI)	1.9 (0.0–4.5)	13.9 (7.1–20.6)	0.0013

14

Source: company submission table 2.13, ERG report table 4.8
Notes
- Transplant rates not reported

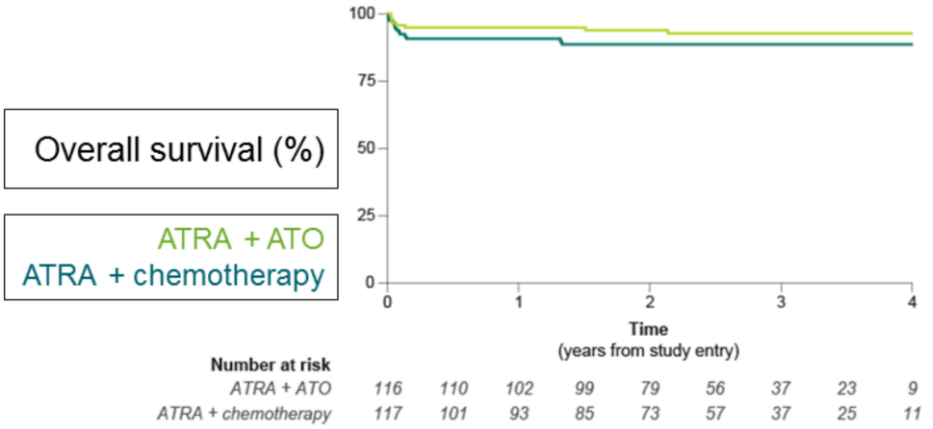


Source: company submission figures 2.6, 2.7, 2.9 and section B2.7.1.4b

AML17 results

Endpoint and time frame	ATRA+ATO (n = 77)	AIDA (n = 79)	Hazard ratio	P value
Event-free survival at 4 years, % (95% CI)	91% (84–95)	70% (56–80)	0.35 (0.18–0.68)	0.002
Overall survival at 4 years, % (95% CI)	93% (86–96)	89% (81–93)	0.60 (0.26–1.42)	0.250

ATO = arsenic trioxide; ATRA = All-trans retinoic acid; AIDA = ATRA + idarubicin; CI=confidence intervals



16

Source: company submission table 2.16 and figure 2.12

Health-related quality of life

APL0406

- Only available for initial patient cohort
  - Long-term analysis in final patient cohort not yet reported
- No baseline assessment performed
- Significant difference between treatment groups (measured on EORTC QLQ-C30 scale) only detected for fatigue (p=0.022)
  - ATRA+ATO associated with lower fatigue severity after induction but not after third consolidation course

AML17

- Measured on EORTC QLQ-C30 scale
- No statistically significant difference detected in the primary outcome of global functioning, but study may have been underpowered
- Small but statistically significant benefits of ATRA+ATO over AIDA seen for cognitive functioning and role functioning

17

Source: Company submission B2.7.1.2e, B2.7.1.4e; ERG report 4.2.5, 4.2.7

Adverse events [1] APL0406 final cohort

Event	Time frame	ATRA + ATO	AIDA	p value
Induction-specific adverse events, n (%)				
Patients with moderate to severe differentiation syndrome	During induction	21 (17)	17 (13)	0.38
Leukocytosis	During induction	56 (43)	NR	NR
Haematological adverse events, n (%)				
Patients with grade 3–4 neutropenia lasting >15 days	During induction	61 (35)	109 (64)	< 0.001
	1st consolidation cycle	8 (16)	40 (67)	< 0.001
	2nd consolidation cycle	7 (7)	90 (92)	< 0.001
	3rd consolidation cycle	5 (15)	28 (85)	< 0.001
Patients with grade 3–4 thrombocytopenia lasting >15 days	During induction	74 (38)	120 (62)	< 0.001
	1st consolidation cycle	6 (26)	17 (74)	< 0.001
	2nd consolidation cycle	6 (7)	77	< 0.001
	3rd consolidation cycle	8 (23)	16 (76)	< 0.001
Fever of unknown origin and infection episodes, n (%)	During induction	30 (23)	75 (55)	< 0.001
	1st consolidation cycle	10 (8)	8 (6)	0.540
	2nd consolidation cycle	4 (3)	46 (38)	< 0.001
	3rd consolidation cycle	2 (1.6)	2 (1.7)	1.000

Source: ERG report 4.2.6, table 4.9

Notes

- ATRA+ATO group: 5 withdrawals due to toxicity (1 during induction (due to severe QTc interval prolongation and electrolyte abnormalities) and 4 during consolidation, (detail of toxicities not reported))
- AIDA group: 10 withdrawals (6 during consolidation (1 patient discontinued due to cardiac toxicity) and 4 during maintenance (2 patients discontinued due to myelosuppression lasting more than 50 days). Details for other 7 patients not reported.)
- EMA commented that “due to the potential synergistic toxicity of ATRA and ATO (i.e. on hepatotoxicity), no direct extrapolation of safety data observed with single-agent ATO is considered adequate”
- EMA has recommended that the company conduct a post-authorisation long term safety cohort study to explore the long-term safety of ATRA+ATO for people with newly diagnosed APL

Adverse events [2] APL0406 final cohort

Event	Time frame	ATRA + ATO	AIDA	p value
Non-haematological adverse events, n (%)				
Patients with QTc prolongation	During induction	11 (8.5)	1 (0.7)	0.002
	1st consolidation cycle	3 (2)	0	0.110
	2nd consolidation cycle	3 (2)	0	0.110
	3rd consolidation cycle	2 (1.5)	0	0.230
Patients with grade 3–4 hepatic toxicity	During induction	51 (40)	4 (3)	< 0.001
	1st consolidation cycle	5 (4)	1 (0.7)	0.110
	2nd consolidation cycle	1 (0.8)	0	0.490
	3rd consolidation cycle	0	0	NA
Patients with grade 3–4 gastrointestinal toxicity	During induction	3 (2)	25 (18.2)	< 0.001
	1st consolidation cycle	0	1 (0.8)	1.000
	2nd consolidation cycle	0	6 (4.9)	0.03
	3rd consolidation cycle	0	0	1.000
Patients with grade 3–4 cardiac function abnormalities	During induction	0	5 (3.7)	0.060
	1st consolidation cycle	0	0	NA
	2nd consolidation cycle	0	0	NA
	3rd consolidation cycle	0	0	NA

Source: ERG report 4.2.6, table 4.9

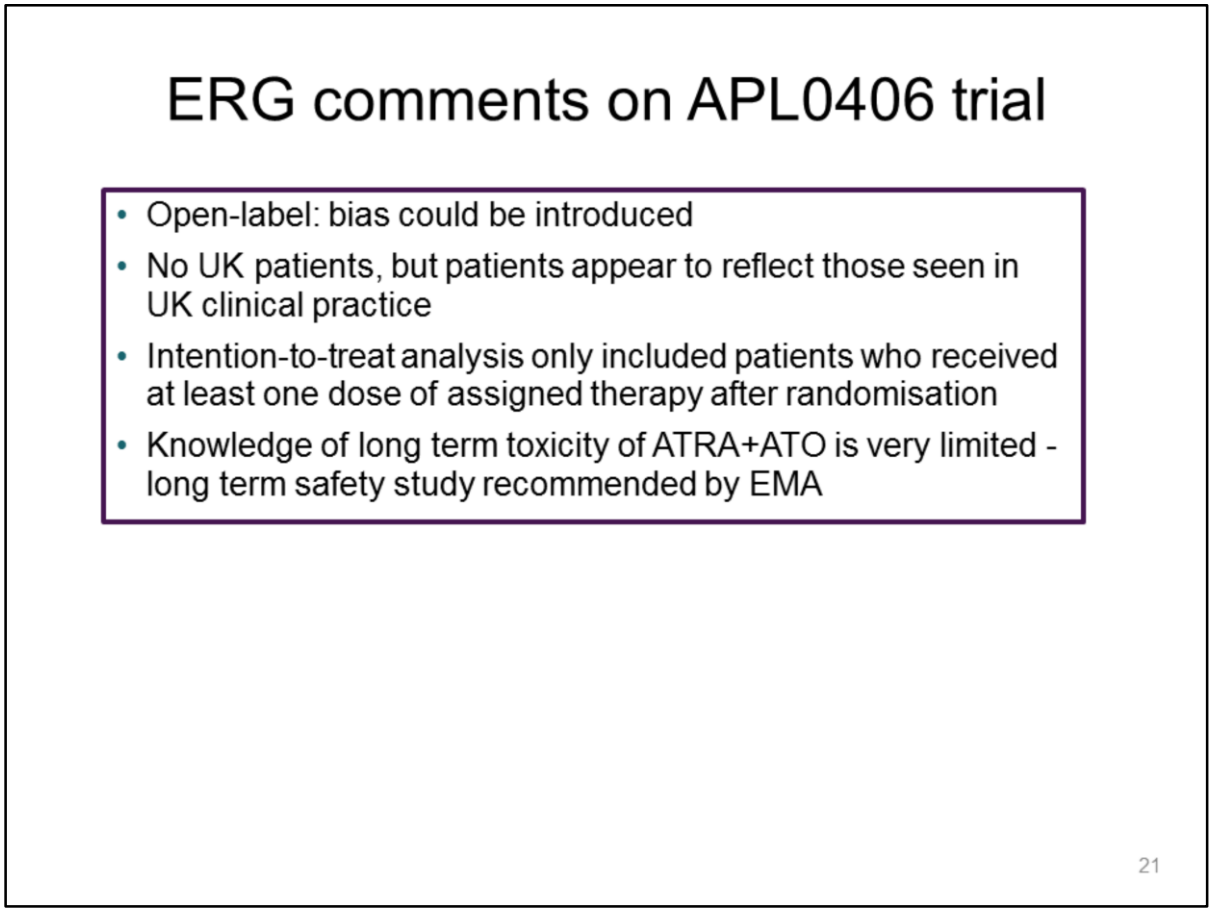
Adverse events [3] APL0406 final cohort

Event	Time frame	ATRA + ATO	AIDA	p value
Neurotoxicity (all grades), n (%)	During induction	1 (0.7)	0	0.480
	1st consolidation cycle	5 (4.2)	0	0.020
	2nd consolidation cycle	6 (5)	0	0.010
	3rd consolidation cycle	7 (5.9)	0	0.006
Hypercholesterolemia, n (%)	During induction	14 (10)	12 (8.7)	0.550
	1st consolidation cycle	19 (16)	12 (9.6)	0.130
	2nd consolidation cycle	19 (16)	12 (9.7)	0.140
	3rd consolidation cycle	16 (14)	11 (9.0)	0.270
Hypertriglyceridemia, n (%)	During induction	29 (22)	29 (22)	0.760
	1st consolidation cycle	22 (18.4)	19 (15.2)	0.490
	2nd consolidation cycle	17 (14.4)	10 (8)	0.120
	3rd consolidation cycle	16 (14)	13 (11)	0.500

Company notes that adverse events in the trials were mostly managed with temporary treatment discontinuation and supportive care, with few permanent discontinuations reported.

20

Source: ERG report 4.2.6, table 4.9



Source: ERG report section 4.2

Innovation and equality considerations

Innovation

- Company comments
  - Offers a chemotherapy-free treatment option for people newly diagnosed with low- to intermediate-risk APL
    - reduces toxicity
    - option for people not suitable for chemotherapy
- Royal College of Pathologists/British Society for Haematology
  - Reduces risk of relapse
  - Reduces need for bone marrow transplant

Equality considerations

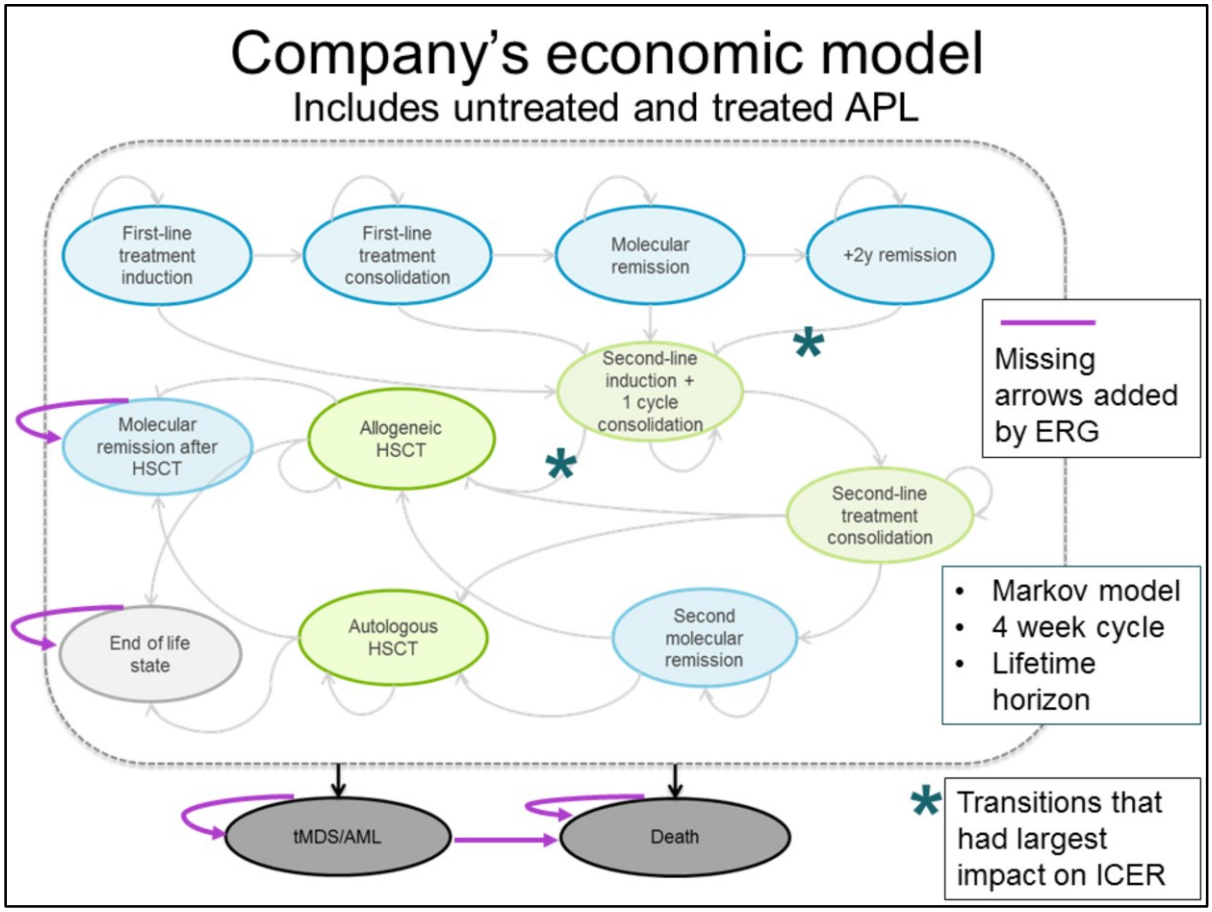
- Are there any equalities issues?
  - Company highlighted that older people who can not have chemotherapy would be eligible for treatment with ATO
  - Company highlighted that ATO+ATRA may decrease blood transfusions compared to AIDA, which may be more acceptable to people who are Jehovah's Witnesses and cannot have blood transfusions

22

Source: Company submission B1.13, B1.4; submission from Royal College of Pathologists/British Society for Haematology

Cost effectiveness

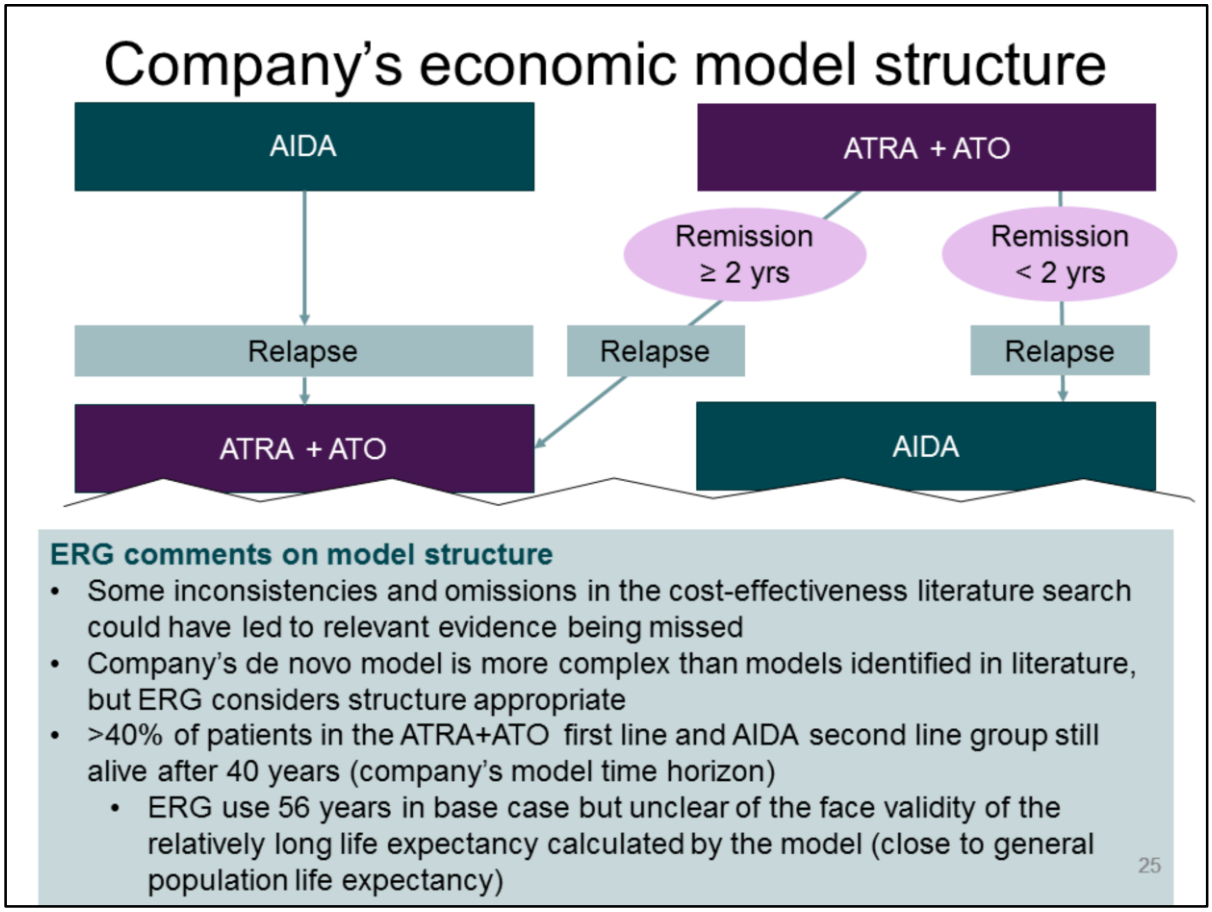
23



Source: company submission B.3.2.2

Notes

- tMDS/AML = treatment-related myelodysplastic syndrome or acute myeloid leukaemia
- HSCT = haematopoietic stem cell transplant



Source: ERG report sections 5.2.2, 5.2.5

Company's economic model inputs

Treatment effectiveness and extrapolation

- Efficacy in newly diagnosed APL estimated through remission rates and rate of relapse in APL0406 trial
- Rate of relapse was higher for the first 2 years of remission (molecular remission state) and lower in '+2y remission' state, where rate was constant until death
- Efficacy data for relapsed/refractory APL derived from studies by Raffoux *et al.*, Tallman *et al.*, Russell *et al.*, and Platzbecker *et al.* clinical expert opinions and an existing US cost-effectiveness model
- Safety data from APL0406 and Raffoux *et al.* trials
- Except for cardiac events, adverse events did not lead to a change of treatment, but impacted on costs and quality of life

ERG comments on treatment effectiveness and adverse events

- Transition probabilities and evidence sources not clearly described
- Transitions from second line health states: evidence weak and method of obtaining probabilities not transparently reported
- Should have used conditional probabilities for relapse after first line treatment
- Justification for sources of adverse events mostly unclear e.g. some sources for duration of adverse effects are over 25 years old.
- Unclear why specific adverse events chosen, e.g. why reversible arrhythmia not considered

Source: company submission B3.3.1, B3.3.2; ERG report section 5.2.6, 5.2.7

Notes

- Outcomes from APL0406 used in the model:
  - Haematological remission rate after induction
  - Proportion of patients evaluable with a PCR test after consolidation
  - Complete molecular remission rate after consolidation
  - Probability (cumulative incidence) of relapse at 24 and 50 months
  - Median time to relapse
  - Proportion of patients experiencing adverse events by treatment phase

Company's economic model inputs

Utility values

- Utility values obtained from literature for other diseases (e.g. chronic lymphocytic leukaemia and acute myeloid leukaemia) because data from EQ-5D not available for APL
  - Adjusted for age (average age of modelled population is 45 years)
  - Adjusted for the utility representing perfect health
  - Disutilities for adverse events included in induction and consolidation

ERG comments on utility values

- Company's method of selecting utility values unclear
- Unclear why values from chronic lymphocytic leukaemia chosen
- Adjustments made by company unjustified
  - Need for age-adjustment is unclear as the impact of disease would outweigh the impact of age
  - No evidence to support method of adjusting for utility representing perfect health – over time, utility values are higher than in general population
  - Adjustments not applied to all health states and rationale for this is unclear

27

Source: company submission B2.4.1, B3.3.1, ERG report table 5.2

Utility values in the model

State	Mean utility value	Reference
First-line induction treatment	0.739	Woods et al., 2012 Szende et al., 2014
First-line consolidation treatment	0.739	Woods et al., 2012 Szende et al., 2014
First molecular remission	0.773	Beusterien et al., 2010 Szende et al., 2014
First long-term molecular remission (>2 years)	0.849	Szende et al., 2014
Second-line induction + 1 cycle consolidation	0.673	Woods et al., 2012 Beusterien et al., 2010
Second-line treatment consolidation	0.702	Beusterien et al., 2010
Second molecular remission	0.849	Szende et al., 2014
Allogeneic HSCT	0.687	Breitscheidel L., 2008
Autologous HSCT	0.687	Breitscheidel L., 2008
Allogeneic HSCT molecular remission	0.849	Szende et al., 2014
End of life state	0.4	Morton et al., 2009
tMDS/AML	0.4	Cooperberg et al., 2013
Hospitalisation	-0.01	Assumption

Source: company submission B3.4.9

Company's economic model inputs

Treatment costs

Phase	ATRA+ATO			AIDA		
	ATRA	ATO	**Total ATRA +ATO**	ATRA	Chemo (IDA+MTZ)	**Total AIDA**
First line: Induction	£464	£16,079	£16,542	£507	£2,097	£2,604
First line: Consolidation	£1,521	£40,196	£41,718	£652	£1,723	£2,375
First line: Total	£1,985	£56,275	**£58,260**	£1,159	£3,820	**£4,979**
Second line: Induction	£362	£12,561	£12,924	£507	£2,097	£2,604
Second line: Consolidation	£1,521	£12,561	£14,083	£652	£1,723	£2,375
Second line: Total	£1,884	£25,123	**£27,006**	£1,159	£3,820	**£4,979**

29

Source: company submission table 3.7

Company's economic model inputs

Medical costs

Items	Value	Reference
Cost per follow-up appointment	£52.50	Personal Social Services Research Unit (PSSRU)
Cost per polymerase chain reaction monitoring test	£280.00	Expert opinion: Guy's Hospital tariff (NHS Foundation Trust)
Allogeneic haematopoietic stem cell transplant (HSCT) remission costs (annual)	£21,585.75	Leunis et al., 2013
Cost per allogeneic HSCT	£27,907.53	National Schedule of Reference Costs
Cost per autologous HSCT	£7,122.97	National Schedule of Reference Costs
End of life costs per month	£4,670.68	Marie Curie Cancer Care

30

Source: company submission table 3.10

Company's economic model inputs

Resource use [1]

	Items	Cost	ATRA+ATO	AIDA	References for resource use
Induction	Bed days per patient	£396.47	First line: 32 Second line: 25	35	AATO: First line: Lo-Coco et al., 2013 Second line: Douer et al., 2005 AIDA: Lo-Coco et al., 2013
	Supportive care transfusions	0	15	22	Burnett et al., 2015
	Annual PCR tests	£280.00	5	4	Expert opinion
Consolidation	Bed days per patient	£396.47	0	4	ATRA+ATO: Expert opinion AIDA: assumption based on treatment schedule
	Ambulatory days per patient	£162.00	First line: 10 Second line: 12.5	0	ATRA+ATO: Expert opinion AIDA: Inpatient treatment assumed
	Number of days of antibiotics	£1.65	1	2	Burnett et al., 201513
	Annual PCR tests	£280.00	5	4	Expert opinion

31

Source: ERG report table 5.13

Company's economic model inputs

Resource use [2]

	Items	Cost	ATRA +ATO	AIDA	Reference for resource use
Molecular remission (first, second, allo- and auto-HSCT)	Duration of follow-up	£210	3	3	First remission: Platzbecker et al., 2015 Others: Expert opinion
	Annual appointments	£52.50	4	4	First remission: Platzbecker et al. Others: Expert opinion
	Annual PCR tests	£280	4 (0 at first remission)	4	First remission: ATRA+ATO: Expert opinion AIDA: Platzbecker et al. Others: Expert opinion
Allo HSCT	Hospitalisation duration	£27,907.53	4 weeks	4	Expert opinion
Auto HSCT	Hospitalisation duration	£7,122.97	3 weeks	3	Expert opinion

ERG comments on costs and resource use

- Lack of justification of sources for cost and resource data
- Monitoring of haematological response costs not included – costs would be higher in ATRA+ATO group because more frequent relapses in second line for AIDA group so less monitoring needed

Source: ERG report table 5.13, section 5.2.9

Total costs – company's base case

ATRA+ATO for untreated APL

Cost category	ATRA+ATO	AIDA	ATRA+ATO vs. AIDA
Total treatments	£60,336	£21,604	£38,731
Administration	£25,402	£31,660	-£6,259
Supportive care and antibiotics	£3,575	£6,487	-£2,912
Follow-up and monitoring	£2,991	£10,389	-£7,398
Adverse events	£4,142	£12,378	-£8,236
Myelodysplastic syndrome	£0	£226	-£226
HSCT	£7,645	£48,326	-£40,681
Palliative care	£906	£5,196	-£4,290
Total	**£104,996**	**£136,267**	**-£31,270**

- Costs generated by the model for the average patient over a lifetime horizon
- Largest cost offsets are for HSCT, adverse events and monitoring

33

Source: company submission table 3.15

Notes

- Discounted costs
- Costs for 'supportive care and antibiotics' largely consists of the cost of transfusions

Company's base case results

ATRA+ATO for untreated APL

Discounted deterministic base case

	Total costs (£)	Total QALYs	Inc. costs (£)	Inc. QALYs	ICER (£/QALY)	NMB (£)
AIDA	136,267	13.72	-	-	-	-
ATRA+ATO	104,996	16.34	-31,270	2.62	Dominant	109,871

Inc., incremental; NMB, net monetary benefit (calculated by NICE technical team based on a £30,000/QALY threshold)

- Undiscounted life years:
  - AIDA = 26.84
  - ATRA+ATO = 33.22
  - Life years gained = 6.38

34

Source: company submission B3.7.3 tables 3.14, 3.16, 3.18

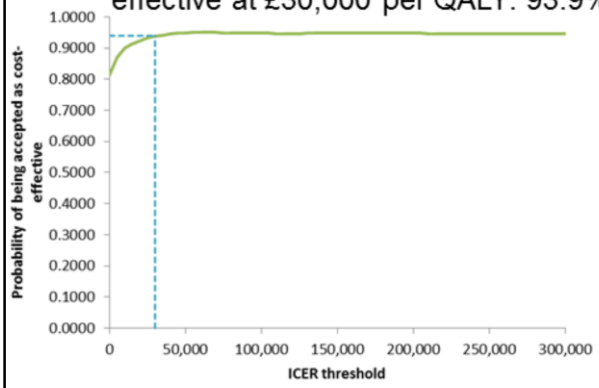
Company's probabilistic sensitivity analysis results

ATRA+ATO for untreated APL

ATRA+ATO vs AIDA	Incremental costs	Incremental QALYs	ICER (£/QALY)
Mean	-£31,088	2.55	Dominant

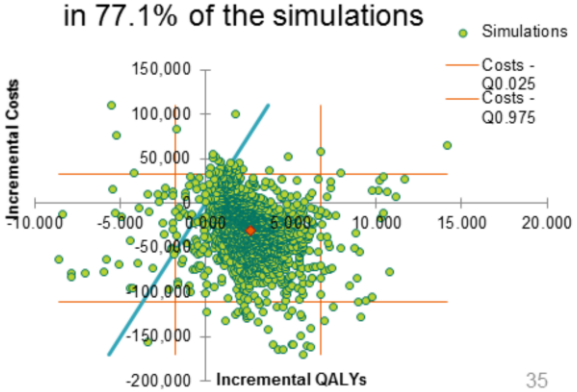
Cost-effectiveness acceptability curve for ATRA+ATO vs AIDA

- Probability of ATRA+ATO being cost-effective at £30,000 per QALY: 93.9%



Incremental cost-effectiveness plane for ATRA+ATO vs AIDA

- ATRA+ATO dominated AIDA in 77.1% of the simulations



35

Source: company submission B.3.8.1 figure 3.2, 3.3, table 3.19

Company's deterministic sensitivity analysis

ATRA+ATO for untreated APL

- ICER only computable in 4 cases (in all other cases, ATRA+ATO dominant over AIDA):

Parameters	Base case value	Lower case		Higher case	
		ICER (£/ QALY)	NMB (£)	ICER (£/ QALY)	NMB (£)
Company base case	-	*Dominant*	*109,870*	-	
Time horizon (5 to 30 years)	40 years	148,179	-17,628	Dominant	87,308
Relapse after remission (48 months) – AIDA (0.082-0.209)	0.139	25,658	4,299	Dominant	179,946
CHR rate, first line - ATRA+ATO (0.4922–1.0000)	0.9845	11,927	13,919	Dominant	109,727
CMR rate, first line - ATRA+ATO (0.5–1.0)	1.0	1,472	31,375	Dominant	109,870

Inc, incremental; CHR, complete haematological remission; CMR, complete molecular remission; NMB, net monetary benefit calculated by NICE technical team based on a £30,000 threshold

Source: company submission B.3.8.2, appendix J

Company's scenario analyses (untreated APL)

Scenario	Description
a. AIDA used in second line following both first-line treatments	To investigate impact of subsequent treatments on cost-effectiveness
b. Utility values from Tallman *et al.*	Values used in an existing model
c. AML17 protocol	Schedule, dosage, efficacy and safety inputs taken from AML17
d. 'Worst-case' scenario	Includes unfavourable inputs for the ATRA+ATO group
e. Probability of undergoing HSCT reflecting clinical practice	Lower proportion of patients undergo autologous HSCT and allogeneic HSCT is reserved for patients not in molecular remission after second line induction
f. Including disease-related mortality	For induction and consolidation phases
g. Including maintenance treatment	2 years of maintenance in AIDA group
h. 26 cycles in consolidation state	Cycle length is 4 weeks so 1 year is 13 cycles
i. Time horizon of 56 years	40 years used in base case
j. Not assuming probability of relapse was the same at 48 months as 50	In base case, probability of relapse at 48 months in first molecular remission assumed to be equal to that at 50 months

Source: company submission B3.8.3, appendix J.4; company's response to clarification, B2, B19, B20 and B21; ERG report 5.2.11 and table 5.20

Notes

- Scenario d: 'Worst-case' scenario:
    - utility values set to minimise difference between first line treatment and relapse health states
    - low values of the CI for utilities concerning first line treatment and high values for second line and HSCT health states were used
    - utility values of the tMDS/AML and end of life health states increased by 25%
    - costs of follow-up ignored in both arms to minimise the cost in the AIDA strategy
    - average weight of the UK population was used to maximise treatment acquisition costs
    - probabilities to undergo allogeneic HSCT from the second line health states set to 0 (except no molecular remission at the end of the induction phase where all patients received allogeneic SCT)
    - probabilities to receive autologous HSCT reduced to 0.25
    - switches due to cardiac events not considered

Company's scenario analyses results

ATRA+ATO vs AIDA for untreated APL

Scenario	Inc. costs (£)	Inc. QALYs	ICER (£/QALY)	NMB (£)
Company's base case	-31,270	2.62	Dominant	109,870
a. AIDA used in second line following both first-line treatments	-21,593	2.72	Dominant	103,193
b. Utility values from Tallman *et al.*	-31,270	2.93	Dominant	119,170
c. Schedule, dosage, efficacy and safety inputs from AML17	-66,384	3.39	Dominant	168,084
d. 'Worst-case' scenario	-9,986	1.58	Dominant	57,386
e. Probability of undergoing HSCT reflecting clinical practice	-28,664	2.43	Dominant	101,564
f. Including disease-related mortality	-21,099	3.80	Dominant	135,099
g. Including maintenance treatment	-33,012	2.62	Dominant	111,612
h. 26 cycles in consolidation state	-31,813	2.63	Dominant	110,713
i. Time horizon of 56 years	-32,922	2.83	Dominant	117,822
j. Not assuming probability of relapse was the same at 48 months as 50	-28,555	2.53	Dominant	104,455

38

Source: company submission B3.8.3, appendix J.4; company's response to clarification, B2, B19, B20 and B21; ERG report 5.2.11 and table 5.20

Notes

- NMB, net monetary benefit, calculated by NICE technical team based on a £30,000 threshold

ERG base case – main changes

- Time horizon
  - Used 56 years instead of 40 years in company's base case
- Alternative utility values
  - Removed utility adjustments and used same value (0.70) for first and second induction and consolidation
  - Utility values capped so as not to exceed general population
- Alternative remission probabilities
  - Based remission probability for all patients on APL0406 trial data and used molecular remission rate to inform probability of transitioning to remission for patients who could be evaluated with PCR testing

39

Source: ERG report sections 5.2.2, 5.2.5, 5.2.6, 5.2.8

ERG base-case results - Untreated APL

Including summary of exploratory analyses

Deterministic results	ATRA+ATO vs. AIDA			
	Inc. costs	Inc. QALYs	ICER incremental (£/QALY)	Net monetary benefit (£)
Company base case	**-£31,270**	**2.62**	**Dominant**	**£109,870**
1. Company base case – errors corrected	-£25,914	2.43	Dominant	£98,814
2. Time horizon 56 years	-£27,540	2.63	Dominant	£106,440
3. Alternative utility values	-£25,914	2.41	Dominant	£98,214
4. Utility values capped	-£25,914	2.26	Dominant	£93,714
5. Alternative remission probabilities	-£21,853	2.27	Dominant	£89,953
ERG base case (1-5 combined)	**-£23,502**	**2.25**	**Dominant**	**£91,002**
Net monetary benefit calculated by NICE technical team based on a £30,000 threshold				

40

Source: ERG report table 6.1

Notes

- ERG unable to perform probabilistic analyses but company's deterministic and probabilistic results are similar, so ERG considers that ATRA+ATO likely to remain dominant if probabilistic results could be produced for ERG base case.

ERG scenario analyses

Untreated APL

Issue	ERG comment
6. Disease-related mortality	No disease-related mortality modelling during on treatment and remission phases. ERG considers that mortality risk likely to be higher than in the general population (consistent with evidence from AML17 trial).
7. Stem cell transplant (HSCT)	In the model, patients can have autologous or allogeneic HSCT. Clinical expert stated that allogeneic HSCT is generally not recommended in APL in the UK.
8. Transition from second line molecular remission to HSCT	Adjusted using the median time to relapse following second line remission. Unadjusted probabilities seem high but unsure of justification for adjustment.
9. Reversible arrhythmia	Expert opinion suggested 2% of patients on ATRA+ATO experience reversible arrhythmia and switch treatment – not modelled.
10. Extrapolation of treatment effectiveness	Company's model assumes treatment benefits are maintained for entire time horizon, e.g. relapse transition probability in first line is constant from 2 years after remission.

41

Source: ERG report 5.2.2, 5.2.3, 5.2.6

ERG scenario analyses – results

ATRA+ATO vs AIDA for untreated APL

Scenario	Inc. costs	Inc. QALYs	ICER (£/QALY)	NMB (£)
ERG base case	**-£23,502**	**2.254**	**Dominant**	**91,122**
6. Adding disease-related mortality in induction phases (first and second line)	-£17,066	2.682	Dominant	97,526
7. Replacing transitions to allogeneic HSCT for transitions to autologous HSCT	-£9,865	1.624	Dominant	58,585
8a. Transitions to HSCT states from second line remission removed	-£24,848	2.281	Dominant	93,278
8b. Transitions to HSCT states from second line remission 'uncorrected'	-£22,723	2.242	Dominant	89,983
9. Incorporating 2% cardiac events for ATRA+ATO in induction phase	-£23,606	2.285	Dominant	92,156
10. Assuming equal relapse probability for all treatments 2 years after first-line remission	£20,407	1.034	£19,734	10,613
All of the above scenarios (except 8b)	**£27,067**	**1.252**	**£21,622**	**10,493**

NMB, net monetary benefit calculated by NICE technical team based on a £30,000 threshold

Source: ERG report 5.3.2 and table 6.2.

Relapsed or refractory APL

Company's results

Treatment	Total costs (£)	Total QALYs	Incremental costs (£)	Incremental QALYs	ICER incremental (£/QALY)
AIDA	188,877	8.84	-	-	-
ATRA+ATO	198,959	9.44	£10,082	-0.60	£16,733

- Provided following request for clarification
- Implemented by changing health states representing first line therapy to second line, and neutralising states representing second line (no transitions to these states were possible)
- Company state that if ATO is used for untreated APL, the number of relapses will decrease so a very small population will have relapsed/refractory APL
- No efficacy data found for ATO alone in relapsed or refractory disease so not modelled
  - Clinical experts stated ATO is rarely used alone in relapsed or refractory disease
- Clinical experts stated that best supportive care is not a relevant comparator – analysis not carried out

43

Source: company submission B2.14, table 1.1; company's response to clarification question B5, table 4

ERG analysis of relapsed/refractory APL

- ERG unsure how company's analysis performed
  - No detail about sources of transition probabilities
- ERG's analysis implemented by removing first line health states
  - Based on ERG base-case model
- ERG state that this analysis should be considered explorative given the concerns with the evidence informing the second line health states

Relapsed /refractory APL	Total costs	Total QALYs	Incremental costs	Incremental QALYs	ICER incremental (£/QALY)
AIDA	£191,158	8.620	-	-	-
ATRA+ATO	£209,365	9.204	£18,207	0.584	£31,184

44

Source: ERG report section 5.3.3 and table 6.3

Key issues – cost effectiveness

- Is the company's model appropriate for decision making?
- Are the model inputs used plausible (transition probabilities, extrapolation of treatment effectiveness)?
- Is arsenic trioxide with ATRA cost-effective in newly diagnosed APL?
- Should best supportive care and stem cell transplant be included as comparators for relapsed or refractory APL?
- Should arsenic trioxide alone be modelled in relapsed or refractory APL?
- Is arsenic trioxide with ATRA cost-effective in relapsed or refractory APL?

45

Authors

- **Kirsty Pitt**
  Technical Lead
- **Alex Filby**
  Technical Adviser
- with input from the Lead Team (Mike Chambers, Andrew Renehan and Judith Wardle)

46

NATIONAL INSTITUTE FOR HEALTH AND CARE EXCELLENCE

Single technology appraisal

Arsenic trioxide for treating acute promyelocytic leukaemia – TA10216

Document B

Company evidence submission from Teva

7 December 2017

File name	Version	Contains confidential information	Date
		No	**7 December 2017**

Executive summary

Background

Acute promyelocytic leukaemia (APL) is a rare subtype of acute myeloid leukaemia (AML) characterised by excessive bleeding, resulting in a high risk of death before treatment and at its early stages. Although when left untreated APL is the most rapidly fatal leukaemia, if promptly diagnosed and treated it becomes the most frequently curable AML subtype.

The incidence of APL is just over 1 in 1,000,000 people[1, 2] and the disease constitutes 3–7.4% of AML cases[1-3]. Thus, APL can be classified as a rare disease according to the European Medicines Agency (EMA), which defines rare diseases as those that affect no more than 5 in 10,000 people in the European Union[4]. In 2014, 2,590 people were diagnosed with AML in England according to the final NICE scope; from this figure and the aforementioned APL incidence rates, we estimated that no more than 187 patients are diagnosed with APL in England per year.

Before the approval of ATO in first line, patients with newly-diagnosed APL were commonly treated with the combination of all-trans retinoic acid (ATRA) and anthracycline-based chemotherapy (often idarubicin, which forms the basis of the AIDA regimen). Treatment of patients with APL is divided into three phases. The first phase is induction therapy, which aims to achieve complete haematological remission (CR). This is followed by two phases of post-remission therapy (consolidation and maintenance) intended to maintain CR. However, in the UK, maintenance treatment is usually omitted.

Although recent studies show that chemotherapy-based treatment is effective in as many as 70% of APL patients[5] (80% if only those of low- to intermediate risk are considered[6]), a considerable proportion of patients relapse and/or die. Furthermore, anthracycline-based treatment is associated with a number of adverse events, including considerable haematological toxicity which puts patients at risk of serious infections. In the longer term, anthracycline-treated patients may experience heart failure[7, 8] or develop a treatment-related secondary leukaemia, which is associated with poor prognosis[9, 10].

Patients who relapse following treatment with ATRA and an anthracycline are commonly treated with arsenic trioxide (ATO) with the aim to achieve a second CR. Approximately 60% of patients subsequently receive a haematopoietic stem cell transplant (HSCT), which appears to reduce the risk of another relapse and improve survival[11].

ATO was initially approved in March 2002 for the treatment of patients with relapsed or refractory APL previously treated with retinoids and chemotherapy, and has since been established in the UK (and around the globe) as the standard and widely recommended therapy in this setting. In November 2016, this initial indication was extended as ATO was approved by the EMA for use in adult patients with previously untreated low-to-intermediate risk APL.

Efficacy of arsenic trioxide in APL

Newly-diagnosed, low- to intermediate-risk APL

The marketing authorisation in newly-diagnosed low- to intermediate-risk APL was primarily based on the international multicentre APL0406 trial, in which 266 eligible adult patients with newly-diagnosed APL were randomised to receive ATO combined with ATRA (ATRA + ATO) or chemotherapy combined with ATRA (the AIDA regimen). This important trial showed that ATRA + ATO significantly improved overall survival (OS) at 50 months compared with AIDA (99.2% vs 92.6% respectively, p=0.007)[6]. The primary endpoint of this trial was event-free survival (EFS) at two years in the initial cohort of 156 patients (97% with ATRA + ATO vs 86% with AIDA, p<0.001 for non-inferiority, p=0.02 for superiority)[12]. EFS was significantly better in the ATRA + ATO group across all subsequent analyses[6, 13] to reach 97.3% at 50 months in the full cohort of 266 patients, compared with just 80.0% in the AIDA group (p<0.001)[6]. The primary source of the observed EFS benefit was a reduction in the number of relapses with ATRA + ATO – at 50 months, the cumulative incidence of relapse was as little as 1.9% in the ATRA + ATO group compared with 13.9% in the AIDA group (p=0.0013)[6].

The use of ATRA + ATO in patients with newly-diagnosed APL is also supported by the primarily UK-based AML17 trial, which included 235 eligible patients of any risk

group. Compared with the APL0406 trial, AML17 used an off-label attenuated ATO schedule and a slightly modified AIDA regimen with no maintenance therapy and an addition of gemtuzumab ozogamicin in high-risk patients. Although the trial failed to meet its primary endpoint (quality of life [QoL]), possibly due to not fulfilling its patient enrolment[5], the efficacy results were positive. This trial confirmed the EFS benefit of ATRA + ATO over AIDA (4-year EFS of 91% vs 70%, p=0.002), particularly in low-risk patients (4-year EFS was 92% in the ATRA + ATO group [n=86] vs 71% in the AIDA group [n=92], p=0.008)[5]. Similarly to the APL0406 trial, the difference in EFS resulted largely from the significantly lower incidence of relapse with ATRA + ATO. The 4-year cumulative incidence of haematological relapse was 18% in the AIDA arm and 1% in the ATRA + ATO arm (p=0.0007)[5]. In this trial, patients were closely monitored for molecular relapse and many were treated before progression into a full haematological relapse, so that the cumulative incidence of molecular relapse at 4 years was 27% in the AIDA group and 0% in the ATRA + ATO group (p<0.0001)[5].

The two aforementioned randomised controlled trials (RCTs) provided robust clinical evidence supporting ATO approval by the EMA. In addition, our systematic review of non-randomised studies evaluating the efficacy and safety of ATO in conditions closer to routine clinical practice generally supported a favourable benefit-risk ratio of ATO in the treatment of APL.

Relapsed/refractory APL

The efficacy of ATO in patients with relapsed or refractory APL was demonstrated in two single-arm studies conducted in the US[14, 15], with no additional European studies supporting the EMA approval in this indication. Our systematic literature review identified a single small RCT[16] conducted in this setting in Europe. This paucity of high-quality data should be viewed in light of the rarity of APL and the fact only up to a third of patients relapse. Furthermore, with first-line ATO use the number of patients who relapse will substantially diminish, so that the relapsed/refractory APL indication will become even smaller. Although RCT data are scarce, a wealth of evidence has been accumulated over the last 15 years confirming the findings of the initial pivotal studies. The RCT by Raffoux, et al. demonstrated that treatment with ATRA + ATO allows as many as 80% of patients to achieve a second remission[16],

which enables them to be considered for a potentially curative HSCT. However, there is a growing body of evidence that ATRA+ATO may be an effective second-line treatment option even if not followed by transplantation[11, 17]. Indeed, recently a 96% 3-year post-relapse OS was reported in the updated analysis of the AML17 trial, despite less than half of the patients (11 out of 25) receiving a transplant[17].

Safety of arsenic trioxide

Overall, treatment with ATO is well-tolerated both in monotherapy and when ATO is combined with ATRA. Compared with AIDA, the combination of ATRA and ATO provides an important advantage of reducing haematological toxicity[6]. Furthermore, no cases of secondary myelodysplastic syndrome or AML were observed among ATO-treated patients in the APL0406 and AML17 trials, compared with a total of 3 chemotherapy-treated patients across both trials[5, 6]. While the combination of ATRA + ATO is associated with adverse events including QTc prolongation, hepatotoxicity, leukocytosis and differentiation syndrome, in clinical trials these were mostly managed with temporary treatment discontinuation and supportive care, with few permanent discontinuations being reported. Finally, although ATO has only recently been approved for the treatment of patients with newly-diagnosed APL, its safety profile is well-established through clinical studies conducted in the relapsed/refractory setting and the substantial volume of post-marketing data collected by Teva since 2000 when ATO was first approved in the US. Furthermore, as second line patients tend to have more health issues than first-line patients, it can be assumed that the current pharmacovigilance data provides reassurance on the safety profile of ATO usage in first-line patients

Economic analysis

With APL being a rare disease, very few economic studies are available and none specific to the UK. A *de novo* cost-utility analysis was therefore conducted for ATO+ATRA in the treatment of newly-diagnosed adult patients with low- to intermediate-risk APL.

Considering the very low relapse rate observed with ATO in RCTs, it is expected that once National Health Service (NHS) funding is available clinical practice will shift

towards the use of ATO as standard of care in newly-diagnosed patients. Therefore, the use of second-line treatments (including ATO) will decrease and will be driven mostly by the type of prior (first-line) therapy received. Consequently, the relapsed/refractory APL population was not evaluated separately, but rather analysed in relation to the newly-diagnosed population. Thus, the model provides an overall ICUR for ATO use in both first and second line.

In the base-case scenario, the combination of ATRA + ATO was associated with an incremental gain of 2.62 QALYs and provided a saving of £31,270 compared with AIDA over a lifetime horizon. Thus, the combination of ATRA + ATO was dominant and no base-case ICUR was calculated. The results of deterministic and probabilistic sensitivity analyses supported the robustness of these conclusions.

This economic analysis suggests that the use of ATRA + ATO in accordance with its licensed indication is a clinically-effective and cost-effective allocation of NHS resources in England and Wales.

References

1 Sant M, et al. Incidence of hematologic malignancies in Europe by morphologic subtype: results of the HAEMACARE project. *Blood*. 2010;116(19):3724-34.

2 Visser O, et al. Incidence, survival and prevalence of myeloid malignancies in Europe. *Eur J Cancer*. 2012;48(17):3257-66.

3 Dores GM, et al. Acute leukemia incidence and patient survival among children and adults in the United States, 2001-2007. *Blood*. 2012;119(1):34-43.

4 European Medicines Agency. Committee for Orphan Medicinal Products (COMP). [cited Oct, 2017]. Available from: http://www.ema.europa.eu/ema/index.jsp?curl=pages/about_us/general/general_content_000263.jsp&mid=WC0b01ac0580028e30.

5 Burnett AK, et al. Arsenic trioxide and all-trans retinoic acid treatment for acute promyelocytic leukaemia in all risk groups (AML17): results of a randomised, controlled, phase 3 trial. *The Lancet Oncology*. 2015;16(13):1295-305.

6 Platzbecker U, et al. Improved Outcomes With Retinoic Acid and Arsenic Trioxide Compared With Retinoic Acid and Chemotherapy in Non-High-Risk Acute Promyelocytic Leukemia: Final Results of the Randomized Italian-German APL0406 Trial. *J Clin Oncol*. 2017;35(6):605-12. English.

7 Kremer LC, et al. Anthracycline-induced clinical heart failure in a cohort of 607 children: long-term follow-up study. *J Clin Oncol*. 2001;19(1):191-6. eng.

8 Thomas X, et al. Anthracycline-related toxicity requiring cardiac transplantation in long-term disease-free survivors with acute promyelocytic leukemia. *Ann Hematol*. 2002;81(9):504-7. eng.

9 Godley LA, Larson RA. Therapy-related myeloid leukemia. *Semin Oncol*. 2008;35(4):418-29.

10 Iland HJ, et al. Have all-trans retinoic acid and arsenic trioxide replaced all-trans retinoic acid and anthracyclines in APL as standard of care. *Best Pract Res Clin Haematol*. 2014;27(1):39-52. eng.

11 Lengfelder E, et al. Arsenic trioxide-based therapy of relapsed acute promyelocytic leukemia: registry results from the European LeukemiaNet. *Leukemia*. 2015;29(5):1084-91.

12 Lo-Coco F, et al. Retinoic acid and arsenic trioxide for acute promyelocytic leukemia. *N Engl J Med*. 2013;369(2):111-21.

13 Lo-Coco F, et al. Targeted Therapy Alone for Acute Promyelocytic Leukemia. *N Engl J Med*. 2016;374(12):1197-8.

14 Soignet SL, et al. United States multicenter study of arsenic trioxide in relapsed acute promyelocytic leukemia. *J Clin Oncol*. 2001;19(18):3852-60. eng.

15 Soignet SL, et al. Complete remission after treatment of acute promyelocytic leukemia with arsenic trioxide. *N Engl J Med*. 1998;339(19):1341-8. eng.

16 Raffoux E, et al. Combined treatment with arsenic trioxide and all-trans-retinoic acid in patients with relapsed acute promyelocytic leukemia. *J Clin Oncol*. 2003;21(12):2326-34. English.

17 Russell NH, et al. Long Term Follow up from the NCRI AML17 Trial of Attenuated Arsenic Trioxide and ATRA Therapy for Newly Diagnosed and Relapsed Acute Promyelocytic Leukaemia. *Blood*. 2016;128(22):897.

Contents

List of Tables

List of Figures

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

B 1.1 Decision problem

This technology appraisal evaluates clinical and cost-effectiveness of arsenic trioxide (ATO) for the treatment of acute promyelocytic leukaemia (APL). It covers the full marketing authorisation for ATO, that is induction of remission, and consolidation in adult patients with:

- newly diagnosed low-to-intermediate risk APL (white blood cell count ≤10×10^3/µl) in combination with all-trans retinoic acid (ATRA) *(also referred to as first-line treatment)*
- relapsed/refractory APL (previous treatment should have included a retinoid and chemotherapy) *(also referred to as second-line treatment)*

characterised by the presence of the t(15;17) translocation and/or the presence of the promyelocytic leukaemia/retinoic-acid-receptor-alpha (PML/RAR-alpha) gene.

No technology appraisal guidance in APL has been published in the UK to date. ATO has never been assessed by the NICE, so it was requested that the submission should cover not only the newly-approved (November 2016) first-line indication, but also the second- line indication, in which ATO has been approved in Europe since March 2002. The full decision problem is described in ***Table 1.1***, along with any differences between this submission and the Final Scope published by NICE, and the rationale for these differences.

Table 1.1. The decision problem

	Final scope issued by NICE	Decision problem addressed in the company submission	Rationale if different from the final NICE scope
Population	Adults with: • untreated low-to-intermediate risk acute promyelocytic leukaemia • relapsed/refractory acute promyelocytic leukaemia (APL)	Adults with: • untreated low-to-intermediate risk acute promyelocytic leukaemia • relapsed/refractory acute promyelocytic leukaemia (APL) characterised by the presence of the t(15;17) translocation and/or the presence of the promyelocytic leukaemia/retinoic-acid-receptor-alpha (PML/RAR-alpha) gene.	None
Intervention	ATO (with or without ATRA)	• First-line treatment: ATO combined with ATRA; both administered according to the APL0406[1] protocol. AML17[2] protocol was studied as a scenario. See section ***B 2.3.1*** for the differences between these two protocols. • Second-line treatment: ATO administered according to the SPC + ATRA administered according to the APL0406[1] protocol (as in first line). The AML17 protocol[2] was studied in a scenario analysis.	In line with both the pivotal APL0406 trial[1, 3] and the AML17 trial[2], ATO is authorised for use in newly-diagnosed patients in combination with ATRA. No treatment combinations are specified for use in relapsed/refractory patients, although in the AML17 trial treatment with ATRA+ATO (administered as in first line) was used in patients who relapsed[4]. Based on clinical expert opinion, it appears ATO alone (without ATRA) is now rarely used in the relapsed/refractory setting. Thus, for both first- and second-line treatment, only the ATRA+ATO combination was considered in the economic analysis.
Comparator(s)	• AIDA regimen (ATRA in combination with idarubicin) • haematopoietic stem cell transplantation (HSCT) (people with relapsed or refractory APL)	• Following a relapse, the choice of therapy strongly depends on prior treatments the patient has received. It is therefore difficult to separate first- and second-line indications of ATO, as they're	• In the second-line indication, HSCT was not considered as a direct comparator, since administration of ATRA+ATO usually precedes transplantation rather than replaces it. Upon relapse, ATRA+ATO can be used to induce remission, which, if possible, would be consolidated with HSCT[5, 6]. Although additional ATO (+ ATRA) cycles may be used in patients who

	Final scope issued by NICE	Decision problem addressed in the company submission	Rationale if different from the final NICE scope
	• best supportive care (people with relapsed or refractory APL)	closely linked. To optimally reflect the treatment pathway of APL patients in the UK, Teva has decided to submit a single model which evaluates of the cost-effectiveness of ATO (+ATRA) in newly-diagnosed patients (first-line indication) with second-line treatments included, rather than presenting a separate cost-effectiveness evaluation of ATO as a second-line treatment. • For first-line treatment, AIDA was the comparator considered in both the pivotal APL0406 trial[1, 3] and in the economic analysis • For the second-line part of the model, we considered a situation where ATO was available first-line and some of the patents who received ATO first line switched to AIDA in second line, so that AIDA was retained as the comparator.	do not undergo a transplant[6, 7], ATO-based maintenance treatment is not included in the licensed administration schedule, and was therefore not considered in the economic analysis. Furthermore, other maintenance treatment options are also available to APL patients who do not undergo transplantation[6], and it would be difficult to include all of them without overtly complicating the analysis. We therefore took a simplified approach of not modelling second-line maintenance treatment, especially given that the number of patients concerned would be very small. • Best supportive care was not considered as a direct comparator in the second-line indication. Following ATO-based treatment of first APL relapse, Lengfelder et al. reported 3-year EFS of ≥45%[7], suggesting that attempting curative treatment may be most appropriate in patients with relapsed/refractory APL. Given the severity of APL, best supportive care can be seen as a palliative approach, and thus expected to be used where the disease is refractory to all other treatments, including ATO in second (or subsequent) treatment lines. Thus, it is unlikely that best supportive care will be considered an alternative to ATO or AIDA (see below) for treatment of relapsed APL. It is, however, worth noting that the economic analysis does take into account best supportive care – upon failure of second-line treatment, patients in the model progressed to an end-of-life state, where they received palliative care. • The choice of second-line treatment is largely determined by the first-line therapy that the patient has received, and ATO (usually + ATRA) is the standard treatment for APL relapses after first-line treatment containing ATRA and an anthracycline (e.g. AIDA). However, the choice of optimal salvage treatment in patients who relapse following first-line ATO use is less clear. This is largely due to the absence of established guidelines, as many treatment guidelines in APL (e.g. from the European LeukemiaNet[8] and ESMO[9]) precede the approval of

	Final scope issued by NICE	Decision problem addressed in the company submission	Rationale if different from the final NICE scope
			ATO for first-line use. In the economic analysis, treatment of relapses following first-line ATO use was therefore based on clinical expert opinion. It was assumed that patients who remained in remission for ≥2 years following first-line ATRA+ATO treatment were re-treated with ATRA+ATO upon relapse. However, patients who achieved only a short (<2 years) remission after first-line treatment with ATRA+ATO, were assumed to be treated with AIDA upon relapse. Thus, AIDA was considered as a comparator also in the relapsed/refractory APL setting. See section ***B 1.3.2*** for more details of APL treatment pathway.
Outcomes	• Overall survival (OS) • Progression-free survival (PFS) • Response rates (bone marrow remission) • Adverse effects of treatment • Health-related quality of life (HRQoL)	• OS • Event-free survival (EFS) • Complete haematological and molecular remission rates • Cumulative incidence of relapse (CIR) • Disease-free survival (DFS) or relapse-free survival (RFS) • Adverse effects of treatment • HRQoL	• PFS was not an endpoint in the pivotal APL0406 trial[1, 3] or in the AML17 trial[2], and is thus not presented. Instead, the manufacturer presented data on EFS – the primary endpoint of the APL0406 trial[1, 3]. It is, however, worth noting that in the APL0406 trial patients failing treatment were those who did not achieve remission, relapsed, or died (see ***section B 2.4***), which is similar to what would be considered treatment failure when analysing PFS. In the AML17 trial, an additional event of treatment-related myelodysplastic syndrome or acute myeloid leukaemia (tMDS/AML) was also included in the EFS analysis; however, only a single patient in this study developed tAML[2], so that inclusion of this event in EFS evaluation could be considered to have little effect on the overall result. In conclusion, although EFS rather than PFS is presented, the two outcomes are similar, so this does not represent a major deviation from the scope. • In addition to the outcomes listed in the Final Scope, the manufacturer will also present data on cumulative incidence of relapse and DFS (or RFS), if available. Given the curative intent of APL treatment, these endpoints are of particular importance, as they provide information on the proportion of patients who remain disease-free.

B 1.2 Description of the technology being appraised

B 1.2.1 Description of the technology

ATO (Trisenox®, Teva Pharma B.V.) is a form of naturally occurring arsenic believed to have multiple mechanisms of action in APL, including inducing cell death by damaging or degrading the PML/RARα fusion protein – the product of the genetic mutation characterising APL. The use of ATO offers a chemotherapy-free treatment option for patients with APL. ATO is administered by intravenous infusion and should be used under the supervision of a physician who is experienced in the management of acute leukaemias; monitoring procedures outlined in the SPC (see ***Appendix C***) should be followed.

Table 1.2. Technology being appraised

UK approved name and brand name	Arsenic trioxide (Trisenox®)
Mechanism of action	The mechanism of action of ATO is not completely understood[10]. ATO causes morphological changes and deoxyribonucleic acid (DNA) fragmentation characteristic of apoptosis in NB4 human promyelocytic leukaemia cells in vitro[10]. It also causes damage or degradation of the PML/RAR alpha fusion protein[10].
Marketing authorisation/CE mark status	ATO has been approved in the US for the treatment of relapsed or refractory patients since 2000[11] and in the EU since 2002[12]. In November 2016 it was approved in the EU for the treatment of newly-diagnosed patients with low- to intermediate-risk APL. ATO received orphan designation for the treatment of APL from the FDA in March 1998[11] and from the EMA in October 2000[12]. The period of market exclusivity (related to orphan drug status) has since ended in both the US and the EU.
Indications and any restriction(s) as described in the summary of product characteristics (SmPC)	ATO is indicated for induction of remission, and consolidation in adult patients with: • Newly diagnosed low-to-intermediate risk acute promyelocytic leukaemia (APL) (white blood cell count, ≤10×10^3/µl) in combination with all-trans-retinoic acid (ATRA) • Relapsed/refractory APL (Previous treatment should have included a retinoid and chemotherapy) characterised by the presence of the t(15;17) translocation and/or the presence of the Pro-Myelocytic Leukaemia/Retinoic-Acid-Receptor-alpha (PML/RAR-alpha) gene. The response rate of other acute myelogenous leukaemia subtypes to arsenic trioxide has not been examined.
Method of administration and dosage	ATO must be administered under the supervision of a physician who is experienced in the management of acute

	leukaemias. Special monitoring procedures apply: • Prior to initiating therapy with ATO, a 12-lead electrocardiogram (ECG) must be performed and serum electrolytes (potassium, calcium, and magnesium) and creatinine must be assessed. Pre-existing electrolyte abnormalities must be corrected and, if possible, medicinal products that are known to prolong the QT interval must be discontinued. • Patients with risk factors for QTc prolongation or risk factors for torsade de pointes should be monitored with continuous cardiac monitoring (ECG). • Patient's electrolyte and glycaemia levels, as well as haematological, hepatic, renal and coagulation parameters must be monitored at least twice weekly (and more frequently for clinically unstable patients) during the induction phase, and at least weekly during the consolidation phase. • Patients should be monitored for the appearance of adverse events associated with ATO, such as differentiation syndrome and hyperleukocytosis (occurring mainly during induction therapy), and hepatotoxicity. ATO dosing and administration schedule depends on the specific indication (first- or second-line) in which it is used: ***Newly diagnosed low-to-intermediate risk APL*** Induction treatment schedule • ATO must be administered intravenously at a dose of 0.15 mg/kg/day, given daily until complete remission is achieved. If complete remission has not occurred by day 60, dosing must be discontinued. Consolidation schedule • ATO must be administered intravenously at a dose of 0.15 mg/kg/day, 5 days per week. Treatment should be continued for 4 weeks on and 4 weeks off, for a total of 4 cycles. ***Relapsed/refractory APL*** Induction treatment schedule • ATO must be administered intravenously at a fixed dose of 0.15 mg/kg/day given daily until complete remission is achieved (<5% blasts present in cellular bone marrow with no evidence of leukaemic cells). If complete remission has not occurred by day 50, dosing must be discontinued. Consolidation schedule • Consolidation treatment must begin 3 to 4 weeks after completion of induction therapy. ATO is to be administered intravenously at a dose of 0.15 mg/kg/day for 25 doses given 5 days per week, followed by 2 days interruption, repeated for 5 weeks.
Additional tests or investigations	ATO is indicated for the treatment of APL characterised by the presence of the t(15;17) translocation and/or the presence of the Promyelocytic Leukaemia/Retinoic-Acid-Receptor-alpha (PML/RAR-alpha) gene. This translocation accounts for up to 98% of APL cases; however, other translocations involving the *RARA* gene have also been identified in APL[13]. It is widely accepted that the diagnosis of APL (as opposed to other types of AML) should be confirmed through molecular testing for PML-RARA. Although the pivotal APL0406 trial accepted a

	number of methods through which genetic confirmation of APL diagnosis could be established [1], the diagnostic tests that appear most feasible for routine use are polymerase chain reaction (PCR) and fluorescent *in situ* hybridisation (FISH). APL patients also undergo repeated bone marrow biopsies and the collected material is PCR-tested for the presence of PML-RARA, which allows the treating clinician to establish how the patient responds to treatment (i.e. if molecular remission has been achieved or if minimal residual disease can be detected), and to monitor the patient for molecular relapse (i.e. the reappearance of PML-RARA in the bone marrow), which allows second-line treatment to be administered early, before the patient progresses into a full haematological relapse that may be life-threatening. The frequency of monitoring depends on treatment choice.
List price and average cost of a course of treatment	£2,920 / pack of 10 ampoules. On average, a patient requires approximately 12/13 packs of Trisenox over the full course of treatment (induction and four consolidation cycles), amounting to £35k–£38k per patient. Note that treatment should be completed within less than a year, with induction therapy taking less than 2 months (up to 60 days) and the full consolidation schedule spanning 28 weeks (6.4 months).
Patient access scheme (if applicable)	Not applicable

B 1.2.2 Impact on healthcare services

ATO has been approved in the relapsed/refractory APL setting for the last 15 years and can be considered the standard of care for second-line APL treatment in the UK. Consequently, any impact on service delivery from introducing ATO-based treatment for relapsed/refractory APL in the UK would already have taken place. Based on a European registry of relapsed APL, Lengfelder, et al.[7] reported that ATO (± ATRA) can successfully induce a second remission in nearly 90% of patients presenting with haematological relapse, and that approximately 60% of relapsed APL patients undergo HSCT as post-consolidation therapy. Thus, the use of ATO in second-line likely increased the eligibility for HSCT procedures, due to a number of patients successfully achieving a second remission and receiving further treatment to consolidate it. However, with the introduction of the ATRA+ATO combination in first-line, it is estimated that the number of patients who relapse will decrease, consequently reducing the need for HSCT procedures (and the associated costs, see section ***B 3.7***).

Another way in which the use of ATRA+ATO in first line may impact healthcare services is through reduced need for inpatient treatment. During the induction phase, patients require particularly close monitoring and are generally hospitalised, whether they receive ATO- or chemotherapy-based treatment. During consolidation, however, patients are able to receive treatment with ATO primarily in the day-care

setting, unless otherwise indicated. The UK-based AML17 trial reported a significantly shorter duration of hospital stay in patients receiving ATRA+ATO during both induction and the first consolidation cycle compared with AIDA-treated patients[2]. During the initial consolidation cycle, the median (interquartile range) length of hospital stay was 1 (0–10) days in the ATRA+ATO group compared with 5 (2–10) in the AIDA group[2], implying a sizeable proportion of patients receiving ATRA+ATO required no hospitalisation, or only a short hospital stay during the first consolidation cycle.

Further, the APL0406 trial showed that, in first-line APL treatment, the combination of ATRA+ATO produces less haematological toxicity than the AIDA regimen[1, 3]. In the UK, the AML17 trial did not directly report haematological toxicity, but showed a reduced requirement for both blood and platelet transfusions among ATRA+ATO treated patients (compared with those treated with AIDA) during induction and the first consolidation cycle[2], suggesting patients receiving ATO-based first-line treatment are likely to need less intensive transfusions in clinical practice.

In terms of drug administration, during the consolidation phase ATO requires more frequent dosing than chemotherapy (80 infusions as opposed to 10 in the AIDA regimen). Although, as mentioned above, many patients should be able to receive ATO in the day clinic setting without the need for prolonged inpatient stay, the additional intravenous infusions will need to be accommodated for. In addition, patients would typically have an ECG before each ATO dose or every week – this is directly related to QTc prolongation frequently observed with ATO (see ***section B 2.11***).

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

B 1.3.1 APL overview

APL is a distinct subtype of acute myeloid leukaemia (AML), classified by the WHO within the category of AML with recurrent genetic abnormalities[14] and as M3 according to the French-American-British (FAB) classification[15]. At the genetic level, APL is caused by a translocation between chromosomes 15 and 17, abbreviated as t(15;17), fusing the *PML* gene with the *RARA* gene, which results in formation of the PML-RARα fusion protein[13]. Although up to 98% of APL cases are caused by *PML-RARA* gene fusion, other translocations involving the RARA gene have been identified in some cases of APL[13].

APL is a rare disease; however, precise incidence estimates vary between reports. In an analysis of 2000–2002 data from 44 cancer registries across Europe, Sant et al. reported that APL and other AMLs with recurrent genetic abnormalities jointly constituted just under 4% of all AML cases, with an overall annual crude incidence rate of 0.14 per 100,000[16]. Similarly, Visser et al. extracted data from the European Cancer Registry-based EUROCARE-4 study, analysing information on patients diagnosed between 1995 and 2002 from 64 European cancer registries, and found that 3% of observed AML cases could be attributed to APL[17], resulting in a crude annual incidence rate of 0.11 per 100,000 people[17]. In the same study, complete prevalence of APL was estimated at 0.6 per 100,000, with just over 3,200 people affected across the EU27[17]. Other estimates of APL incidence are, however, somewhat higher. A study by Dores, et al. based on the US Surveillance, Epidemiology and End Results (SEER) Program registry reported that age-adjusted incidence of APL was 0.27 per 100,000 person-years, with the disease accounting for 7.4% of AML cases[18].

In terms of the age at which patients are most often affected, Dores, et al. reported a median age of 47 years at diagnosis[18]. The incidence of APL was low in children under the age of approximately 10 years, but rose steeply during the teen years to remain almost constant through adult working years, and decreased again in the elderly[18]. Similar findings were seen in a UK study by Vickers et al., which analysed 159 APL cases from four Regional Leukaemia Registries[19]. This age distribution is a key difference between APL and most other AML types, which are diagnosed at a median age exceeding 60 years[18]. Thus, APL is likely to pose a considerable societal burden, affecting people of working age. Regarding other patient characteristics, men and women are equally affected by the disease[17, 18], but the incidence does appear to vary by ethnicity, being higher in Hispanics[18].

Although initial symptoms of APL, such as fatigue, abnormal bruising and bleeding may be initially inconspicuous, the disease can progress rapidly with very poor survival prognosis. Retrospective analyses report that 10–29% of patients die within 30 days of hospital admission or diagnosis[20-23] (i.e. usually during induction therapy); 31–55% of these deaths result from haemorrhage (CNS[20-22] and pulmonary[20, 22]), the risk of which is considerable in APL, due to coagulopathy frequently associated with the disease. Even in the setting of recent randomised controlled trials (RCTs), mortality during induction treatment can be considerable. Four deaths were recorded during induction therapy among 136 evaluable patients (3%) in the chemotherapy arm of the APL0406 trial [3] (see section ***B 2.7.1***). In the AML17 trial (see section ***B 2.7.1***), 11 of 119 patients randomised to ATRA + chemotherapy (9%) died by day 60

(i.e. by the end of the induction phase), as did six of 116 patients (5%) randomised to ATRA+ATO[2]. Treatment guidelines from the European LeukemiaNet recommend that diagnostic suspicion of APL should be considered a medical emergency[8].

Relapse risk stratification plays an important role in determining the most appropriate treatment options for APL patients. Definition of relapse risk categories and development of risk-adapted treatment strategies are considered significant advances in the management of APL[24]. Assessment of relapse risk in APL is primarily based on white blood cell (WBC) count at presentation, with patients whose WBC count exceeds 10×10^9/L generally predicted to have a higher risk of relapse. Risk stratification was developed through a joint analysis of two multicentre trials (AIDA0493 and LPA96)[25]. Univariate analysis revealed that remission duration was significantly affected by WBC at presentation and showed a tendency towards being affected by platelet count[25]. In a multivariate analysis of DFS, WBC count $\leq10\times10^9$/L and platelet count $>40\times10^9$/L were significantly associated with a favourable prognosis[25]. Conversely, WBC count $>10\times10^9$/L and platelet count $\leq40\times10^9$/L were significantly associated with unfavourable prognosis[25]. Consequently, the following relapse risk categories were identified in a simplified predictive model for RFS: low-risk (WBC $\leq10\times10^9$/L and platelet count $>40\times10^9$/L), intermediate-risk (WBC and platelet counts $\leq10\times10^9$/L and $\leq40\times10^9$/L, respectively) and high-risk (WBC count $>10\times10^9$/L)[25]. This risk definition plays an important role when making treatment decisions in APL, as treatment is commonly risk-stratified. Especially, ATO is currently only licensed for the treatment of patients with newly-diagnosed low- to intermediate-risk APL, as well as all patients with relapsed or refractory disease.

B 1.3.2 Treatment strategies in newly-diagnosed and relapsed/refractory APL

Compared with many other subtypes of AML, APL requires a different treatment approach[9] and is associated with a more favourable prognosis, with over 70% of recently-treated patients achieving long-term (≥4-year) EFS[2] (>80% in the low- to intermediate-risk group[26]), compared with just up to around 40% of AML patients who can be considered cured[27]. Thus, treatment of APL has a curative aim, that is to achieve and maintain molecular remission and provide the patient with chances of long-term disease-free survival. Long-term survival in APL remission can be achieved in ≥70% of newly-diagnosed patients following chemotherapy-based treatment and >90% following ATRA+ATO[2, 3]. Although these figures were traditionally lower in relapsed or refractory patients (≥45% following second-line ATO-based treatment in a European registry of relapsed APL[7]), recent studies suggest treatment outcomes in this setting are improving. A report on the 25 patients

who relapsed following AIDA treatment in the AML17 trial and received ATRA+ATO mentioned three second relapses and two deaths in this group[4], so that the remaining 20 (80%) of patients remained alive in second CR at the time results were published.

B 1.3.2.1 European treatment guidelines in APL

First-line therapy in APL generally consists of three consecutive treatment phases: induction, consolidation and maintenance, although maintenance is usually omitted in the UK clinical practice with the aim of minimising the risk of treatment-related myelodysplastic syndrome or acute myeloid leukaemia (tMDS/AML). ***Table 1.3*** provides an overview of European guidelines for first-line treatment of APL. It is worth noting, that the treatment guidelines described below precede the approval of ATO for first-line treatment of APL. Consequently, the 2013 guidelines from the European Society for Medical Oncology (ESMO)[9] and the 2009 European LeukemiaNet guidelines[8] do not explicitly recommend an ATO-based first-line treatment regimen for wider use, acknowledging the paucity of available data at the time of guideline preparation[9] or restricting it to patients with contraindications to chemotherapy and clinical trials[8]. However, German guidelines from both the Deutsche Gesellschaft für Hämatologie und Medizinische Onkologie (DGHO)[28] and the German Intergroup[5] do list the ATRA+ATO combination as an option for treating newly-diagnosed low- to intermediate-risk patients.

In relapsed or refractory disease, ATO is the most widely recommended treatment across the four aforementioned sets of guidelines. The ESMO guidelines mention durable remissions achieved with ATO in this setting[9]. Similarly, the European LeukemiaNet guidelines recommend ATO-based regimens as the treatment of choice for patients with relapsed APL, although they also state ATRA in combination with chemotherapy may be used in this setting[8]. The German Intergroup and DGHO guidelines recommend ATO-based treatment for induction and consolidation in second line, with the exception of patients treated with ATO in first line, who should be ***switched*** to ATRA + anthracycline-based chemotherapy (with the addition of Ara-C in consolidation)[5, 28]. However, the DGHO guidelines mention that ***re-treatment*** with ATO may also successfully induce a second remission, albeit ATO efficacy in this case may be reduced[28].

Patients with relapsed or refractory APL may receive a HSCT to consolidate second remission[5, 8, 28], especially if they are considered at risk of additional relapses[9]. According to the European LeukemiaNet guidelines, autologous transplantation is an

option in patients who are PCR-negative (no evidence of PML-RARA on bone marrow PCR), while allogeneic HSCT is recommended in those failing to achieve a second molecular remission, as this transplant modality offers a greater antileukaemic activity due to graft-versus-leukaemia effect[8]. In patients who are not transplant candidates, additional ATO cycles (with or without other treatments, i.e. ATO or chemotherapy) may be used[8]. The German recommendations on transplantation in relapsed APL are broadly similar[5, 28]. In general, however, the choice of post-consolidation therapy depends not only on PCR status, but also on donor availability, age, clinical condition, and other considerations. Detailed recommendations can be found in the European recommendation for salvage therapy of relapsed APL[6].

Table 1.3. Phases of first-line APL treatment

Induction	Consolidation	Maintenance
Aim of treatment phase[5, 28]		
Induce APL remission.	Stabilise and maintain remission.	Stabilise and maintain remission*.
Regimens used during treatment phase[5, 8, 9, 28]		
ATRA + anthracycline-based chemotherapy	2–3 cycles of anthracycline-based chemotherapy – the addition of ATRA during consolidation appears to provide a clinical benefit. Addition of Ara-C appears to provide a benefit in high-risk patients.	Non-myeloablative chemotherapy – notably MTX, 6-MP, and ATRA.
ATRA+ATO (low- to intermediate-risk only)	4 cycles of ATRA+ATO	Not applicable
Duration of treatment phase[5, 28]		
Until CR is achieved, or up to 60 days.	Consolidation cycles should be separated by enough time to allow haematological recovery from the previous cycle.	Generally administered for 2 years.

6-MP=6-mercaptopurine; APL=acute promyelocytic leukaemia; Ara-C=cytarabine; ATO=arsenic trioxide; ATRA=all-trans retinoic acid; CR=complete remission; MTX=methotrexate
*Evidence for the clinical benefit of maintenance treatment remains inconclusive.

B 1.3.2.2 Treatment pathway in APL

To investigate actual treatment paradigms in APL, which may differ from available guidelines, Teva commissioned primary market research conducted in 2015 across seven European countries (Austria, France, Germany, Italy, Spain, Switzerland and the UK)[29]. This market research suggested that patients newly-diagnosed with APL

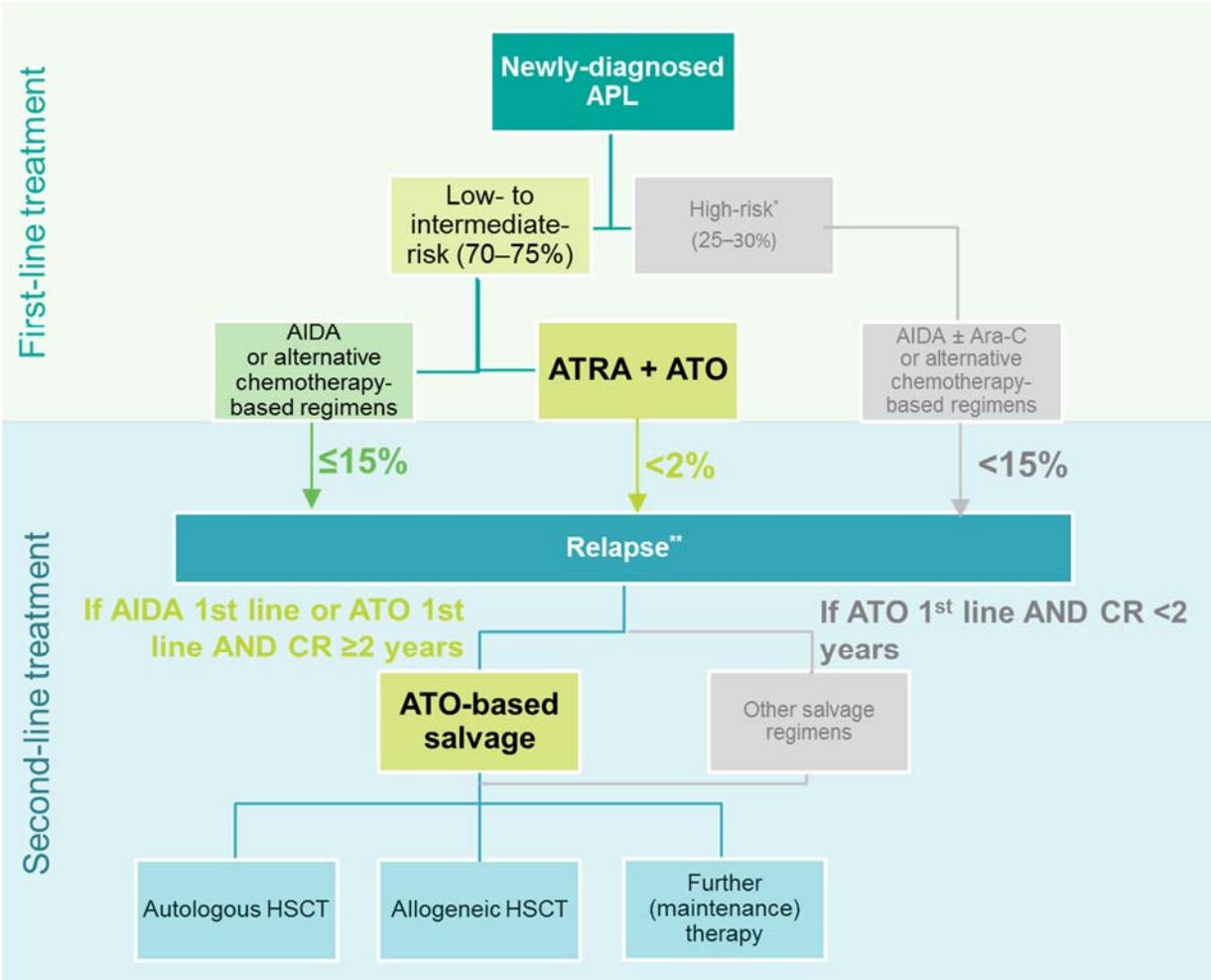
in the UK are commonly treated according to clinical trial protocols (MRC AML trials), as they are recommended to enrol in ongoing trials upon diagnosis[29]. AIDA (although generally administered without maintenance treatment which is rarely used in the UK[29]) is the standard chemotherapy-based treatment approach, based on the results of the AML15 trial, which showed no benefit of additional chemotherapy beyond idarubicin/mitoxantrone[30]. Recently, both clinical practice paradigms and some of the available guidelines suggest ATRA+ATO is becoming well-recognised as the new standard of care, replacing the combination of ATRA and chemotherapy for the treatment of newly-diagnosed APL in Europe.

According to expert opinion, patients in the UK are treated as soon as molecular relapse is detected and before the patient progresses into a haematological relapse. The risk of relapse is considered highest in the first two years, and patients are usually monitored during that time. However, as the incidence of relapse with ATRA+ATO is low (see ***section B 2.3***), patients treated with this combination in first line would not usually require PCR monitoring for molecular relapse, unlike those treated with chemotherapy. One a relapse is detected, UK patients would normally be ***switched*** from their first line therapy to an alternative (e.g. from AIDA to ATRA+ATO and from ATRA+ATO to AIDA)[29], so that the choice of second-line treatment is driven by the type of prior (first-line) therapy the patient has received. ATO-based second-line treatment is the standard approach to treating APL relapsed following first-line treatment including ATRA and chemotherapy[29]. However, there is a lack of well-established paradigms or guidelines for second-line treatment following ATRA+ATO administration in first line, and the field is constantly evolving with growing experience of first-line ATO use. Therefore, informed by a clinical expert opinion, we used a ***mixed re-treatment/switch approach*** in the economic analysis, which assumed that patients who remained in remission for 2 years or longer following first-line ATRA+ATO treatment were re-treated with ATRA+ATO upon relapse, while patients who achieved only a short (<2 years) remission after first-line treatment with ATRA+ATO were treated with AIDA. It is also worth noting that the model also included another mode of treatment switch, which could be prompted by the occurrence of a cardiac serious adverse event (SAE).

In eligible UK patients, second remission is often consolidated with a HSCT[29]. According to clinical expert opinion, allogeneic HSCT is generally used in patients who enter haematological remission following second-line treatment but fail to achieve molecular remission; in patients who achieve a second molecular remission, allogeneic HSCT is rarely considered due its associated risks. It is worth noting that clinical expert opinion suggests patients salvaged with ATO do not necessarily need

transplantation, while those salvaged with chemotherapy generally do. Indeed, the clinical expert mentioned that, based on the results of the AML17 study[4], ATRA+ATO is increasingly being applied without transplantation as a curative approach for UK patients with relapsed APL and no CNS involvement. However, in the absence of clear guidelines, the decision on transplantation is likely to be made individually for each patient; hence, HSCT was a possibility for all patients in the model following second-line treatment, regardless of the nature of salvage therapy received.

Figure 1.1 Simplified treatment pathway in APL showing the licensed indications for ATO



*Note that treatment of APL is frequently risk-adapted, with high-risk patients receiving more intensified treatment. ATO is currently not licensed for use in newly-diagnosed, high-risk patients.
**According to clinical expert opinion, patients in the UK are treated as soon as a relapse is detected – this is frequently at the molecular level, before the patient progresses into a haematological relapse.
The percentage of patients relapsing after first-line treatment including chemotherapy is based on the following trials: LPA 2005[31], AIDA-2000[32], German AMCLG[33] and APL2000 trials[34], while the percentage relapsing after ATO-based first-line treatment in based on the APL0406 trial[3]. The choice of ATO-based vs. other salvage treatment was based on an assumption.

The perceived role of ATO in the treatment of both newly-diagnosed and relapsed/refractory patients with APL is presented in ***Figure 1.1***, which is also broadly aligned with the structure of the economic model described in ***section B 3.2.2***, although treatment of high-risk patients and second-line maintenance treatment were not modelled. Note, that the relapse incidence estimates in this figure are likely to be rather conservative – in the AIDA arm of the recent UK AML17 trial including patients of all risk groups, 4-year cumulative incidence of haematological relapse was 18% (vs 1% in the ATO arm) and that of molecular relapse was even higher at 27% (0% in the ATO arm)[2].

B 1.3.3 Ultra-orphan considerations

With an estimated incidence of just over 1 in a million people[16, 17] and complete prevalence of 6 in a million[17], APL is classified as a rare disease according to the European Medicines Agency (EMA), which defines rare diseases as those that affect no more than 5 in 10,000 people in the European Union (EU)[35]. ATO received orphan designation for the treatment of APL from the FDA in March 1998[11] and from the EMA in October 2000[12]. The period of market exclusivity (related to orphan drug status) has since ended in both the US and the EU.

The final NICE scope reported 2,590 people were diagnosed with AML in England in 2014. However, ATO is only licensed in adult patients. Clinical expert opinion suggests AML in children is rare, with approximately 70 cases diagnosed per year, so that there are an estimated 2,520 adult AML patients in England. According to the studies identified in section ***B 1.3.1***, APL constitutes up to 7.4% of AML cases, corresponding to 187 adult patients with APL diagnosed per year in England. As ATO is only licensed for use in low- to intermediate-risk patients, who accounted for 75.7% of patients in the recent UK-based AML17 trial, 142 patients are likely to meet the criteria for ATO-based treatment every year. Consequently, APL can be classified as an ultra-orphan disease according to the definition used by NICE, whereby an ultra-orphan condition affects fewer than 1,000 patients in the UK[36]. It is also worth noting that 10–29% of APL patients die within 30 days of hospital admission or diagnosis[20-23]. Those suffering an early death would only receive some of their induction treatment, but no consolidation, and as many as 21% of these patients have been reported to receive no treatment at all[20]. Thus 133–139 patients in England could be expected to receive ATRA+ATO treatment in first-line; although there may be some patients who are intolerant to ATO, so that the actual number is likely to be smaller.

With no first-line ATO treatment and assuming a relapse rate of 15% (see ***Figure 1.1***) amongst newly-diagnosed patients treated with ATRA and chemotherapy, approximately 28 patients would be expected to receive second-line treatment with ATO per year (15% of 187). However, if the relapse incidence from the AIDA arm of the UK AML17 study is used, this figure would rise to 51 patients (27%[2] of 187). In contrast, treating newly-diagnosed low- to intermediate-risk APL patients with ATRA+ATO results in a relapse incidence of up to 2%[2, 3], so that approximately 10-16 patients per year are expected to be treated for relapsed disease (depending on whether a relapse rate of 15% or 27% is assumed in high-risk patients) following National Health Service (NHS) funding for first-line ATO-based treatment.

B 1.4 Equality considerations

We have not identified any specific studies evaluating equality of access to APL therapy in the UK. However, previous studies have shown that age and gender are linked to inequality of access to anti-cancer drugs[37]. Although further studies are needed to evaluate the impact of those factors on access to APL treatments, elderly patients who are not eligible for chemotherapy would still be eligible for treatment with ATO. Therefore, making ATO available on the NHS is likely to allow a greater number of elderly patients to be treated, which may be an important step towards addressing the topical issue of under-treatment among elderly oncology patients.

Another group for which equality can be a concern are Jehovah's Witness patients, as one of the most significant teachings of the Jehovah's Witness church is abstinence from receiving blood transfusions. However, with the low prevalence of APL and the relatively small number of Jehovah's Witness in the UK (estimated at around 1 in 450 people) this may only concern 1 patient or less per year. However, as ATO+ATRA may decrease the number of transfusion compared to AIDA, an ATO-based treatment approach may be more acceptable to this patient group.

B.2 Clinical effectiveness

The trials described in ***sections 2.4–2.8*** (i.e. APL0406, AML17 and Raffoux, et al.) were not sponsored by Teva, so the manufacturer only has access to published data. The regulatory approval of the ATRA+ATO combination in newly-diagnosed low- to intermediate-risk APL was based primarily on the APL0406 trial which was investigator-initiated, and supported by the UK-based AML17 trial sponsored by the National Cancer Research Institute (NCRI). A comprehensive summary of the published data from these trials is presented in this section.

B 2.1 Identification and selection of relevant studies

In order to address the decision problem presented in section ***B 1.1***, we conducted two systematic literature reviews gathering evidence on clinical effectiveness and safety of ATO and other available treatments for APL (see ***section B 2.2*** and ***Appendix D*** for details). Expecting a limited evidence base in the first-line indication due to the rarity of APL and the fact ATO has only recently been approved in first line, the manufacturer decided to include non-RCTs in the review in order to provide the widest possible range of data. Therefore, while one search aimed to provide information on RCT-based efficacy and safety, the other collected data from non-randomised studies.

B 2.2 Systematic literature review

B 2.2.1 Search strategy

We searched the following databases for both RCTs and non-RCTs:

- Medline® (Ovid MEDLINE® In-Process & Other Non-Indexed Citations and Ovid MEDLINE® 1946–Present),
- EMBASE® (Ovid EMBASE),
- ASCO abstracts (2011–2017)
- American Society of Hematology (ASH) Annual Meeting Abstracts (2011–2017)
- European Hematology Association (EHA) congress abstracts (2011–2017)

In addition, the following databases were also searched specifically for RCTs:

- ClinicalTrials.gov
- Cochrane Central Register of Controlled Trials

Initially, searches were performed on July 20, 2016; all were subsequently updated on October 10, 2017. No changes to the search strategy were made for the update. All search terms and the relationships between them (e.g., Boolean operators) are

presented in ***Appendix D***. Although the precise search strategies differed from database to database, in line with the decision problem they were all designed to capture studies (RCTs and non-RCTs) enrolling adult (aged ≥16 years) patients with APL. While the search for RCTs included terms specifically designed to capture studies with this type of design, a range of designs was incorporated into the terms when searching for non-RCTs, including observational, prospective cohort (non-RCT), cross-sectional, and case-control studies, as well as patient registries and case series.

B 2.2.2 Study selection

The selection processes for studies identified in the systematic review are listed in ***Table 2.1*** (RCTs) and ***Table 2.2*** (non-RCTs), according to the population, intervention, comparators, outcomes and study design (PICOS) criteria. The same eligibility criteria were used during the initial search and the subsequent update, and no language or geographical scope restrictions were applied.

Table 2.1. Eligibility criteria used in the search strategy for RCTs

PICOS	Inclusion criteria	Exclusion criteria
Population	• Adult participants with APL, aged ≥16 years, of both genders	• Paediatric-only population • High-risk newly diagnosed APL • Significant cardiac comorbidities • Significant pulmonary comorbidities • Active non-APL malignancy • Pregnant women • Women who were breastfeeding during the time of the study
Intervention	• Any intervention	• None
Comparators	• Any comparator	• None
Outcomes	• Adverse events • OS • EFS • DFS or RFS • Cumulative incidence of relapse • Response rates (complete haematological and molecular remission rates)	• None
Study design	• RCTs, Phase II/III studies, systematic literature reviews of RCTs, or meta-analyses	• Opinion, editorial letter
Other	• None	• Old conference abstracts:

		conference abstracts published prior to 2014 were excluded, based on an assumption that if they were of good quality, the study would have been published as a paper by the time of the search, and thus would be captured anyway. • No full text available online: if the link to the full text of the study was not found after searching in multiple databases in relevant languages, the study was excluded. • Chinese articles published in non-core journals were excluded, due to their frequently poor quality. Furthermore, Trisenox® is not marketed in China, so Chinese studies may be expected to report on the use of other ATO formulations. Nonetheless, relevant Chinese articles that met the inclusion criteria are summarised in ***Appendix L***.

APL=acute promyelocytic leukaemia; ATO=arsenic trioxide; DFS=Disease-free survival; EFS= Event-free survival; OS=Overall survival; RCT=Randomised Controlled trial; RFS=Relapse-free survival

The PICOS criteria applied to non-RCTs were similar and are presented in ***Table 2.2*** below. However, the search for non-RCT evidence was focused on the use of ATO in the recently approved first-line indication. This was motivated by the well-established and widespread use of ATO in relapsed/refractory APL, and the fact it has long been considered first-choice therapy for induction and consolidation in this setting. A relevant review published as early as 2005 concluded ATO was the therapy of choice for patients in relapsed/refractory APL, as it resulted in high complete and molecular remission rates and lower toxicity compared to ATRA + chemotherapy[38]. A similar stance was taken by authors of a more recent review[39], supporting ongoing ATO use in the second-line indication. Furthermore, as described in ***section B 1.3.2***, ATO is widely recommended by European clinical guidelines for the treatment of patients with relapsed/refractory APL. Thus, only studies discussing first-line ATO use were included when searching for relevant non-RCTs.

Two independent reviewers screened the list of unique titles and abstracts identified through the search, in order to determine the eligibility of each study based on inclusion and exclusion criteria. The two lists of selected references were then compared and all disagreements were solved by discussion, or if persistent, by a

third reviewer. If abstracts of potentially relevant publications were unavailable, full text publications were retrieved and screened to check if they meet the eligibility criteria.

Table 2.2. Eligibility criteria used in the search strategy for non-RCTs

PICOS	Inclusion criteria	Exclusion criteria
Population	• Adult participants with APL, aged ≥16 years, of both genders	• Paediatric-only population (aged ≤15 years)
Intervention	• Any intervention	• None
Comparators	• Any comparator	• None
Outcomes	• Adverse events • OS • EFS • DFS or RFS • Cumulative incidence of relapse • Response rates (complete haematological and molecular remission rates)	• None
Study design	• Observational study • Cohort study • Prospective study (non-RCT) • Patient registry • Cross sectional study • Case-control study • Cases series including 6 cases or more	• Opinion, editorial letter • RCTs • Case reports • Case series with ≤5 cases
Other	• None	• Old conference abstracts: conference abstracts published prior to 2014 were excluded, based on an assumption that if they were of good quality, the study would have been published as a paper by the time of the search, and thus would be captured in the search anyway. • Studies including a population of <50 patients • Studies that did not include ATO in first line were excluded

APL=acute promyelocytic leukaemia; ATO=Arsenic trioxide; DFS=Disease-free survival; EFS= Event-free survival; OS=Overall survival; RCT=Randomised Controlled trial; RFS=Relapse-free survival

Full text screening was also performed by two independent reviewers. The primary reason for exclusion was recorded, based on the criteria outlined in ***Table 2.1*** and ***Table* 2.2**. Selected publications were screened for duplicates, which were excluded.

Systematic reviews identified through the search were retrieved, and their lists of references were screened against the included studies.

One reviewer extracted relevant data from included studies and the results were reviewed by a senior manager to control the quality. PRISMA flow diagrams present the numbers of RCTs (***Figure 2.1***) and non-RCTs (***Figure 2.2***) included and excluded at each systematic review stage. A complete list of included and excluded studies is available in ***Appendix D***, alongside full details of the process and methods used to identify and select the relevant clinical evidence.

Network meta-analysis based on the RCTs identified through the literature review was not feasible, due to the lack of a mutual comparator. Details of feasibility assessment are provided in ***section 0***.

Figure 2.1. PRISMA flow diagram indicating the number of RCTs included and excluded at each review stage

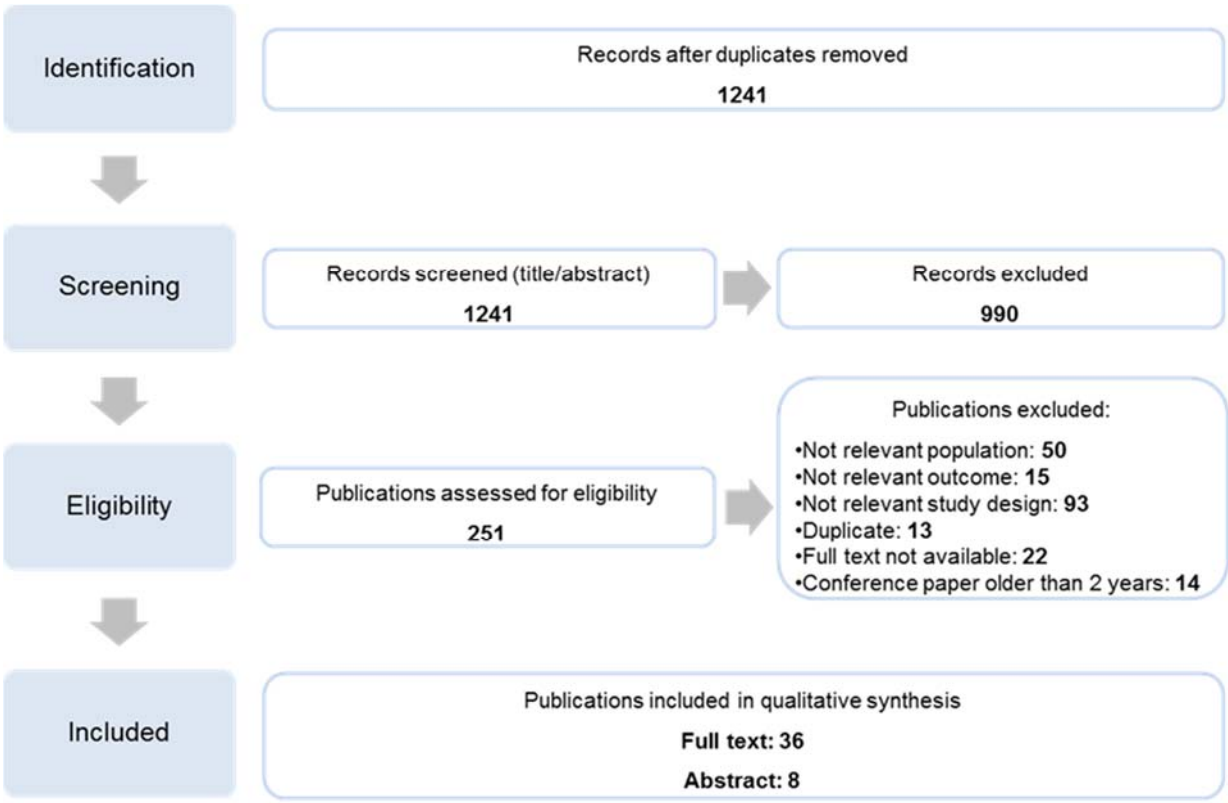
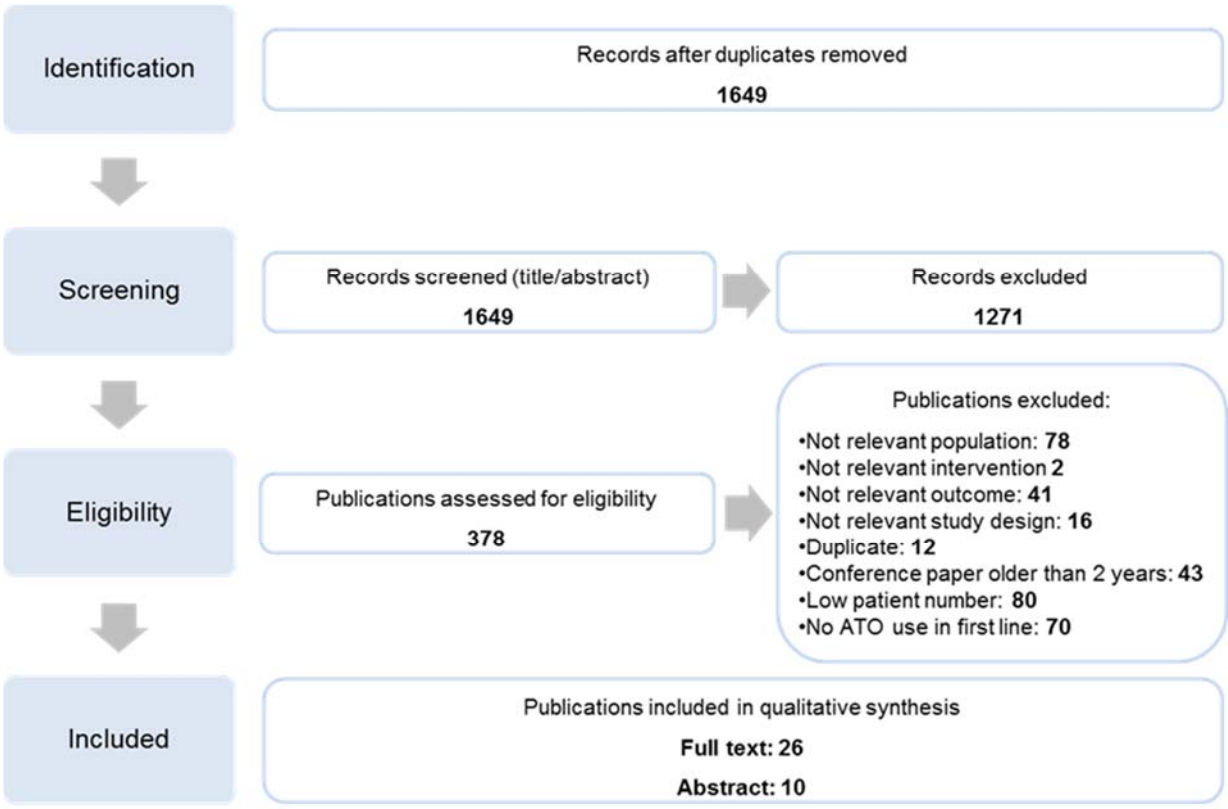


Figure 2.2. PRISMA flow diagram indicating the number of non-RCTs included and excluded at each review stage



B 2.3 List of relevant clinical effectiveness evidence

B 2.3.1 Newly-diagnosed APL

B 2.3.1.1 RCTs

Two studies (APL0406 and AML17) that provided evidence on the treatment of newly-diagnosed APL with ATRA+ATO are described below; both of these informed the economic model. In ***Appendix L*** we provided a brief overview of eight additional RCTs and one meta-analysis that were identified through the systematic literature review as using the ATRA+ATO combination for the treatment of newly-diagnosed APL. Originating in China, these studies may not represent UK or European treatment paradigms well; hence, they were not included in the economic model. However, they do provide additional information on ATRA+ATO use in newly-diagnosed APL. It is also worth noting that it was in China that the results of ATO-based APL treatment were first published[24].

B 2.3.1.1.a) APL0406

Table 2.3. Study overview - APL0406

Study	***APL0406***
Publications	• Lo-Coco, et al. 2013[1] – results from the initial cohort included in the study • Lo-Coco, et al. 2016[26] – updated results from the initial study cohort • Efficace, et al. 2014[40] – quality of life (QoL) results from the initial study cohort • Platzbecker, et al. 2014[41] – preliminary results from the extended and final study cohort (ASH abstract) • Platzbecker, et al. 2017[3] – results from the extended and final study cohort • The protocol of this study was published as a supplementary appendix to Lo Coco, et al. 2013 and is available online[42]
Study design	Prospective, randomised, multicentre, open-label, phase III non-inferiority trial
Population	Patients with newly-diagnosed, low- to intermediate-risk APL aged 18–71. Initial and final cohorts included, respectively, 156 and 266 patients with genetically confirmed APL who received at least one dose of assigned therapy. **The initial patient cohort formed part of the final cohort** (see ***section B 2.5.1.1.b***) for details of trial expansion).
Intervention(s)	ATRA+ATO (n=77 and n=129 in the initial and final cohorts, respectively): • Induction: o Oral ATRA (45 mg/m^2/day) + IV ATO (0.15 mg/kg/day) o Both continued until CR or up to 60 days • Consolidation: o Cycles 1–3: oral ATRA (45 mg/m^2/day for 15 days, two weeks on, two weeks off) + IV ATO (0.15 mg/kg/day 5 days per week, four weeks on, four weeks off) o Cycle 4: oral ATRA (45 mg/m^2/day for 15 days) + IV ATO (0.15 mg/kg/day 5 days per week for four weeks) • No maintenance phase with ATRA+ATO • See ***Figure 2.3*** for further details of the treatment regimen
Comparator(s)	AIDA (n=79 and n=137 in the initial and final cohorts, respectively): • Induction: o Oral ATRA (45 mg/m^2/day until CR or up to <60 days) + o IV idarubicin (12 mg/m^2/day for a total of 4 doses) • Consolidation: o 1st cycle: oral ATRA (45 mg/m^2/day for 15 days) + IV idarubicin (5 mg/m^2/day for a total of 4 doses) o 2nd cycle: oral ATRA (45 mg/m^2/day for 15 days) + IV mitoxantrone (10 mg/m^2/day for a total of 5 days) o 3rd cycle: oral ATRA (45 mg/m^2/day for 15 days) + IV idarubicin

Study	*APL0406*				
	(12 mg/m^2/day for 1 dose) • Maintenance: o Oral ATRA (45 mg/m^2/day for 15 days every 3 months for 2 years, for a total of 6 courses) alternating with o intramuscular or oral methotrexate MTX (15 mg/m^2/week) + oral 6-MP (50 mg/m^2/day) for a total of 7 courses • See ***Figure 2.3*** for further details of the treatment regimen				
Median follow-up	• 34.4 months in the initial cohort, with an updated analysis after a median of 53 months[26] • 40.6 months in the final cohort				
Indicate if trial supports application for marketing authorisation	Yes	*X*	**Indicate if trial used in the economic model**	Yes	*X*
	No			No	
Rationale for use/non-use in the model	*Pivotal trial supporting the approval of ATO in combination with ATRA for the treatment of newly-diagnosed low- to intermediate-risk APL*				
Reported outcomes specified in the decision problem	• Primary endpoint: o EFS at 2 years after diagnosis • Secondary endpoints: o Rate of haematological CR after induction o Rate of molecular CR after 3 consolidation cycles o Probability of OS o Cumulative incidence of relapse o Toxic effects o QoL[40]				
All other reported outcomes	• Kinetics of minimal residual disease				

6-MP=6-mercaptopurine; APL=Acute promyelocytic leukaemia; ASH=American Society of Hematology (annual meeting); ATO=Arsenic trioxide; ATRA=All-trans retinoic acid; CR=Complete remission; EFS=Event-free survival; IV=intravenous; MTX=Methotrexate; OS=Overall survival; QoL=Quality of life; WBC=White blood cell

B 2.3.1.1.b) AML17

Table 2.4. Study overview - AML17

Study	***AML17***
Publications	Burnett et al. 2015[2] Russell et al. 2016 (ASH abstract)[4]
Study design	Randomised, controlled, phase III open-label multicentre trial
Population	Patients with newly-diagnosed APL of any risk group, aged 16 or over (no upper age limit). A total of 235 patients with genetically confirmed APL were randomised.
Intervention(s)	ATRA+ATO (n=116) • Induction: o Oral ATRA (45 mg/m^2/day until CR or for up to 60 days) + IV ATO (0.3 mg/kg on days 1–5 and 0.25 mg/kg twice-weekly in weeks 2–8) o Gemtuzumab ozogamicin (6 mg/m^2 single IV infusion within days 1–4).

<table>
<tr><th>Study</th><th colspan="5">AML17</th></tr>
<tr><td></td><td colspan="5">Note that gemtuzumab ozogamicin (GO) was an optional treatment in high-risk patients randomised to ATRA+ATO. Of 30 high-risk patients in this group, 28 (93%) received GO, with the remaining two patients given an anthracycline instead. Additionally, seven low- to intermediate-risk patients in this study received GO to counteract rising WBC counts.
• Consolidation:
o Cycles 1–3: oral ATRA (45 mg/m^2/day for 15 days, two weeks on, two weeks off) + IV ATO (0.3 mg/kg on days 1–5 and 0.25 mg/kg twice-weekly in weeks 2–4)
o Cycle 4: oral ATRA (45 mg/m^2/day for 15 days) + IV ATO (0.3 mg/kg on days 1–5 and 0.25 mg/kg twice-weekly in weeks 2–4)
• No maintenance phase</td></tr>
<tr><td>Comparator(s)</td><td colspan="5">AIDA (n=119)
• Induction:
o Oral ATRA (45 mg/m^2/day until CR or up to 60 days)
o IV idarubicin (12 mg/m^2/day for a total of 4 doses)
• Consolidation:
o 1st cycle: oral ATRA (45 mg/m^2/day for 15 days) + IV idarubicin (5 mg/m^2/day for a total of 4 doses)
o 2nd cycle: oral ATRA (45 mg/m^2/day for 15 days) + IV mitoxantrone (10 mg/m^2/day for a total of 4 days)
o 3rd cycle: oral ATRA (45 mg/m^2/day for 15 days) + IV idarubicin (12 mg/m^2/day for 1 dose)
• No maintenance phase</td></tr>
<tr><td>Median follow-up</td><td colspan="5">30.5 months
53.4 months in the updated analysis[4]</td></tr>
<tr><td rowspan="2">Indicate if trial supports application for marketing authorisation</td><td>Yes</td><td>X</td><td rowspan="2">Indicate if trial used in the economic model</td><td>Yes</td><td>X (scenario)</td></tr>
<tr><td>No</td><td></td><td>No</td><td></td></tr>
<tr><td>Rationale for use/non-use in the model</td><td colspan="5">The licensed indication uses the dosing schedule from trial APL0406. While Teva does not wish to convey any support or encouragement for off-label use, for completeness NICE should be aware that certain KOLs/experts/medical community have indicated that they may wish to consider alternative dosage regimens.</td></tr>
<tr><td>Reported outcomes specified in the decision problem</td><td colspan="5">• Primary endpoint:
o QoL assessed using EORTC QLQ-C30 and HADS
• Secondary endpoints:
o OS
o RFS
o EFS
o Incidence of relapse (morphological and molecular)</td></tr>
<tr><td>All other reported outcomes</td><td colspan="5">• Incidence of death without relapse
• Incidence of tMDS-AML</td></tr>
</table>

APL=Acute promyelocytic leukaemia; ASH = American Society of Hematology (annual meeting); ATO=Arsenic trioxide; ATRA=All-trans retinoic acid; CR=Complete remission; EFS=Event-free survival; GO=Gemtuzumab ozogamicin; IV=intravenous; OS=Overall survival; QoL=Quality of life; RFS=Relapse-free survival; tMDS-AML=Therapy-related myelodysplastic syndrome or acute myeloid leukaemia; WBC=White blood cell

B 2.3.1.2 Non-RCTs

A total of 36 non-RCTs were identified through the systematic literature search; however, they were not selected as sources of information on clinical effectiveness for the economic model because of the potential bias associated with non-randomisation. Hence, these studies are not extensively described here; however, as they are generally supportive of the effectiveness of ATO, an overview of non-RCTs identified through the literature search is provided in ***Appendix L***.

B 2.3.2 Relapsed or refractory APL

B 2.3.2.1 RCTs

Only a single RCT (Raffoux, et al. 2003[43]) enrolling patients with relapsed/refractory APL provided clinical evidence that informed the economic model. This is presented below.

B 2.3.2.1.a) Raffoux et al. 2003

Table 2.5. Study design - Raffoux et al.

Study	**Raffoux, et al. 2003**
Publications	Raffoux, et al. 2003[43]
Study design	Randomised study
Population	Twenty patients (10 in each arm) with APL in first or subsequent relapse, aged ≥12 years, and with no contraindication to arsenic therapy. All patients were previously treated with ATRA and anthracycline-based chemotherapy.
Intervention(s)	ATRA+ATO (n=10) • Induction: o Oral ATRA (45 mg/m^2/day until CR) + IV ATO (0.15 mg/kg/day for up to 56 days) o Patients with a WBC count >30×10^9/L (either at baseline or during therapy), received chemotherapy consisting of 3 consecutive days of daunorubicin (60 mg/m^2/day) or amsacrine (90 mg/m^2/day). • Consolidation: o Allogeneic or autologous HSCT was generally offered to patients who achieved CR. Post-remission therapy was not initially specified in the protocol. Consolidation cycles of ATO were considered later on during the trial. Four patients received 2 consolidation cycles of ATRA+ATO (15 mg/kg/day) for 28 consecutive days, and an additional patient received a single consolidation cycle. Consolidation cycles were separated by 21 days.

Study	Raffoux, et al. 2003				
Comparator(s)	ATO (n=10) • Induction: o IV ATO (0.15 mg/kg/day for up to 56 days) o Patients with a WBC count >30×10^9/L (either at baseline or during therapy), received chemotherapy consisting of 3 consecutive days of daunorubicin (60 mg/m^2/day) or amsacrine (90 mg/m^2/day). • Consolidation: o As in the intervention arm, but with five patients receiving 2 consolidation cycles of ATO (15 mg/kg/day) for 28 consecutive days. • Maintenance: o One patient in this group received four 28-day cycles of ATO in addition to ATRA, MTX and 6-MP-based maintenance treatment.				
Median follow-up	21 months				
Indicate if trial supports application for marketing authorisation	Yes		**Indicate if trial used in the economic model**	Yes	*X*
	No	*X*		No	
Rationale for use/non-use in the model	*As one of few randomised trials evaluating ATRA+ATO in relapsed or refractory APL, this study provided valuable information on the efficacy of this combination for second-line APL treatment.*				
Reported outcomes specified in the decision problem	• Secondary endpoints: o Safety o Molecular response				
All other reported outcomes	• Primary endpoint: o Time necessary to reach CR				

6-MP=6-mercaptopurine; APL=Acute promyelocytic leukaemia; ATO=Arsenic trioxide; ATRA=All-trans retinoic acid; CR=Complete remission; HSTC=Hematopoietic stem cell transplantation; IV=Intravenous; MTX=Methotrexate; WBC=White blood cell

An additional RCT (Wang, et al. 2015[44]) was identified through the systematic literature review. However, as it was conducted in China its relevance to UK clinical practice was uncertain and it was not used in the economic model. This study is presented in ***Appendix L.***

B 2.3.2.2 Non-RCTs

The systematic search for non-RCTs focused on the first-line indication and none of the non-RCTs identified in the literature search was used as a source of inputs related to clinical effectiveness of ATO in second line. ***Appendix L*** provides an overview of two additional non-RCT studies[45, 46] that supported the regulatory approval of ATO for the treatment of relapsed/refractory APL.

B 2.4 Summary of methodology of the relevant clinical effectiveness evidence

B 2.4.1 Newly-diagnosed APL

Methodology of the two relevant RCTs enrolling patients with newly-diagnosed APL (APL0406 and AML17) is compared in ***Table 2.6***, and details of the treatment protocols are shown in ***Figure 2.3***. Major methodological differences between the two trials include different treatment schedules (see ***Figure 2.3***) and primary endpoints (EFS in APL0406 and QoL in AML17). Please note, that the manufacturer sponsored neither the APL0406 nor the AML17 trial, and as such only published information on study methodology and results is available. In terms of baseline patient characteristics, the most notable difference between the APL0406 and AML17 trials was the inclusion of high-risk patients in the latter. In contrast, the pivotal APL0406 trial only included low- to intermediate-risk patients. Another relevant difference relates to the age of the included patients – while the APL0406 trial enrolled patients aged 18–71, the AML17 trial was open to patients aged 16 or over, with no upper age limit. Baseline characteristics of patients included in both trials are presented in

Table 2.7.

Table 2.6. Methodology of the APL0406 and AML17 trials

Study	***APL0406***[1, 3, 40, 42]	***AML17***[2]
Study design	Prospective, randomised, multicentre, open-label, phase III non-inferiority trial. Randomisation was centralized and stratified according to institution.	Randomised, controlled, phase III, open-label multicentre trial for patients with AML (including APL) and high-risk MDS. Results of patients with APL were reported separately. Eligible participants were randomised between arms at a 1:1 ratio through web-based computer minimisation hosted by Cardiff University (Cardiff, UK). Minimisation parameters were age (0–15, 16–29, 30–39, 40–49, 50–59, or ≥60 years), WHO performance status (0, 1, 2, 3, or 4), and *de-novo* or secondary disease.
Population	The study included patients with newly-diagnosed, low- to intermediate-risk APL. Initial and final cohorts included, respectively, 156 and 266 patients with genetically confirmed APL who received at least one dose of assigned therapy. All patients provided written informed consent according to IGH/EU/GCP and national local laws. • Inclusion criteria: o Age 18–71 years o Newly-diagnosed APL o Low- to intermediate-risk APL (WBC count at diagnosis ≤10×10^9/L) o Genetic confirmation of diagnosis required after initial enrolment* o WHO performance status score ≤2 o Creatinine level ≤3.0 mg/dL (≤265 µmol/L) o Bilirubin level ≤3.0 mg/dL (≤51 µmol/L) • Exclusion criteria: o Age <18 and ≥71 o WBC count at diagnosis >10×10^9/L o Other active malignancy at time of study entry o Lack of diagnostic confirmation at genetic level o Significant arrhythmias, ECG abnormalities** or neuropathy o Cardiac contraindications for intensive chemotherapy (L-VEF <50%) o Uncontrolled, life-threatening infections o Severe uncontrolled pulmonary or cardiac disease o Pregnancy*** or breastfeeding o Concomitant severe psychiatric disorder	The study randomised 235 patients with genetically confirmed APL of any risk group. All patients provided signed informed consent. • Inclusion criteria: o Age ≥16 o Genetic confirmation of APL diagnosis by a reference laboratory o Previously untreated APL • Exclusion criteria: o Concurrent active malignancy o Substantial cardiac arrhythmia, ECG abnormalities or neuropathy o LVEF ≤50% o Uncontrolled life-threatening disease o Severe uncontrolled pulmonary or cardiac disease o Pregnancy or breastfeeding

Study	*APL0406*[1, 3, 40, 42]	*AML17*[2]
	o HIV positivity o Use of other investigational drugs at the time of enrolment or within 30 days before study entry	
Enrolment period	• October 2007 – September 2010 for patients included in the initial cohort (results in Lo-Coco, et al. 2013[1]) • October 2007 – January 2013 for the extended and final cohort (results in Platzbecker, et al. 2017[3])	• May 2009 – October 2013
Settings and locations where the data were collected	• 40 centres from Gruppo Italiano Malattie Ematologiche dell'Adulto (GIMEMA) • 27 centres from Study Alliance Leukemia and German–Austrian Acute Myeloid Leukemia Study Group (SAL–AMLSG) • Patients were enrolled in two countries – Germany and Italy.	• 81 hospitals in the UK, Denmark and New Zealand
Trial drugs	Intervention: ATRA+ATO (n=77 and n=129 in the initial and final cohorts, respectively). Comparator: AIDA (n=79 and n=137 in the initial and final cohorts, respectively) See ***Figure 2.3*** for details of the treatment protocol. Dose titrations were recommended in the study protocol in case of the following adverse events: • Differentiation syndrome: ATRA and/or ATO were temporarily discontinued. Upon improvement of symptoms and patient's clinical condition, ATRA and/or ATO was resumed at 50% of the previous dose for the first 7 days, with full dosage used thereafter if the previous toxicity did not worsen. In case of reappearance of symptoms, ATRA and ATO were reduced to the previous dosage. • QTc prolongation: ATO was discontinued together with any medication known to prolong the QTc interval and electrolytes were repleted. Once QTc normalised, ATO was resumed at a dose of 0.075 mg/kg (50% of the dose) for the first 7 days, and subsequently at 0.11 mg/kg for another 7 days. Thereafter, if QTc prolongation did not reoccur, ATO was resumed at full dose.	Intervention: ATRA+ATO (n=116) Comparator: AIDA (n=119) See ***Figure 2.3*** for details of the treatment protocol. No formal guidance on treatment or dose modification was part of the protocol. Although clinicians could discuss such modifications with clinical coordinators, no treatment modifications were actually needed.

<table>
<tr><th>Study</th><th>APL0406[1, 3, 40, 42]</th><th>AML17[2]</th></tr>
<tr><td></td><td>
• Hepatotoxicity: ATRA and/or ATO were temporarily discontinued, and were resumed when serum bilirubin, and/or AST, and/or alkaline phosphatase were reduced to <4 times the ULN, treatment with ATRA and/or ATO was resumed at 50% of the previous dose for 7 days. Thereafter, if the previous toxicity did not worsen, ATRA and/or ATO were resumed at full dosage. In case of reappearance of hepatotoxicity, treatment was to be definitely discontinued.
• Other non-haematological toxicities: ATO and/or ATRA were reduced by one dose level for grade 2 non-haematological toxicity. For grade 3-4 non haematological toxicity, therapy was held until resolution to grade <2, then restarted at a dose reduced by two levels, according to the following table of dose reduction.
<table>
<tr><td>Dose Level</td><td>ATO (mg/kg)</td><td>ATRA (mg/m^2)</td></tr>
<tr><td>Start level</td><td>0.15</td><td>45</td></tr>
<tr><td>-1</td><td>0.11</td><td>37.5</td></tr>
<tr><td>-2</td><td>0.10</td><td>25</td></tr>
<tr><td>-3</td><td>0.075</td><td>20</td></tr>
</table>
• Myelosuppression:
o In case of significant myelosuppression, defined as absolute neutrophil count <1×10^9/L and platelet count <50×10^9/L for >5 weeks after the start of a treatment course, one dose level reduction was recommended. If myelosuppression lasted for 50 days or more, or occurred on two consecutive courses, bone marrow aspirate was collected and specimens were sent for RT-PCR evaluation. If molecular remission was detected, treatment was resumed at a dose reduced by one level.
o For myelosuppression during maintenance treatment (AIDA arm only), the doses of MTX and 6-MP were reduced by 50% if WBC counts were between 2.5 and 3.5×10^9/L; MTX and 6-MP were temporarily discontinued if WBC count fell below 2.5×10^9/L.
</td><td></td></tr>
<tr><td>Concomitant medications</td><td>
• Supportive care during induction (both arms):
o Prednisone, 0.5 mg/kg/day from day 1 until the end of induction
</td><td>
• If QTc prolongation occurred, clinicians were advised to ensure that electrolyte levels, including that of magnesium, were corrected.
</td></tr>
</table>

Study	*APL0406*[1, 3, 40, 42]		*AML17*[2]	
	to prevent differentiation syndrome. o Platelet concentrate transfusions to maintain platelets >30×10^9/L during the first 10 days. o After day 10, platelet concentrates were transfused when platelet count was <20×10^9/L or in the presence of haemorrhagic symptoms. o Packed red cell concentrates to maintain haemoglobin levels >8 g/dL. o Prophylactic and therapeutic antibiotics and anti-fungal drugs according to institutional protocols. o Supplemental electrolytes administered intravenously to maintain electrolyte levels within the normal range. • Management of leucocytosis (both arms): o Hydroxyurea administered at 500 mg four times a day if WBC count was 10–50×10^9/L and at 1 g four times a day if WBC count exceeded 50×10^9/L. Hydroxyurea was discontinued when WBC count fell below 10×10^9/L.		• Guidance was provided for blood product and platelet support, and intervention for suspected differentiation syndrome, as set out in the British Committee for Standards in Haematology guidelines[47]. • No prophylaxis for differentiation syndrome was recommended, but dexamethasone was to be used promptly on clinical suspicion of the syndrome.	
Outcome measures and their definitions	• Primary endpoint: o EFS at 2 years after diagnosis, with treatment failure defined as any of the following: 1) no achievement of hematologic CR after induction; 2) no achievement of molecular CR after three consolidation courses; 3) molecular relapse; 4) haematological relapse, or 5) death. Following the protocol amendment to expand the trial (see ***section B 2.5.1.1.b***), EFS in the extended cohort (276 patients) was added as a secondary endpoint. • Secondary endpoints: o Rate of haematological CR after	The following outcomes were used in the model: • Haematological remission rate after induction • Proportion of patients evaluable with a PCR test after consolidation • Complete molecular remission rate after consolidation • Probability (cumulative incidence) of relapse at 24 and 50 months • Median time to relapse • Proportion of patients	• Endpoint definitions used revised International Working Group criteria[48]. • Primary endpoint: o Quality of life, assessed with the European Organisation for Research and Treatment of Cancer (EORTC) QLQ-C30 questionnaire and the Hospital Anxiety and Depression Scale (HADS). Patients completed the questionnaires at baseline and at 3, 6, 12, and 24 months after randomisation. • Secondary endpoints: o OS o RFS o EFS was defined as time from randomisation to death, treatment-	The following outcomes were used in the model: • Haematological remission rate after induction • Proportion of patients avaluable with PCR test after consolidation • Complete molecular remission rate after consolidation • Probability (cumulative

Study	*APL0406*[1, 3, 40, 42]		*AML17*[2]	
	induction. CR was defined according to the NCI workshop criteria[48]. o Rate of molecular remission after 3 consolidation cycles. Molecular remission was defined as described previously[49]. o Probability of OS: OS was defined as the time from entry onto the study to death from any cause, censoring patients alive at last follow-up. Defined according to the revised International Working Group criteria[48]. o Cumulative incidence of relapse: Cumulative incidence of relapse was defined as the time from achievement of haematological complete remission to relapse (molecular or haematological, whichever was detected first), persistence of PCR-positivity after the third consolidation cycle, or to date of last follow-up for patients alive in first molecular remission, using the cumulative incidence method and considering death in remission as competing risk. Defined according to the revised International Working Group criteria[48]. o Toxic effects, graded using NCI CTCAE, version 3. o QoL at the end of induction phase and at the end of the 3rd consolidation cycle, measured	experiencing adverse events by treatment phase: o Thrombocytopenia (grade 3–4, >15 days) o Neutropenia (grade 3–4, >15 days) o Infection o Leukocytosis o Hepatic toxicity o Neurotoxicity o Differentiation syndrome o Cardiac events o QTc prolongation o Myelodysplastic syndrome	related myelodysplastic syndrome or acute myeloid leukaemia, or morphological relapse for patients entering remission. Patients who did not achieve complete remission were defined as experiencing an event on day 1. o Incidence of relapse (both morphological and molecular). Cumulative incidence of molecular relapse was defined only for patients with confirmed molecular negativity as time to any relapse (haematological or molecular), with death or treatment-related myelodysplastic syndrome or acute myeloid leukaemia as competing risks. For cumulative incidence of haematological relapse, tMDS-AML and death were competing risks. o Death without relapse. o tMDS-AML. The cumulative incidence of tMDS-AML had competing risks of death or relapse. • Other relevant reported outcomes: o Toxicity, recorded using the NCI-CTCAE version 3.0	incidence) of relapse at 24 and 50 months • Proportion of patients experiencing adverse events by treatment phase: o Differentiation syndrome o QTc Prolongation

Study	*APL0406*[1, 3, 40, 42]		*AML17*[2]	
	using the EORTC QLQ-C30 questionnaire. • Other relevant reported outcomes: o DFS, defined as the time from achievement of haematological complete remission to relapse (either molecular or haematological), persistence of PCR positivity after consolidation therapy, or death, whichever occurred first. Data on patients who were still alive and in first molecular complete remission were censored at the last follow-up visit. Note that DFS, although presented in the study publications[1, 3], was not mentioned as an endpoint the APL0406 trial protocol[42], suggesting it may have been analysed post-hoc.			
Pre-planned subgroups	None		In addition to overall analyses, the study included pre-specified exploratory analyses stratified by: • Age • Sex • WBC count (including high-risk and low-risk patients) • Diagnosis • Performance status • Reverse transcript (RARA–PML) status • PML breakpoint	

6-MP=6-mercaptopurine; APL=Acute promyelocytic leukaemia; ATO=Arsenic trioxide; ATRA=All-trans retinoic acid; CR=Complete remission; DFS= Disease-free survival; ECG=electrocardiogram; EFS=Event-free survival; HIV=human immunodeficiency virus; L-VEF= left-ventricular ejection fraction; MTX=Methotrexate; NCI CTCAE= National Cancer Institute Common Terminology Criteria for Adverse Events; OS=Overall survival; PCR=Polymerase chain reaction; QoL=Quality of life; RFS=Recurrence-free survival; RT-PCR= Reverse transcription polymerase chain reaction; tMDS-AML=Therapy-related myelodysplastic syndrome or acute myeloid leukaemia; ULN= upper limit of normal; WBC=White blood cell; WHO=World Health Organisation

*Confirmation of diagnosis at genetic level was required for patient eligibility. However, to avoid delay in treatment initiation, patients were randomised on the basis of morphologic diagnosis only, before the results of genetic tests were available. APL diagnosis was genetically confirmed by one or more of the following methods: 1) detection of the PML–RARA fusion gene by RT-PCR, 2) demonstration of the t(15;17) translocation by conventional karyotyping or FISH, 3) evidence of a microspeckled PML pattern by indirect immunofluorescence assay

** Including: 1) congenital long QT syndrome, 2) history or presence of significant ventricular or atrial tachyarrhythmia, 3) clinically significant resting bradycardia (<50 beats per minute), 4)QTc >450 ms on screening EKG, 5) Right bundle branch block plus left anterior hemiblock, bifascicular block

*** Women who were either pregnant or breast feeding, or of child-bearing potential were excluded, defined as all women physiologically capable of becoming pregnant, unless they meet one of the following definitions: amenorrhea; post-surgical bilateral oophorectomy with or without hysterectomy; using a highly effective method of birth control (defined as those which result in a failure rate less than 1% per year) when used consistently and correctly, such as implants, injectables, oral contraceptives, IUDs, sexual abstinence or vasectomized partner.

Table 2.7. Key baseline patient characteristics in the APL0406 and AML17 trial[1-3]

Study population	APL0406 initial cohort		APL0406 final cohort		AML17	
Treatment arm	ATRA+ATO (n=77)	AIDA (n=79)	ATRA+ATO (n=129)	AIDA (n=137)	ATRA+ATO (n=116)	AIDA (n=119)
Age, years; median (range)	44.6 (19.1–70.2)	46.6 (18.7–70.2)	46.6 (18.8–70.2)	46.6 (18.0–70.3)	47 (16–75)	47 (16–77)
Male gender; n (%)	50 (52%)	36 (46%)	60.0 (46.5%)	70.0 (51.1%)	60 (52%)	60 (50%)
WBC count, $\times 10^9$/L; median (range)	1.49 (0.32–10.00)	1.60 (0.30–9.61)	1.4 (0.3–10.0)	1.5 (0.3–9.6)	3.0 (0.4–100.9)	2.2 (0.4–78.2)
Platelet count, $\times 10^9$/L; median (range)	31 (3–224)	27 (3–236)	36.5 (3–224)	31.5 (3–236)	Not reported	Not reported
Low risk, n (%)	33 (43%)	27 (34%)	57.0 (45.2%)	55.0 (41.3%)	86 (74%)	92 (77%)
Intermediate risk, n (%)	44 (57%)	52 (66%)	69 (54.7%)	78 (58.6%)	Not reported	Not reported
High risk, n (%)	Not applicable	Not applicable	Not applicable	Not applicable	30 (26%)	27 (23%)

ATRA=All-trans retinoic acid; ATO=Arsenic trioxide; WBC=White blood cell

Figure 2.3. Treatment regimens used in the APL0406 and AML17 trials[1, 2, 42]

6-MP=6-mercaptopurine; APL=Acute promyelocytic leukaemia; ATO=Arsenic trioxide; ATRA=All-trans retinoic acid; MTX=Methotrexate
*During induction in both the APL0406 and the AML17 trial, ATRA was administered for a maximum of 60 days or until CR.
**Gemtuzumab ozogamicin was an optional treatment in high-risk patients randomised to ATRA+ATO. Additionally, seven low- to intermediate-risk patients in this study received gemtuzumab ozogamicin during induction to counteract rising WBC.
Note that all drugs shown in this figure were administered intravenously, with the exception of ATRA (administered orally in both trials), 6-MP (administered orally, APL0406 only) and MTX (administered intramuscularly or orally, APL0406 only)

B 2.4.2 Relapsed or refractory APL

The methodology of the RCT published by Raffoux, et al. is presented in ***Table 2.8***; treatments administered to individual patients are listed in ***Table 2.10*** and patient characteristics in ***Table 2.9***. Please note, that this trial was also not sponsored by the manufacturer, so that only published information is available on the details of study methodology and results.

Table 2.8. Methodology of the study by Raffoux, et al.

Study	***Raffoux, et al. 2003[43]***
Study design	Prospective, randomised study.
Population	Twenty patients were included in total. • Eligibility criteria: o APL in first or subsequent relapse o Age ≥12 years o No visceral contraindication to arsenic therapy. o Previously treated with ATRA and anthracycline-based chemotherapy.
Enrolment period	• September 1998–January 2002. The study was terminated early as the first planned interim analysis (with a total of 20 patients included) showed no anticipated benefit of simultaneous ATRA+ATO administration.
Settings and locations where the data were collected	• Details are not reported. • Patients were referred onto the study from 17 hospitals in France.
Trial drugs	• Patients were randomised to receive induction treatment consisting of ATO (0.15 mg/kg/day for up to 56 days) with (intervention group) or without (comparator group) ATRA (45 mg/m^2/day until CR). During induction, ATO was administered for a maximum of 56 days, that is until: o CR achievement, o Severe toxicity (grade 2–4, depending on the organ concerned), o Serum arsenic concentration reaching ≥10^{-5} M. o After three patients had been included, the response-based stopping criteria were amended to stop ATO administration 7 days after bone marrow blast clearance. • Post-remission therapy was not initially specified in the protocol. Allogeneic or autologous HSCT was generally offered to patients who achieved CR. • Consolidation cycles of ATO were considered later on during the trial, following publication of relevant studies from the US[45, 46] (see ***Appendix L*** for their overview). Across both treatment groups, nine patients received two 28-day consolidation cycles of ATO (15 mg/kg/day) ± ATRA according to initial randomisation, and an additional patient received a single consolidation cycle. Consolidation cycles were separated by 21 days. • Details of the treatments that each patient received are provided in ***Table 2.10***.
Concomitant medications	• To prevent potential arsenic-related neurotoxicity, all patients received vitamin B1 (250 mg/day) and clobazam (10–30 mg/day) during treatment. • Patients with a WBC count >30×10^9/L (either at baseline or during therapy), received chemotherapy consisting of 3 consecutive days of

Study	*Raffoux, et al. 2003*[43]	
	daunorubicin (60 mg/m^2/day) or amsacrine (90 mg/m^2/day). • In patients with clinical symptoms of differentiation syndrome, dexamethasone was initiated at a dose of 10 mg every 12 hours for at least 3 days	
Outcome measures and their definitions	• The authors considered it difficult to demonstrate a significant improvement in outcome in the limited population of patients with relapsed APL, and chose a potential surrogate marker as the primary objective of the study, based on a significant reduction in the time needed to reach CR being associated with a prolonged survival in an animal study. • Thus, the primary objective of this study was a reduction by 2 weeks of the time needed to obtain haematological CR. o Haematological response was evaluated according to the National Cancer Institute criteria[50] on days 14 and 28, 42 and 56 and at the time of peripheral-blood CR criteria achievement. • Secondary objectives were safety and molecular response. The latter was evaluated using RT-PCR, as described previously[51]. • Other relevant reported outcomes: o OS calculated from the time of first ATO administration until death. Patients who were still alive were censored at the time of last contact. o DFS was calculated from the date of CR achievement until first relapse or death in CR. Patients alive in CR were censored at the time of last contact.	Outcomes used in the economic model: • Complete remission rate after induction with ATRA+ATO
Pre-planned subgroups	Not reported	

APL=Acute promyelocytic leukaemia; ATO=Arsenic trioxide; ATRA=All-trans retinoic acid; CR=Complete remission; HSTC=Hematopoietic stem cell transplantation; WBC=White blood cell

Table 2.9. Key patient characteristics in the study by Raffoux, et al.[43]

Patient No.	Age, years	Sex	Prior relapses	Prior AHSCT	WBC count, 10^9/L	Platelet Count, 10^9/L	Differentiation syndrome therapy	Response to induction
ATO group								
1101	37	M	2	No	1.1	39		CR
1102	47	F	2	No	5.2	47		CR
1105	40	M	2	Yes	1.8	79		CR
1106	64	M	1	No	2.2	163		CR
1109	55	F	3	No	19.5	18		ED
1111	51	M	2	Yes	8.2	44	DXM, Amsa	CR
1113	67	M	2	No	4.8	54	DXM, Amsa	ED
1114	46	M	1	No	2.5	128		CR
1119	59	F	2	No	7.7	64		CR
3120	25	F	1	No	11	66	DXM, Amsa	CR
ATRA+ATO group								
1103	37	M	1	No	0.7	46		CR
1104	25	M	2	Yes	4.6	12		RD
1107	32	F	1	No	26.6	69	DXM, Amsa	CR
1110	40	M	2	No	0.9	21	DXM	CR
1112	40	M	2	Yes	5.5	26	DXM, Amsa	CR
1115	30	F	1	No	1.3	149		CR
1116	50	F	2	No	20.8	63	DXM, DNR	CR
1117	52	M	2	No	1.8	62		RD
1118	46	M	1	No	2.5	105		CR
3108	69	F	2	No	2.2	100		CR

AHSCT=Autologous haematopoietic stem cell transplantation; Amsa=Amsacrine; ATRA=All-trans-retinoic acid; ATO=Arsenic trioxide; CR=Complete remission; DNR=Daunorubicin; DXM= Dexamethasone; ED=Early death; WBC=white blood cell

Table 2.10. Patient treatments in the study by Raffoux, et al. 2003[43]

Previous treatments						First CR Duration, months	Time from last ATRA exposure, months	Post-remission Treatments		
Patient No.	No. of Relapses	First Line*	Second Line**	Third Line	HSCT in Second CR			Consolidation	Maintenance	HSCT
ATO group										
1101	2	APL-93	ATRA-EMA		None		3	DA	NA (relapse)	None
1102	2	ATRA-I	EMA	ATRA***	None		1	LTF	LTF	LTF
1105	2	ATRA-DA	ATRA-A	ATRA-A†	Autologous		1	ATO (2 cycles)	None	Allogeneic
1106	1	APL-93			NA	43	16	ATO (2 cycles)	ATRA-MP	None
1109	3	APL-93	ATRA-IA	ATRA	None		16	NA	NA	NA
1111	2	APL-93	ATRA-EMA		Autologous		8	ATO (2 cycles)	ATO-ATRA-MP	None
1113	2	APL-93	ATRA-Amsa		None		1	NA	NA	NA
1114	1	APL-93			NA	36	29	ATO (2 cycles)	None	Allogeneic
1119	2	APL-93	ATRA-DA		None		8	ATO (2 cycles)	NA (relapse)	None
3120	1	ATRA-DA			NA	26	2	DA	None	None
ATRA+ATO group										
1103	1	APL-93			NA	16	11	None	None	Allogeneic
1104	2	APL-93	ATRA-EMA		Autologous		6	NA	NA	NA
1107	1	APL-93			NA	22	22	ATO/ATRA (2 cycles)	None	Allogeneic‡
1110	2	APL-93	ATRA-EMA		None		2	None	None	Allogeneic
1112	2	APL-93	ATRA-EMA		Autologous		9	ATO/ATRA (2 cycles)	None	Allogeneic§
1115	1	APL-93			NA	15	2	ATO/ATRA (2 cycles)	None	Autologous
1116	2	APL-93	ATRA-IA		None		5	ATO/ATRA (2 cycles)	NA (relapse)	None
1117	2	APL-93	ATRA-DA		None		5	NA	NA	NA
1118	1	APL-93			NA	19	3	ATO/ATRA (1 cycle)	None	Allogeneic
3108	2	ATRA-IA	ATRA		None		3	Amsa	MP	None

A=cytarabine; Amsa=amsacrine; APL=acute promyelocytic leukaemia; ATO=arsenic trioxide; ATRA=all-trans-retinoic acid; CR= complete remission; D=daunorubicin; HSCT=hematopoietic stem cell transplantation; I=idarubicin; LTF=lost to follow-up; NA=not applicable; MP=methotrexate + 6-mercaptopurine.
*For APL-93 trial, see Fenaux et al.[52]
**For ATRA-EMA trial, see Thomas et al.[53]
***ATRA was successfully reintroduced for a molecular relapse during the second haematological CR period.
†This patient was in refractory second relapse at enrolment.
‡Matched-unrelated-donor allogeneic HSCT.
§Non-myeloablative allogeneic HSCT.

B 2.5 Statistical analysis and definition of study groups in the relevant clinical effectiveness evidence

Statistical considerations related to the APL0406, AML17 and Raffoux, et al. 2003 studies are summarised in ***Table 2.11***.

B 2.5.1 Newly-diagnosed APL

B 2.5.1.1 APL0406[1, 3, 42]

All efficacy analyses in the APL0406 trial were based on the intention-to-treat (ITT) principle, comparing groups according to the randomly assigned treatment. Specifically, the ITT analysis set included all patients who received at least one dose of assigned therapy following randomisation (n=156 in the initial cohort, n=266 in the final cohort). A per-protocol non-inferiority analysis was also carried out for the primary efficacy endpoint (EFS at 2 years). The per-protocol analysis set included 229 patients with sufficient follow up (>24 months).

Designed as a non-inferiority trial, the APL0406 study aimed to prove equivalence between the two treatment arms, understood as the experimental (ATRA+ATO) arm being at most 5% inferior to the control (AIDA) arm in terms of the percentage of patients who were alive and failure-free at 2 years (EFS at 2 years). The goal was therefore to reject the null hypothesis that the between-arm difference favouring the control arm was equal to or higher than 5%, in favour of the alternative hypothesis that it was lower than 5%. Non-inferiority was assessed by estimating the two-sided 95% confidence interval for the between-group difference in crude rates of 2-year EFS and was confirmed if the lower bound was ≥-5%. The robustness of the results was confirmed with a sensitivity analysis that addressed all relevant scenarios for the patients who could not be evaluated, assuming poor outcome for all patients, favourable outcome for all patients, or poor outcome for patients in the ATRA+ATO group and favourable outcome for those in the AIDA group.

EFS was assessed by comparing Kaplan–Meier curves, taking into account time to treatment failure and loss to follow-up. Survival distributions (EFS, OS and DFS) were estimated with the use of the Kaplan–Meier product-limit estimator and compared between groups using log-rank test. Cumulative incidence of relapse was estimated using the proper non-parametric method, with Gray K-sample test used for between-group comparisons. Differences of percentages (response rates, toxicity) were evaluated using Fisher exact test. More broadly, non-parametric tests were used for comparisons between groups (Chi-Squared and Fisher Exact test for

categorical variables, Mann-Whitney and Kruskal-Wallis test for continuous variables). All tests were two-sided.

B 2.5.1.1.a) Sample size calculation[42]

Expected 2-year EFS was 85% in the AIDA arm, based on the AIDA-2000 trial[32], and 95% in the ATRA+ATO arm, based on an earlier non-randomised study by Ravandi et al., which used ATRA+ATO to treat newly-diagnosed APL patients[54]. According to the Farrington and Manning formulas[55], evaluating 73 patients per treatment arm would allow to determine that ATRA+ATO was not more than 5% inferior to AIDA, with a type I error probability of 5% and a power of 92%. The target sample was increased to 162 patients to allow for a 10% expected rate of loss.

B 2.5.1.1.b) Expansion of the study[42]

The APL0406 study reached its target accrual in September 2010, at which point randomisation and enrolment were closed. However, based on a preliminary analysis of available data, compliance with QoL assessment proved suboptimal. As the effects of arsenic-based treatment on patients' QoL were unknown at the time, performing an appropriate QoL analysis was considered important. Consequently, the protocol was amended to increase the target accrual to 276 patients (57 additional patients per arm) to reach optimal QoL compliance. The results of the initial (Lo-Coco, et al. 2013[1] and final (Platzbecker, et al. 2017[3]) cohorts were published separately.

B 2.5.1.1.c) QoL assessment[40]

Compliance was computed for each time point as the percentage of valid questionnaires returned out of those expected from patients still on study at that time. Differences in patient characteristics were assessed between patients who completed the questionnaire after induction and those who did not, using Fisher's exact or Wilcoxon Mann-Whitney test, as appropriate (α=0.05). Primary analysis involved estimating mean scores and differences in these mean scores between treatment arms for all EORTC QLQ-C30 scales over time. A repeated measures linear mixed model with an unrestricted covariance structure was used. The model included treatment, assessment time, and treatment–time interaction. For each EORTC QLQ-C30 scale, the null hypothesis tested by an overall F statistic was that the estimated difference between the two treatment arms was equal to 0 at all time points. If this test was significant (α=0.05), the estimated differences between treatment arms were tested separately for each single time point using a t-test

(α=0.05). The means and standard deviations (SDs) estimated from the model, differences in means between treatment arms, and the corresponding 95% CIs were reported for all health-related quality of life (HRQoL) assessments. P values were not adjusted for multiple testing due to the exploratory nature of the study. Clinical relevance of the estimated mean score differences between treatment arms was interpreted according to the evidence-based guidelines for EORTC QLQ-C30[56].

Missing data were investigated to check their possible impact on results. Mean observed HRQoL scores after induction were compared between patients who returned the questionnaire after consolidation therapy and those who did not, in order to investigate relationships between dropout and outcome. The impact of key sociodemographic, biological, and clinical characteristics of patients on the probability of not completing the HRQoL questionnaire at any time point was investigated using logistic regression analysis. Robustness of the final results was assessed using the explicit regression model approach, which identified a linear regression model predicting a value to be imputed for each missing HRQoL scale. The linear regression model included variables related to both the missing data mechanism and the HRQOL values to be imputed. The linear mixed model described above was subsequently applied to the augmented data set.

B 2.5.1.2 AML17[2]

All analyses were performed on the ITT population with confirmed APL. No follow-up data for survival or relapse were available for two patients in the AIDA group (one high-risk and one low-risk). QoL was assessed using a multilevel models repeated measures analysis. The questionnaires used were the EORTC QLQ-C30 questionnaire, the Leukaemia Specific Module, and the Hospital Anxiety and Depression Scale (HADS). Effect sizes were calculated in such a way that a positive value represented a benefit for ATRA+ATO. Categorical secondary endpoints were compared using Mantel-Haenszel tests to yield Peto odds ratios and 95% CIs. Continuous or scale variables were analysed using Wilcoxon rank-sum tests, and time-to-event outcomes were analysed using log-rank tests and Kaplan-Meier or cumulative incidence curves. Appropriate tests for interaction were performed when conducting exploratory subgroup analyses. In the results displayed in ***section 0***, odds ratios (ORs) or HRs <1 indicate a benefit for ATRA+ATO over AIDA. All p values were two-sided.

B 2.5.1.2.a) Sample size calculation

With 300 patients enrolled, the trial would have had more than 80% power to detect a difference of a third of an SD on the primary outcome (QoL), equating to 6–7 points (out of 100) on the global health scale of the EORTC QLQ-C30 questionnaire based on the data from the earlier Medical Research Council AML15 trial. However, the AML17 trial closed randomisation after recruiting 235 eligible patients as no further drug supply was available. As a result, the trial had 72% power to detect a difference of a third of an SD, or 80% power to detect a difference of three-eighths of an SD (7.5 points on the global health scale of the EORTC QLQ-C30 questionnaire).

B 2.5.2 Relapsed or refractory APL

B 2.5.2.1 Raffoux, et al. 2003[43]

The primary objective of this study was achieving a 2-week reduction in the time needed to obtain haematological CR. Binary variables were compared using Fisher's exact test and continuous variables were compared using Mann-Whitney test. Data on CR achievement and outcome were estimated using the Kaplan-Meier method, with log-rank test used for comparisons. Details of sample size calculation were not reported.

Table 2.11. Summary of statistical analyses in the included studies

Trial	Hypothesis objective	Statistical analysis	Sample size, power calculation	Data management, patient withdrawals
APL0406	• Null hypothesis: the between-arm difference in 2-year EFS in favour of the control arm was ≥5%. • Alternative hypothesis: the between-arm difference in 2-year EFS in favour of the control arm was <5%.	• EFS was assessed by comparing Kaplan–Meier curves. Non-inferiority was assessed by estimating the two-sided 95% confidence interval for the between-group difference in crude rates of 2-year EFS and was confirmed if the lower CI bound was ≥-5%.	• See ***section B 2.5.1.1.a)*** for details of sample size calculation.	• Of 156 patients included in the ITT population, six (4%) could not be evaluated in the primary analysis, due to insufficient follow-up or molecular testing not being performed after the third consolidation cycle. Thus, 150 patients were included in the ITT population for the primary analysis.
AML17	• Primary objectives: To compare QoL, toxicity and resource usage of patients receiving AIDA with those receiving chemo-free treatment with ATRA+ATO. • Secondary objectives: o To compare CR, OS and relapse rates in the two arms o To compare the kinetics of MRD in the two arms. o To correlate plasma arsenic levels with disease response and treatment-related toxicities in APL patients allocated to receive ATO	• QoL was assessed using a multilevel models repeated measures analysis	• See ***section B 2.5.1.2.a)*** for details of sample size calculation.	• A total of 671 completed QoL forms were received during the AML17 study: o 156 at baseline (76 in the AIDA group and 80 in the ATRA+ATO group) o 137 at 3 months (64 in the AIDA group and 73 in the ATRA+ATO group) o 139 at 6 months (70 in the AIDA group and 69 in the ATRA+ATO group) o 136 at 12 months (64 in the AIDA group and 72 in the ATRA+ATO group) o 103 at 24 months (49 in the AIDA group and 54 in the ATRA+ATO group)
Raffoux, et al. 2003	• The primary objective of the study was achieving a 2-week reduction in the time needed to obtain haematological CR	• Mann-Whitney test	• Not reported	• The primary objective was evaluated in 16 patients who achieved CR. • Out of the 20 patients included in the study, 4 were not evaluated: o Two patients (both from the ATO group) died early during the induction cycle – one patient, with a prior history of CNS haemorrhage, died of septic shock with seizures at day 14. The other patient died at day 16 from differentiation

Trial	Hypothesis objective	Statistical analysis	Sample size, power calculation	Data management, patient withdrawals
				syndrome with hyperleukocytosis which did not respond to treatment. • Two patients from the ATRA+ATO group were alive with resistant disease after the induction cycle.

APL=Acute promyelocytic leukaemia; ATO=Arsenic trioxide; ATRA=All-trans retinoic acid; CNS=Central nervous system; CR=Complete remission; ITT=Intention-to-treat; EFS=Event-free survival; MRD=Minimal residual disease; OS=Overall survival; QoL=Quality of life

B 2.6 Quality assessment of the relevant clinical effectiveness evidence

B 2.6.1 APL0406 trial

Of the five publications reporting the results of the APL0406 trial (***see Table 2.3***), the two that reported results from the initial and the final study cohorts (Lo Coco, et al. 2013[1] and Platzbecker, et al. 2017[3], respectively) were assessed for quality.

Both publications reported results from a prospective, randomised, multicentre, open-label, phase 3 non-inferiority trial. The publication on the initial cohort[1] reported that written consent was obtained before study entry from all patients and that the trial was conducted in accordance with the Declaration of Helsinki. In this trial, 67 centres took part by enrolling at least one patient, and the institutional review board of each participating centre reviewed and approved the study. Participants were randomly assigned to receive ATRA+ATO or AIDA and the assignment was stratified by institution. No significant differences in baseline characteristic were observed between two treatment groups. The authors did not include information on randomisation method. In this initial cohort, efficacy analysis in the ITT population was performed, but this population was incomplete (i.e., did not include all patients who received at least one dose of the assigned therapy after randomisation) as 4% of patients could not be evaluated. In addition to the ITT analysis, a per-protocol analysis for the primary efficacy endpoint of EFS non-inferiority was carried out, including patients who received protocol treatment as scheduled.

The publication by Platzbecker, et al. 2017 reported results of the extended and final study cohort of the APL0406 trial[3]. According to the protocol amendment (see ***section B 2.5.1.1.b)***, a total of 276 patients were recruited, which included the initial patient cohort. This modification was approved by the ethics committees of all participating centres, and all enrolled patients provided written informed consent. As for the initial cohort, random assignment was stratified by institution. Patients were randomised in blocks of four. No further information about randomisation was provided. Almost all of the patients (98.9%) were analysed following the ITT principle.

B 2.6.2 AML17 trial

Results of the AML17 trial concerning patients with APL were published in at least two publications: Burnett et al. 2015 (full publication)[2] and Russel et al. 2016 (abstract)[4]. Both of them were assessed for quality.

The AML17 was a randomised, controlled, open-label multicentre trial. Participants were randomised between two treatment arms (ATRA+ATO or AIDA) at a 1:1 ratio. Treatment allocation was conducted by web-based computer minimisation hosted by Cardiff University (Cardiff, UK). Minimisation parameters were: age, WHO performance status and de novo versus secondary disease. No significant differences in baseline characteristic were observed between two treatment groups. All analyses were done according to the ITT principle.

B 2.6.3 Raffoux et al. 2003

The study was approved by the Ethics Committee of Hôpital Pitie-Salpetriere (Paris, France), and all patients gave signed informed consent. Participants were randomly assigned into one of two treatment groups: ATO alone or ATRA+ATO. No further information about randomisation method, masking or concealment was provided. There was no significant difference in baseline characteristics between treatment groups. No information was available on whether an ITT or per-protocol analysis was conducted.

Quality assessment of all three studies is summarised in ***Table 2.12***.

Table 2.12. Quality assessment of trials included in the clinical effectiveness section

Trial number (acronym)	APL0406		AML17 (Burnett et al. 2015 and Russel et al. 2016)	Raffoux et al. 2003
	Platzbecker et al. 2017	Lo Coco et al. 2013		
Was randomisation carried out appropriately?	Not clear	Not clear	Yes	Not clear
Was the concealment of treatment allocation adequate?	Not clear	Not clear	N/A	N/A
Were the groups similar at the outset of the study in terms of prognostic factors?	Yes	Yes	Yes	Yes
Were the care providers, participants and outcome assessors blind to treatment allocation?	No	No	No	N/A
Were there any unexpected imbalances in drop-outs between groups?	No	No	No	No
Is there any evidence to suggest that the authors measured more outcomes than they reported?	Not clear	Not clear	Not clear	Not clear
Did the analysis include an ITT analysis? If so, was this appropriate and were appropriate methods used to account for missing data?	Not clear*	Not clear*	Yes	Not performed

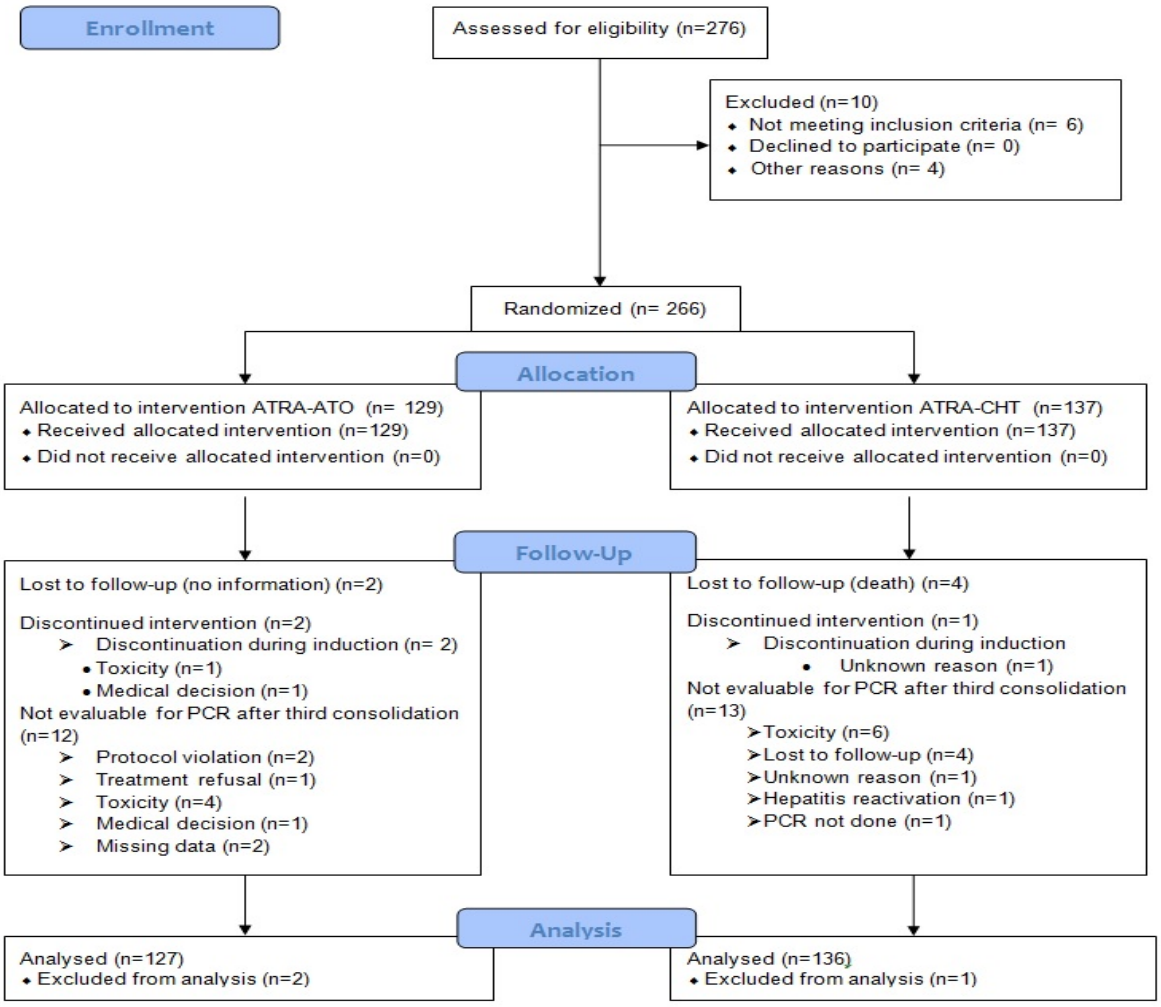
Adapted from Systematic reviews: CRD's guidance for undertaking reviews in health care[57]
*The ITT population was described as including all patients who received at least one dose of assigned therapy after randomisation, i.e. 266 and 156 patients in the final[3] and initial[1] cohorts, respectively. However, the ITT analysis for the primary endpoint actually included 263 and 150 patients, respectively, and the available information is insufficient to conclude if this analysis was appropriate and if appropriate methods were used to account for missing data.

B 2.7 Clinical effectiveness results of the relevant trials

B 2.7.1 Newly-diagnosed APL

B 2.7.1.1 Participant flow in the APL0406 trial[3]

Figure 2.4. Participant flow in the final cohort of the APL0406 trial. Note that the initial cohort is included.



ATO=Arsenic trioxide; ATRA=All-trans retinoic acid; CHT=chemotherapy; PRC=Polymerase chain reaction

B 2.7.1.2 Key efficacy results from the APL0406 trial

A summary of key efficacy results from the initial and final patient cohorts of the APL0406 trial is presented in ***Table 2.13163***. Note, that the final cohort included the initial cohort, so that in subsequent sections, only final cohort results are presented, with the exception of EFS at 2-years in the initial cohort which was the primary endpoint of the APL0406 study. The remaining results of the initial APL0406 cohort are presented in ***Appendix M***.

Table 2.13. APL0406 trial: Key clinical efficacy results

Endpoint	Initial cohort[1, 26]				Final cohort[3]			
	Time frame	ATRA+ATO (n=77)	AIDA (n=79)	p value	Time frame	ATRA+ATO (n=129)	AIDA (n=137)	p value
EFS; % (95% CI)	2 years	97%	86%	<0.001 for non-inferiority; 0.02 for superiority	2 years	98.3 (95.9–100)	86.8 (81.1–92.8)	<0.001
	50 months (4.2 years)[26]	96% (92–100)	81% (73–91)	0.003	50 months	97.3 (94.3–100)	80.0 (72.9–88.0)	
OS; % (95% CI)	2 years	99% (96–100)	91% (85–97)	0.020	2 years	99.2 (97.7–100)	94.8 (91.1–98.6)	0.007
	50 months (4.2 years)[26]	99% (96–100)	88% (81–96)	0.006	50 months	99.2 (97.7–100)	92.6 (87.9–97.5)	
DFS, % (95% CI)	2 years	97% (94–100)	90% (84–97)	0.110	2 years	98.3 (95.9–100)	89.4 (84.1–95.0)	<0.001
					50 months	97.3 (94.3–100)	82.6 (75.6–90.3)	
Haematological CR rate; %	after induction	100%	95%	0.120	After induction	100.0	97.0	0.120
Molecular CR rate; %	after the 3rd consolidation cycle (n=145)	75 (100%)	70 (100%)	Not reported	after the 3rd consolidation cycle (n=234)	115 (100%)	117 (98.3%)	Not reported
CIR, % (95% CI)	2 years	1% (0–4)	6% (0–11)	0.240	2 years	0.9 (0–2.7)	8.2 (3.3–13.2)	0.0013
					50 months	1.9 (0.0–4.5)	13.9 (7.1–20.6)	

ATO=Arsenic trioxide; ATRA=All-trans retinoic acid; CIR=Cumulative incidence of relapse; CR=Complete remission; DFS=Disease-free survival; EFS= Event-free survival; OS=Overall survival

B 2.7.1.2.a) Response rates

Response to induction therapy could be evaluated in 263 of 266 patients in the final cohort (127 of 129 and 136 of 137 patients in the ATRA+ATO and AIDA groups, respectively)[3]. The assigned treatment was terminated early in two patients who started induction in the ATRA+ATO arm – in one patient this was due to a major protocol violation, and in the other due to toxicity[3]. The protocol violation involved an early evaluation of bone marrow aspirate, resulting in an inappropriate diagnosis of resistant disease; this patient was alive at 2 years from diagnosis[3]. The other patient had his assigned treatment permanently withdrawn by the treating physician as a result of toxicity (QTc prolongation and electrolyte abnormalities); the patient was lost to follow-up[3]. One patient in the AIDA group discontinued induction treatment for an unknown reason; this patient was lost to follow-up[3]. Following induction therapy, the rates of haematological CR among patient evaluable for induction response were similar in the ATRA+ATO and AIDA groups (127 [100%] vs 132 [97%], respectively, p=0.12)[3]. The 4 patients in the AIDA arm who did not achieve CR died during induction from differentiation syndrome (n=2), ischaemic stroke (n=1) and bronchopneumonia (n=1); all of these deaths were recorded in the initial cohort[1, 3].

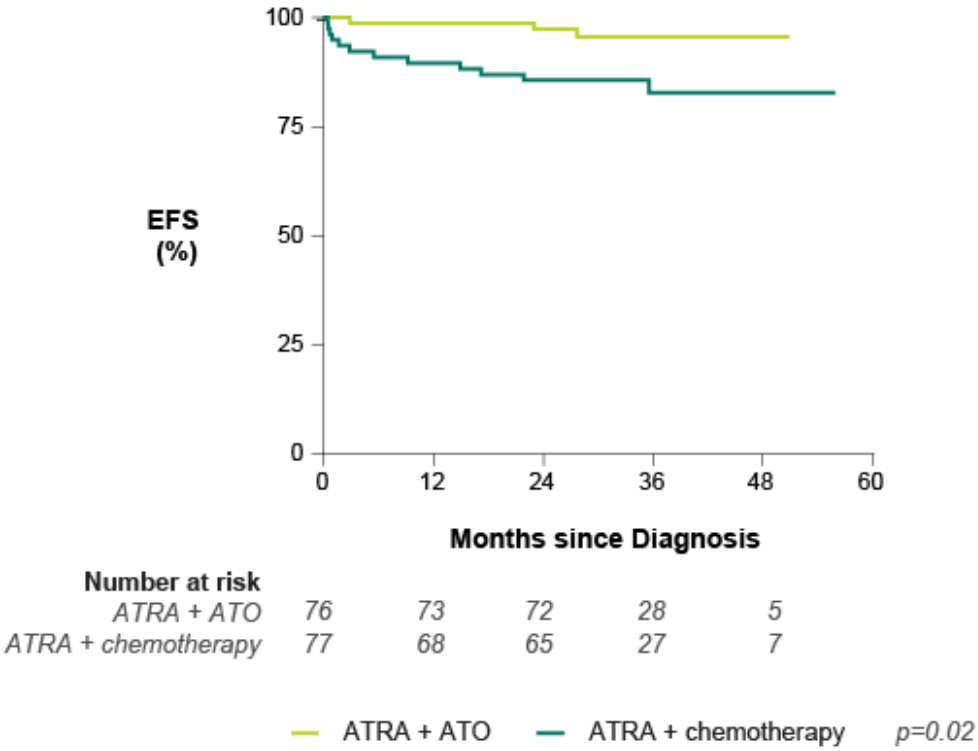
Molecular response was evaluated after the third consolidation cycle and a total of 234 patients (115 in the ATRA+ATO arm and 119 in the AIDA arm) were evaluable[3]. See ***Figure 2.4*** for patient flow, including detailed reasons for non-evaluation of molecular response. All 115 patients (100%) in the ATRA+ATO arm and 117 of 119 patients (98.3%) in the AIDA arm showed PCR negativity after the third consolidation course[3]. The remaining 2 patients in the AIDA group tested PCR-positive and were considered to have molecularly resistant disease[3]. In one patient, a second PCR performed at the end of consolidation did not confirm PCR positivity, and the patient proceeded to receive maintenance therapy subsequently remaining in remission for >24 months[3]. The other patient received ATO-based salvage therapy followed by allogeneic HSCT off-protocol and remained in second CR 14 months after transplantation[3].

B 2.7.1.2.b) Event-free survival

EFS at 2 years in the initial patient cohort was the primary endpoint of the APL0406 trial. Of 156 patients included in the ITT population, 6 (4%) could not be assessed for EFS at 24 months, due to insufficient follow-up or lack of PCR evaluation after the third consolidation cycle[1]. Of the remaining 150 patients, 72 of 74 (97%) in the ATRA+ATO group and 65 of 76 patients (86%) in the AIDA group were alive and free

of events at 24 months[1]. Thus, the between-group difference was 11% (95% CI: 2–22), and, since the lower bound of the 95% CI for the difference in EFS rates was no lower than -5%, non-inferiority of ATRA+ATO to AIDA was confirmed (p<0.001)[1]. Analysis of the EFS curves (***Figure 2.5***) with log-rank test indicated the superiority of ATRA+ATO over AIDA (p=0.02)[1]. In the per-protocol population of 138 patients, 64 of 66 patients (97%) in the ATRA+ATO group and 61 of 72 patients (85%) in the AIDA group were alive and event-free at 24 months (difference of 12%, 95% CI: 2–23, p<0.001 for non-inferiority)[1].

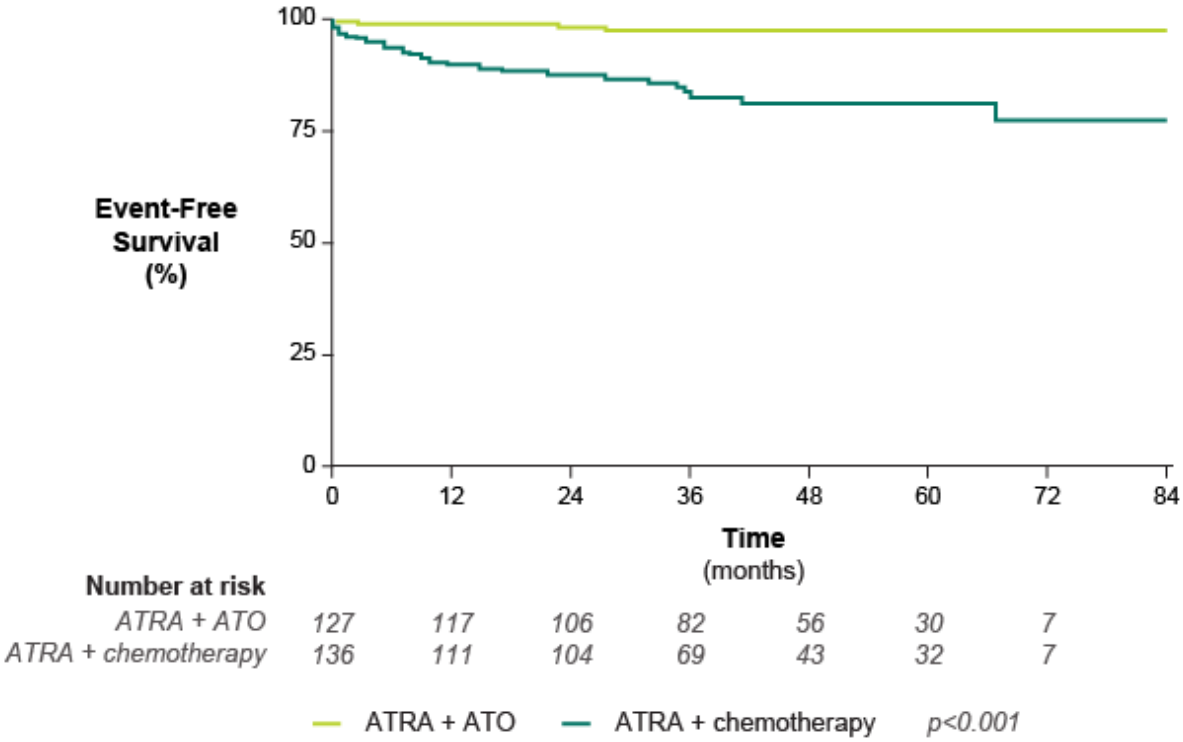
Figure 2.5. EFS in the initial cohort of the APL0406 trial at a median follow-up of 34.4 months[1]



Upon protocol amendment to expand the trial (see ***section B 2.5.1.1.b***), EFS in the final patient cohort was added to the protocol as a secondary endpoint[42]. In the final cohort, 263 of 266 patients in the ITT population (127 of 129 in the ATRA+ATO arm and 136 of 137 in the ATRA arm) were evaluable for EFS[3]. In the ATRA+ATO group, significantly more patients (98.3% [95% CI: 95.9%–100%]) were event-free at 24 months from randomisation compared with the AIDA group (86.8% [95% CI: 81.1%–92.8%]; p<0.001)[3]. At 50 months, EFS estimates were 97.3% (95% CI: 94.3%–100%) in the ATRA+ATO group compared with 80.0% (95% CI, 72.9% to 88.0%) in the AIDA group[3]. The Kaplan-Meier plot for EFS is shown in ***Figure 2.6***. Non-inferiority analysis was carried out in 229 patients with sufficient follow-up (>24

months), of whom 98.15% in the ATRA+ATO group (106 of 108 patients) were alive and event-free at 24 months, compared with 85.95% in the AIDA group (104 of 121 patients)[3]. The difference in EFS rates at 2 years was 12.2% (95% CI: 4.3%–20.1%) and, with the lower bound of the 95% CI for this difference exceeding -5%, the non-inferiority of ATRA+ATO to AIDA was confirmed (p<0.001)[3].

Figure 2.6. EFS in the final cohort of the APL0406 trial at a median follow-up of 40.6 months[3]

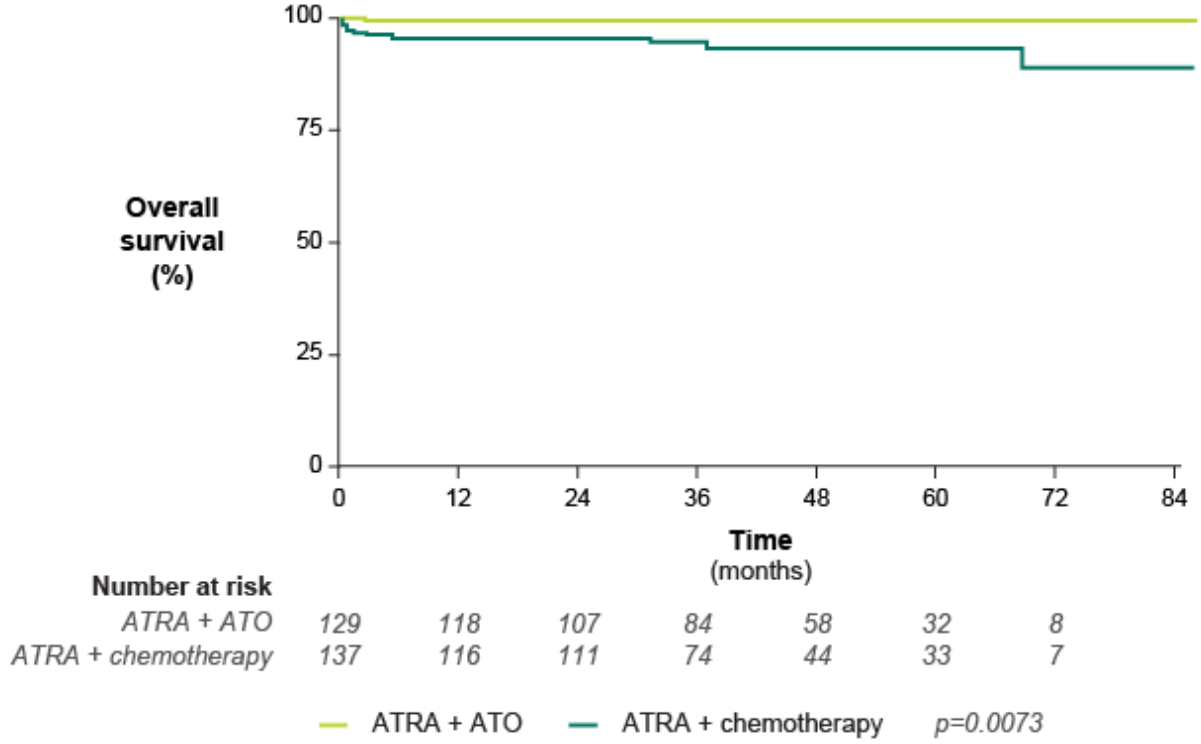


B 2.7.1.2.c) Overall survival

ATRA+ATO provided an OS benefit vs AIDA in the extended and final APL0406 cohort of 266 patients (***Figure 2.7***). OS probability at 24 months was 99.2% (95% CI: 97.7%–100%) in patients treated with ATRA+ATO, compared with 94.8% (95% CI: 91.1%–98.6%) in patients treated with AIDA (p=0.0073)[3]. As no further deaths occurred in the ATRA+ATO arm with longer follow-up, the Kaplan-Meyer OS estimate at 50 months was the same as at 24 months (99.2% [95% CI: 97.7%–100%]), while in the AIDA group OS probability at this time point was 92.6% (95% CI: 87.9%–97.5%)[3], so that the difference in OS between the two treatment arms increased over time. In total, 9 patients in the AIDA group died, with 4 induction deaths (see section ***B 2.7.1.2.a***) and 5 deaths in CR, from haemorrhagic shock (n=1), pulmonary embolism (n=1), bronchopneumonia (n=2) and secondary MDS

(n=1)[3]. Apart from a single patient who died in CR of bronchopneumonia no other deaths were noted in the ATRA+ATO arm[3].

Figure 2.7. OS in the final APL0406 cohort at a median follow-up of 40.6 months[3]



B 2.7.1.2.d) Disease-free survival and cumulative incidence of relapse

Of 259 patients in the final APL0406 cohort who achieved CR, 17 subsequently relapsed[3]. Two of the relapses occurred among 127 CR patients in the ATRA+ATO arm (at 22 and 27 months) and 15 occurred among 132 CR patients in the AIDA arm at a median of 14.0 months (range: 2.5–39.8 months)[3]. Four relapses were detected at the molecular level before progression to a hematologic relapse, and these patients were administered pre-emptive salvage therapy[3]. One patient in the ATRA+ATO arm died in CR of bronchopneumonia, and five deaths in CR occurred in the AIDA arm (see ***section* Error! Reference source not found.**[3]. In this final patient cohort, both DFS and cumulative incidence of relapse were significantly better with ATRA+ATO than with AIDA. DFS at 2 years was 98.3% (95% CI: 95.91%–100%) in the ATRA+ATO group and 89.4% (95% CI: 84.1%–95.0%) in the AIDA group (p<0.001); at 50-months DFS was estimated at 97.3% (95% CI: 94.3%–100%) and 82.6% (95% CI: 75.6%–90.3%), respectively, in the two groups[3] (see ***Figure 2.8***).

Figure 2.8. DFS in the final APL0406 cohort at a median follow-up of 40.6 months[3]

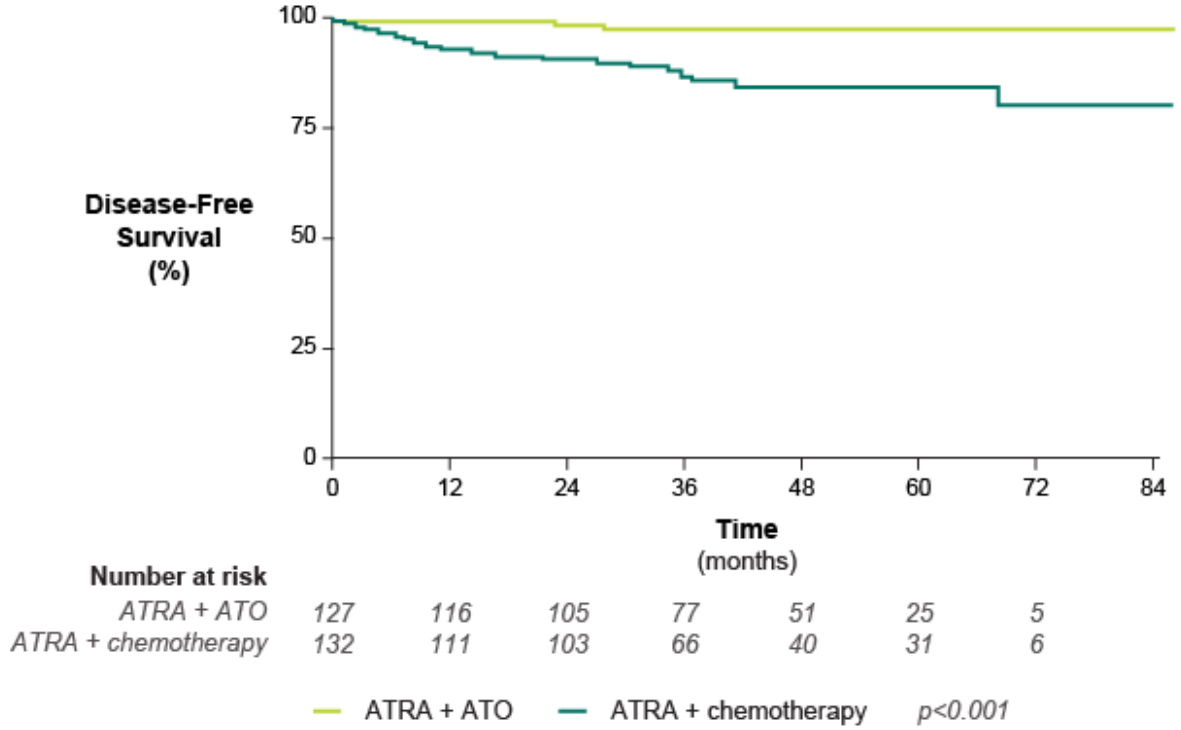
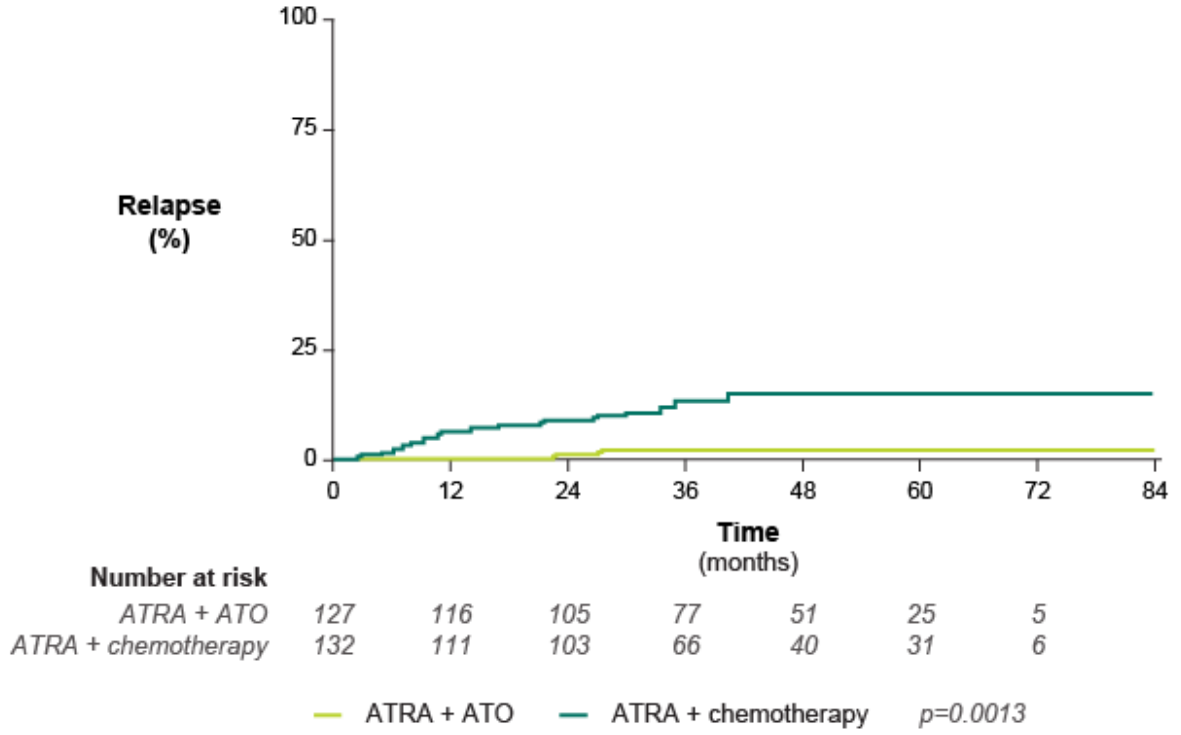


Figure 2.9. Cumulative incidence of relapse in the final APL0406 cohort at a median follow-up of 40.6 months[3]



Cumulative incidence of relapse (***Figure 2.9***) at 24 months was 0.9% (95% CI: 0–2.7%) in the ATRA+ATO group, compared with 8.2% (95% CI: 3.3%–13.2%) in the

AIDA group (p=0.0013)[3]. The Kaplan-Meier curves diverged further over time, with 50-month cumulative incidence of relapse estimated at 1.9% (95% CI: 0–4.5%) in the ATRA+ATO group and 13.9% (95% CI: 7.1%–20.6%) in the AIDA group[3].

B 2.7.1.2.e) Quality of life

QoL results from the APL0406 trial are available only for the initial patient cohort (156 patients comprising the ITT population)[40]. HRQoL in the APL0406 trial was assessed at end of induction and following the third consolidation course; however, no pre-treatment baseline assessment was performed. HRQoL was analysed using the EORTC QLQ-C30 (version 3) – a validated and widely used (>3,000 studies worldwide) questionnaire developed to assess the QoL of cancer patients[58]. Several validated disease-specific modules of this questionnaire exist and additional ones are in development[58].

EORTC QLQ-C30 has been shown to be reliable and responsive to change; it is brief to fill in (mean time of 11 minutes) and consists of 30 questions grouped into scales (***Table 2.14***)[59].

Table 2.14. Scales of the EORTC QLQ-C30 questionnaire[59]

Functional scales (16 questions):	Symptom scales (6 questions):	Single items (6 questions):
• Physical • Role • Cognitive • Emotional • Social	• Fatigue • Pain • Nausea/vomiting • Appetite	• Constipation • Diarrhoea • Sleep • Dyspnoea • Financial impact

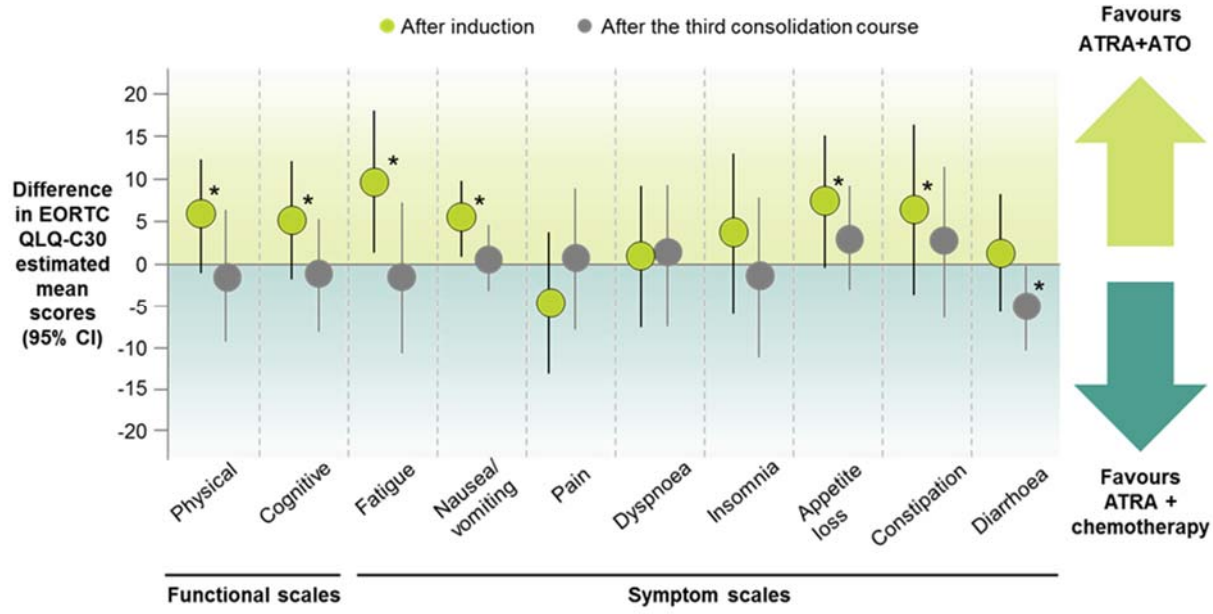
Note that financial impact was not investigated in the APL0406 study[40].

Of 150 patients eligible for HRQoL assessment at the end of induction therapy, 115 returned HRQoL forms (77% compliance), with a slightly higher compliance after the third consolidation cycle (84%, 119 forms returned from 142 eligible patients)[40]. Compliance rates did not differ significantly between the two treatment arms[40]. Among all the EORTC QLQ-C30 scales, a significant overall difference between treatment arms was only detected for fatigue (p=0.022)[40]. Comparison of scores at individual time points revealed that ATRA+ATO was associated with significantly lower fatigue severity after induction (p=0.034) but not after the third consolidation course (p=0.66)[40]. Other small but clinically relevant benefits were observed after induction in the severity of nausea/vomiting, constipation and appetite loss, as well as physical and cognitive functioning; all favoured ATRA+ATO over AIDA[40]. After the

third consolidation course, only the severity of diarrhoea showed a clinically relevant difference (albeit small), favouring AIDA over ATRA+ATO[40]. No meaningful between-group differences in the estimated mean scores were detected for other symptoms and functional scales of the EORTC QLQ-C30 questionnaire[40]. QoL results of the APL0406 study are summarised in **Figure 2.10** and **Table 2.15**.

A long-term QoL analysis in the final APL0406 patient cohort remains to be reported[3]. However, the publication by Platzbecker, et al. reported that QoL results in the extended cohort broadly confirm the findings from the initial cohort described above[3]. Notably, fatigue severity was significantly lower in the ATRA+ATO group than the AIDA group in this larger patient population (p =0.008)[3].

Figure 2.10. Between-group differences in mean EORTC QLQ-C30 scores observed in the APL0406 trial[40]



Asterisks mark small clinically relevant differences, according to evidence-based guidelines for EORTC QLQ-C30[56]. Of these, only the difference in fatigue severity was statistically significant.
ATRA=all-trans retinoic acid; CI=confidence interval; EORTC= European Organisation for Research and Treatment of Cancer; QLQ-C30=Quality of Life Questionnaire – Core 30.

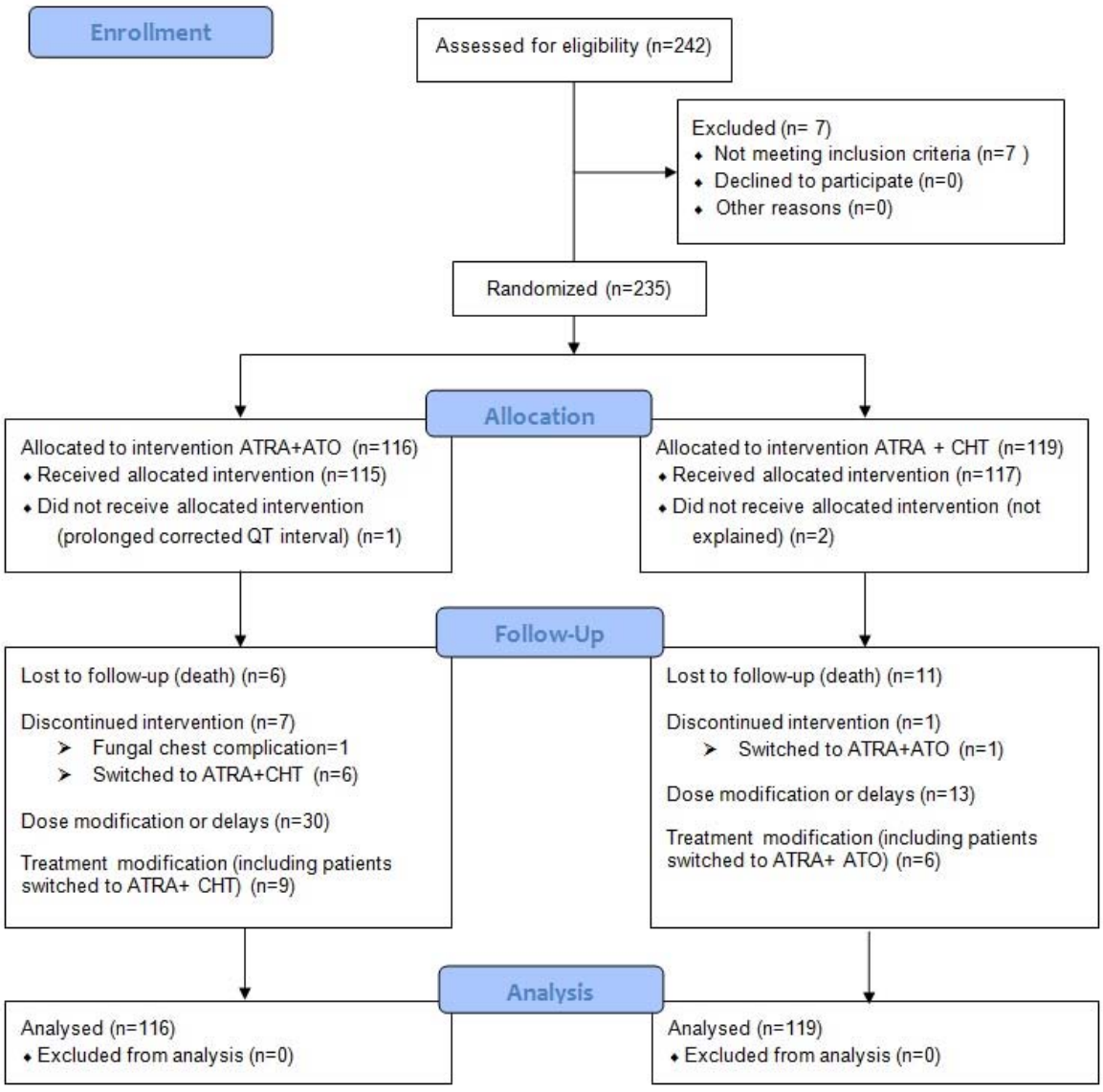
Table 2.15. APL0406 QoL results[40]

Scale	Estimated EORTC QLQ-C30 Mean Score (±SD)		Mean score difference between treatment arms, Δ (95% CI)	p value (F-test)	Clinical relevance
	ATRA+ATO	AIDA			
After induction					
Physical functioning	80.9 ± 21.5	75.6 ± 23.5	Δ=5.3 (-1.9 to 12.4)	0.147	Small, favours ATRA+ATO
Role functioning	68.0 ± 34.8	62.5 ± 38.1	Not reported	0.280	Trivial
Emotional functioning	81.2 ± 21.0	76.8 ± 23.0	Not reported	0.088	Trivial
Cognitive functioning	87.2 ± 21.1	81.4 ± 23.1	Δ=5.9 (-1.2 to 12.9)	0.089	Small, favours ATRA+ATO
Social functioning	68.8 ± 30.2	72.4 ± 33.1	Not reported	0.778	Trivial
Global health status/ QoL	67.2 ± 21.8	64.7 ± 24.1	Not reported	0.761	Trivial
Fatigue	29.1 ± 25.7	38.4 ± 28.1	Δ=-9.3 (-17.8 to -0.7)	0.022	Small, favours ATRA+ATO
Nausea/ vomiting	3.1 ± 13.9	8.3 ± 15.2	Δ=-5.1 (-9.7 to -0.5)	0.095	Small, favours ATRA+ATO
Pain	16.1 ± 25.3	11.1 ± 27.8	Not reported	0.467	Trivial
Dyspnoea	15.8 ± 24.8	16.3 ± 27.1	Not reported	0.978	Trivial
Insomnia	16.2 ± 28.7	19.4 ± 31.4	Not reported	0.614	Trivial
Appetite loss	5.6 ± 22.6	12.7 ± 24.8	Δ=-7.1 (-14.6 to 0.5)	0.183	Small, favours ATRA+ATO
Constipation	14.5 ± 30.5	20.6 ± 33.4	Δ=-6.1 (-16.3 to 4.0)	0.489	Small, favours ATRA+ATO
Diarrhoea	8.6 ± 20.6	9.5 ± 22.5	Not reported	0.106	Trivial
After the third consolidation course					
Physical functioning	80.6 ± 20.6	81.9 ± 21.5	Not reported	0.147	Trivial
Role functioning	72.4 ± 29.7	75.8 ± 31.0	Not reported	0.280	Trivial
Emotional functioning	74.7 ± 24.0	79.6 ± 25.1	Not reported	0.088	Trivial
Cognitive functioning	80.8 ± 24.5	82.4 ± 25.8	Not reported	0.089	Trivial
Social functioning	77.5 ± 26.5	78.0 ± 27.6	Not reported	0.778	Trivial
Global health status/ QoL	72.7 ± 22.4	72.9 ± 23.4	Not reported	0.761	Trivial
Fatigue	29.8 ± 26.4	27.9 ± 27.4	Not reported	0.022	Trivial
Nausea/ vomiting	5.3 ± 13.1	5.7 ± 13.5	Not reported	0.095	Trivial
Pain	16.9 ± 26.3	17.2 ± 27.4	Not reported	0.467	Trivial
Dyspnoea	16.7 ± 26.6	17.5 ± 27.6	Not reported	0.978	Trivial
Insomnia	21.2 ± 30.0	19.4 ± 31.2	Not reported	0.614	Trivial
Appetite loss	5.3 ± 19.9	8.1 ± 21.0	Not reported	0.183	Trivial
Constipation	5.3 ± 27.6	7.7 ± 28.6	Not reported	0.489	Trivial
Diarrhoea	7.8 ± 15.9	2.3 ± 16.3	Δ=5.5 (-0.4 to 10.6)	0.106	Small, favours AIDA

*Clinical relevance was determined using the 2011 evidence-based guidelines for EORTC QLQ-C30[56], which provide mean differences in scores for each EORTC QLQ-C30 scale that correspond to trivial, small, medium and large difference between the test groups. Large difference was defined as representing unequivocal clinical relevance, medium as likely to be clinically relevant (albeit to a lesser extent), and small as subtle but clinically relevant. Trivial was used to describe no difference, or one unlikely to have any clinical relevance.

B 2.7.1.3 Participant flow in the AML17 trial[2]

Figure 2.11. Participant flow in the AML17 trial



ATO=Arsenic trioxide; ATRA=All-trans retinoic acid; CHT=chemotherapy

B 2.7.1.4 Key efficacy results from the AML17 trial

A summary of key efficacy results from the AML17 trial is presented in ***Table 2.16.***

Table 2.16. AML17 trial: Key clinical efficacy results[2]

Endpoint	Time frame	ATRA+ATO (n=116)	AIDA (n=119*)	HR or OR (95% CI)	p value
Haematological CR; n (%)	Not reported	109 (94%)	106 (89%)	OR 0.54 (0.21–1.34)	0.180
Molecular CR; n (%)	Not reported	106 (91%)	105 (88%)	OR 0.71 (0.31–1.65)	0.430
OS; % (95% CI)	4 years	93% (86–96)	89% (81–93)	HR 0.60 (0.26–1.42)	0.250
Early mortality; % (95% CI)	30 days	4% (2–10)	6% (3–12)	HR 0.72 (0.23–2.31)	0.560
	60 days	5% (2–11)	9% (5–16)	HR 0.55 (0.21–1.43)	0.220
EFS; % (95% CI)	4 years	91% (84–95)	70% (56–80)	HR 0.35 (0.18–0.68)	0.002
Haematological RFS; % (95% CI)	4 years	97% (90–99)	78% (63–88)	HR 0.24 (0.09–0.60)	0.004
Molecular RFS; % (95% CI)	4 years	98% (91–99)	70% (62–83)	HR 0.17 (0.08–0.39)	<0.001
Cumulative incidence of death in remission; % (95% CI)	4 years	2% (1–9)	1% (0.2–8)	HR 1.72 (0.18–16.6)	0.640
Cumulative incidence of haematological relapse; % (95% CI)	4 years	1% (0.1–7)	18% (10–34)	HR 0.16 (0.06–0.46)	<0.001
Cumulative incidence of molecular relapse; % (95% CI)	4 years	0%	27% (18–45)	HR 0.12 (0.05–0.30)	<0.001
Cumulative incidence of tMDS-AML	4 years	0%	3% (0.4–17)	HR 0.15 (0.003–7.48)	0.340

ATRA=all-trans retinoic acid; CI=confidence interval; CR=complete remission; EFS=event-free survival; HR=hazard ratio; NR=not reported; OR=odds ratio; OS=overall survival; RFS=recurrence-free survival; tMDS-AML=treatment-related acute myeloid leukaemia or myelodysplastic syndrome
*No data was available for survival or relapse for two patients in the ATRA + chemotherapy group (one low-risk and one high-risk)

B 2.7.1.4.a) Response rates

In total, 215 of 235 (91%) patients included in the trial achieved haematological CR after induction[2]. Similarly to the APL0406 trial[1, 3], CR rates after induction in the AML17 trial were comparable between the two treatment groups (94% for ATRA+ATO vs. 89% for ATRA + chemotherapy, odds ratio [OR]=0.54 [95% CI: 0.21–1.34]; p=0.18)[2]. Several patients died during induction therapy, with 5 deaths in the ATRA+ATO group and 7 deaths in the AIDA group occurring in the first 30 days[2]. By day 60 (corresponding to the longest permitted duration of induction therapy), 6

patients had died in the ATRA+ATO group compared with 11 in the AIDA group[2]. The causes of death by day 60 days were:[2]

- ATRA+ATO group: 3 cardiac events, 1 infection, 1 renal failure, and 1 death from several causes.
- AIDA group: 3 haemorrhages, 3 infections, 2 pulmonary causes, 1 renal cause, and 2 cases of progressive disease.

Neither 30-day (4% [95% CI: 2%–10%] in the ATRA+ATO group vs 6% [95% CI: 3%–12%] in the AIDA group, hazard ratio [HR]=0.72 [95% CI: 0.23–2.31]; p=0.56) nor 60-day (5% [95% CI: 2%–11%] in the ATRA+ATO group vs 9% [95% CI: 5%–16%] in the AIDA group, HR=0.55 [95% CI: 0.21–1.43]; p=0.22) mortality differed significantly between treatment arms[2]. Of the 215 patients in haematological CR, molecular remission was achieved in 211 patients – 106 (91%) in the ATRA+ATO group and 105 (88%) in the AIDA group (OR=0.71 [95% CI: 0.31–1.65]; p=0.43)[2]. Across both treatment arms, 4 patients had molecularly resistant disease (3 in the ATRA+ATO group and 1 in the AIDA group)[2].

B 2.7.1.4.b) Event-free survival

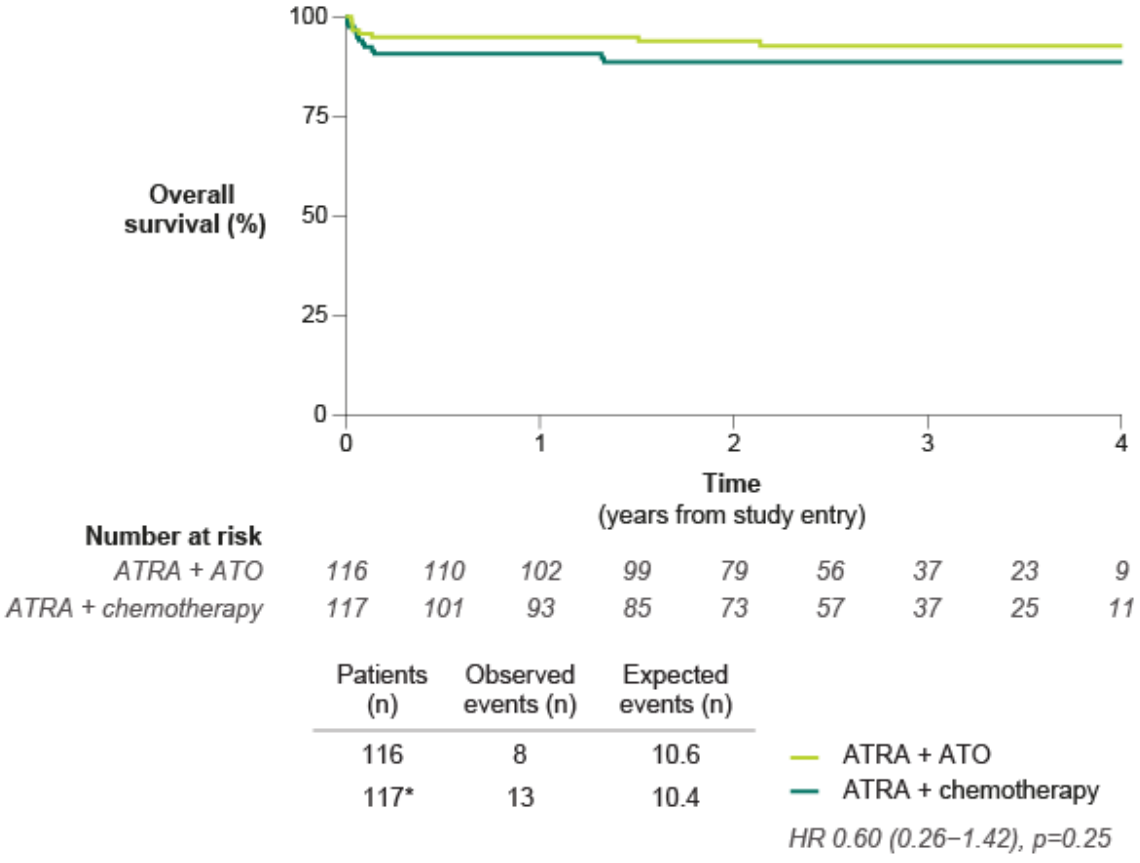
Unlike in the APL0406 trial[1, 3], where the EFS definition was close to that of PFS, development of tMDS-AML was one of the events considered in the AML17 trial[2]. One patient in the AIDA group developed tAML (trisomy 11) at 39.5 months after APL diagnosis[2]. In addition, a total of 23 patients across both treatment groups experienced either disease persistence or recurrence[2]. Three patients in the ATRA+ATO group had molecularly persistent disease (with CNS involvement in one case) compared with one patient in the AIDA group[2]. The number of relapses was much higher in the AIDA group (n=19) than in the ATRA+ATO group (n=1)[2]. Four-year EFS was significantly better among patients treated with ATRA+ATO than among those treated with AIDA (91% [95% CI: 84%–95%] vs 70% [95% CI: 56%–80%], HR=0.35 [95% CI: 0.18–0.68]; p=0.002)[2].

B 2.7.1.4.c) Overall survival

By day 60, 6 patients in the ATRA+ATO and 11 in the AIDA group had died (see ***section 0*** for causes of death)[2]. Two further patients (one in each arm) died following a relapse[2]. One, assigned to AIDA, died before salvage treatment could be given, while the other patient (the only patient in the ATRA+ATO group who relapsed) died in second CR from an infection 78 days after receiving a sibling allograft[2]. The cumulative incidence of death in remission was low and did not differ

significantly between treatment groups (2% [95% CI: 1%–9%] in the ATRA+ATO group vs 1% [95% CI: 0.2%–8%] in the AIDA group, HR=1.72 [95% CI: 0.18–16.6]; p=0.64)[2]. No follow-up data for survival or relapse were available for 2 patients in the AIDA group (1 high-risk and 1 low-risk)[2]. Four-year OS (***Figure 2.12***) was not significantly different between the two treatment arms: 93% (95% CI: 86%–96%) in the ATRA+ATO arm compared with 89% (95% CI: 81%–93%) in the AIDA arm (HR=0.60 [95% CI: 0.26–1.42]; p=0.25)[2]. In the updated analysis of the trial, 5-year survival was 93% with ATRA+ATO and 87% with AIDA (HR=0.61 [95% CI: 0.27–1.35]; p=0.2) after a median follow-up of 53.4 months[4].This lack of a significant difference in OS between treatment arms is likely due to the effect of ATO-based salvage treatment on survival in the AIDA arm. See ***section B 2.8*** for details on OS in selected patient subgroups and ***section B 2.14*** for the discussion of the apparent discrepancy between survival outcomes in the AML17 and APL0406 trials.

Figure 2.12. OS in the full population of the AML17 trial[2]



HR=Hazard ratio
*Note that no follow-up data available for survival or relapse for available for 2 patients in the AIDA group (one high risk and one low risk)

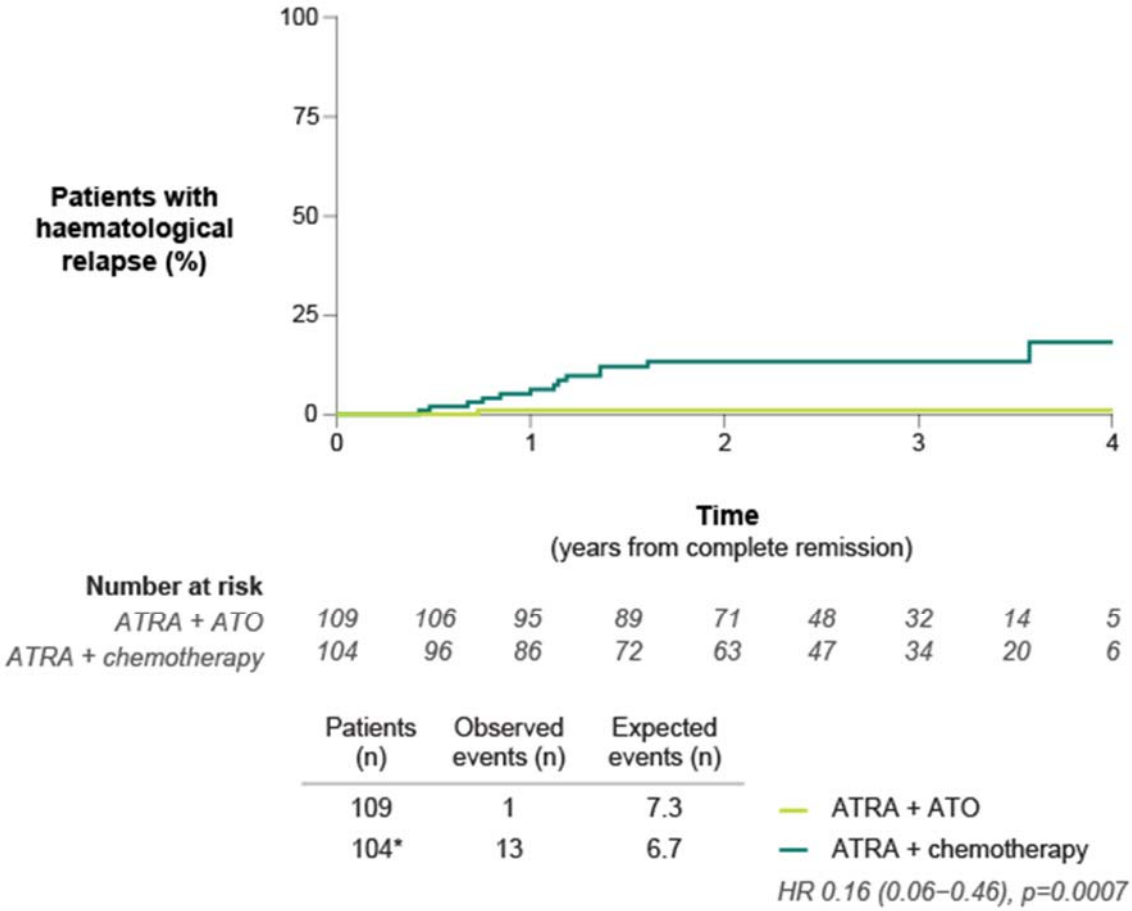
B 2.7.1.4.d) Recurrence-free survival and cumulative incidence of relapse

Considering both treatment arms together, 23 patients had persistent or recurrent disease[2]. Four patients had molecularly persistent disease (1 in the AIDA group and 3 in the ATRA+ATO group), with one of these patients subsequently progressing into a CNS relapse[2]. Six relapses were detected at the molecular level and haematological relapse occurred in 8 patients, with 2 further patients experiencing both haematological and CNS relapse[2]. Three patients had isolated extramedullary relapses[2]. Extramedullary involvement of the skin (1 patient in the AIDA group) or CNS (1 patient in the ATRA+ATO group and 3 in the AIDA group) was always accompanied by detection of bone marrow PCR positivity[2].

Cumulative incidence of death in remission was low and similar across treatment groups (2% [95% CI: 1%–9%] in the ATRA+ATO group vs 1% [95% CI: 0.2%–8%] in the AIDA group, HR=1.72 [95% CI: 0.18–16.6]; p=0.64)[2], so that differences in RFS between treatment arms were primarily driven by differences in relapse rates. In the AML 17 trial, the cumulative incidences of haematological (morphological) and molecular relapse were quantified separately, with the latter quantified only in patients who achieved molecular remission[2]. Of the 20 relapses occurring in the 215 patients who reached CR across both treatment groups 19 were detected in the AIDA group and 1 in the ATRA+ATO group, so that 4-year cumulative incidence of haematological relapse (***Figure 2.13***) was 18% (95% CI: 10%–34%) in the AIDA arm and 1% (95% CI: 0.1%–7%) in the ATRA+ATO arm (HR=0.16 [95% CI: 0.06–0.46]; p=0.0007)[2]. Four-year haematological RFS was significantly higher in the ATRA+ATO group at 97% (95% CI: 90%–99%) compared with 78% (95% CI: 63%–88%) in the AIDA group (HR=0.24 [95% CI: 0.09–0.60]; p=0.004)[2].

The only patient in the ATRA+ATO group who relapsed did not become molecularly negative; thus, the cumulative incidence of molecular relapse (***Figure 2.14***) at 4 years was 27% (95% CI: 18%–45%) in the AIDA group and 0% in the ATRA+ATO group (HR 0.12 [95% CI: 0.05–0.30], p<0.0001)[2]. Four-year molecular RFS, also quantified in patients who achieved molecular remission, was significantly better in the ATRA+ATO group (98% [95% CI: 91%–99%]) than in the AIDA group (70% [95% CI: 62%–83%], HR=0.17 (95% CI: 0.08–0.39); p<0.0001)[2]. No follow-up data for survival or relapse were available for two patients in the AIDA group (one high-risk and one low-risk)[2].

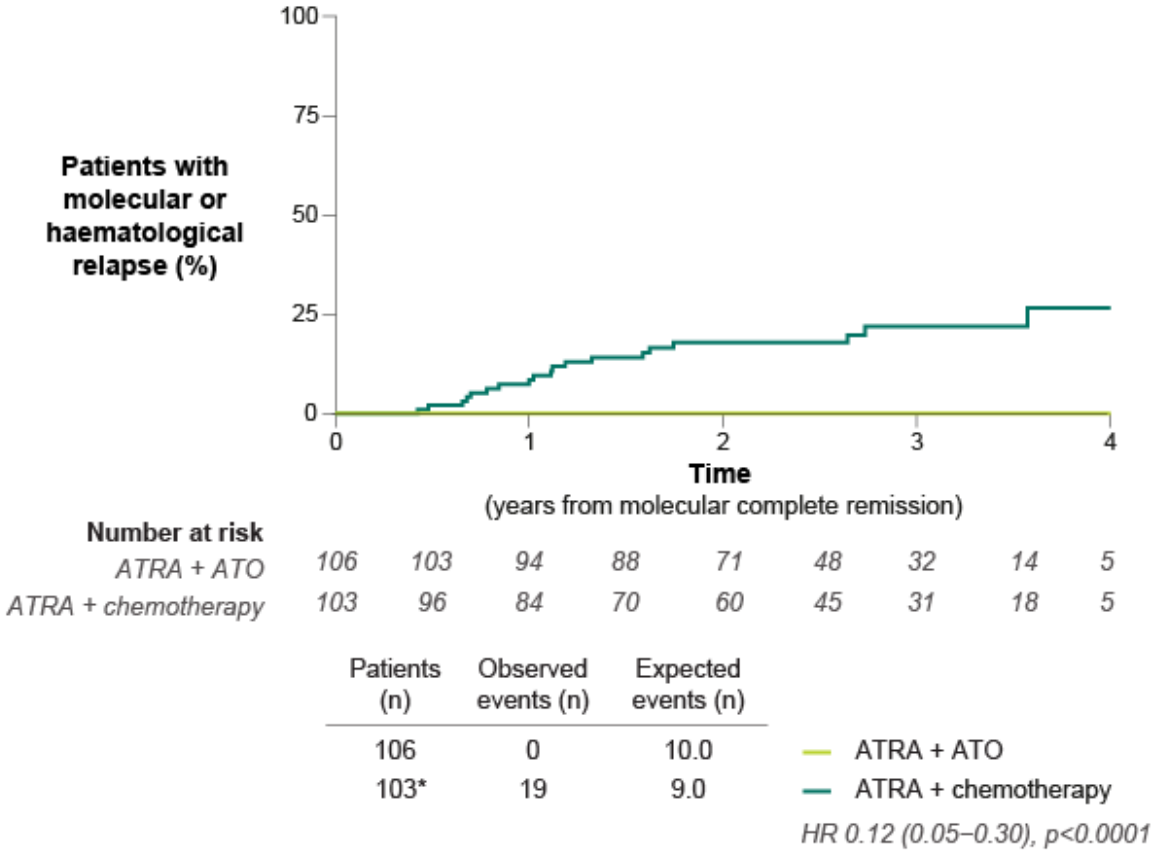
Figure 2.13. Cumulative incidence of haematological relapse in the AML17 trial[2]



HR=Hazard ratio
*Note that no follow-up data available for survival or relapse for available for 2 patients in the AIDA group (one high risk and one low risk)

In the updated analysis, ATRA+ATO continued to show a significant reduction in the incidence of relapse compared with AIDA (2% vs 16%, respectively, at 5 years, HR=0.19 [95% CI: 0.09–0.45]; p=0.0005), which translated into a continuing significant RFS benefit (96% vs 82%, respectively, HR=0.30 [95% CI: 0.13–0.67]; p=0.004)[4].

Figure 2.14.Cumulative incidence of molecular or haematological relapse in the AML17 trial[2]



HR=Hazard ratio
*Note that no follow-up data available for survival or relapse for available for 2 patients in the AIDA group (one high risk and one low risk)

B 2.7.1.4.e) Quality of life

Quality of life – the primary endpoint of the AML17 trial – was investigated using the EORTC QLQ-C30 questionnaire[2] (the same questionnaire that was used in the APL0406 study[40]; see ***section B 2.7.1.2.e***) for its brief description). The Hospital Anxiety and Depression Scale (HADS) was used in addition to the EORTC QLQ-C30 in the AML17 trial[2]. HADS includes two scales, one measuring depression and the other measuring anxiety; the scales detect these disorders in hospital patients and measure their severity[60].

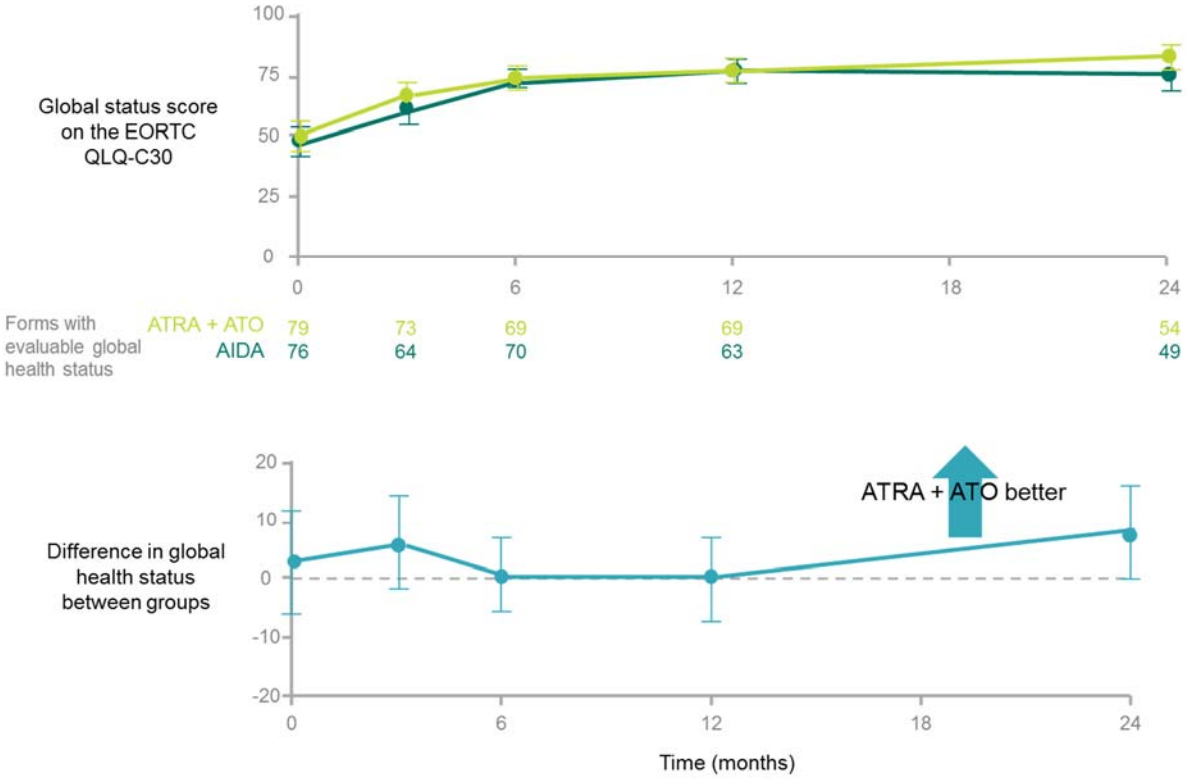
Patients enrolled in the AML17 trial returned a total of 671 completed QoL forms:[2]

- 156 at baseline (76 in the AIDA group and 80 in the ATRA+ATO group)
- 137 at 3 months (64 in the AIDA group and 73 in the ATRA+ATO group)
- 139 at 6 months (70 in the AIDA group and 69 in the ATRA+ATO group)

- 136 at 12 months (64 in the AIDA group and 72 in the ATRA+ATO group)
- 103 at 24 months (49 in the AIDA group and 54 in the ATRA+ATO group)

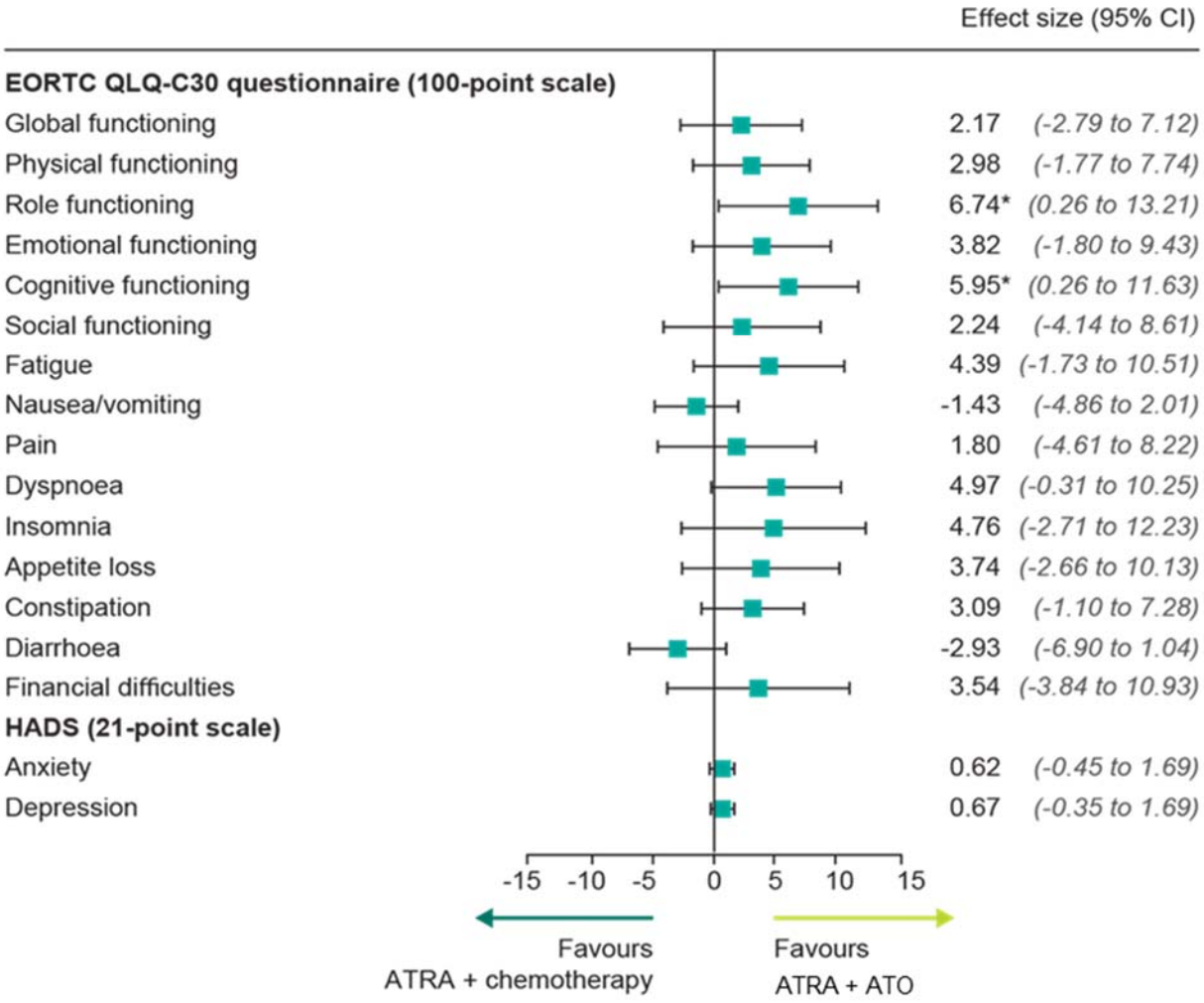
No statistically significant difference was detected in the primary outcome of global functioning (effect size=2.17 [95% CI: 2.79–7.12]; p=0.39, ***Figure 2.15***)[2]. Based on the power calculation, the confidence intervals ruled out a minimally clinically important disadvantage of six points for ATRA+ATO compared with AIDA[2]. For other measures, including fatigue which was significantly better with ATRA+ATO than AIDA in the APL0406 trial[40] (see ***section B 2.7.1.2.e***), point estimates tended to favour ATRA+ATO over AIDA, although any benefits were of modest size[2]. Small but statistically significant benefits of ATRA+ATO over AIDA were detected only for cognitive functioning (effect size=5.95 [95% CI: 0.26–11.63]; p=0.04) and role functioning (effect size=6.74 [95% CI: 0.26–13.21]; p=0.04)[2]. Quality of life results of the AML17 study are summarised in ***Figure 2.16*** and ***Table 2.17***.

Figure 2.15. EORTC QLQ-C30 global health status over time in the AML17 trial[2]



EORTC=European Organisation for Research and Treatment of Cancer; HADS=Hospital Anxiety and Depression Scale

Figure 2.16. QoL results obtained in the AML17 trial using EORTC QLQ-C30 and HADS[2]



ATRA=all-trans retinoic acid; CI=confidence interval; EORTC= European Organisation for Research and Treatment of Cancer; HADS= Hospital Anxiety and Depression Scale; QLQ-C30=Quality of Life Questionnaire – Core 30

Table 2.17. QoL results obtained in the AML17 trial[2]

Scale	Effect size (95% CI) Positive effect size indicates a benefit of ATRA+ATO over AIDA		p value
EORTC QLQ-C30			
Global functioning	2.17	(-2.79–7.12)	0.39
Physical functioning	2.98	(-1.77–7.74)	0.22
Role functioning	6.74	(0.26–13.21)	0.04
Emotional functioning	3.82	(-1.80–9.43)	0.18
Cognitive functioning	5.95	(0.26–11.63)	0.04
Social functioning	2.24	(-4.14–8.61)	0.49
Fatigue	4.39	(-1.73–10.51)	0.16
Nausea/ vomiting	-1.43	(-4.86–2.01)	0.41
Pain	1.80	(-4.61–8.22)	0.58
Dyspnoea	4.97	(-0.31–10.25)	0.06
Insomnia	4.76	(-2.71–12.23)	0.21
Appetite loss	3.74	(-2.66–10.13)	0.25
Constipation	3.09	(-1.10–7.28)	0.15
Diarrhoea	−2.93	(-6.90–1.04)	0.15
Financial difficulties	3.54	(-3.84–10.93)	0.35
HADS			
Anxiety	0.62	(-0·45–1.69)	0.26
Depression	0.67	(-0.35–1.69)	0.19

APL=acute promyelocytic leukaemia; ATRA=all-trans retinoic acid; CI=confidence interval; EORTC=European Organisation for Research and Treatment of Cancer; HADS=Hospital Anxiety and Depression Scale; HRQoL=health-related quality of life; ITT=intention-to-treat; QoL=quality of life

B 2.7.2 Relapsed or refractory APL

B 2.7.2.1 Key efficacy results from Raffoux, et al. 2003

B 2.7.2.1.a) Response rates and time to achieve CR

Following induction therapy, haematological CR rate was 80%, with 16 out of 20 patients (8 patients in the ATRA+ATO and 8 patients in the ATO only group) achieving CR[43]. Of the remaining 4 patients, 2 (both in the ATO only group) suffered an early death, and the other 2 (both in the ATRA+ATO group) were alive with resistant disease after induction[43]. The early deaths were caused by a septic shock with seizures (in a patient who had a previous history of CNS haemorrhage) and differentiation syndrome with hyperleukocytosis which did not respond to treatment with dexamethasone and amsacrine[43].

The median time needed to reach haematological CR among the 16 patients was 42 days in both treatment groups (p=0.58)[43]. Kaplan-Meier cumulative percentage of CR was similar across treatment groups (p=0.74, by the log-rank test)[43].

Only 3 out of 16 patients who reached CR achieved molecular remission after induction (2 patients in the ATO group and 1 patient in the ATRA+ATO group)[43]. All of these patients received amsacrine to counteract hyperleukocytosis occurring during treatment[43].

Ten CR patients received consolidation therapy; all of them were PCR-positive after the first consolidation cycle (including the 2 patients who were PCR–negative after induction)[43]. Of 8 patients tested after the second consolidation cycle, 2 were PCR-negative[43]. One patient, who remained PCR-positive after two consolidation cycles converted to PCR-negative during maintenance therapy which included ATO[43].

B 2.7.2.1.b) Other outcomes

Half of the 16 patients who achieved CR subsequently proceeded to receive HSCT; 2 patients underwent transplantation immediately after CR achievement with ATRA+ATO, 1 patient after a single consolidation cycle with ATRA+ATO and the remaining 5 patients were transplanted after two consolidation cycles[43]. There were a total of 5 relapses in 16 patients who achieved CR, including 2 relapses following HSCT[43]. Four of the patients who relapsed died from APL progression and one additional patient died in CR from sepsis following consolidation chemotherapy. Of the 9 remaining patients, 6 received allogeneic HSCT and remained in CR after a median follow-up of 18 months (range: 10–26 months)[43]. Two further patients became PCR-negative after additional chemotherapy and one patient became PCR-negative during maintenance treatment[43]. Thus, across both study groups, the estimated 2-year OS was 59% (95% CI: 35%–77%;

Figure 2.17) and the estimated 2-year DFS was 59% (95% CI: 29%–80%; ***Figure 2.18***)[43]. Both OS and DFS were similar in the two treatment groups[43].

Figure 2.17. OS in the study by Raffoux, et al. 2003

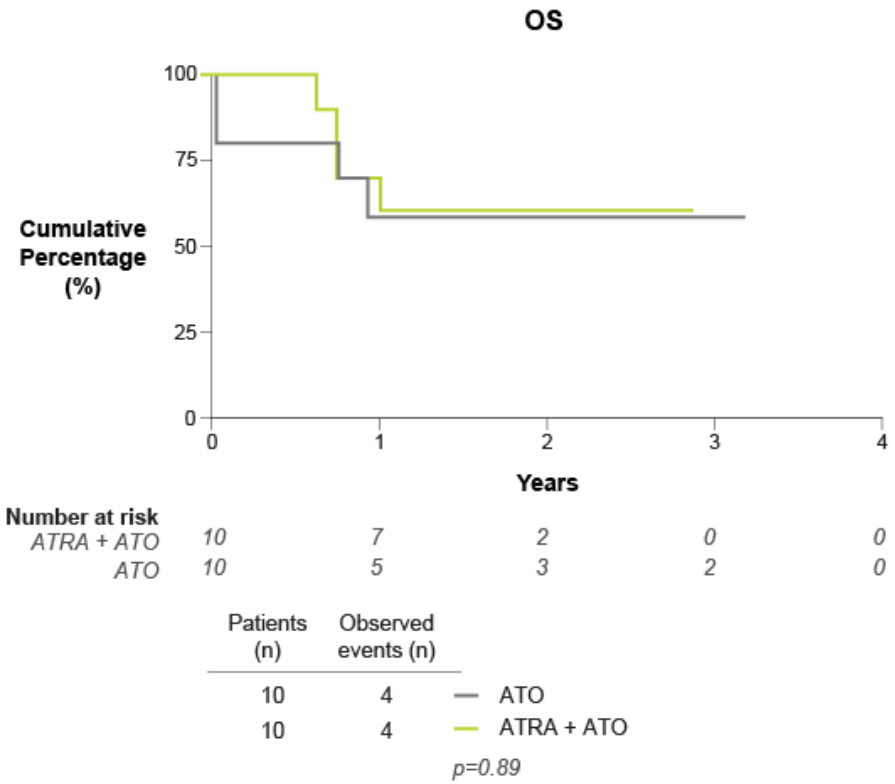
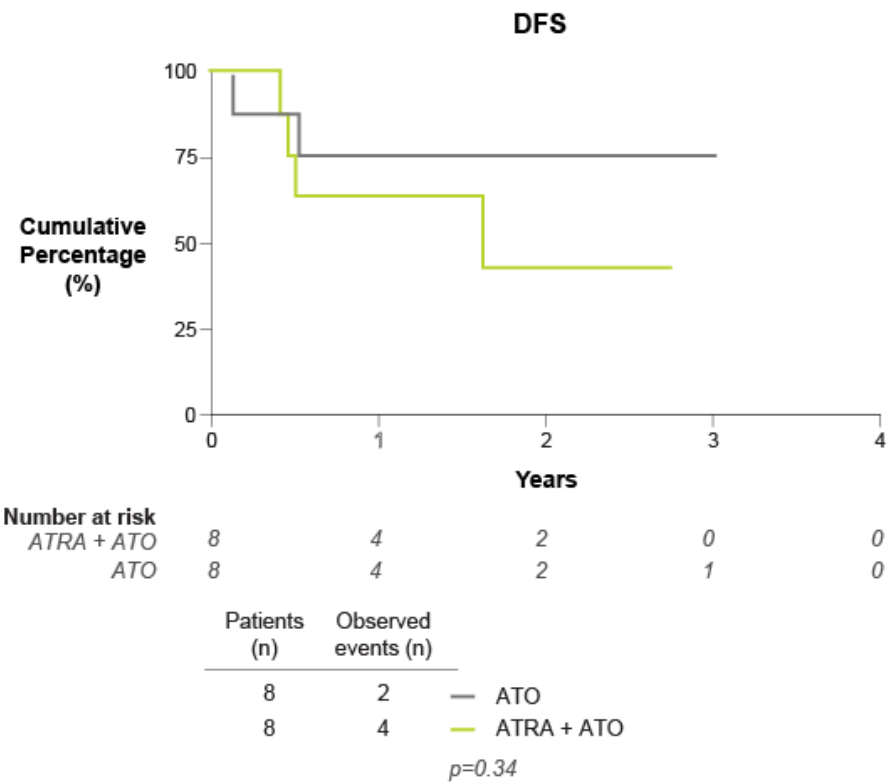


Figure 2.18. DFS in the study by Raffoux, et al. 2003



B 2.8 Subgroup analysis

Of the three studies presented here, a formal subgroup analysis was performed only in the AML17 trial, where pre-specified exploratory analyses stratified for age, sex, WBC count (including high-risk and low-risk patients), diagnosis, performance status, reverse transcript (RARA–PML) status, and PML breakpoint were conducted[2]. However, not all subgroup analyses were reported. As such, there is no ***Appendix E*** accompanying this submission.

B 2.8.1 Analysis by risk group

Amongst the 235 participants in the AML17 trial, 57 had high-risk APL; 30 of those patients were assigned to ATRA+ATO and 27 to AIDA. Risk status appeared to affect EFS in the AML17 trial, with a significant benefit of ATRA+ATO apparent in low- to intermediate-risk (4-year EFS of 92% [95% CI: 84%–97%] in the ATRA+ATO group vs 71% [95% CI: 55%–83%] in the AIDA group; HR=0.34 [95% CI: 0.15–0.75]; p=0.008), but not in high-risk patients (87% [95% CI: 68%–95%] vs 64% [95% CI: 42%–79%] in the ATRA+ATO and AIDA groups, respectively; HR=0.34 [95% CI: 0.11–1.08], p=0.07), despite very similar hazard ratios and no interaction between EFS and risk (p=1.0) [2]. It is therefore possible, that the relatively small number of high-risk patients (n=57, 24.3% of the trial population) contributed to the between-group EFS difference not being significant among these patients[2].

OS, which was not significantly different between treatment arms in the full study population (see ***section B 2.7.1.4.c)*** for detailed data and ***section B 2.14*** for discussion), was also similar across treatment arms in both risk groups[2]. Four-year OS in low- to intermediate-risk patients was 95% (95% CI: 86%–98%) in the ATRA+ATO arm and 90% (95% CI: 81%–95%) in the AIDA arm (***Figure 2.19***)[2]. In high-risk patients, 4-year OS was 87% (95% CI: 68%–95%) and 84% (95% CI: 63%–94%) in the two treatment arms, respectively (***HR=Hazard*** ratio

*Note that no follow-up data available for survival or relapse for available for 2 patients in the AIDA group (one high risk and one low risk)

Figure 2.20)[2]. There was no interaction between treatment and risk group (p=0.5)[2]. In the updated analysis, RFS was significantly better in the ATRA+ATO arm than the AIDA arm[4]. This was observed both in low- to intermediate-risk (95% vs 86%, respectively, HR=0.45 [95% CI: 0.17–1.20]; p=0.11) and high-risk patients (100% vs 69%, respectively, HR=0.10 [95% CI: 0.02–0.46]; p=0.003), p=0.11 for heterogeneity[4]. While the results obtained with ATRA+ATO in high-risk patients

enrolled in the AML17 trial appear promising, the data is currently considered insufficient and ATO is not licensed in this population.

Figure 2.19. OS in low- and intermediate-risk patients enrolled in the AML17 trial[2]

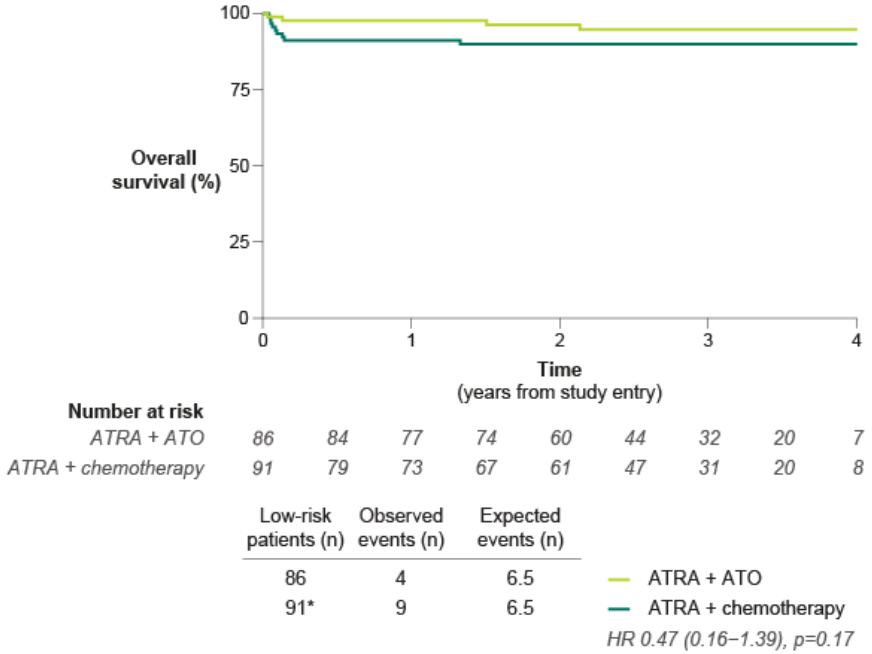
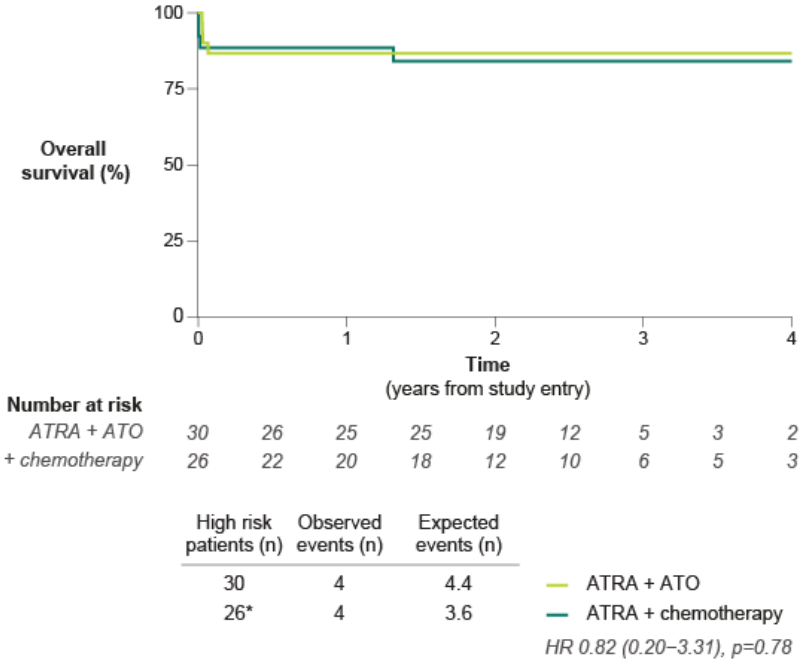


Figure 2.20. OS in high-risk patients enrolled in the AML17 trial[2]

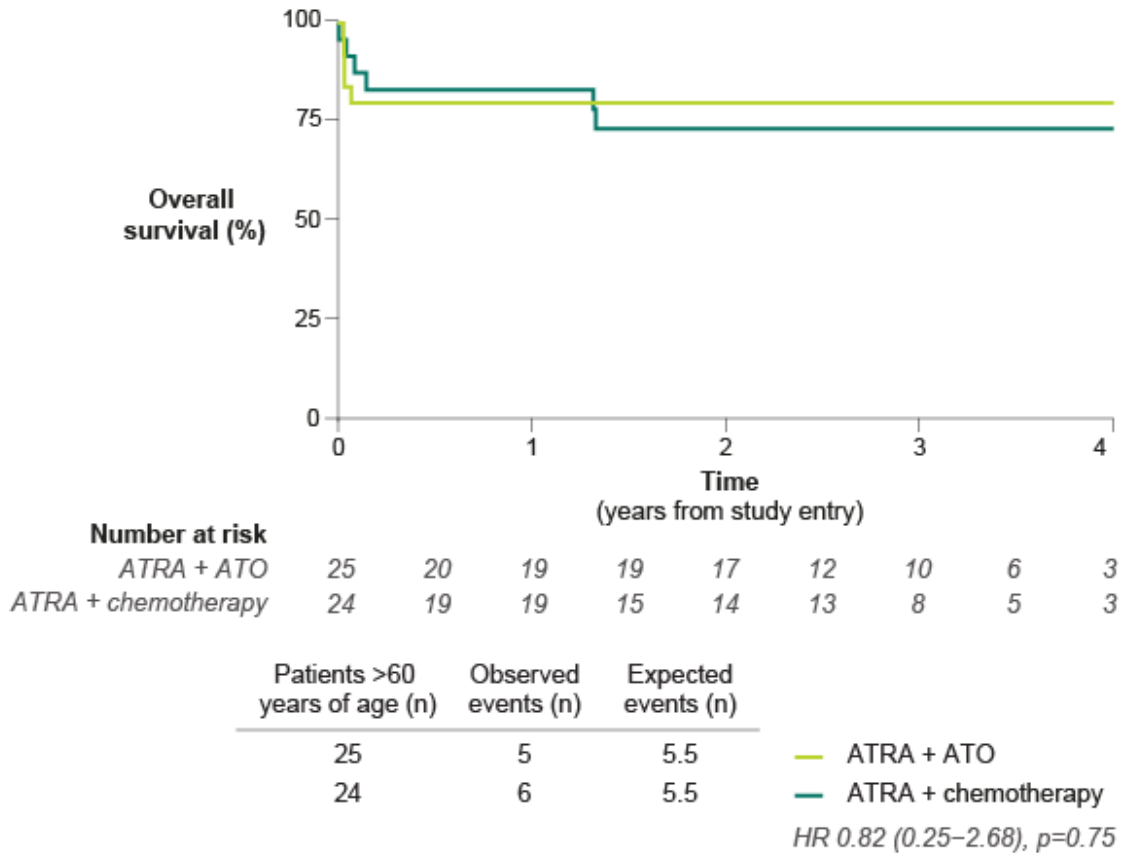


HR=Hazard ratio
*Note that no follow-up data available for survival or relapse for available for 2 patients in the AIDA group (one high risk and one low risk)

B 2.8.2 Analysis by age

Although the AL0406 trial did not include pre-specified subgroup analyses, this trial (as well as the AML17 trial[2]) provided some information on results obtained in elderly patients. The APL0406 study included patients younger than 71 years, and the results of 35 elderly patients (aged 60–70 years) included in the initial cohort of this study were published as scientific correspondence[61]. Sixteen elderly patients were assigned to ATRA+ATO and 19 to AIDA, and 2-year EFS rates were 100% and 84.2% in the two groups, respectively (p=0.40)[61]. As the EFS rates in elderly patients were similar to the results obtained in the full study population[1] (see ***section B 2.7.1.2.b***), the lack of statistical significance is likely caused by the small size of the elderly subgroup. In the ATRA+ATO group, no patients died and a single patient relapses at 27 months[61]. Three elderly patients died in the AIDA group – one during induction from differentiation syndrome and two during consolidation from pulmonary embolism and bronchopneumonia[61].

Figure 2.21. OS in the elderly subgroup of the AML17 trial[2]



The AML17 trial included 49 patients aged 60–77 years, of whom 37 were low- or intermediate-risk and 12 were high-risk[2]. Amongst the elderly patients, 25 received

ATRA+ATO and 24 received AIDA[2]. OS rates in the elderly subgroup (***Figure 2.21***) were lower than in the total study population for both treatments (see ***section B 2.7.1.4.c***), with no significant difference between treatment arms (80% [95% CI: 58%–91%] in the ATRA+ATO group, compared with 74% [95% CI: 50%–87%] in the AIDA group)[2].

B 2.9 Meta-analysis

The type of meta-analysis considered suitable in the context of available data was a network meta-analysis (NMA), in which multiple treatments are compared using both direct comparisons of interventions within RCTs and indirect comparisons. The trials captured in the systematic literature review (see ***Appendix D*** for details) used different comparators; thus, the NMA would synthesize the outcomes reported in those studies.

To ensure validity of findings from NMAs, similarity, homogeneity and consistency of included studies should be considered. Across studies, there are many sources of heterogeneity that are important, including clinical heterogeneity (i.e. different patient characteristics), or methodologic heterogeneity (i.e. studies not conducted in a similar fashion).

Studies identified from the literature review were grouped in a network if they reported a similar outcome measure, in a consistent way, and over similar time periods. All of the included studies were evaluated for feasibility of performing an NMA. After appraisal of the different study combinations, it was concluded that an NMA was not feasible for any of the outcomes. This was because the studies which were comparable in terms of time points and reported the same outcome did not have a mutual comparator.

Although the studies identified in the SLR did not allow us to perform an NMA, a recently published meta-analysis of 3 studies including a total of 585 newly-diagnosed APL patients (317 in ATRA+ATO ± chemotherapy group and 268 in ATRA + chemotherapy group) supported the use of ATRA+ATO as standard of care, particularly in low- to-intermediate risk APL patients[62]. Similarly, a Bayesian NMA of 14 studies (the majority conducted in Asia) that included a total of 1407 newly-diagnosed APL patients concluded that arsenic-based treatments (including ATO) may be considered optimal therapy in newly-diagnosed APL[63].

See ***section B 2.14*** for a qualitative summary and discussion of clinical results from studies reported in ***section B 2.3***.

B 2.10 Indirect and mixed treatment comparisons

As explained in ***section 0***, conducting an NMA proved not to be feasible; therefore, direct, indirect and mixed comparisons were not performed.

B 2.11 Adverse reactions

B 2.11.1 Summary of the safety profile of ATO (including the ATRA+ATO combination)

Although ATO has only recently been approved for the treatment of patients with newly-diagnosed APL, its safety profile is well-established through the clinical studies conducted in relapsed/refractory APL and the post-marketing data collected by Teva since 2000 when ATO was first approved in the US. Furthermore, a post-authorisation safety study (PASS) is in the process of approval by the European Medicines Agency. The study will start in 2018 and run for 5 years to evaluate long-term safety in newly-diagnosed APL patients treated with ATRA+ATO.

Safety data available to date indicate that notable adverse events associated with ATO include differentiation syndrome (also a known adverse effect of ATRA), QTc interval prolongation, hepatotoxicity, and leukocytosis. While all of these are potentially serious, adverse events occurring with ATRA+ATO in the APL0406 and AML17 trial were generally managed with temporary treatment discontinuation and subsequent temporary dose adjustment followed by a return to the full dose, as well as with appropriate supportive care[1-3]. Perhaps the most important safety advantage of the ATRA+ATO combination over AIDA was the considerable reduction in haematological toxicity observed in the APL0406 trial[1, 3]. In addition, patients receiving ATRA+ATO experienced less alopecia[2] and mucositis[1], both of which may be considered bothersome, and fewer serious adverse events (SAEs)[2, 3] than AIDA-treated patients.

Furthermore, data from the latest global Periodic Safety Update Report (PSUR) provide reassurance on the real-world safety of ATO. At the time of this PSUR, which covered data collected up to September 2016, products containing ATO as active substance had been registered by Teva Group in 39 countries around the world. Based on sales data it was estimated that interval (October 2015–September 2016) exposure to Teva Group products containing arsenic trioxide was approximately 3,649 patients. The estimated overall cumulative exposure to Teva Group products containing ATO was approximately 13,855 patients, with an estimated 363 patients exposed to ATO in 6 clinical trials sponsored by Teva Group.

Cumulatively, from post-marketing data sources (including non-serious reactions originating from solicited reports), there were 1022 case reports concerning ATO, of which 94 cases were received during the period covered by the latest PSUR. From the efficacy point of view, no new significant efficacy or effectiveness information was revealed in the reporting period. During the PSUR period, the following safety issues were analysed, as requested by regulatory authorities: renal failure, cerebral haemorrhage, cerebral infarction, rhabdomyolysis, hepatic disorders (in particular acute hepatitis and hepatic failure), myocardial toxicity, fatal cases, use of ATO in paediatric and elderly populations, and off-label use. Overall, based on post-marketing data sources and studies, no signals were identified that would necessitate changing the reference safety information.

As APL is a curable disease, potential long-term toxicities of treatment may be considered as important as its acute adverse effects. The use of ATRA+ATO may offer a long-term safety advantage over chemotherapy, as currently available data suggest fewer cases of tMDS-AML with this combination. In the final cohort of the APL0406 trial, one of 137 patients in the AIDA arm died from secondary MDS, while another patient presented with tAML and remained in CR 14 months after receiving an allogeneic stem cell transplant[3]. In the AML17 trial, one of 119 patients in the AIDA arm developed tAML[2]. No patient in the ATRA+ATO arm of either trial developed tMDS-AML[2, 3] and any reduction in the risk of tMDS-AML is an important advantage, since MDS or AML arising as a complication of prior cytotoxic treatment is associated with a poor prognosis[64, 65]. However, data on the incidence of secondary haematological malignancies following ATO-based treatment is at present very limited, with post-marketing surveillance expected to provide additional long-term safety information. Another potential long-term advantage of ATRA+ATO over AIDA arises from the fact anthracycline-induced cardiotoxicity may be avoided. While ATO causes a specific type of cardiac adverse event – QTc prolongation, which is an abnormality of the heart rhythm – this adverse effect is of relatively short duration with no long-term sequelae. In contrast, with anthracycline-based therapy, the risk of developing heart failure reaches 5% after 15 years from treatment[66], potentially necessitating a cardiac transplant[67]. Thus this toxicity of anthracyclines may severely impact patients who survive in long-term APL remission, so that avoiding it is another important advantage of the ATRA+ATO regimen.

The list of adverse events reported in the current SPC for ATO is shown in ***Appendix F***. The full list of AEs reported in the AML17 trial (including Grade 1–2 events) is also presented in ***Appendix F***. Adverse events reported in the three studies outlined in ***section B 2.3*** are presented in ***Table 2.18*** (APL0406), ***AEs were***

not spontaneously reported in the AML17 study. A limited list of AEs occurring during induction and the first consolidation cycle was presented, including: nausea, vomiting, alopecia, oral toxicity, diarrhoea, cardiac function, liver function (AST, ALT and bilirubin), renal function (creatinine, proteinuria, haematuria)[2].

Table 2.19 (AML17) and ***Table 2.20*** (Raffoux, et al. 2003). The next sections provide an overview of key safety issues related to ATO and their management in the APL0406 and AML17 trials. For details of any applicable dose titrations in the APL0406 study see ***section B 2.4.1***.

B 2.11.2 Safety evaluation in the reported trials

The safety evaluation in the published studies focused on toxicities rather than comprehensive reporting of all adverse events. All AEs, adverse drug reactions (ADRs), SAEs and serious unexpected adverse reactions were recorded during the treatment period in the APL0406 study. No long-term safety data were collected. The protocol stated that AEs were defined in accordance with the GCP definitions (2001/20/EC). A full list of SAEs was reported in the publication on the final cohort included in the study[3] and can be found in ***Appendix F***. Results on grade 3 or 4 haematological toxicity (neutropenia or thrombocytopenia), liver toxicity, neurotoxicity, GI toxicity, QTc prolongation, cardiac function, hypercholesterolaemia and hypertriglyceridaemia at induction and after each consolidation cycle were reported for the final patient cohort[3]. For the initial cohort the only results reported were a comparison of the incidence of haematological and non-haematological toxicity episodes during treatment[1], which was a secondary endpoint of the study; details of these AEs are available in ***Appendix F***.

Table 2.18. Adverse events occurring in the final patient cohort of the APL0406 trial[3]

Adverse Event	Time frame	ATRA+ATO	AIDA	p value
Induction-specific adverse events, n (%)				
Patients with moderate to severe differentiation syndrome	During induction	21 (17%)	17 (13%)	0.38
Leukocytosis*	During induction	56 (43%)	NR	NR
Haematological adverse events				
Patients with grade 3–4 neutropenia lasting >15 days	During induction	61 (35%)	109 (64%)	<0.001
	1st consolidation cycle	8 (16%)	40 (67%)	<0.001
	2nd consolidation cycle	7 (7%)	90 (92%)	<0.001
	3rd consolidation cycle	5 (15%)	28 (85%)	<0.001
Patients with grade 3–4	During induction	74 (38%)	120 (62%)	<0.001

Adverse Event	Time frame	ATRA+ATO	AIDA	p value
thrombocytopenia lasting >15 days	1st consolidation cycle	6 (26%)	17 (74%)	<0.001
	2nd consolidation cycle	6 (7%)	77 (93%)	<0.001
	3rd consolidation cycle	8 (23)%	26 (76)%	<0.001
FUO and infection episodes; n	During induction	30 (23%)	75 (55%)	<0.001
	1st consolidation cycle	10 (8%)	8 (6%)	0.540
	2nd consolidation cycle	4 (3%)	46 (38%)	<0.001
	3rd consolidation cycle	2 (1.6%)	2 (1.7%)	1.000
Non-haematological adverse events				
Patients with QTc prolongation**	During induction	11 (8.5%)	1 (0.7%)	0. 002
	1st consolidation cycle	3 (2%)	0	0.110
	2nd consolidation cycle	3 (2%)	0	0.110
	3rd consolidation cycle	2 (1.5%)	0	0.230
Patients with grade 3–4 hepatic toxicity	During induction	51 (40%)	4 (3%)	<0.001
	1st consolidation cycle	5 (4%)	1 (0.7%)	0.110
	2nd consolidation cycle	1 (0.8%)	0	0.490
	3rd consolidation cycle	0	0	NA
Patients with grade 3–4 gastrointestinal toxicity	During induction	3 (2%)	25 (18.2%)	<0.001
	1st consolidation cycle	0	1 (0.8%)	1.000
	2nd consolidation cycle	0	6 (4.9%)	0.03
	3rd consolidation cycle	0	0	1.000
Patients with grade 3–4 cardiac function abnormalities	During induction	0	5 (3.7%)	0.060
	1st consolidation cycle	0	0	NA
	2nd consolidation cycle	0	0	NA
	3rd consolidation cycle	0	0	NA
Neurotoxicity (all grades)	During induction	1 (0.7%)	0	0.480
	1st consolidation cycle	5 (4.2%)	0	0.020
	2nd consolidation cycle	6 (5%)	0	0.010
	3rd consolidation cycle	7 (5.9%)	0	0.006
Hypercholesterolemia	During induction	14 (10%)	12 (8.7%)	0.550
	1st consolidation cycle	19 (16%)	12 (9.6%)	0.130
	2nd consolidation cycle	19 (16%)	12 (9.7%)	0.140
	3rd consolidation cycle	16 (14%)	11 (9.0%)	0.270
Hypertriglyceridemia	During induction	29 (22%)	29 (22%)	0.760
	1st consolidation cycle	22 (18.4%)	19 (15.2%)	0.490
	2nd consolidation cycle	17 (14.4%)	10 (8%)	0.120
	3rd consolidation cycle	16 (14%)	13 (11%)	0.500

ATO=Arsenic trioxide; ATRA=all-trans retinoic acid; FUO=Fever of unknown origin

*Leukocytosis was defined as WBC count >10 × 10^9/L
**Defined as QTc increased to >450 msec in males and >460 msec in females.

AEs were not spontaneously reported in the AML17 study. A limited list of AEs occurring during induction and the first consolidation cycle was presented, including: nausea, vomiting, alopecia, oral toxicity, diarrhoea, cardiac function, liver function (AST, ALT and bilirubin), renal function (creatinine, proteinuria, haematuria)[2].

Table 2.19. AML17 trial: Adverse events[2]

Events	**Notes**	**Time frame**	**ATRA+ATO**	**AIDA**
Number of SAEs*		NR	64	82
Patients experiencing SAEs*		NR	46 (39.7%)	53 (44.5%)
Patients with differentiation syndrome**	Low-risk	NR	23 (26.7%)	15 (16.3%)
	High-risk	NR	7 (23.3%)	10 (37%)
Patients with grade 3–4 toxicity		Induction or 1st consolidation cycle	40	57

Specific events occurring during induction or 1st consolidation course						
Events n/n returning data for each event (%)	**Grade*** **	**Time frame**	**ATRA+ATO**		**AIDA**	
Nausea	Grade 3	Induction	0/110		5/115	(4%)
		1st consolidation cycle	0/93		1/101	(1%)
	Grade 4	Induction	0/110		0/115	
		1st consolidation cycle	0/93		0/101	
Alopecia	Grade 3	Induction	3/95	(3%)	13/98	(13%)
		1st consolidation cycle	0/77		11/89	(12%)
	Grade 4	Induction	2/95	(2%)	10/98	(10%)
		1st consolidation cycle	2/77	(3%)	14/89	(16%)
		1st consolidation cycle	3/93	(3%)	3/101	(3%)
Diarrhoea	Grade 3	Induction	1/109	(1%)	7/115	(6%)
		1st consolidation cycle	1/93	(1%)	1/101	(1%)
	Grade 4	Induction	0/109		0/115	
		1st consolidation cycle	0/93		0/101	
		1st consolidation cycle	0/94		5/101	(5%)
Oral toxicity	Grade 3	Induction	1/109	(1%)	17/115	(15%)
		1st consolidation cycle	0/94		0/101	
	Grade 4	Induction	0/109		5/115	(4%)
		1st consolidation cycle	0/94		0/101	
		1st consolidation cycle	2/92	(2%)	0/99	
Cardiac toxicity	Grade 3	Induction	1/107	(1%)	5/110	(5%)
		1st consolidation cycle	3/92	(3%)	0/99	
	Grade 4	Induction	1/107	(1%)	1/110	(1%)
		1st consolidation cycle	0/92		0/99	
		1st consolidation cycle	0/38		0/48	
Raised liver ALT	Grade 3	Induction	22/109	(20%)	9/108	(8%)
		1st consolidation cycle	0/38		0/48	

	Grade 4	Induction	5/109	(5%)	2/108	(2%)
		1st consolidation cycle	0/38		0/48	
		1st consolidation cycle	9/93	(10%)	10/98	(10%)
Raised liver AST	Grade 3	Induction	2/46	(4%)	2/51	(4%)
		1st consolidation cycle	2/93	(2%)	2/98	(2%)
	Grade 4	Induction	0/46		0/51	
		1st consolidation cycle	0/93		0/98	
		1st consolidation cycle	0/93		1/101	(1%)
Hyperbilirubinaemia	Grade 3	Induction	1/110	(1%)	6/114	(5%)
		1st consolidation cycle	0/93		0/101	
	Grade 4	Induction	0/110		2/114	(2%)
		1st consolidation cycle	0/93		0/101	
		1st consolidation cycle	0/93		2/101	(2%)
Raised creatinine	Grade 3	Induction	1/110	(1%)	0/114	
		1st consolidation cycle	0/93		0/101	
	Grade 4	Induction	0/110		1/114	(1%)
		1st consolidation cycle	0/93		0/101	
		1st consolidation cycle	1/82	(1%)	0/73	
Proteinuria	Grade 3	Induction	0/82		0/87	
		1st consolidation cycle	0/82		0/73	
	Grade 4	Induction	0/82		1/87	(1%)
		1st consolidation cycle	0/82		0/73	
		1st consolidation cycle	0/67		0/76	
Haematuria	Grade 3	Induction	0/82		0/90	
		1st consolidation cycle	0/67		0/76	
	Grade 4	Induction	0/82		1/90	(1%)
		1st consolidation cycle	0/67		0/76	

ALT=alanine transaminase; AST=aspartate aminotransferase; ATO=Arsenic trioxide; ATRA=all-trans retinoic acid; CTCAE=Common Terminology Criteria for Adverse Events; NCI=National Cancer Institute; NR=not reported; SAE=serious adverse event
*Including death, which was not routinely assessed in relation to study treatment.
**The overall p value for the comparison was 0.38.
***Toxicity was recorded and graded using the NCI-CTCAE version 3.0

Table 2.20. Raffoux, et al. 2003: Treatment-related adverse events during induction[43]

Adverse event	ATO (n, %)	ATRA+ATO (n, %)	All patients	
			n	%
Gain in weight	6 (60%)	6 (60%)	12	60
ALT/AST elevation, grade ≥2	5 (50%)	4 (40%)	9	45
Hypokalaemia	4 (40%)	3 (30%)	7	35
Headaches	1 (10%)	5 (50%)	6	30
Hyperglycaemia	3 (30%	2 (20%)	5	25
Nausea, vomiting	2 (20%)	3 (30%)	5	25
QT prolongation	3 (30%)	2 (20%)	5	25
Diarrhoea	1 (10%)	3 (30%)	4	20
Peripheral neuropathy	0	2 (20%)	2	10
Deep venous thrombosis	2 (20%)	0	2	10

Differentiation syndrome	3 (30%)	4 (40%)	7	35
With hyperleukocytosis	3	3	6	
With respiratory failure	2	1	3	

ALT=alanine transaminase; AST=aspartate aminotransferase; ATO=Arsenic trioxide; ATRA=all-trans retinoic acid

B 2.11.3 Study withdrawals

In the final cohort of the APL0406 study, there were 5 withdrawals due to toxicity in the ATRA+ATO group (1 during induction and 4 during consolidation) and 10 in the AIDA group (6 during consolidation and 4 during maintenance)[3]. In the ATRA+ATO group severe QTc interval prolongation and electrolyte abnormalities resulted in premature (day 3) termination of induction therapy in one patient[1,3], who was subsequently lost to follow-up[3]. No information was provided on the toxicities experienced during consolidation therapy that caused 4 patients among the final APL0406 cohort to go off-protocol before being evaluated for molecular remission status[3]; however, based on the earlier publication, one patient did not proceed to consolidation therapy due to repetitive tachycardia[1]. Details of the toxicities causing study withdrawals in the AIDA arm were not reported in the publication on the final APL0406 cohort[3] but information on the nature of 3 of these toxicities is available in the earlier (initial cohort) publication. Specifically, 1 patient in the AIDA group did not receive consolidation therapy due to a cardiac toxic effect and 2 patients did not complete maintenance therapy, due to prolonged myelosuppression lasting more than 50 days[1]. No information is available from the AML17 study on the number of patients discontinuing ATRA+ATO due to safety reasons. However, it was noted that 1 patient withdrew from the trial during induction because of fungal chest complications, while during the first consolidation course, 2 patients in the ATRA+ATO arm were not given ATO due to QTc prolongation[2]. Overall, across both studies there were very few discontinuations of ATRA+ATO due to safety reasons. No information on safety-related withdrawals is available from Raffoux, et al. 2003[43].

B 2.11.4 APL Differentiation Syndrome

All patients in the APL0406 study received differentiation syndrome prophylaxis with prednisone (0.5 mg/kg/day) from day 1 until the end of induction treatment. In contrast, no prophylaxis for differentiation syndrome was recommended in the AML17 study. In both studies, at the earliest manifestations of suspected differentiation syndrome (e.g., unexplained respiratory distress) temporary discontinuation of ATRA and/or ATO treatment and prompt administration of dexamethasone was recommended. An important insight into the risk of

differentiation syndrome comes from the study by Raffoux, et al., where the incidence of differentiation syndrome was similar with ATO alone and ATRA+ATO[43], suggesting this adverse effect of both medications is non-additive. Indeed, co-administration of ATO with ATRA does not appear to increase the risk of differentiation syndrome relative to ATO alone, with similar rates observed in the AIDA and ATRA+ATO arms of both the APL0406 and AML17 trials[1-3].

B 2.11.5 ECG abnormalities

In the APL0406 trial, QTc prolongation was defined as QTc (corrected using the Framingham formula) exceeding 450 ms for men and 460 ms for women[3]. It is unclear how it was defined in the AML17 trial; however both the APL0406 and the AML17 studies required that in the event of QTc prolongation ATO should be temporarily discontinued and electrolytes should be repleted.

B 2.11.6 Hepatotoxicity

In both studies, hepatotoxicity was defined as serum bilirubin, AST, or alkaline phosphatase exceeding 5 times the upper limit of normal (ULN). In addition, in the AML17 trial ALT exceeding 5 times the ULN was also considered a sign of hepatotoxicity. Both the APL0406 and the AML17 trials recommended temporary discontinuation of ATRA and/or ATO in the event of hepatotoxicity. Overall, the incidence of hepatotoxicity was lower in the AML17 than the APL0406 trial – a possible benefit of the less frequent dosing regimen.

To better understand the mechanism of hepatotoxicity occurring with ATO, Teva requested an expert review from Professor Minotti (Professor of Clinical Pharmacology, University Campus Bio-Medico of Rome). He concluded that, while the mechanism of this hepatotoxicity is multifactorial and only partially understood, oxidative stress appears to play an important role and there is substantial variability between individuals in terms of susceptibility to hepatotoxic effects, most likely related to genetic differences. Professor Minotti also highlighted that, even though the results of the APL0406 study may suggest that the intensity of oxidative stress increases with ATRA co-administration, accumulation of arsenic and prolonged hepatic damage is unlikely, as hepatotoxicity was reversible and did not reoccur after ATO was reintroduced.

B 2.11.7 Hyperleukocytosis

The APL0406 study included only low- to intermediate-risk patients (WBC count at baseline $<10\times10^9$/L). Hydroxyurea was the recommended treatment in the event of

leukocytosis, defined as WBC count rising above 10×10^9/L. Upon WBC count returning below that threshold, hydroxyurea was discontinued.

In the AML17 study, high-risk patients were recommended to receive a single infusion of gemtuzumab ozogamicin administered in the first four days of induction treatment. Of the 30 high-risk patients, 28 received gemtuzumab ozogamicin and the remaining 2 were given an anthracycline instead[2]. In addition, gemtuzumab ozogamicin was administered to 7 low- and intermediate-risk patients for a rising WBC count[2].

B 2.11.8 Haematological toxicity

In the initial cohort of patients enrolled in APL0406, grade 3–4 haematological toxicity (neutropenia or thrombocytopenia) lasting for over 15 days was significantly more common in the AIDA group than in the ATRA+ATO group at each time point investigated (i.e., during induction, and following each consolidation course)[1]. Infections and fever of unknown origin occurring during induction or consolidation were also more common in the AIDA group than in the ATRA+ATO group (59 vs 26 episodes, p<0.001) [1]. The results were similar in the final cohort of patients enrolled in the APL0406 trial, and patients treated with AIDA experienced grade 3–4 neutropenia or thrombocytopenia lasting more than 15 days significantly more often than those treated with ATRA+ATO[3]. This pattern was observed during induction as well as after each consolidation course[3]. Infections or fever of unknown origin were less frequent in the ATRA+ATO arm than in the AIDA arm during induction (30 vs 75 episodes, p<0.001) and during the second consolidation course (4 vs 46 episodes, p<0.001); however, during the first (10 vs 8 episodes, p=0.54) and third consolidation courses (2 episodes in each group, p=1.0) the rates were similar between the two treatment arms[3]. See ***Table 2.18.*** Adverse events occurring in the final patient cohort of the APL0406 trial for detailed rates of haematological toxicity in the APL0406 study. Haematological toxicity was not reported in the AML17 trial.

B 2.12 Ongoing studies

As mentioned above, a PASS is in the process of approval by the European Medicines Agency. The study will start in 2018 and run for 5 years to evaluate long-term safety in APL patients treated first-line with ATRA+ATO; however, the protocol of this study has not yet been approved. Furthermore, the investigator-sponsored APOLLO study is currently recruiting participants with high-risk APL[68]. This is an open-label, randomised parallel-assignment phase 3 trial comparing standard AIDA-based treatment with the ATRA+ATO combination in patients with high-risk APL[68].

Two doses of idarubicin will be added to ATRA+ATO during induction with the aim to control hyperleukocytosis and achieve better long-term disease control[68].

B 2.13 Innovation

Although ATO is not a new chemical entity, its impact on the treatment of APL is profound. ATO has been available as a treatment for relapsed or refractory APL patients for over 15 years, but its true innovative potential lies in offering a chemotherapy-free treatment option to newly-diagnosed low- to intermediate-risk APL patients. In fact, data from the AML17 trial suggest that the ATRA+ATO combination can also be used to successfully treat high-risk APL patients while minimising the need for chemotherapy. However, data in this latter population is scarce at present, and the use of ATO in high-risk APL and the optimal treatment combinations remain to be investigated further.

The main advantage of chemotherapy-free treatment is avoiding the related toxicity; this is especially true for haematological toxicity (see section ***B 2.11.8***). Furthermore, intensification of chemotherapy appears to increase the number of deaths in CR[30]. Indeed, although the numbers were low in both groups, numerically fewer deaths in CR occurred in the ATRA+ATO than the AIDA arm of the APL0406 trial (1 vs 5 deaths). Another important benefit of ATRA+ATO, which provides excellent efficacy combined with a manageable safety profile, is that it offers an effective treatment option for patients with contraindications to chemotherapy, who would not be eligible for chemotherapy-based treatment. Indeed, ATRA+ATO is already being used in these patients in UK clinical practice. Finally, the reduced incidence of relapse observed with the ATRA+ATO combination compared with the standard chemotherapy-based AIDA regimen means that the majority of treated patients can remain disease-free (i.e., “cured”) and avoid the need for further therapy beyond first-line treatment.

B 2.14 Interpretation of clinical effectiveness and safety evidence

APL can be considered a curable disease, and the most important outcome that treatment can offer to affected patients is long-term disease-free survival. In newly-diagnosed APL, the combination of ATRA and ATO provides an advantage over AIDA in that it significantly improves both EFS and DFS/RFS – a benefit observed in both the APL0406 and the AML17 trial. With few deaths in CR observed in either trial, and the number of patients achieving CR similar between ATRA+ATO and AIDA, reduction in relapse rate can be considered the primary source of the EFS advantage observed with ATO. Indeed, in both the APL0406 and the AML17 trials,

cumulative incidence of relapse in the ATRA+ATO arm did not exceed 2% and was significantly lower than in the AIDA arm. Considering the very low number of patients relapsing following first-line ATRA+ATO therapy, it can be expected that with widespread use of this combination in newly-diagnosed patients the need for second-line treatment will diminish.

In both the APL0406 and AML17 trials, QoL results were similar between treatment arms, suggesting that the combination of ATRA+ATO improves outcomes without adversely impacting QoL. Finally, the APL0406 showed a significant improvement in OS with ATRA+ATO compared with AIDA. This OS benefit was not reproduced in the AML17 trial, although the results did numerically favour ATRA+ATO. The non-significant difference in OS in the AML17 trial can likely be attributed to the impact of ATO-based salvage therapy on OS estimates. Of the 20 relapses in the AML17 trial, 19 occurred in the AIDA arm and 18 of these patients (15% of all patients randomly assigned to AIDA) received ATO as second-line treatment upon relapse. Of the 20 patients in both groups who relapsed, only two died (including the only patient in the ATRA+ATO group who relapsed). Thus, the impact of salvage therapy was particularly evident in the AIDA arm of AML17, with successful second-line treatment effectively “boosting” survival in this treatment arm.

Although both trials used the ITT principle, so that patients who relapsed remained in their randomly assigned treatment group for OS analysis even if they subsequently received other treatments, the impact of salvage treatment was more profound in the AML17 study than in the APL0406 study. While ATO could also be used (off-protocol) as a salvage treatment in the APL0406 trial, only one patient was reported to be treated with ATO at relapse and it is unclear if any additional patients received ATO-based salvage treatment. Furthermore, testing for molecular relapse was more rigorous in the AML17 trial than in the APL0406 trial, leading to more relapses being detected during the study period (20 in 215 patients at a median follow-up of 30.5 months, compared with 17 in 259 patients at a median follow-up of 40.6 months in APL0406). With the higher number of relapses detected in the AML17 trial and the almost uniform use of ATO at relapse, the impact of salvage treatment on OS estimate was much more substantial in the AML17 trial than the APL0406 trial.

In the relapsed/refractory APL population, cure is still a valid aim of treatment. Treatment with ATRA+ATO allows a considerable proportion of patients (80% in the study by Raffoux, et al.[43]) to achieve a second remission, allowing patients to be considered for potentially curative allogeneic or – if molecular remission is achieved – autologous HSCT. However, available real-world evidence from a European

registry of relapsed APL suggest that ATO-based treatment may provide patients with chances of long-term survival even if not followed with a HSCT, with 3-year OS rates approaching 60% in this patient group[7]. More recently, 96% 3-year OS has been reported in the updated analysis of the AML17 trial[4]. Amongst the 25 patients in the AIDA arm who relapsed after a median follow-up of 53.4 months, 24 (including 5 with concomitant CNS involvement) were treated with ATRA+ATO and 11 subsequently underwent HSCT (8 autologous and 3 allogeneic) [4]. Three patients relapsed again (one of them post-transplant)[4]. Two patients who relapsed subsequently died – one in haematological relapse before treatment could be initiated and the other at 37 months following transplantation[4].

In terms of safety, the combination of ATRA and ATO provides an important advantage of reducing haematological toxicity compared with AIDA. Furthermore, its long-term safety profile may also prove more favourable, with no cases of tMDS-AML observed in either trial. Although the combination of ATRA+ATO is associated with important adverse events, including QTc prolongation, hepatotoxicity, leucocytosis and differentiation syndrome, adverse events in clinical trials were mostly managed with temporary treatment discontinuation and supportive care, with few permanent discontinuations being reported. Importantly, it appears that the incidence of differentiation syndrome – a potentially fatal adverse effect of both ATRA and ATO – is not increased through concomitant administration of the two medications.

Overall, the combination of ATRA+ATO provides newly-diagnosed APL patients with an effective treatment option, characterised by a manageable safety profile and a lack of adverse impact on QoL. For relapsed or refractory patients, it offers a bridge to potentially curative HSCT, but may also be a valid treatment option without subsequent transplantation.

APL is a rare disease, which is reflected in the relatively small number of available RCTs, especially in relapsed/refractory disease. Furthermore, the population of the only RCT in relapsed/refractory APL presented here comprised as few as 20 patients in total[43]. Both studies in newly-diagnosed APL enrolled over 200 patients and used robust methodology, although both were open-label. One could, however, argue that blinding would not have been feasible in these studies, given the very different administration schedules and adverse event profiles of the investigated treatments. The validity of the APL0406 trial is supported by the fact it was considered as the pivotal source of clinical data for the approval of ATRA+ATO in newly-diagnosed low- to intermediate-risk APL. The AML17 study, although providing only supporting

regulatory information, was conducted primarily in the UK, and market research commissioned by Teva suggested that patients were encouraged to participate in this trial upon diagnosis. Thus, during its course, AML17 can be considered as having effectively constituted the UK approach to APL treatment, at least for those patients who were eligible to enrol.

ATO has been marketed in the UK and in Europe for more than 10 years, with over 10,000 of patients treated worldwide and over 1,000 patients in the UK. Thus, in addition to the strong evidence from the RCTs reported in this section, the effectiveness and safety of ATO are supported by many non-randomised studies conducted in patient populations more closely resembling those encountered in everyday clinical practice, and by evidence of routine use demonstrated by a stable market share in the UK.

Given the high rates of overall survival achieved with APL treatments, ATO is unlikely to meet the end-of-life criteria. For the discussion on the expected number of patients to be treated, see ***section B 1.3.3***.

B.3 Cost effectiveness

B 3.1 Published cost-effectiveness studies

A systematic literature review was performed with the objective of identifying published cost-effectiveness evaluations of APL treatments. The selection process of relevant studies is detailed in ***Appendix G***.

Five studies (presented in 6 publications) describing economic analyses of APL treatments fulfilled all of the inclusion criteria listed in ***Appendix G*** and are summarised in ***Table 3.1***. Note, that the two publications by Tallman, et al. present the same study; however, the abstract[69] presented a cost-effectiveness comparison of ATRA+ATO vs ATRA + Ara-C + CT, which was not available in the full publication[70]. We hypothesise this is because the full publication only compared cost-effectiveness of regimens for which a direct clinical trial comparison was available (i.e. AIDA vs ATRA + Ara-C + CT and ATRA+ATO vs AIDA; to the best of our knowledge no study compared ATRA+ATO vs ATRA + Ara-C + CT). As the abstract provides additional information of interest (comparison of cost effectiveness between ATRA+ATO vs ATRA and Ara-C + CT), both the abstract and the full publication by Tallman, et al. are presented in ***Table 3.1.***

All of the studies presented below used the Markov approach to calculate the cost-effectiveness of analysed treatments. Across studies, the population consisted of

adult patients newly diagnosed with low- to intermediate-risk APL patients (i.e. those with a WBC count ≤10x10^9/L).In general, all of the included studies considered the cost-effectiveness of ATRA+ATO or ATO alone, compared to the combination of ATRA and chemotherapy. In all cases, the number of QALYs was higher in the groups receiving ATO than in the comparators group. Mean total costs of ATO alone or ATRA+ATO were higher than the costs of comparator treatments. Incremental cost effectiveness ratios (ICERs) differed between studies, which could be related to a number of methodological factors, including the fact that the studies concerned different countries.

Table 3.1. Summary list of published cost-effectiveness studies

Study	Year	Summary of model	Patient population (average age in years)	QALYs (intervention, comparator)	Costs (currency) (intervention, comparator)	ICER (per QALY gained)
Lachaine, et al. 2015a[71]	2015	A Markov model simulating the course of newly diagnosed APL patients receiving induction therapy followed by consolidation therapy in case of a complete response (CR). The model was developed with the following health states: event-free survival, treatment failure, post-treatment failure, and death. The length of each Markov cycle was 1 month during the 48 month study period and 1 year thereafter.	The population considered included 162 adult patients with newly diagnosed, genetically confirmed, low- to intermediate-risk (WBC count ≤10x10^9/L) APL. The mean age of the cohort was 46 years.	The intervention was ATO+ATRA (14.68 QALYs) and the comparator was AIDA (13.24 QALYs). The applicable discount rate was 5%.	• Total cost of ATO+ATRA treatment from the MoH perspective: 145,962 Can$ • Total cost of ATO+ATRA treatment from the societal perspective: 168,043 Can$ • Total cost of AIDA treatment from the MoH perspective: 73,768 Can$ • Total cost of AIDA treatment from the societal perspective: 95,640 Can$ The applicable discount rate was 5%.	• ICER from the MoH perspective: $CAD 50,193 (€35,665) • ICER from the societal perspective: $CAD 50,338 (€35,769)
Tallman, et al. 2015a[70]	2015	A Markov model was developed with 4 health states: stable disease (1st line treatment), disease event/stable disease (2nd line treatment), disease event and death. The length of each cycle was 1 month.	Newly-diagnosed, adult (aged ≥18 years) patients with low- to intermediate- risk (WBC count ≤10x10^9/L) APL.	The model considered the following interventions: • ATO+ATRA (14.33 QALYs) • AIDA (8.13 QALYs) • ATRA+AraC+CT (6.71 QALYs) The applicable discount rate was 3%.	Total costs of treatments from the US third party payer perspective: • Total cost of ATO+ATRA treatment: 136,170 $ • Total cost of AIDA treatment: 101,396 $ • Total cost of ATRA+AraC+CT treatment: 96,940 $ The applicable discount rate was 3%.	• ICER for ATO+ATRA (vs AIDA) was $5,614 • ICER for AIDA (vs ATRA+Ara-C+CT) was $3,122
Tallman, et al.	2015	A Markov model with 4 health states	Newly-diagnosed,	The model	Not reported	• ICER for

Study	Year	Summary of model	Patient population (average age in years)	QALYs (intervention, comparator)	Costs (currency) (intervention, comparator)	ICER (per QALY gained)
2015b[69]		was developed: 1st line stable disease, 2nd line stable disease, 2nd line disease event and death. Each cycle lasted for 1 month. Patients in the model began treatment at age 45 and were followed until death. Eight months treatment duration of ATO+ATRA was compared to either 15 months of treatment with ATRA+Ara-C+chemotherapy or 33 months of treatment with. AIDA	low- to intermediate-risk adult APL patients.	considering following intervention: • ATO+ATRA and comparator: • ATRA+AraC+CT • AIDA Discount rate not available.		ATO+ATRA vs ATRA + Ara-C + CT: $5,900 • ICER for ATRA+ATO vs. AIDA: weakly dominant
Kruse, et al. 2015[72]	2015	A Markov model was constructed with 3 health states: stable disease, disease event and death. After a disease event, patients discontinued initial treatment and switched to the other regimen as second-line therapy. Each cycle lasted 1 month and the time horizon was 3 years.	Newly diagnosed, Adult APL patients (aged ≥18 years) with low- to intermediate- risk disease based on WBC count.	Intervention was ATO+ATRA, and the comparator was AIDA. Number of QALYs wasn't calculated as this was a budget impact model.	• The cumulative cost of the ATRA+ATO regimen over 3 years (including pharmacy, medical, adverse event and disease event costs) was €60,000. • The cumulative cost of the AIDA regimen over 3 years (including pharmacy, medical, adverse event and disease event costs) was € 38,800. No discounting was applied.	• ICER not available. • Yearly budget impact of introducing ATRA+ATO calculated from the Italian healthcare perspective in the final (3rd) year of the model was €477,800
Lachaine, et al. 2015b[73]	2015	A Markov model simulating the course of relapsed/refractory APL patients receiving induction therapy followed by consolidation therapy if they achieved CR. The model considered 5 health states: induction, complete remission after relapse, treatment failure, post-	Patients with relapsed/refractory APL, diagnosed by bone marrow morphology. The mean age of the cohort was 45	Intervention was ATO (14.08 QALYs) and the comparator was ATRA+CT followed by HSCT (12.21 QALYs).	• Total cost of ATO treatment from the MoH perspective: 168,849 Can$ • Total cost of the ATO treatment from societal perspective: 196,848 Can$	• ICER, MoH perspective: $CAD 20,551 • ICER, societal perspective: $CAD 22,219

Study	Year	Summary of model	Patient population (average age in years)	QALYs (intervention, comparator)	Costs (currency) (intervention, comparator)	ICER (per QALY gained)
		treatment failure, and death. The length of each cycle was 1 month during the 24 month study period, and 1 year thereafter.	years.	The applicable discount rate was 5%.	• Total cost of ATRA + CT from the MoH perspective: 130,460 Can$ • Total cost of ATRA + CT from the societal perspective: 155,343 Can$ The applicable discount rate was 5%.	
Schonsteiner et al., 2017[74]	2017	A Markov model was constructed to access health care costs. The patient cohort included German participants of the APL0406 Trial treated with ATRA+ATO or AIDA. Patients received ATRA+ATO until complete remission followed by 4 cycles of consolidation therapy; whereas patients treated with AIDA received induction, consolidation and 2-year maintenance therapy. The model presented mean QALYs in treatment for event free survival and overall survival in both treatment arms.	Adult (≥18 and ≤70 years) newly-diagnosed APL patients with low- to intermediate-risk disease (WBC count at diagnosis ≤10x10^9/L)	Treatments included: •ATO+ATRA (mean QALYs for EFS: 1.55, mean QALYs for OS: 1.56) •AIDA (mean QALYs for EFS: 1.44, mean QALYs for OS: 1.53) Discount rate not available.	Mean total costs of treatment in ATO+ATRA group:* • Considering EFS: 89,038.80 • Considering OS: 91,846.32 Mean total costs of treatment in AIDA group:* • Considering EFS: 39,065.36 • Considering OS: 59,642.60 Discount rate not available.	• ICER for progression-free state 35,220* • ICER for the whole treatment 74,925*

Abbreviations: Ara-C=cytarabine; APL=Acute promyelocytic leukaemia; ATO=Arsenic trioxide; ATRA=All-trans retinoic acid; EFS=Event-free survival; ICER=Incremental cost-effectiveness ratio; MoH=Ministry of Health; OS=Overall survival; QALY=quality-adjusted life years; WBC=White Blood Cell
*Note that currency was not specified in this abstract; however, as the analysis applies to Germany one could sepeculate that the costs are expressed in €

B 3.2 Economic analysis

A new cost-effectiveness model was prepared for the purpose of this submission, in order to fully capture the impact of making ATO available to UK APL patients treated within the NHS. Previous cost-effectiveness analyses were conducted for the US[69, 70] and Canadian[71, 73] markets. The methodology and results of the relevant economic analyses conducted in other countries are summarised in ***section B 3.1*** and details are provided in ***Appendix G***. Briefly, in the US the ATRA+ATO regimen proved to be highly cost-effective compared to AIDA or ATRA + daunorubicin + cytarabine in the treatment of newly-diagnosed low- to intermediate-risk APL patients[69, 70], and a similar outcome was seen in Canada in this setting when ATRA+ATO was compared with AIDA[71]. Another Canadian study identified ATO as a cost-effective alternative to ATRA + conventional chemotherapy for the treatment of relapsed APL[73]. However, the applicability of US and Canadian findings to the UK healthcare setting is potentially low. An Italian economic analysis is also available[72], but focuses only on budget impact, without investigating cost-effectiveness. A study of cost-effectiveness of ATRA+ATO in Germany was also identified[74]; however, the data are scarce as only the conference abstract is accessible. Thus, conducting a de novo cost-effectiveness analysis was considered the most appropriate approach for the purpose of this submission. The methodological details of the US model[69, 70] were available to Teva; however, as this model differed from the relevant NICE guidance in several aspects, it did not inform the design of the current model.

B 3.2.1 Patient population

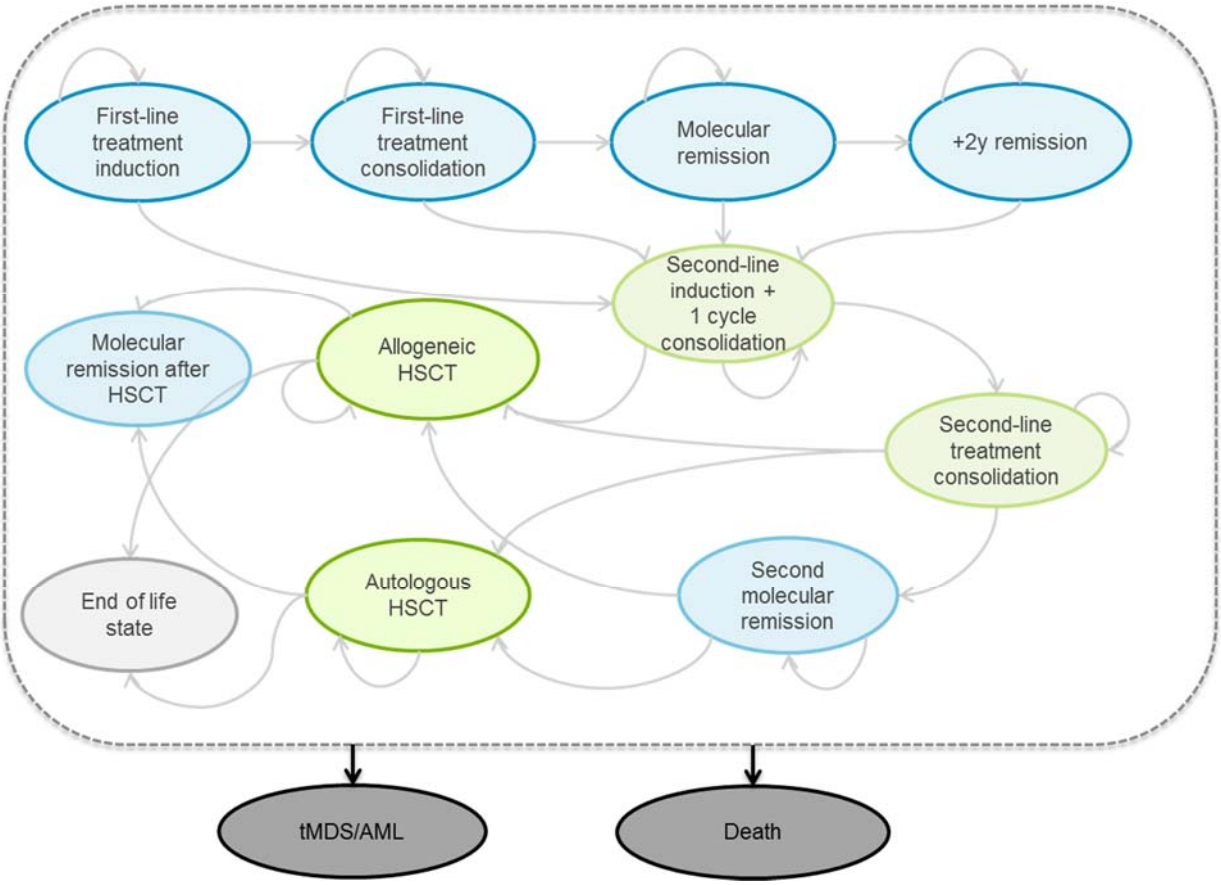
In line with the SPC[10] and the APL0406 trial[1, 3], the population of interest were newly diagnosed adult (≥18 years) patients with low- to intermediate-risk APL. With the approval of ATO in first line and the very low relapse rate observed in the relevant RCTs[2, 3], it is expected that clinical practice will shift towards the use of ATO as standard of care in newly-diagnosed patients. Therefore, the use of second line treatments (including ATO) will decrease and will be driven by the type of first-line treatment the patients received. Consequently, the relapsed/refractory APL population was not evaluated separately in the model, but rather analysed in relation to the newly-diagnosed population, so that the model provides an overall ICUR for ATO use in both first and second line. Thus, after their initial diagnosis, patients commenced first line therapy (AIDA or ATRA+ATO) and upon disease relapse or a critical adverse event they moved onto second line treatment.

B 3.2.2 Model structure

The analysis was performed using a Markov cohort model, programmed with a 4-week (Markov) cycle length to approximate the treatment schedule. Thirteen cycles represented a full year spent in the model.

The model included a total of 14 health states (***Figure 3.1***). Patients entered the model in "First-line treatment induction phase", during which they were hospitalised. This health state was compounded of three tunnels (sub-health states representing a period of 4 weeks spent in the health state), ensuring that patients could not remain in the first-line induction state for more than 12 weeks (3 cycles of 4 weeks). Cardiac events prompted a treatment switch and affected patients moved to the "Second-line induction + 1 cycle consolidation" health state.

Figure 3.1. Markov model structure with 14 health states



HSCT=haematopoietic stem cell transplant; tMDS/AML=treatment-related myelodysplastic syndrome or acute myeloid leukaemia; y=year

Patients moved from "First-line treatment induction" to "First-line treatment consolidation" at a rate that depended on the average (or median) time necessary to achieve CR. During consolidation, it was assumed that whatever treatment the patients received, they followed the whole consolidation course except if they

experienced a cardiac event leading to a treatment switch (which resulted in patients moving to the “Second-line induction” state). This health state consisted of ten tunnel states, allowing the consolidation phase to comprise up to five treatment cycles and assuring that patients remained in this phase for the right amount of time. Each treatment cycle included one 4-week model cycle on-treatment and one 4-week model cycle off-treatment. Patients in the induction and consolidation health states could experience an increased risk of mortality (from bleeding, infection, etc.), depending on the treatment they received and the treatment phase, although this increased mortality risk was not applied in the base-case analysis.

During the induction and consolidation phases, in addition to adverse events (some of which led to treatment switch), patients could experience tMDS/AML. An absorbent state was dedicated to these patients, in order to take into account their poor quality of life, high mortality risk and increased costs.

At the end of the consolidation phase, patients were either in molecular remission (MR) and moved to the “Molecular remission” state, or the treatment failed and they moved to the “Second line induction” state. As in the APL0406 trial[3], a small proportion of patients could not be evaluated for remission. Based on expert opinion (see ***Appendix M***), those patients also moved to the “Molecular remission” state. In other words, if a patient was not evaluable for molecular remission, we assumed that they did not switch treatment, as there was no evidence of treatment failure. Since the number of non-evaluable patients was very similar between the arms of the APL0406 trial (12 in the ATRA+ATO arm vs 13 in the AIDA arm)[3], this assumption does not favour ATRA+ATO.

The “Molecular remission” health state represented the first two years of remission during which patients could receive maintenance treatment, depending on the first-line regimen they received. This state consisted of 24 tunnel states, so that patients could not remain in it for more than 24 months. During these first 2 years of remission, the probability of relapse was higher than in the following years. In case of a relapse, patients moved to the “Second-line induction” state; otherwise, they went through all tunnel states and after 24 Markov cycles they moved to the “+2y remission” state.

Patients moved to “+2y remission” state after 2 years spent in molecular remission. It was assumed that patients in this state were at a lower risk of relapse than in the first two years of remission. In case of a relapse, patients moved to the “Second-line induction” state, otherwise, they remained in this health state until death.

Patients began second-line treatment in the “Second-line induction + 1 cycle consolidation” state. The following 3 events could lead patients to receive second-line treatment:

- failure of first-line treatment after consolidation cycles (i.e., lack of haematological or molecular remission),
- cardiac event requiring a treatment switch,
- haematological or molecular relapse.

The “Second-line induction + 1 cycle consolidation” health state grouped the induction phase, which was built in the same way as first-line induction (3 tunnel states to ensure a maximum of 90 days spent in this phase with one cycle of consolidation. The latter was modelled as 2 tunnel states allowing a treatment cycle of 4 weeks on treatment followed by 4 weeks off-treatment. Within this state, the monthly transition probability of transferring from the induction phase to the first consolidation cycle was based on the average time needed to achieve remission). This construction was selected based on expert opinion, ensuring that patients always followed at least one cycle of consolidation, regardless of whether they reached molecular remission at the end of the induction.

The outcome of treatment covered by the “Second-line induction + 1 cycle consolidation” state determined the health state into which patients transitioned next:

- No remission: all patients who did not achieve molecular remission after “Second line induction + 1 cycle consolidation” underwent allogeneic HSCT.
- Complete molecular remission: among those who achieved molecular remission at the end of the first consolidation cycle, some underwent allogeneic SCT, others underwent autologous SCT, and the remaining patients continued consolidation treatment. Further (maintenance) treatment for patients not undergoing a HSCT was not included in the model. According to expert opinion, allogeneic HSCT is very rarely used in UK patients who achieve a second CR, so we analysed a scenario where this proportion was equal to 0 (see section ***0***).
- Discontinuation/treatment stop: it was assumed that patients who experienced a cardiac adverse event discontinued their second-line treatment and underwent an allogeneic HSCT. This represents a limitation of the model as, according to UK expert opinion, patient pathway would vary depending on the treatment they received in second line, as the nature of cardiac events differs between modelled treatments. Patients experiencing an AIDA-induced cardiac event are unlikely to be able to undergo HSCT and would switch to ATRA+ATO, while patients having

received ATRA+ATO in second line would undergo autologous HSCT. Nevertheless, this was considered a minor limitation due to the small number of affected patients, and the expert agreed on the fact that it represented a conservative assumption as patients receiving AIDA in second line were only present in the ATRA+ATO arm. This means that simulating a switch from AIDA to HSCT generated higher costs and lower quality of life only in the ATRA+ATO arm. Furthermore, the expert confirmed that patients generally do not discontinue ATO in second line due to cardiac adverse events, which they were able to do in the model. Nevertheless, no patients experienced cardiac events in the ATRA+ATO arm of the APL0406 study[3] which was used to estimate the probability of these events in second line, assuming that the safety in second line would be the same as in first line.

After “Second-line induction + 1 cycle consolidation”, it was assumed that patients who did not undergo HSCT following the first consolidation cycle continued consolidation treatment until its completion, except if they experienced a cardiac event requiring treatment discontinuation. The “Second-line treatment consolidation” health state was programmed with ten tunnel states providing the possibility to consider five additional consolidation cycles of 4 weeks on treatment and 4 weeks off-treatment. Thus, up to a total of 6 consolidation cycles could be included in the model.

At the end of the consolidation phase, most patients underwent HSCT (allogeneic or autologous), while the remaining patients (i.e., those who could not receive a HSCT) moved directly to the “Second molecular remission” state with no further (maintenance) treatment. Patients could stay in the “Second molecular remission” health state until their death; nevertheless, they were at a non-negligible risk of relapse and could also undergo allogeneic or autologous HSCT.

Allogeneic and autologous HSCT were represented by distinct health states. The ”Allogeneic HSCT” health state represented the critical periods of hospitalisation and subsequent monitoring and was programmed as a 6-tunnel state, allowing the monitoring to last up to approximately six months. During this phase, patients were at risk of developing acute graft-versus-host disease (GvHD) and at a higher risk of mortality. At the end of the ”Allogeneic HSCT” state, patients who remained in molecular remission moved to the “Allogeneic HSCT remission” state, while the remaining patients transitioned to the “End of life” state.

After a successful allogeneic HSCT, patients remained in the absorbent “Allogeneic HSCT remission” state until their death. In order to account for the generated costs

and the lower QoL, patients in this health state had an increased mortality risk compared to the general population and were also at risk of developing chronic GvHD.

With shorter hospitalisation and monitoring periods compared to allogeneic HSCT, the autologous HSCT procedure was modelled through a 3-tunnel state allowing a maximum duration of approximately 3 months for the phase. As in the “Allogeneic HSCT” health state, a specific probability of death was applied in order to reflect the increased mortality risk over the HSCT period. At the end of the “Autologous HSCT” state, patients in molecular remission moved to the “Autologous HSCT remission” health state, and those for whom the transplant failed proceeded to the “End of life” state.

Compared to those having undergone an allogeneic SCT, patients in remission after an autologous HSCT incurred lower costs, and had a better quality of life and no risk of chronic GvHD. Nevertheless, they presented the same increased mortality risk until the end of their life. The “Autologous HSCT remission” health state was programmed as an absorbent health state, where patients remained until they died.

Patients for whom the HSCT procedure failed proceeded to the absorbent “End of life” health state. Patients in this state experienced a low QoL and required extensive palliative care, generating high costs. Furthermore, patients in this health state had a very high probability of death.

Similarly to first-line treatment, patients receiving second-line treatment could experience tMDS/AML or die. This absorbent “tMDS/AML” health state received patients having developed a therapy-induced myelodysplastic syndrome or acute myeloid leukaemia during one of the APL treatment phases (i.e., first- or second-line induction and consolidation). The possible treatment pathways were simplified in a single tMDS/AML health state in which the overall costs and QoL of patients with the condition were estimated.

In order to quantify mortality related to APL, the “Death” health state was divided into two compartments, one dedicated to APL-related deaths and the other to background mortality.

NICE has not appraised other technologies used in the treatment of APL, so that a comparison of the current model with earlier economic analyses was not feasible. ***Table 3.2*** compares the present analysis with the methods specified in the NICE reference case.

Table 3.2. Comparison of the economic analysis with the NICE reference case

Element of health technology assessment	Reference case	Comparison of current model and NICE reference case
Defining the decision problem	The scope developed by NICE	• Clear statement of the decision problem provided in ***section B.1*** • CE model reflects the scope of treatment in the pathway of care (first and second line APL) Selection of comparators for the first-line indication was in line with the scope of treatment; for discussion on second-line comparators see section ***B 1.1***
Comparator(s)	As listed in the scope developed by NICE	AIDA (see ***Table 1.1***)
Perspective on outcomes	All direct health effects, whether for patients or, when relevant, carers	As reference case
Perspective on costs	NHS and PSS	As reference case
Type of economic evaluation	Cost–utility analysis with fully incremental analysis	As reference case
Time horizon	Long enough to reflect all important differences in costs or outcomes between the technologies being compared	Life time
Synthesis of evidence on health effects	Based on systematic review	As reference case
Measuring and valuing health effects	Health effects should be expressed in QALYs. The EQ-5D is the preferred measure of health-related quality of life in adults.	QALYs – various source and assumptions due to rareness of APL
Source of data for measurement of health-related quality of life	Reported directly by patients and/or carers	Various source and assumptions due to rareness of APL
Source of preference data for valuation of changes in health-related quality of life	Representative sample of the UK population	Various source and assumptions due to rareness of APL
Equity considerations	An additional QALY has the same weight regardless of the other characteristics of the individuals receiving the health benefit	As reference case
Evidence on resource use and costs	Costs should relate to NHS and PSS resources and should be valued using the prices relevant to the NHS and PSS	As reference case
Discounting	The same annual rate for both costs and health effects (currently 3.5%)	As reference case

NICE=National Institute for Health and Care Excellence; NHS= National Health Service; PSS=personal social services; QALYs=quality-adjusted life years; EQ-5D=standardised instrument for use as a measure of health outcome.

B 3.2.3 Intervention technology and comparators

The intervention technology was ATO (combined with ATRA), which was compared with AIDA in newly-diagnosed APL; the same comparator was also retained for the second-line part of the model. In the newly-diagnosed setting, AIDA was the only comparator in the pivotal APL0406 trial[1, 3], and the primary comparator in the AML17 trial (which also used gemtuzumab ozogamicin in some patients, see ***section B 2.4***)[2]. However, the publication of many treatment guidelines in APL preceded the approval of ATO for first-line use, and treatment protocols for relapse following first-line ATO administration are still an area of discussion. Based on clinical expert opinion, the economic analysis assumed that patients who remained in remission for ≥2 years following first-line ATRA+ATO treatment were re-treated with ATRA+ATO upon relapse. However, patients who achieved only a short (<2 years) remission after first-line treatment with ATRA+ATO were assumed to be treated with AIDA upon relapse. Thus, AIDA was also considered as a comparator in the relapsed or refractory APL part of the model. Dr Dillon stated that there is no general agreement in the UK regarding the decision mode for second-line treatment; he explained that both approaches ("Chemo-free" and "With chemo") coexist in the UK. However, from his point of view, it is more likely that patients receive ATRA+ATO in second line. Finally, as this concerns very few patients the impact on results can be considered insignificant.

Similar to first-line treatment, in the relapsed/refractory setting ATO was used only in combination with ATRA. Based on clinical expert opinions from Professor Lo-Coco, Dr Cicconi and Dr Dillon, ATO alone (without ATRA) was confirmed to be rarely used in the relapsed/refractory setting in clinical practice. Thus, only the ATRA+ATO combination was considered in the base-case economic analysis, although the use of ATO alone in second line was modelled as a scenario.

B 3.3 Clinical parameters and variables

B 3.3.1 Outcomes

The outcomes related to the use of ATRA+ATO and AIDA in newly-diagnosed APL were mainly estimated based on the head-to-head APL0406 clinical trial (see ***section B 2.4.1*** for the list of outcomes that was used in the model). A scenario analysis was also conducted with the treatment schedule and outcomes from the AML17 clinical trial. Efficacy was estimated through remission rates and rate of relapse, and safety through the proportion of patients experiencing adverse events. However, no head-to-head data vs AIDA were available for second-line treatment

and, in addition to the RCT by Raffoux, et al. (see ***section B 2.4.2*** for the list of outcomes used in the model), the efficacy data in the second-line part of the model were informed by clinical expert opinions and a previous cost-effectiveness model developed for ATRA+ATO in the US[70], as the structure of this model was comparable. When available in the RCT from Raffoux et al., safety data reported therein were used, otherwise, the safety profile of second-line treatments was assumed to be the same as in first line, and data from the APL0406 RCT were used. Details of clinical effectiveness inputs used in the model are listed in ***Table 3.3***.

Table 3.3. Inputs related to clinical effectiveness and their sources

Input	ATRA+ATO	Source	AIDA	Source
Haematological CR rate (1st line)	98.45%	APL0406 (Platzbecker et al. 2016)[3]	96.35%	APL0406 (Platzbecker et al. 2016) [3]
Proportion of patients evaluable with PCR test	90.55%	APL0406 (Platzbecker et al. 2016) [3]	90.15%	APL0406 (Platzbecker et al. 2016) [3]
Molecular remission rate (among eligible) after first line	100%	APL0406 (Platzbecker et al. 2016) [3]	98.32%	APL0406 (Platzbecker et al. 2016) [3]
Molecular remission rate (among eligible) after second line induction + 1 cycle consolidation	80%	Complete remission rate in Raffoux et al. 2003[43]	70%	Expert opinion
Molecular remission rate (among eligible) after second line consolidation	100%	Assumption: patients in second line consolidation are already in molecular remission	100%	Assumption: patients in second line consolidation are already in molecular remission
Molecular remission rate (among eligible) after allogeneic HSCT	72.24%	Holter Chakrabarty et al. 2014[75]	72.24%	Holter Chakrabarty et al. 2014[75]
Molecular remission rate (among eligible) after autologous HSCT	98.11%	Holter Chakrabarty et al. 2014[75]	98.11%	Holter Chakrabarty et al. 2014[75]
Probability of relapse at 24 months for patients in first remission	0.009	APL0406 (Platzbecker et al. 2016) [3]	0.082	APL0406 (Platzbecker et al. 2016) [3]
Probability of relapse at 48 months for patients in first remission	0.019	APL0406 (Platzbecker et al. 2016) [3]	0.139	APL0406 (Platzbecker et al. 2016) [3]

Probability of relapse for patients in second remission	0.0110	Tallman et al. 2015[70]	0.0110	Conservative assumption: same as in the ATRA+ATO arm
Time to relapse for patients in second remission (in months)	24.5	Median time to relapse in APL0406 (Platzbecker et al. 2016) [3]	14	Median time to relapse in APL0406 (Platzbecker et al. 2016) [3]

Key assumptions related to clinical outcomes in the model include the following:

- In first line, patients who did not undergo PCR testing (thus, their molecular remission status could not be evaluated) were considered to be in molecular remission. This was based on the way these patients were treated in the APL0406 trial, i.e., they did not receive second-line treatment but instead continued until the end of the consolidation courses.
- In second line, haematological CR rate for ATRA+ATO was used instead of molecular remission. This assumption was made due to the lack of clear data on molecular remission in second line, and was validated by experts.
- The rate of molecular remission following second-line treatment with AIDA was not available in the literature and was based on expert opinion.
- Rates of molecular remission after HSCT were extracted from a publication selected by clinical experts[75]. The rates were read from the DFS Kaplan-Mayer curves at the time point corresponding to the end of the hospitalisation period (3 months for autologous HSCT, 6 months for allogeneic HSCT).
- For patients in first molecular remission, the probability of relapse at 48 months was assumed to be equal to that at 50 months. This assumption related to both treatment arms and was validated by experts
- Probability of relapse in patients who were in second remission was taken from the cost-effectiveness analysis published by Tallman et al[70]. As the publication reported only the probability of a second event following ATRA+ATO use in second line, in order for the model to remain conservative it was assumed that the probability of relapse following second-line treatment with AIDA would be the same.
- The average time to relapse for patients in second remission was assumed to be equivalent to that observed among patients in first remission. This assumption was validated by experts (for more details on expert see ***section B 3.3.5***).

B 3.3.2 Extrapolation of clinical data and costs beyond the trial follow-up period

Methods related to data extrapolation beyond the follow-up period of relevant clinical trials (APL0406, AML17 and Raffoux, et al.) are described below. For details of internal and external validation of the model see ***section B 3.10***.

The model assumed that the rates of relapse were different during the first two years of remission and thereafter (years 2+). We assumed constant rates of relapse between remission and two years post-remission, and after two years from remission, which was confirmed by clinical experts.

Costs per cycle were assumed to remain constant over time in the health states. This could be important for those health states in which patients are likely to stay for a long time, which include: +2years remission, molecular remission after second line, molecular remission after allo-HSCT, molecular remission after auto-HSCT, end of life state and tMDS/AML.

In the UK, patients are followed only for the first three years after molecular remission, this means that patients in the +2y remission state incurred only costs of follow-up and monitoring for one year. Therefore, the costs of these three years of intensive monitoring were spread in the tunnels of the first two years in remission, in order to avoid this cost being applied for the patient's lifetime. Nevertheless, we assumed patients who achieved remission following second-line treatment incurred a constant lifetime cost.

Molecular remission after allogeneic or autologous HSCT involved costs of follow-up and monitoring, including those of long-term follow-up. Due to the need for long-term monitoring and the risk of chronic GvHD following allogeneic HSCT, this assumption seemed to be reasonable. Nevertheless, DSA showed that ICER was not dependent on the costs of remission following allo- or auto-HSCT.

Regarding tMDS/AML, these conditions usually last for the rest of the patients' life. Therefore, assuming that costs associated with this state are constant over time appears justified. The end of life health state was associated with costs for palliative care, which were considered lifetime. The high mortality rate of patients in this health state reflected their short life expectancy.

B 3.3.3 Transition probabilities

B 3.3.3.1 Time-dependent probabilities

As the model was programmed with tunnel states, in order to respect the time spent in the different health states (especially the duration of treatment phases), time-dependent probabilities were widely used. Indeed, the probabilities had to be computed over a time frame in an equation that incorporated the cycle length of the model. “Per cycle” probabilities were calculated using the following formula: $\mathrm{P}_{\mathrm{t1}}(\mathrm{A}) = 1 - (1 - \mathrm{P}_{\mathrm{t}}(\mathrm{A}))^{\mathrm{t1}/\mathrm{t}}$ With $\mathrm{P}_{\mathrm{t}}(\mathrm{A})$ being the probability of the event A measured at the time point t, and $\mathrm{t1}$ representing the time point at which the probability is used in the model (i.e, 4 weeks or 28 days in this case).

B 3.3.3.2 Treatment schedule

The treatment schedule determined the rates at which patients went through the treatment pathway, more precisely, it determined the way they moved between the induction and consolidation health states (first- and second-line).

In line with the data presented in ***section B.2***, the schedule for each treatment phase was as follows:

- Induction:
  - Average duration was based on median time to complete haematological CR observed in the clinical trials (APL0406[3] or AML17[2], depending on the scenario). Four-week probability to achieve CR was computed assuming an exponential distribution (see ***section B 3.3.3.1*** for the formula) and considering a probability of 50% of patients achieving CR in the average duration provided to the model. Following each cycle, patients achieving remission moved to the consolidation phase, while the remainder stayed in the induction health state until they achieved molecular remission or until the maximum duration of the phase.
  - The maximum duration of the induction phase was based on the treatment protocol described in the SPC for ATO[10] (or the AML17 trial[2] in the corresponding scenario) and determined the number of tunnels used for the induction phase. If patients did not achieve remission in this period of time, they were assumed to transition to the consolidation phase.
- Consolidation:
  - Patients went through the full consolidation phase. The total duration of the consolidation phase was calculated based on the trial parameters (APL0406[3] or

AML17[2], depending on the scenario) listed below, and used to determine the number of tunnels needed to model the phase.

- The number of cycles of ATRA and ATO (for the ATRA+ATO strategy)
- Duration of on-treatment period in a cycle
- Duration of the off-treatment period in a cycle.
- For ATRA+ATO, the duration of the consolidation phase was computed for each drug and the maximum duration was used in the model.

- Maintenance:
  - The maintenance phase was available in the model, since in most EU countries the maintenance phase follows consolidation in chemotherapy regimens. However, since maintenance is rarely used in UK clinical practice, all values related to costs and resource use in the maintenance phase were set to zero in the model for the UK setting.

B 3.3.3.3 Treatment efficacy

The efficacy of each treatment strategy was considered through the following parameters:

- Haematological CR rate in first line, expressed among all patients receiving the treatment.
- Complete molecular remission rate represented the main treatment outcome, as it determined if the patients were considered “cured” (i.e., long-term survivors) or not. Complete molecular remission was assessed in the following manner after each treatment phase:
  - First line: evaluated at the end of the consolidation phase and expressed among evaluable patients.
  - Second-line induction + 1 cycle consolidation: expressed among all patients receiving the strategy in second line.
  - Second-line consolidation: the molecular remission rate was fixed to one, considering that patients were already in molecular remission if they continued into the consolidation phase. As specified in ***section B 3.2.2***, patients who did not achieve molecular remission at the end of the first consolidation cycle underwent allogeneic transplantation.
  - After allogeneic and autologous HSCT: the molecular remission rate was derived from DFS observed at the end of the monitoring period following HSCT (i.e. six months for allogeneic and three months for autologous transplantation).

Data were extracted from the Kaplan-Meyer curves published by Holter Chakrabarty et al.[75], as advised by clinical experts.

- The proportion of patients evaluable for molecular remission varied in the different clinical trials. This has an impact on the complete molecular remission rate when calculated among the total number of patients receiving the treatment (e.g., following second line induction and first consolidation cycle).
- Probability of relapse: alongside complete molecular remission rate, this was another important parameter differentiating the modelled treatment strategies. At different time points of the patient pathway, the probability of relapse was considered in the following manner:
  - The probability of relapse for patients in first remission was calculated separately at 24 months and at 48 months, to account for the different probabilities of relapse during the first two years of remission and after 24 months. The probability of relapse after 24 months was computed as the proportion of patients relapsing between 24 and 48 months: $\mathrm{P}(\mathrm{relapse} > 2\mathrm{y}) = \mathrm{P}(\mathrm{relapse}\ 48\mathrm{m}) - \mathrm{P}(\mathrm{relapse}\ 24\mathrm{m})$.
  - For ATRA+ATO, the probability of relapse at 24 months determined which second-line treatment was administered to patients who relapsed. Based on expert opinion, patients who remained in remission for at least 24 months received ATRA+ATO in second line, while those who relapsed after less than 24 months from achieving remission received AIDA. The clinical expert stated that there is no consensus on this approach in the UK, but thought it was a reasonable assumption considering the small number of patients relapsing after having received ATRA+ATO in first line.
  - Following second line remission, the time point (expressed in months) at which the probability of relapse was evaluated had to be supplied to the model in order to compute the “per cycle” probability.

B 3.3.3.4 Haematopoietic stem cell transplant

The probabilities of undergoing an autologous or allogeneic HSCT procedure were estimated at the following steps of the patient pathway:

- After second-line induction and the first consolidation cycle for patients achieving complete molecular remission; all patients who did not achieve molecular remission proceeded to allogeneic HSCT.
- At the end of the second-line consolidation phase.
- In the “second-line remission” health state.

In order to determine the number of necessary tunnels ensuring that the patients stay in the HSCT health states for the correct time period, the duration (in months) of the monitoring phase following HSCT was fixed for both types of transplant (see ***section B 3.2.2***).

B 3.3.3.5 Safety

The only safety parameter taken into account when calculating transition probabilities was the rate of cardiac adverse events related to treatment, as they generally reflect intolerance to ATO or cardiotoxicity of the chemotherapy used in APL, requiring a treatment switch according to the opinions of Professor Lo-Coco, Dr Cicconi and Dr Dillon (see ***section B 3.3.5*** for details on experts). Patients experiencing such an event were moved to the subsequent treatment line (i.e., from first line to second line and from second line to allo-HSCT, see ***section B 3.2.2***).

B 3.3.4 Mortality

Two types of mortality were considered in the model:

- Background mortality identified using UK death rates by age and gender from the life tables published by the Office for National Statistics[76], which reflected non-APL mortality. For each age, an average "per cycle" death probability was computed, taking into account the gender ratio in the population of interest. In each cycle, the appropriate death probability was applied to each health state, according to the current age of the patients.
- Disease-related mortality was incorporated in the model by applying different death probabilities in the different health states of the patient pathway:
  - For the full duration of allogeneic and autologous HSCT states, specific deaths probabilities were applied representing the increased risk of death during the HSCT procedure and the subsequent monitoring phase.
  - Death probability following HSCT failure applied in the "End of life" health state represented the short life expectancy of patients receiving palliative care.
  - The risk of fatal, often delayed, complications following HSCT was incorporated as an increased probability of dying in the HSCT remission states and the "End of life" health state.
  - Patients in the tMDS/AML health state had an increased probability of death, representing the relatively high mortality risk of patients who develop tMDS/AML.

In addition to these mortality-related probabilities, the proportion of patients with fatal events occurring during the induction and consolidation treatment phases can be

considered in the model. However, as the difference in on-treatment mortality between arms was not tested in the APL0406 trial, we assumed zero deaths during the induction and consolidation treatment phases. As the number of deaths was higher in the AIDA arm than the ATRA+ATO arm of the APL0406 trial, this assumption can be considered conservative. No scenario analysis was conducted using the proportion of deaths observed during treatment phases in the two arms of the APL0406 trial as it would lead to better cost-effectiveness results for ATRA+ATO which would not be so informative regarding the uncertainty.

B 3.3.5 Sources of clinical expert opinion

Professor Francesco Lo-Coco (one of the APL0406 study coordinators) and Dr Laura Cicconi, both from Tor Vergata University, Rome, Italy, were involved in the development of the global model, providing insights on patient management and validating assumptions and inputs. Dr Richard Dillon reviewed, amended and validated the structure, assumptions and inputs of the model to ensure that the version submitted to NICE is relevant to UK clinical practice. Dr Dillon is a Consultant Haematologist at Guy's Hospital and a Senior Lecturer in Cancer Genetics at King's College London. Details of the expert selection process and the methods of collecting expert opinion are provided in ***Appendix M***.

B 3.4 Measurement and valuation of health effects

As APL may have a substantial effect on life expectancy, potentially reducing it to as little as few days from diagnosis if left untreated, the economic analysis considered both quality-adjusted life years (QALYs) and cost-effectiveness (LY).

B 3.4.1 Life years

The number of life years (LYs) that patients receiving each treatment accumulated over the time horizon of the model was considered an important measure, due to both the detrimental effect untreated APL has on patients' life expectancy, and the differences in efficacy and safety between treatments that may influence how long patients survive. LYs were calculated simply by counting the proportion of patients still alive (i.e,, present in every health state except death) over the model time horizon. The proportions of patients alive per cycle and per health state were summed for every year across health states to get the total number of patients alive per year. Subsequently, the total numbers of patients alive per year were summed over the model time horizon to get the total number of LYs.

B 3.4.2 Quality of life

The disease and its different treatment phases had a strong impact on patients' QoL. Hence, the most important health outcome measured was the number of QALYs accumulated over the time horizon with each treatment strategy. QALYs were measured by relating the QoL of patients in the different health sates (expressed in terms of utility) to the time that patients spent in these states.

B 3.4.3 Remissions following first-line treatment

Amongst the outcomes of interest reported in the model were the proportion of patients achieving molecular remission after first-line treatment and the proportion of patients reaching long-term (>2 years) remission. These outcomes were estimated with the number of new patients achieving remission (i.e., new patients entering the remission health state) accumulated over the time horizon.

B 3.4.4 MDS

The proportion of patients experiencing tMDS/AML was reported for each treatment strategy through the cumulative number of "new cases" (i.e. new patients entering the tMDS/AML health state) over the time horizon.

B 3.4.5 Discounting

A discount rate was applied to health outcomes reported in the model, according to the following formula:

$Discounted\ O_n = \frac{O_n}{(1+r)^{n-1}}$ with O_n being the value of the outcome for the year n, and r the discount rate.

In order to estimate the impact of the discounting, both discounted and not discounted total values measured over the time horizon were reported.

B 3.4.6 Health-related quality-of-life data from clinical trials

No clinical trials reporting QoL outcomes measured in APL patients using the EQ-5D were identified. While both the APL0406 and the AML17 trials investigated QoL, they used the EORTC QLQ-C30 instrument and not the EQ-5D. Details of QoL results from both studies are presented in sections 0 (APL0406) and **Error! Reference source not found.** (AML17).

B 3.4.7 Health-related quality-of-life studies

Our systematic literature search revealed two studies reporting utilities associated with the treatment of APL; both of these are cost-effectiveness analyses already described in ***section B 3.1*** and ***Appendix G***. In terms of the reported utilities, in the study by Tallman, et al. these were based on studies in chronic lymphocytic leukaemia and adjusted for age and country, as APL-specific utilities were not available[70]. The Tallman, et al. model included the following utilities: first line stable disease=0.78, first line disease event/second line stable disease=0.65, second line disease event=0.47[70]. The other study identified in the systematic literature search was the Canadian cost-effectiveness analysis from Lachaine, et al. which considered the following health states: event-free survival (base-case utility of 0.9, lower bound [LB]=0.8, upper bound [UB]=1.0), treatment failure, i.e., relapse (base-case utility of 0.5, LB=0.2, UB=0.8), and post-treatment failure, i.e., second CR (base-case utility of 0.8, LB=0.4, UB=0.95)[71]. Utilities for these health states were based on an earlier study in adult AML patients, as QoL data in APL was deemed insufficient[71]. The study also reported disutilities associated with the following adverse events: neutropenia (base-case utility of -0.135, LB= -0.3, UB= -0.09), thrombocytopenia (base-case utility of -0.095, LB= -0.108, UB= -0.081), fever episodes (base-case utility of -0.088, LB= -0.195, UB= 0), hepatotoxicity (base-case utility of 0, LB= -0.136, UB= 0) and QTc prolongation (base-case utility of 0, LB= -0.136, UB= 0)[71]. The authors specified that these disutilities were based on a literature review of published utility values associated with cancer treatment-related AEs[71]. As both of the aforementioned studies presented utilities which were based on conditions other than APL[70, 71] and the APL0406 and AML17 trials presented QoL data collected from APL patients, the results of the studies identified in the literature cannot be readily compared with QoL results presented in ***sections 0*** (APL0406) and *Error! Reference source not found.* (AML17)

B 3.4.8 Adverse reactions

The tolerability and safety aspect of each treatment strategy was considered through the proportion of patients experiencing adverse events and the duration of these events. The following treatment-induced adverse events were considered in the model:

- Thrombocytopenia (grade 3–4, duration >15 days)
- Neutropenia (grade 3–4, duration >15 days)
- Infection

- Leukocytosis
- Hepatic toxicity
- Neurotoxicity
- Differentiation syndrome
- Cardiac events
- QTc prolongation

For each event and each treatment phase (induction and consolidation), the rate per cycle was computed form the proportions of patients experiencing the event and the duration of the treatment phase, in order to determine the proportion of patients experiencing adverse events for each cycle

Except for cardiac events, the adverse events listed above did not lead to a change of treatment, but impacted only the costs and patients' QoL. Duration of each adverse event was used to compute the QALYs lost due to the QoL impairment in patients experiencing the event.

In addition to the adverse events listed above, therapy-induced MDS or AML were also taken into consideration by including the probability of developing the disease during each treatment phase. As tMDS/AML are life-threatening conditions requiring specific treatment protocols, patients experiencing tMDS/AML were moved into a new absorbing health state, characterised by low QoL.

The model was able to consider fatal events occurring during treatment through mortality rates applied during the induction and consolidation phases. As a conservative assumption (see ***section B 3.2.2*** for a detailed explanation) the base case did not consider these probabilities. A scenario analysis was conducted considering the mortality rates observed during treatment in the APL0406 trial, with patients dying during treatment moving to the "APL related death" health state.

B 3.4.9 Health-related quality-of-life data used in the cost-effectiveness analysis

We were not able to identify utility values that were specific to APL from the systematic search. Similarly, previous cost-effectiveness studies in APL[70, 71] used proxy utilities for other conditions that the authors considered to be associated with utilities analogous to APL. We used a similar approach and extracted utility values for each health state from the literature; given the small indication and limited data, the utility values had to be sourced from another leukaemia field (chronic lymphocytic leukaemia, CLL) and adjusted to APL based on age differences. For

some health states, the population in which the utility was measured did not present the same characteristics (age, utility of reference, etc.) as the population relevant to this submission. In these cases, the values were adjusted to reflect the modelled population. Adjustment factors were calculated as the ratio of the utility in the general population sharing the same characteristic as the modelled population to the utility in the general population with the same characteristic as the population in which the utility was assessed. For example, if a utility of 0.91 was reported in a population aged 60 (with a utility of 1 representing perfect health), two adjustment factors were required:

- Adjustment for age: the average age of the modelled population was 45 years. The utility of the general population aged by 45 is 0.849 and that of the general population aged 60 years is 0.804. The adjustment factor was calculated to be $0.849/0.804 = 1.056$
- Adjustment for the utility representing perfect health: utility in the general population aged 60 years is 0.804. The adjustment factor was therefore $0.804/1 = 0.804$

Thus, the adjusted utility was $0.91 \times 1.056 \times 0.804 = 0.773$

The main publication we used was from Woods et al., 2012[77] as it presented utility values for similar health states to those in our model, reflecting the treatment pathway (e.g. treatment phases, remissions/stable disease) and with a logical ranking of the values between health states, which we were able to keep after adjustment. Furthermore, this publication referred to Beusterien et al., 2010[78] which proposed useful utility values. Utility norms were extracted from Szende et al., 2014[79], which is the publication of reference for this type of data. The sources we used, the process for estimating utility values, as well as the computed values were submitted to the experts and validated by all three of them (see **section** Error! Reference source not found. for details on experts).

The lower QoL of patients experiencing AEs was taken into account through specific disutilities applied for the duration of each AE. Disutility values and AE durations were collected from the literature. Acute and chronic GvHD were considered as an AE related to allogeneic HSCT, and a disutility was applied to the proportion of patients experiencing the condition. In addition, to reflect the lower QoL of hospitalised patients, a disutility was applied during hospitalisation periods. The total disutility was computed following each cycle and in each health state. For each AE, the disutility was multiplied by the proportion of patients affected and the duration of the event; the disutilities per event were subsequently summed to get the total

disutility per cycle for each health state. In each cycle, the disutilities were subtracted from the utility of the health state and the result was multiplied by the proportion of patients experiencing adverse events within that heath state. These values were summed over 13 cycles to get the number of QALYs per year. The total number of QALYs accumulated with each treatment strategy over the time horizon was reported. Utilities used in the model are summarised in ***Table 3.4***.

Table 3.4. Summary of utility values for cost-effectiveness analysis

State	Mean utility value	95% confidence interval	Reference	Justification (comment)
First-line induction treatment	0.739	(0.708–0.771)	Woods et al., 2012[77] Szende et al., 2014[79]	• Baseline utility, used for both strategies during active treatment in Woods et al., 2012, mean (95% CI): **0.7** (0.67, 0.73). • General UK population at the age of 45 and 60 has a mean utility of 0.849 and 0.804 respectively. • Age adjustment factor was calculated based on UK general population utility at the average APL population age (45), divided by the general population utility for average CLL population age (60). Hence the adjustment factor was: 0.849/0.804=**1.056** • The adjusted utility, mean (95% CI) is: 0.7 (0.67, 0.73) × 1.056= **0.739 (0.708, 0.771)** • Disutility of hospitalisation was applied in the model to this value.
First-line consolidation treatment	0.739	(0.708–0.771)	Woods et al., 2012[77] Szende et al., 2014[79]	• Utilities were calculated in the same way as for first-line induction; however, since patients were not hospitalised for the full duration of the consolidation phase, the disutility of hospitalisation was not applied to this value and only disutilities due to adverse events were considered.
First molecular remission	0.773	(0.747–0.790)	Beusterien et al., 2010[78] Szende et al., 2014[79]	• Mean utility value for CLL remission was reported at 0.91 (95%CI: 0.88, 0.93). • This value is rather high, because of a reference point for full health at 1. We adjusted this based on the average utility of the general population at the average age of the CLL population, which is 60 years (utility value: 0.804). • Same as for first-line induction, age adjustment was performed to reflect the age of APL patients (adjustment factor: 1.056) • Thus, the utility was calculated as: mean (95%CI): 0.91(0.88, 0.93)×0.804×1.056=**0.773**(0.747; 0.790)
First long-term molecular remission	0.849	NR	Szende et al., 2014[79]	• Utility value was assumed to be the same as for the general population aged 45: 0.849.

State	Mean utility value	95% confidence interval	Reference	Justification (comment)
(>2 years)				
Second-line induction + 1 cycle consolidation	0.673	(0.644–0.702)	Woods et al., 2012[77] Beusterien et al., 2010[78]	• First line utility (based on CLL and used for both strategies during active treatment) reported in Woods et al., 2012, mean (95% CI): **0.7** (0.67, 0.73). • The starting utility value shown above is equal to first-line treatment utility. This was adjusted for second-line treatment using the following values: • CLL stable disease (i.e., symptoms not worsening or improving), mean (95% CI): 0.78 (0.75, 0.82). • Stable CLL + second line treatment (assumed to be post-second line), mean (95% CI): 0.71, (0.68, 0.75). • First- to second-line adjustment factor: 0.71/0.78=**0.91** (i.e., patients after second-line treatment have a utility value that is 91% of that in patients with the same status after first-line treatment). • Same age adjustment factor as first line: 0.849/0.804=**1.056** • Based on the adjustments specified above, the utility value was mean (95% CI): 0.7(0.67, 0.73)×0.91×1.056= **0.673 (0.64, 0.702)**
Second-line treatment consolidation	0.702	NR	Beusterien et al., 2010[78]	• Patients in this state were assumed to be in molecular remission; thus, we used the first-line molecular remission utility of **0.773** as the basis for further calculation, adjusting the utility value for second-line treatment as follows: • No change: CLL- stable disease (i.e., symptoms not worsening or improving): 0.78 (0.75, 0.82). • No change CLL +second line treatment: 0.71 (0.68, 0.75). • First • Same first- to second line adjustment factor as computed above: 0.71/0.78=**0.91** (i.e., patients after second line treatment have a utility value that is 91% of that in patients with the same status after first line treatment). • Based on the aforementioned adjustment, the utility value was:

State	Mean utility value	95% confidence interval	Reference	Justification (comment)
				0.773×0.91=**0.702**
Second molecular remission	0.849	NR	Szende et al., 2014[79]	• Based on expert opinion, the utility was assumed to be the same as for the general population at the age of 45.
Allogeneic HSCT	0.687	NR	Breitscheidel L., 2008[80]	• Mean unadjusted* utility weight in CML after HSCT without GvHD was 0.979. This was applied to the utility in the "second-line molecular remission" state: 0.979×0.702= **0.687** • Disutilities of hospitalisation and possible GvHD were applied to this value.
Autologous HSCT	0.687	NR	Breitscheidel L., 2008[80]	• The utility value was the same as for allogeneic HSCT, with the exception of only disutilities for hospitalisation (and not GvHD) being applied.
Allogeneic HSCT molecular remission	0.849	NR	Szende et al., 2014[79]	• Based on expert opinion, the utility was assumed to be the same as for the general population at the age of 45. • Disutility for chronic GvHD was applied to this value
Autologous HSCT molecular remission	0.849	NR	Szende et al., 2014[79]	• Based on expert opinion, the utility was assumed to be the same as for the general population at the age of 45.
End of life state (Palliative care)	0.4	NR	Morton et al., 2009[81]	• Palliative care for patients with a malignancy
tMDS/AML	0.4	NR	Cooperberg et al., 2013[82]	• Secondary malignancy following treatment for prostate cancer
Death	0	NR	NR	
Hospitalisation	-0.01	NR	Assumption	
Thrombocytopenia (grade 3-4, >15 days)	-0.18	NR	Attard et al., 2014[83]Attard et al., 2014[83]	
	Induc.:	NR	Wolff et al.,	

State	Mean utility value	95% confidence interval	Reference	Justification (comment)
	20 days Cons.: 25 days		1989	
Neutropenia (grade 3-4, >15 days)	-0.18	NR	Attard et al., 2014[83]Attard et al., 2014[83]	
	19 days		Fenaux et al., 1993	Assumed to be the same as for ATRA+DNR+ARA-C
Infection	-0.15	NR	Stevenson et al., 2014[84]	Based on table A1 in Platzbecker 2016[3], most infections are pneumonia. Disutility of pneumonia was considered.
	17 days		Pneumonia – What happens (http://www.webmd.com/lung/tc/pneumonia-what-happens)	Based on table A1 in Platzbecker et al., 2016[3], most infections are pneumonia, thus duration of pneumonia (2-3 weeks) was considered.
Leukocytosis	-0.08	NR	Assumption	
	14 days		Shoenfeld et al., 1981	
Hepatic toxicity	-0.20	NR	Choi et al., 2013[85]	
	10 days		Zhu et al., 2013	“Less than two weeks”
Neurotoxicity	-0.21	NR	Prica et al., 2014[86]	

State	Mean utility value	95% confidence interval	Reference	Justification (comment)
	365 days		Assumption based on Ratnaike, 2003	"Acute poisoning from arsenic can lead to peripheral neuropathy which can last for max 2 years"
Differentiation syndrome	-0.12	NR	Assumption	
	4 days		Breccia et al., 2008	Assumed to be same as AIDA
Cardiac events	-0.16	NR	Nshimyumukiza et al., 2013[87]	• Myocardial infarction (MI)
	1 day		Mathews et al., 2002	Assumed to be same as ATRA+ATO
QTc prolongation	-0.001	NR	Assumption	
	0.5		Siu et al., 2006	Assumed to be same as ATRA+ATO
Acute GvHD	-0.08	NR	Breitscheidel L., 2008[80]	• Mean utility weight after HSCT without GvHD, re-scaled: 0.836 • Mean utility weight after HSCT with GvHD, re-scaled: 0.769 • Disutility of GvHD: 1-(0.769/0.836)= 0.080 • Applied for the duration of the monitoring phase for the proportion of patients experiencing acute GvHD
Chronic GvHD	-0.08	NR	Breitscheidel L., 2008[80]	Assumed to be the same as acute GvHD Applied for a lifetime for the proportion of patients experiencing chronic GvHD

APL=Acute promyelocytic leukaemia; HSCT= Haematopoieticstem cell transplantation CLL=Chronic lymphocytic leukaemia; CML=Chronic myeloid leukaemia; GvHD=Graft-versus-host disease; LB=Lower bound; NR, not reported; tMDS/AML=Treatment-related myelodysplastic syndrome or acute myeloid leukaemia; UB=Upper bound; Induc.: induction; Cons.: consolidation.

* The study multiplied the utility value estimated by physicians (0.979) by 0.854 to reflect the Euro-QoL baseline: 0.979×0.854=0.836. However, the utility value was already adjusted in second-line molecular remission using the same approach; hence the utility weight of 0.979 was used to avoid double re-scaling

B 3.5 Cost and healthcare resource use identification, measurement and valuation

The model included direct costs estimated from the NHS and PSS perspective. The following cost categories were included: treatment acquisition costs, medical costs (treatment administration, supportive care, monitoring and follow-up, HSCT, palliative care), and costs of managing adverse events. For each item, the cost per cycle was computed and applied to the proportion of patients present in the health states and concerned by the type of cost. Costs were discounted at a rate of 3.5% per year. Note, that although a systematic literature review was performed to collect appropriate costs and resource use data for England, none of the 11 studies identified by the SLR was actually used in the present economic model. The decision not to include data from the SLR was in some cases driven by the fact that the information captured was not compatible with that needed to populate the model, and in others by the fact NHS reference costs were preferentially used to ensure relevance to the current situation in England.

B 3.5.1 Intervention and comparators' costs and resource use

Data on the dosage and number of doses of intervention and comparator were extracted from the publications of relevant clinical trials. Drug costs were based on the British National Formulary (BNF)[88]. Except ATO, for which the approved treatment schedules in first and second line include a different number of doses, the same dosage and number of doses were considered for first- and second-line treatment. The number of doses and the dosage were validated by the clinical experts (see **section** Error! Reference source not found. for details on experts). When several container sizes were available on the market, the size minimising the costs and wastage was used to reflect real-life practice. The input values and their sources are presented in ***Table 3.5, Table*** 3.6 ***and Table*** 3.7 for each drug and treatment phase.

Table 3.5 Unit treatment acquisition costs associated with the technologies studied in the economic model – Induction phase

Model parameter	Strategy	Drug	Value	Reference
Number of doses	ATRA+ATO First Line	ATRA	32	Lo-Coco et al., 2013[1]
		ATO	32	
	ATRA+ATO Second Line	ATRA	25	Douer et al., 2005[89]
		ATO	25	
	AIDA	ATRA	35	Lo-Coco et al., 2013[1]

Model parameter	Strategy	Drug	Value	Reference
		IDA	4	
Indicated dose per day	ATRA+ATO	ATRA	45 mg/m^2	Lo-Coco et al., 2013[1]
		ATO	0.15 mg/kg	
	AIDA	ATRA	45 mg/m^2	
		IDA	12 mg/m^2	
Container size	ATRA+ATO	ATRA	10 mg	BNF[88]
		ATO	10 mg	
	AIDA	ATRA	10 mg	
		IDA	10 mg	
Cost per container	ATRA+ATO	ATRA	£1.61	BNF[88]
		ATO	£292.00	
	AIDA	ATRA	£1.61	
		IDA	£174.72	

ATRA=all-trans retinoic acid; ATO=arsenic trioxide; BNF=British National Formulary; IDA=idarubicin

Container size and costs were not reported for the consolidation phase for drugs already presented in ***Table 3.5***.

Table 3.6. Unit treatment acquisition costs associated with the technologies studied in the economic model – Consolidation phase

Model parameter	Intervention	Drug	Value	Reference
Number of doses	ATRA+ATO First Line	ATRA	15	Lo-Coco et al., 2013[1]
		ATO	20	
	ATRA+ATO Second Line	ATRA	15	Lo-Coco et al., 2013[1]
		ATO	25	SPC[10]
	AIDA	ATRA	15	Lo-Coco et al., 2013[1]
		IDA (cycle 1)	4	
		Mitoxantrone (cycle 2)	5	
		IDA (cycle 3)	1	
Indicated dose per day	ATRA+ATO	ATRA	45 mg/m^2	Lo-Coco et al., 2013[1]
		ATO	0.15 mg/kg	
	AIDA	ATRA	45 mg/m^2	
		IDA (cycle 1)	5 mg/m^2	
		Mitoxantrone (cycle 2)	10 mg/m^2	
		IDA (cycle 3)	12 mg/m^2	
Container size	AIDA	Mitoxantrone	20 mg	BNF[88]
Cost per container	AIDA	Mitoxantrone	£100.00	BNF[88]

ATRA=all-trans retinoic acid; ATO=arsenic trioxide; BNF=British National Formulary; IDA=idarubicin

Table 3.7.Costs of technology per treatment phase

Phase	*ATRA+ATO*			*AIDA*		
	ATRA	**ATO**	**Total ATRA+ATO**	**ATRA**	**Chemo (IDA+MTZ)**	**Total AIDA**
First line: Induction	£463.68	£16,078.58	£16,542.26	£507.15	£2,096.64	£2,603.79
First line: Consolidation	£1,521.45	£40,196.44	£41,717.89	£652.05	£1,723.04	£2,375.09
First line: Total	£1,985.13	£56,275.02	£58,260.15	£1,159.20	£3,819.68	£4,978.88
Second line: Induction	£362.25	£12,561.39	£12,923.64	£507.15	£2,096.64	£2,603.79
Second line: Consolidation	£1,521.45	£12,561.39	£14,082.84	£652.05	£1,723.04	£2,375.09
Second line: Total	£1,883.70	£25,122.77	£27,006.47	£1,159.20	£3,819.68	£4,978.88

ATRA=all-trans retinoic acid; ATO=arsenic trioxide; IDA=idarubicin; MTZ=Mitoxantrone

Medical costs and resource use associated with each technology are presented in ***Table 3.8*** and ***Table*** 3.9, respectively.

Table 3.8. Unit medical costs associated with the technology in the economic model

Category	Items	Value	Reference
Administration cost	Cost per bed day	£396.47	National Schedule of Reference Costs, 2014–2015[90] The cost per bed day was assumed to be the same as the national average unit cost of an excess bed day for AML with a CC score of 0–1 (HRG: SA25M) as reported by the NHS
	Cost per outpatient day	£162.00	National Schedule of Reference Costs, 2014–2015[90] Clinical Haematology (Service code: 303) outpatient consultant led
	Cost per IV infusion	0	Assumption: assumed to be included in the costs for inpatient and outpatient day
Supportive care	Cost per supportive care transfusion	£156.58	NHSBT Price list 2015/16[91] Transfusion costs are £120.00 per unit of red blood cells (item code: BC001), and £193.15 per unit of platelets (item unit: BC041). Average cost of transfusion was calculated as (193.15+120) ÷ 2 = £156.58
	Average cost per day of antibiotic treatment	£1.65	Expert opinion and BNF[88]. Ampicillin is a commonly-administered broad-spectrum antibiotic, with a recommended

			dose of 500 mg every 8 hours. The unit cost of a 500 mg vial for injection is £0.55. Therefore, the cost per day of antibiotic treatment was 0.55 x 3 = £1.65
Monitoring and follow-up	Cost per PCR monitoring test	£280	Expert opinion: Guy's Hospital tariff (NHS Foundation Trust)

AML= acute myeloid leukaemia; BNF=British National Formulary; IV=intravenous; NHSBT= National Health Service Blood and Transplant; PCR=polymerase chain reaction; NHS=National Health Service

Table 3.9. Resource use associated with the technology in the economic model

Category	Items	Treatment phase	Intervention	Value	Reference
Administration	Number of bed days per patient	Induction	ATRA+ATO First Line	32	Lo-Coco et al., 2013[1]
			ATRA+ATO Second Line	25	Douer et al., 2005[89]
			AIDA	35	Lo-Coco et al., 2013[1]
		Consolidation (per cycle)	ATRA+ATO	0	Expert opinion
			AIDA	4	Assumption based on treatment schedule
	Number of ambulatory days per patient	Consolidation (per cycle)	ATRA+ATO First Line	10	Expert opinion
			ATRA+ATO Second Line	12.5	Expert opinion
			AIDA	0	Inpatient treatment assumed
Supportive care	Number of supportive care transfusions	Induction	ATRA+ATO	15	Burnett et al., 2015[2]
			AIDA	22	Burnett et al., 2015[2]
	Number of days of antibiotics (per cycle of consolidation)	Consolidation	ATRA+ATO	1	Burnett et al., 2015[2]
			AIDA	2	Burnett et al., 2015[2]
Monitoring and follow-up	Number of annual PCR tests	Induction + consolidation	ATRA+ATO	5	Expert opinion
			AIDA	4	Expert opinion

ATRA=all-trans retinoic acid; ATO=arsenic trioxide; PCR=polymerase chain reaction

B 3.5.2 Health-state unit costs and resource use

Costs associated with each of the health states described in ***section B 3.2.2*** are presented in Table 3.10 and healthcare resources utilised by patients in each of the states are shown in ***Table 3.11***.

Table 3.10. List of health states and associated costs in the economic model

Heath states	Items	Value	Reference
First molecular remission	Cost per follow-up appointment	£52.50	Personal Social Services Research Unit (PSSRU)[92]
	Cost per PCR monitoring test	£280	Expert opinion: Guy's Hospital tariff (NHS Foundation Trust)
Second molecular remission	Cost per follow-up appointment	£52.50	Personal Social Services Research Unit (PSSRU)[92]
	Cost per PCR monitoring test	£280	Expert opinion: Guy's Hospital tariff (NHS Foundation Trust)
Allogeneic HSCT molecular remission	Cost per follow-up appointment	£52.50	Personal Social Services Research Unit (PSSRU)[92]
	Cost per PCR monitoring test	£280	Expert opinion: Guy's Hospital tariff (NHS Foundation Trust)
	Allogeneic HSCT remission costs (annual)	£21,585.75	Leunis et al., 2013[93]
Autologous HSCT molecular remission	Cost per follow-up appointment	£52.50	Personal Social Services Research Unit (PSSRU)[92]
	Cost per PCR monitoring test	£280	Expert opinion: Guy's Hospital tariff (NHS Foundation Trust)
	Autologous HSCT remission costs (annual)	£5.776.01	Leunis et al., 2013[93]
Allogeneic HSCT	Cost per allogeneic HSCT	£27,907.53	National Schedule of Reference Costs[90]
Autologous HSCT	Cost per autologous HSCT	£7,122.97	National Schedule of Reference Costs[90]
End of life state (Palliative care)	Costs per month	£4,670.68	Marie Curie Cancer Care[94]

HSCT= haematopoietic stem cell transplant; PCR=polymerase chain reaction

Table 3.11. List of health states and associated resource use in the economic model

Heath states	Items	Intervention	Value	Reference
First molecular remission	Duration of follow-up (years)	ATRA+ATO	3	Platzbecker et al., 2015[5]
		AIDA	3	Platzbecker et al., 2015[5]
	Number of annual	ATRA+ATO	4	Platzbecker et al.,

Heath states	Items	Intervention	Value	Reference
	appointments			2015[5]
		AIDA	4	Platzbecker et al., 2015[5]
	Number of annual PCR tests	ATRA+ATO	0	Expert opinion
		AIDA	4	Platzbecker et al., 2015[5]
Second molecular remission	Duration of follow-up (years)	ATRA+ATO	3	Expert opinion
		AIDA	3	Expert opinion
	Number of annual appointments	ATRA+ATO	4	Expert opinion
		AIDA	4	Expert opinion
	Number of annual PCR tests	ATRA+ATO	4	Expert opinion
		AIDA	4	Expert opinion
Allogeneic HSCT molecular remission	Duration of follow-up (years)	ATRA+ATO	3	Expert opinion
		AIDA	3	Expert opinion
	Number of annual appointments	ATRA+ATO	4	Expert opinion
		AIDA	4	Expert opinion
	Number of annual PCR tests	ATRA+ATO	4	Expert opinion
		AIDA	4	Expert opinion
Autologous HSCT molecular remission	Duration of follow-up (years)	ATRA+ATO	3	Expert opinion
		AIDA	3	Expert opinion
	Number of annual appointments	ATRA+ATO	4	Expert opinion
		AIDA	4	Expert opinion
	Number of annual PCR tests	ATRA+ATO	4	Expert opinion
		AIDA	4	Expert opinion
Allogeneic HSCT	Hospitalisation duration (weeks)	ATRA+ATO	4	Expert opinion
		AIDA	4	Expert opinion
Autologous HSCT	Hospitalisation duration (weeks)	ATRA+ATO	3	Expert opinion
		AIDA	3	Expert opinion

ATRA=all-trans retinoic acid; ATO=arsenic trioxide; HSCT=haematopoietic stem cell transplant PCR=polymerase chain reaction

B 3.5.3 Adverse reaction unit costs and resource use

For each type of adverse event, the cost per occurrence was firstly searched in the National Schedule of Reference Costs, 2014–2015. If the cost was not available therein, a targeted search was conducted and the cost was retrieved from the literature. Recent publications reporting English or UK costs were preferred; nevertheless, when no publication presenting these characteristics were found, older studies or papers reporting foreign costs were used. In these cases, costs were converted to sterling using the average annual exchange rates of the year the cost related to and uplifted to 2015 using annual inflation rates published by the Office for National Statistics. The values used in this analysis are reported in ***Table 3.12*** below.

Table 3.12. List of adverse reactions and summary of costs in the economic model

Adverse reactions	Value	Reference	Justification
Thrombocytopenia (grade 3-4, >15 days)	£1,746.00	NHS Reference Costs 2014–15[90]	Unit cost of treating thrombocytopenia with no complications as reported by the NHS
Neutropenia (grade 3-4, >15 days)	£2,845.43	Morgan et al., 2007[95]	The cost of treating neutropenia, including hospitalisation and antibiotic administration, was reported to be £2,286 in 2007. This was uplifted to 2015 amounting to £2845.43.
Infection	£253.97	Soini et al., 2016[96]	Infection (grade 3) as an adverse event following treatment for CLL: cost per visit was €309.04 (2014); converted into sterling and inflated to 2015, amounting to £253.97
Leukocytosis	£349.44	Expert opinion	In the UK, leukocytosis is treated with a single dose idarubicin (8mg/m^2). For an average patient (1.95m^2) this represents 1.56 vial of 10mg of idarubicin, which means 2 vials are used. As the cost per vial is £174.72, the cost of treating leukocytosis is £349.44.
Hepatic toxicity	£5.56	Akhtar and Chung, 2014[97]	Patients with hepatotoxicity discontinue treatment until liver function returns to normal, usually for 7–10 days (Expert opinion). During this time period, patients undergo liver function test every 2–3 days – therefore, it is conservatively assumed that patients with hepatotoxicity receive two liver function tests. The cost of a liver function test, as reported by the BMJ, is £2.78. Thus, the cost of treating hepatotoxicity was calculated at £5.56
Neurotoxicity	£675.88	Calhoun et al., 2001[98]	Based on chemotherapy-related toxicities for patients with ovarian cancer, mean direct medical costs of neurotoxicity were $688 per episode (1999) and included GP visits, drugs/devices and phone calls to medical/nursing providers. These costs were converted into pounds and inflated to 2015, amounting to £675.88.
Differentiation syndrome	£1,225.23	Milligan et al., 2006[47]; BNF[88]; National Schedule of Reference Costs[90]	Treatment for differentiation syndrome is 10 mg dexamethasone every 12 hours until the disappearance of symptoms and for a minimum of 3 days. Dexamethasone is available from the BNF as 1 mL vials for injection (3.8 mg/mL). The cost per vial is £1.99 and the cost per bed day is estimated to be £396.47. Total cost was calculated as 3 vials*6 injections+3 bed days = £35.82+£1,189.41=£1,225.23
Cardiac events	£1,104.02	National Schedule of Reference Costs[90]	Costs of a cardiac arrest with a CC Score 0-4
QTc prolongation	£34.50	Expert opinion; NICE clinical guideline, No. 108[99];	Patients with prolonged QTc interval discontinue treatment until the cardiac rhythm normalises. They also receive daily ECG monitoring and one infusion of serum electrolytes (expert opinion). The unit cost of ECG monitoring reported by the NHS is

Adverse reactions	Value	Reference	Justification
		NICE clinical guidelines, No. 174[100]	£32.00, while the cost of administering serum electrolytes is £12.50. The total cost is therefore:= 32 + 12.50 = £34.50
tMDS/AML	£6,207.00	National Schedule of Reference Costs[90]	Average cost of treatment for MDS and AML as reported by the 2014–15 NHS reference costs
Acute GvHD	£34,493.05	Saito et al., 2008[101]	Grade II to IV acute GVHD: $46,414(2004). This was converted to sterling and uplifted to 2015, amounting to £34,493.05
Chronic GvHD	£8,785.25	Jones et al., 2016[102]	Predicted total 10-year medical costs were $5,273,079,941 for a total of 44,450 predicted cases. Ten-year medical costs per patient: $118,630; 1-year cost: $11,863; converted to sterling: £8,785.25

BNF=British National Formulary; CLL= Chronic Lymphocytic Leukaemia; ECG=electrocardiogram; GP=general physician; GvHD=graft-versus-host disease; NH=National Health Service; NICE=National Institute for Health and Care Excellence; tMDS/AML=treatment-related myelodysplastic syndrome or acute myeloid leukaemia

B 3.6 Summary of base-case analysis inputs and assumptions

Key inputs used in the base-case scenario are reported in **Tables 3.3 – 3.13.** All base case values of the model's parameters are listed in **Table 1.1** of **Appendix J**, together with their range in the Deterministic Sensitivity Analysis (DSA) and 95% Confidence Interval (CI) of the distributions used in the Probabilistic Sensitivity Analysis (PSA).

Table 3.13. Summary of variables used in the economic model – base case and sensitivity analyses

Variable	Base case Value	Measurement of uncertainty and distribution: DSA: Low–High (change) PSA: CI (distribution)	Reference
Time horizon (years)	40	DSA: 5–30 Not included in PSA	
Discount rate for costs	3.5%	DSA: 0%–5% Not included in PSA	NICE guideline
Discount rate for health outcomes	3.5%	DSA: 0%–5% Not included in PSA	NICE guideline
Percentage of males	48.7%	DSA: 28.7%–68.7% (± 20 points) PSA: 28.69%-68.93% (Beta)	Lo-Coco et al., 2013[1]
Start age of patients (years)	45	DSA: 35–55 (± 10 years)(Normal)	Lo-Coco et al., 2013[1]
Patient weight			
Proportion of patients <70Kg	27.9%	DSA: 42.3%-14.0% (LC: [1-p(>85Kg)]/2; HC:-50%) PSA: 16.43%-42.35% (Dirichlet)	Health survey for England 2012, Chapter 10[103]
Proportion of patients 70-85Kg	41.4%	DSA: 42.3%-43.0% (LC:[1-p(>85Kg)]/2 ; HC: [1-p(<70Kg)]/2) PSA: 30.60%-53.02% (Dirichlet)	Health survey for England 2012, Chapter 10[103]
Proportion of patients >85Kg	30.7%	DSA: 15.3%-43.0% (LC:-50% ; HC: [1-p(<70Kg)]/2) PSA: 19.19%-44.52% (Dirichlet)	Health survey for England 2012, Chapter 10[103]
Average height (cm)	168.84	DSA:126.63–211.05 (± 25%)(Normal)	Health survey for England 2012, Chapter 10[103]

B 3.7 Base-case results

B 3.7.1 Disaggregated and total costs

Total cost in patients treated with the ATRA+ATO strategy amounted to £104,996 (discounted), with the most important cost driver being the treatment acquisition costs of £60,336 (discounted). In the AIDA arm, total cost was estimated at £136,267 (discounted, with lower treatment acquisition costs than in the ATRA+ATO arm (£21,604 discounted). Costs in all other categories were higher in the AIDA arm than the ATRA+ATO arm, with the highest costs related to transplantation (£48,326 discounted). Not discounted and discounted costs generated over the lifetime horizon of the model are presented in *Error! Reference source not found.* and *Error! Reference source not found.*, respectively.

Cost savings in palliative care generated by the use of ATRA+ATO in first line were relatively important (-£4,290 discounted), representing approximately a month of spending in the end of life health state where patients received palliative care. This is certainly related to the effectiveness of ATRA+ATO, which reduces the risk of treatment failure compared with AIDA.

Substantial savings were also observed in adverse event management costs (-£8,236 discounted), with ATRA+ATO presenting a better haematological safety profile than AIDA characterised by less thrombocytopenia and neutropenia.

Table 3.14. Not discounted disaggregated and total costs - base-case scenario

Cost category	ATRA+ATO	AIDA	ATRA+ATO vs. AIDA
Treatments	£62,727	£28,636	£34,091
Administration	£26,782	£35,710	-£8,928
Supportive care and antibiotics	£3,881	£7,383	-£3,503
Follow-up and monitoring	£4,152	£15,854	-£11,701
Adverse Events	£4,453	£13,293	-£8,840
MDS	£0	£246	-£246
HSCT	£16,576	£90,508	-£73,932
Palliative care	£1,741	£8,083	-£6,342
Total	**£120,312**	**£199,713**	**-£79,401**

ATO=arsenic trioxide; ATRA=all-trans retinoic acid; HSCT=haematopoietic stem cell transplantation; MDS=myelodysplastic syndrome

Table 3.15 Discounted disaggregated and total costs - base-case scenario

Cost category	ATRA+ATO	ATRA+IDA	ATRA+ATO vs. AIDA

Treatments	£60,336	£21,604	£38,731
Administration	£25,402	£31,660	-£6,259
Supportive care and antibiotics	£3,575	£6,487	-£2,912
Follow-up and monitoring	£2,991	£10,389	-£7,398
Adverse Events	£4,142	£12,378	-£8,236
MDS	£0	£226	-£226
HSCT	£7,645	£48,326	-£40,681
Palliative care	£906	£5,196	-£4,290
Total	**£104,996**	**£136,267**	**-£31,270**

ATO=arsenic trioxide; ATRA=all-trans retinoic acid; HSCT=haematopoietic stem cell transplantation; MDS=myelodysplastic syndrome

B 3.7.2 Health outcomes

Over the 40 years of the simulation, patients treated with ATRA+ATO accumulated 19.56 LYs and 16.34 QALYs (discounted), while those treated with AIDA accumulated 16.56 LYs and 13.72 QALYs (discounted). The proportion of patients entering first molecular remission was increased by approximately 11% when using the ATRA+ATO as opposed to AIDA strategy; this increased to 16.7% when long-term remission (over two years) was considered. No tMDS/AML were observed in the ATRA+ATO arm, while 1.39% of patients treated in the ATRA+IDA arm developed the disease. The proportion of APL-related deaths was reduced by around 20% with ATRA+ATO compared to AIDA. Health outcomes generated over the lifetime horizon in the base case settings are presented in *Error! Reference source not found.* (not discounted) and *Error! Reference source not found.* (discounted).

Table 3.16. Not discounted health outcomes in the model

	ATRA+ATO	AIDA	ATRA+ATO vs. AIDA
Number of QALYs	27.91	22.38	5.52
Number of LYs	33.22	26.84	6.38
First remission	99.83%	89.11%	10.72%
First long remission (> 2 years)	99.45%	81.53%	17.92%
MDS	0.00%	1.39%	-1.39%
Death	57.04%	74.13%	-17.09%
APL related death	7.54%	39.38%	-31.85%
Non-APL related death	49.51%	34.75%	14.76%

ATO=arsenic trioxide; ATRA=all-trans retinoic acid; LY=life years; MDS=myelodysplastic syndrome; QALY=quality-adjusted life years

Table 3.17. Discounted health outcomes in the model

	ATRA+ATO	AIDA	ATRA+ATO vs. AIDA
Number of QALYs	16.34	13.72	2.62
Number of LYs	19.56	16.56	3.00
First remission	99.83%	89.11%	10.72%
First long remission (> 2 years)	92.84%	76.11%	16.73%
MDS	0.00%	1.39%	-1.39%
Death	23.36%	38.06%	-14.69%
APL related death	3.74%	23.65%	-19.91%
Background death	19.63%	14.41%	5.22%

ATO=arsenic trioxide; ATRA=all-trans retinoic acid; LY=life years; MDS=myelodysplastic syndrome; QALY=quality-adjusted life years

B 3.7.3 Base-case incremental cost-effectiveness analysis results

In the base-case scenario, the combination of ATRA+ATO was associated with an incremental gain of 2.62 QALYs (3.00 LYs) and provided a saving of £31,270 compared with AIDA. Thus, the combination of ATRA+ATO was dominant and no base-case ICUR was calculated (***Table 3.18***).

Table 3.18. Base-case incremental results (discounted)

Treatment	Total costs (£)	Total LYG	Total QALYs	Incremental costs (£)	Incremental LYG	Incremental QALYs	ICER versus baseline (£/QALY)	ICER incremental (£/QALY)
ATRA+ATO	104,996	19.56	16.34	–	-	-	Dominant	Dominant
AIDA	136,267	16.56	13.72	-£31,270	-3.00	-2.62	-	-

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

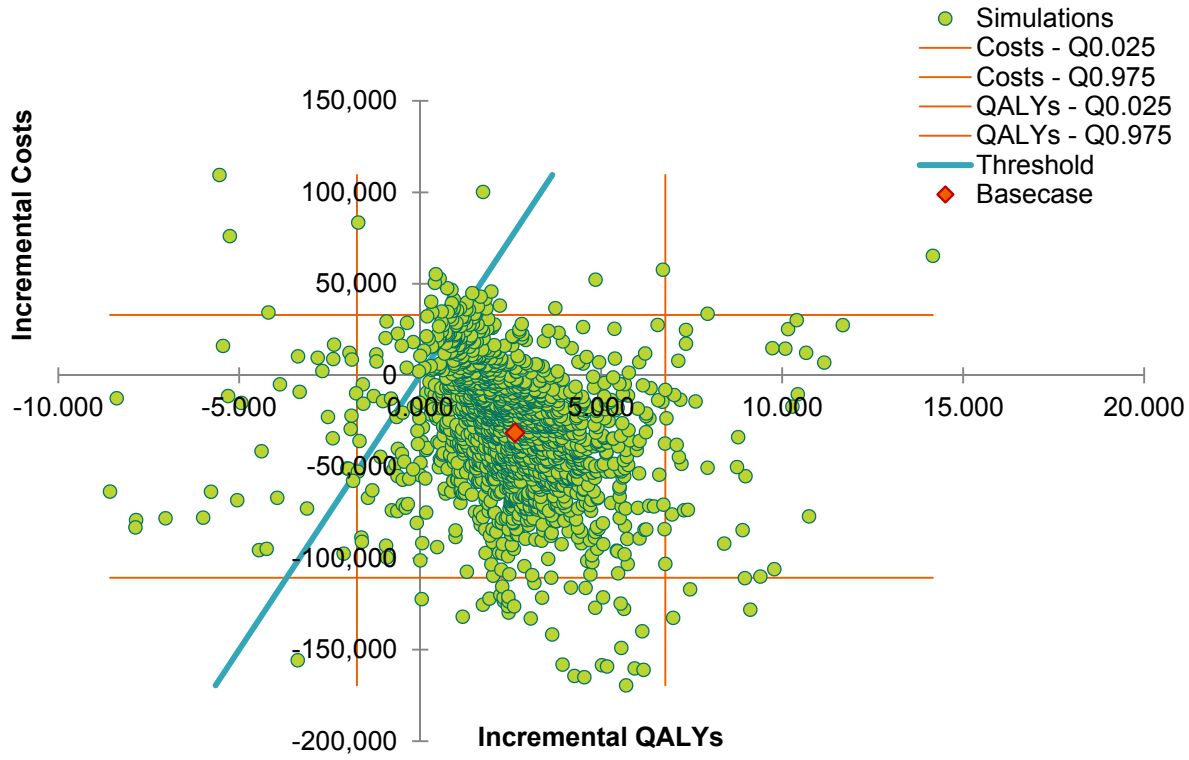
B 3.8 Sensitivity analyses

B 3.8.1 Probabilistic sensitivity analysis

The results of probabilistic sensitivity analysis were based on 1,500 Monte Carlo simulations of different input sets in order to estimate credible limits of the ICER. Having specified distributions for all the model inputs (presented ***Table 3.13***), probabilistic sensitivity analysis was performed by randomly sampling from each of the parameter distributions and calculating the expected costs and expected QALYs for each strategy. Distributions were defined according to the type of parameters in order to reflect the distribution that they generally follow: probabilities and utilities were simulated according to beta distributions, patients' height and weight, as well

as event duration and resource use according to normal distributions, and costs according to gamma distribution. The results of these 1,500 iterations are illustrated on the cost-effectiveness plane (***Figure 3.2***) together with the baseline estimate of the efficient frontier (the blue line representing a threshold of £30,000 per QALY) and a range between two quantiles: the 2.5 centile and the 97.5 centile (the orange lines).

Figure 3.2. Incremental cost-effectiveness plane for ATRA+ATO vs AIDA



The model was robust to simultaneous variation of model inputs. On average, patients treated with ATRA+ATO were expected to gain 2.55 additional QALYs at a cost lower by £31,088 (***Table 3.19***) compared to AIDA. One can notice that the mean simulated values of incremental costs and QALYs differ slightly from the base case results. This can be explained by the fact that the PSA allowed inputs that were considered to be zero in the base case analysis (for example, the proportion of patients with certain AEs and some resource use parameters) to take their values from the distribution with non-zero (positive) skewness.

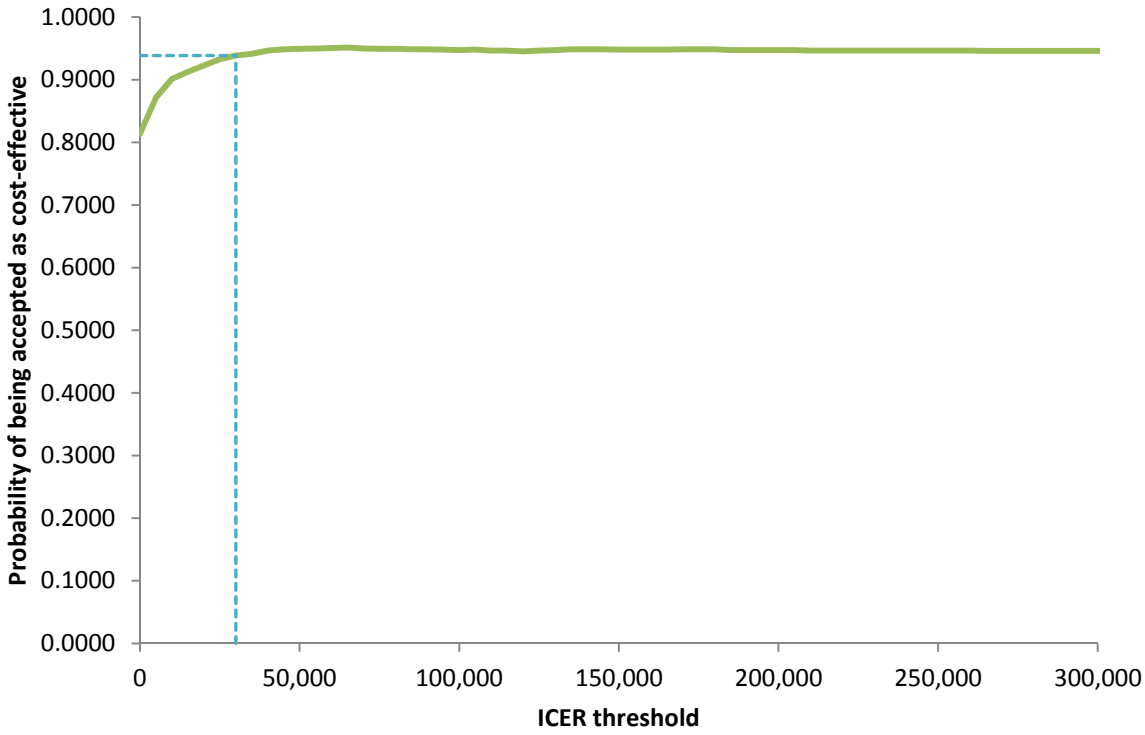
Table 3.19. Probabilistic sensitivity analysis results.

	Incremental costs (£)	Incremental QALYs	ICER (£/QALY)
Mean	-31,088	2.54596	Dominant
Std Deviation	36,400	2.10568	422,020

Median	-28,654	2.43542	10,922
Min	-169,499	-8.56954	28
Q 0.025	-110,732	-1.74592	568
Q 0.975	32,992	6.77110	184,853
Max	109,569	14.16720	7,366,602

PSA results are presented graphically in the cost-effectiveness acceptability curve (CEAC, ***Figure 3.3***), which shows the probability that the intervention is cost-effective at different willingness-to-pay thresholds. The CEAC demonstrated that the probability of ATRA+ATO being cost-effective was around 81.33% at a £0 willingness-to-pay threshold, and increased further when this threshold increased, to reach 93.87% at £30,000 per QALY. Furthermore, ATRA+ATO dominated AIDA in 77.13% of the simulations.

Figure 3.3. Cost-effectiveness acceptability curve for ATRA+ATO vs AIDA



B 3.8.2 Deterministic sensitivity analysis

The impact of varying individual parameter values (ranges of variation are presented ***Table 3.13***) on incremental costs, incremental QALYs and ICUR of ATRA+ATO compared with AIDA was studied with a one-way deterministic sensitivity analysis.

B 3.8.2.1 Incremental costs

The twenty parameters with the greatest impact on incremental cost are displayed in the tornado chart (***Figure 3.4***) and in ***Incremental costs were mostly affected by varying the probability of relapse*** at 48 months after first remission in the AIDA arm, the discount rate for costs, the time horizon, the complete haematological remission rate following AIDA in first line, and the probability of relapse at 24 months after first remission in the AIDA arm. Among parameters related to the safety profile of the studied strategies (proportion of patients experiencing adverse events, costs of adverse events, etc.), only one was identified as a key driver of the incremental costs: the proportion of patients experiencing cardiac events during first-line consolidation with AIDA, which can be explained by the fact that these patients switched to ATRA+ATO.

Table 3.20 below.

Incremental costs were mostly affected by varying the probability of relapse at 48 months after first remission in the AIDA arm, the discount rate for costs, the time horizon, the complete haematological remission rate following AIDA in first line, and the probability of relapse at 24 months after first remission in the AIDA arm. Among parameters related to the safety profile of the studied strategies (proportion of patients experiencing adverse events, costs of adverse events, etc.), only one was identified as a key driver of the incremental costs: the proportion of patients experiencing cardiac events during first-line consolidation with AIDA, which can be explained by the fact that these patients switched to ATRA+ATO.

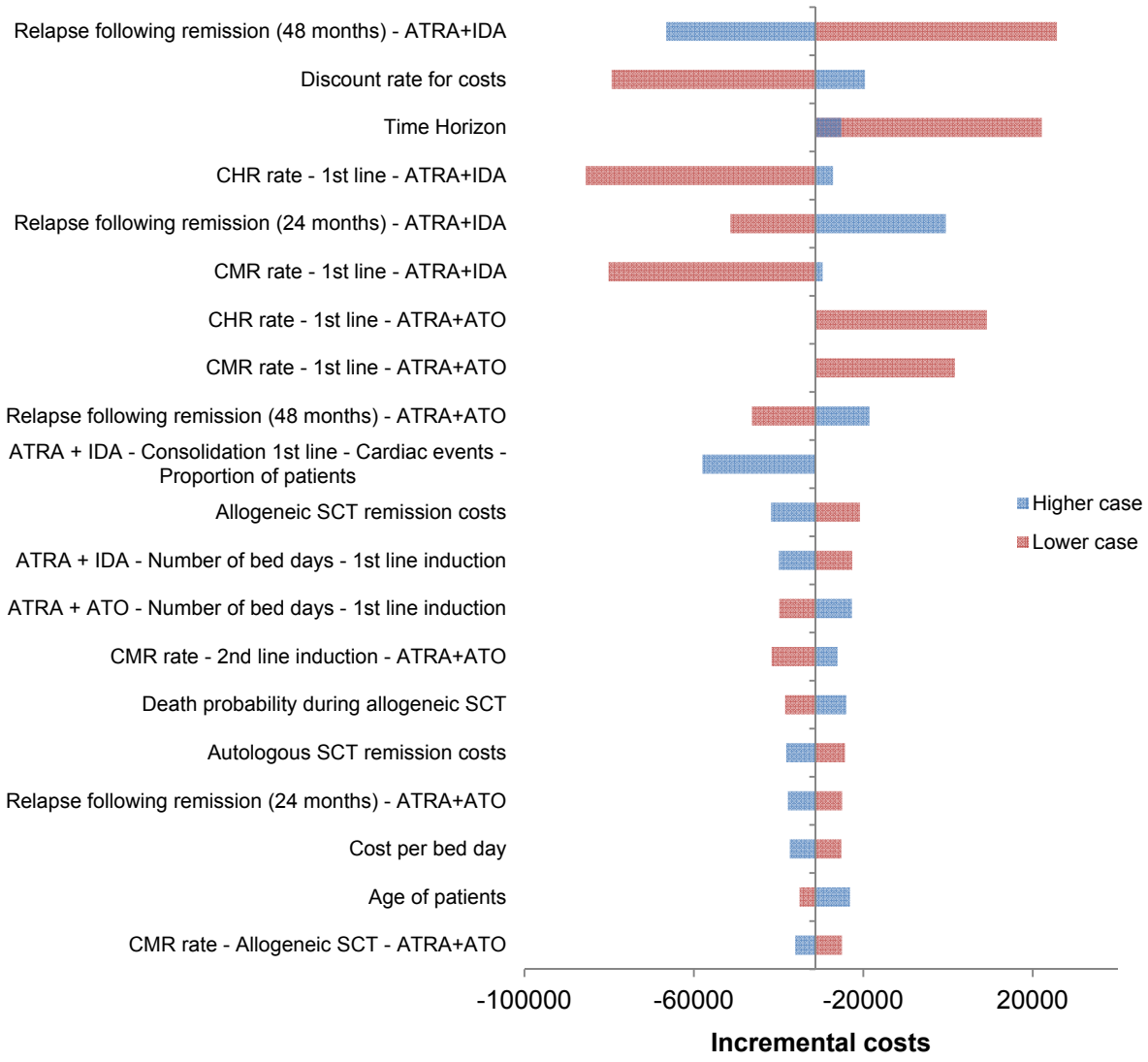
Table 3.20. Deterministic sensitivity analysis – Incremental cost results

Parameters	**Incremental Costs (£)**		
	Lowcase	**Highcase**	**Distance***
Relapse following remission (48 months) - AIDA	25,701	-66,546	92,246
Discount rate for costs	-79,401	-19,602	59,799
Time Horizon	22,128	-25,208	59,461
CHR rate - First line - AIDA	-85,568	-27,157	58,411
Relapse following remission (24 months) - AIDAAIDA	-51,405	-510	50,896
CMR rate - First line - AIDA	-80,138	-29,600	50,538
CHR rate - First line - ATRA+ATO	9,181	-31,127	40,595
CMR rate - First line - ATRA+ATO	1,625	-31,270	32,895
Relapse following remission (48 months) - ATRA+ATO	-46,301	-18,538	27,763

Parameters	Incremental Costs (£)		
	Lowcase	Highcase	Distance*
AIDA - Consolidation First line - Cardiac events - Proportion of patients	-31,270	-57,993	26,723
Allogeneic SCT remission costs	-20,755	-41,785	21,030
AIDA - Number of bed days - First line induction	-22,603	-39,938	17,334
ATRA+ATO - Number of bed days - First line induction	-39,847	-22,694	17,153
CMR rate - Second line induction - ATRA+ATO	-41,667	-26,072	15,596
Death probability during allogeneic SCT	-38,450	-24,046	14,404
Autologous SCT remission costsATRAATRA+ATO	-24,330	-38,211	13,881
Relapse following remission (24 months) - ATRA+ATO	-25,002	-37,811	12,809
Cost per bed day	-25,165	-37,376	12,210
Age of patients	-35,066	-23,154	11,913
CMR rate - Allogeneic SCT - ATRA+ATO	-25,023	-36,071	11,048

ATO=arsenic trioxide; ATRA=all-trans retinoid acid; CHR=complete haematological remission; CMR=complete molecular remission; FU=follow-up; GvHD=Graft-versus-host disease; HSCT=haematopoietic stem cell transplantation; IDA=idarubicin; MDS=myelodysplastic syndrome; SCT=stem cell transplant.
* Distance is ABS(Lowcase – Basecase)+ABS(Highcase – Basecase)

Figure 3.4. Tornado diagram – The twenty parameters with the greatest impact on incremental costs of ATRA+ATO vs AIDA



ATO=arsenic trioxide; ATRA=all-trans retinoid acid; CHR=complete haematological remission; CMR=complete molecular remission; IDA=idarubicin; SCT=stem cell transplant

B 3.8.2.2 Incremental effectiveness

The twenty parameters with the greatest impact on incremental effectiveness are displayed in the tornado diagram (***Figure 3.5***) and in *Error! Not a valid bookmark self-reference.*. The incremental effectiveness (difference in QALYs) was mostly affected by the changes in the discount rate for health outcomes, time horizon, probability of relapse observed at 48 months in the AIDA arm, first line haematological remission rate associated with AIDA treatment and the utility value in the first molecular remission (2> years) health state. Only two utility values were

found among these twenty most influential parameters: utility in the first molecular remission (2> years) health state and utility in molecular remission after autologous HSCT.

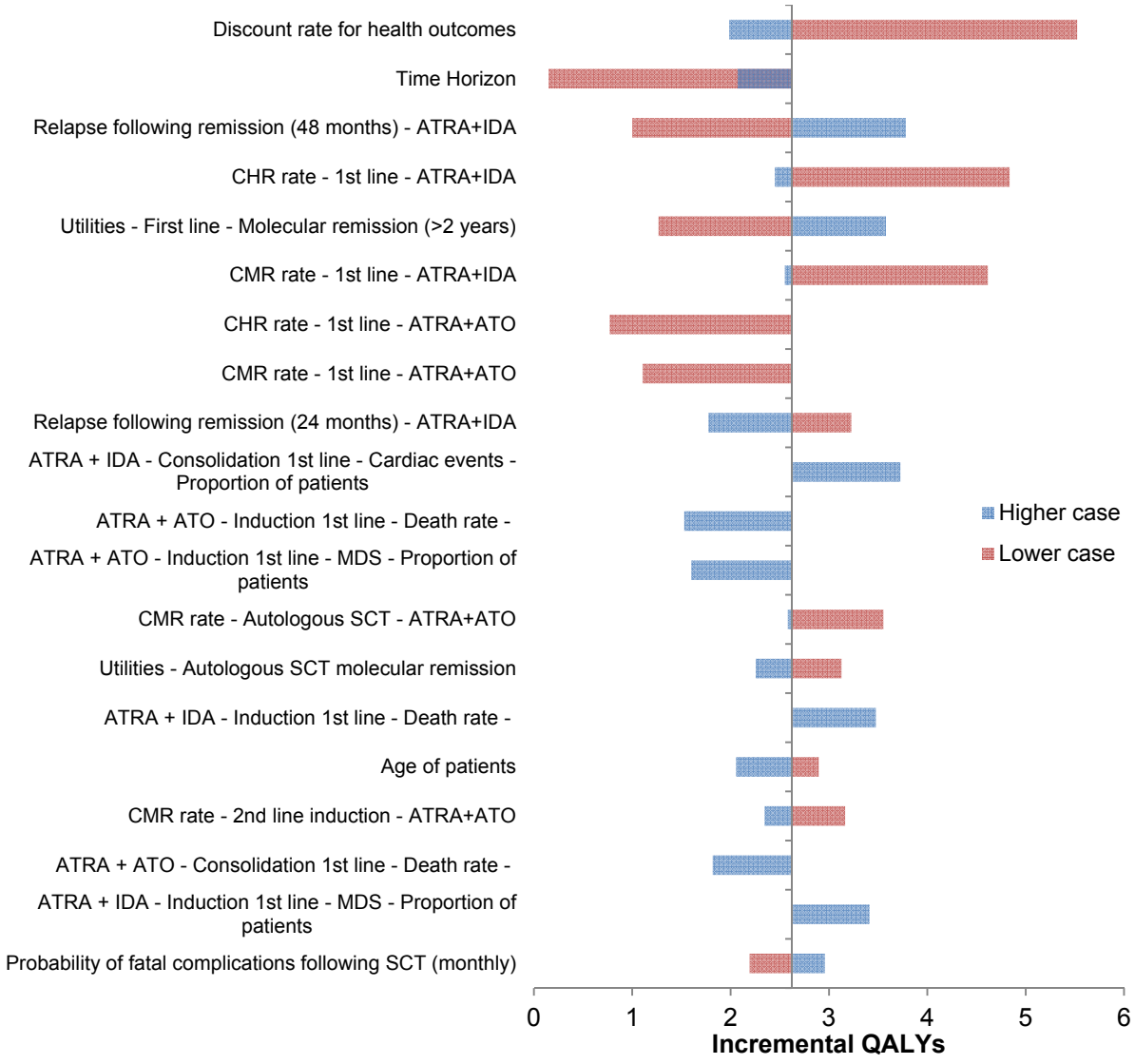
Considering incremental effectiveness, ATRA+ATO always generated a gain in QALYs versus AIDA across all values included in the deterministic sensitivity analysis.

Table 3.21. Deterministic sensitivity analysis – incremental QALYs

Parameters	**Incremental QALYs**		
	Lowcase	**Highcase**	**Distance***
Discount rate for health outcomes	5.52	1.98	3.54
Time Horizon	0.15	2.07	3.02
Relapse following remission (48 months) - AIDA	1.00	3.78	2.78
CHR rate - First line - AIDA	4.84	2.45	2.39
Utilities - First line - Molecular remission (>2 years)	1.27	3.58	2.31
CMR rate - First line - AIDA	4.61	2.55	2.06
CHR rate - First line - ATRA+ATO	0.77	2.62	1.85
CMR rate - First line - ATRA+ATO	1.10	2.62	1.51
Relapse following remission (24 months) - AIDA	3.23	1.77	1.45
AIDA - Consolidation First line - Cardiac events - Proportion of patients	2.62	3.73	1.11
ATRA+ATO - Induction First line - Death rate - ATRA+ATO	2.62	1.53	1.09
ATRA+ATO - Induction First line - MDS - Proportion of patientsATRA+ATO	2.62	1.60	1.02
CMR rate - Autologous SCT - ATRA+ATO	3.55	2.58	0.97
Utilities - Autologous SCT molecular remission	3.13	2.26	0.87
AIDA - Induction First line - Death rate -	2.62	3.48	0.86
Age of patients	2.89	2.06	0.84
CMR rate - Second line induction - ATRA+ATO	3.16	2.35	0.82
ATRA+ATO - Consolidation First line - Death rate - ATRA+ATO	2.62	1.82	0.80
AIDA - Induction First line - MDS - Proportion of patientsAIDA	2.62	3.41	0.79
Probability of fatal complications following SCT (monthly)	2.19	2.96	0.77
Discount rate for health outcomes	5.52	1.98	3.54

ATO, arsenic trioxide; ATRA, all-trans retinoid acid; CHR, complete haematological remission; CMR, complete molecular remission; FU, follow-up; GvHD, Graft-versus-host disease; HSCT, haematopoietic stem cell transplantation; IDA, idarubicin; MDS, myelodysplastic syndrome; SCT, stem cell transplant.
* Distance is ABS(Lowcase – Basecase)+ABS(Highcase – Basecase)

Figure 3.5. Tornado diagram – The twenty parameters with the greatest impact on incremental QALYs of ATRA+ATO vs AIDA



ATO=arsenic trioxide; ATRA=all-trans retinoid acid; CHR=complete haematological remission; CMR=complete molecular remission; IDA=idarubicin; SCT=stem cell transplant

B 3.8.2.3 Incremental cost-effectiveness analysis results

Considering the ICUR, ATRA+ATO was the dominant treatment strategy versus AIDA with most of the values included in the deterministic sensitivity analysis. The ICUR was computable in only four cases, reported in ***Table 3.22***. ATRA+ATO was not cost-effective over a time horizon of five years, due to high treatment acquisition costs occurring in the first year and a time horizon being insufficient to capture the benefits on QoL. In the three other scenarios (lower probability of relapse over 48 months in the AIDA arm, lower haematological remission rate and lower molecular

remission rate with ATRA+ATO in first line) ATRA+ATO was cost-effective against AIDA. With ATRA+ATO being dominant against AIDA in the base case and in most of the tested scenarios, no tornado chart of the ICURs is shown.

Table 3.22. Deterministic sensitivity analysis – Cost-utility results

Parameters	ICER (£/QALY)	
	Lower case	Higher case
Time Horizon	148,179	Dominant
Relapse following remission (48 months) - AIDA	25,658	Dominant
CHR rate - First line - ATRA+ATO	11,927	Dominant
CMR rate - First line - ATRA+ATO	1,472	Dominant

ATO, arsenic trioxide; ATRA, all-trans retinoid acid; CHR, complete haematological remission; CMR, complete molecular remission; FU, follow-up; GvHD, Graft-versus-host disease; HSCT, haematopoietic stem cell transplantation; IDA, idarubicin; MDS, myelodysplastic syndrome; SCT, stem cell transplant.
* Distance is ABS(Lowcase – Basecase)+ABS(Highcase – Basecase)

B 3.8.3 Scenario analysis

In order to explore the uncertainty around specific parameters, several scenario analyses were conducted. In addition of those reported in this section, three scenarios are reported in the **Section J.4** of **Appendix J**:

- AML17 protocol: a scenario using the schedule, dosage, efficacy and safety inputs based on the AML17 clinical study.
- "Worst case" scenario: a scenario accumulating unfavourable inputs for the ATRA+ATO strategy
- Probability of undergoing HSCT reflecting clinical practice, with a lower proportion of patients undergoing autologous HSCT and allogeneic HSCT reserved for patients who did not achieve molecular remission after second-line induction.

B 3.8.3.1 AIDA used in second line following both first-line treatments

In order to check that the cost-effectiveness of ATRA+ATO against AIDA was not driven by the strategy used in second line, a scenario in which all relapsing patients received AIDA was conducted. The results of this scenario are presented in ***Table 3.23,ATO=arsenic*** trioxide; ATRA=all-trans retinoic acid; ICER=Incremental cost-effectiveness ratio; ICUR-Incremental cost-utility ratio; LY=life year; QALY=quality-adjusted life year

Table 3.24 and ***ATO=arsenic*** trioxide; ATRA=all-trans retinoic acid; HSCT=haematopoietic stem cell transplant; MDS=myelodysplastic syndrome

Table 3.25. Despite the slightly lower numbers of LYs and QALYs accumulated in both arms in this scenario than in the base case, incremental effectiveness was somewhat increased (LYs: 3.12 vs. 3.00 compared to the base case; QALYs: 2.72 vs 2.62 compared to the base). Regarding costs, both strategies were less costly than in the base case, with a larger difference in the AIDA arm. This reduced the cost savings generated in the ATRA+ATO arm. As expected, this scenario demonstrated that the cost-effectiveness of ATRA+ATO against AIDA was not driven by its use in second line.

Table 3.23. Incremental cost-effectiveness results in the "no second-line ATO use" scenario

	ATRA+ATO vs. ATRA+IDA					
	Incremental costs		Incremental effectiveness		Incremental ratio	
ICERs	Not discounted	Discounted	Not discounted	Discounted	Not discounted	Discounted
Cost / QALY (ICUR)	-£65,974	-£21,593	5.70	2.72	Dominant	Dominant
Cost / LY	-£65,974	-£21,593	6.60	3.12	Dominant	Dominant

ATO=arsenic trioxide; ATRA=all-trans retinoic acid; ICER=Incremental cost-effectiveness ratio; ICUR-Incremental cost-utility ratio; LY=life year; QALY=quality-adjusted life year

Table 3.24. Cost results in the "no second-line ATO use" scenario

Cost category	ATRA+ATO		AIDA		ATRA+ATO vs. AIDA	
	Not discounted	Discounted	Not discounted	Discounted	Not discounted	Discounted
Treatments	£57,896	£57,631	£7,291	£6,513	£50,605	£51,118
Administration	£26,998	£25,523	£36,666	£32,337	-£9,667	-£6,814
Supportive care and antibiotics	£3,912	£3,593	£7,521	£6,585	-£3,609	-£2,992
Follow-up and monitoring	£4,096	£2,966	£15,543	£10,227	-£11,446	-£7,261
Adverse Events	£5,050	£4,475	£15,928	£14,241	-£10,879	-£9,765
MDS	£33	£17	£396	£325	-£363	-£307
HSCT	£16,775	£7,787	£91,030	£49,001	-£74,255	-£41,214
Palliative care	£1,762	£925	£8,122	£5,283	-£6,360	-£4,357
Total	**£116,522**	**£102,918**	**£182,496**	**£124,510**	**-£65,974**	**-£21,593**

ATO=arsenic trioxide; ATRA=all-trans retinoic acid; HSCT=haematopoietic stem cell transplant; MDS=myelodysplastic syndrome

Table 3.25. Effectiveness results in the "no second-line ATO use" scenario

	ATRA+ATO		ATRA+IDA		ATRA+ATO vs. AIDA	
	Not discounted	Discounted	Not discounted	Discounted	Not discounted	Discounted

Number of QALYs	27.86	16.32	22.16	13.60	5.70	2.72
Number of LYs	33.17	19.54	26.57	16.42	6.60	3.12
First line remissions	99.83%	99.83%	89.11%	89.11%	10.72	10.72
First line long remissions (>2 years)	99.45%	92.84%	81.53%	76.11%	17.92	16.73
MDS	0.20%	0.11%	2.27%	2.01%	-2.07	-1.90
Death	57.17%	23.48%	74.54%	38.67%	-17.37	-15.19
APL related death	7.78%	3.89%	40.25%	24.44%	-32.48	-20.55
Background death	49.39%	19.59%	34.29%	14.23%	15.10	5.36

APL=Acute promyelocytic leukaemia; ATO=arsenic trioxide; ATRA=all-trans retinoic acid; LY=life year; MDS=myelodysplastic syndrome; QALY=quality-adjusted life year

B 3.8.3.2 Utilities from Tallman et al.

To address the uncertainty around the utility values in the different health states, a scenario analysis was conducted with utilities already published and used in an existing cost-effectiveness model. The utility values reported by Tallman et al.[70] were:

- First-line stable disease = 0.7828; this value was used for first remission health states (≤2 years and long-term remission) in this scenario
- First-line disease event/second line stable disease = 0.6529, this value was used for first-line treatment phases, second line and HSCT remission.
- Second-line disease event = 0.4729, used for second-line treatment phases and HSCT.
- The utility values for tMDS/AML and end of life were kept at their base-case values (0.4)

In this scenario, the number of QALYs accumulated over the time horizon was reduced by1.3 discounted QALYs in the ATRA+ATO arm and 1.6 discounted QALYs in AIDA arm compared to the base case, leading to a slightly increased number of QALYs gained (2.93 vs. 2.62 in the base case). The number of QALYs accumulated in each arm over the 40-year time horizon in this scenario is presented in **Table 3.26** below.

Table 3.26. Incremental QALYs in the "utilities from Tallman, et al." scenario

	ATRA+ATO		**AIDA**		**ATRA+ATO vs. AIDA**	
	Not discounted	Discounted	Not discounted	Discounted	Not discounted	Discounted
Number of QALYs	25.63	15.07	19.60	12.13	6.03	2.93

ATO=arsenic trioxide; ATRA=all-trans retinoic acid; QALY=quality-adjusted life year

B 3.8.3.3 Societal perspective analysis

Societal perspective included the costs of lost productivity (absenteeism and presenteeism). As expected, cost savings with the ATRA+ATO strategy in this scenario increased compared with the base case (£32,833 saved compared to £31,270, respectively). Discounted indirect costs were estimated at £30,771 with ATRA+ATO and £32,333 with AIDA.

B 3.8.4 Summary of sensitivity analyses results

Sensitivity analyses demonstrated the robustness of the model and its results. Despite the wide variability of parameter values used in the DSA, ATRA+ATO dominated AIDA in almost all cases; the only case where the strategy was not cost-effective being that where the time horizon was too short to capture all the benefits of ATRA+ATO. PSA confirmed the expected cost savings and QALYs gained that ATRA+ATO could generate when used first-line. The high proportion of PSA simulations where ATRA+ATO was cost-effective against AIDA reduces the uncertainty around the ICUR.

The different scenario analyses conducted showed that the cost-effectiveness of ATRA+ATO against AIDA was driven neither by the strategy used in the second and "third" (HSCT) line, nor by the utility values considered in the model. The scenario conducted with the most unfavourable inputs for the studied strategy confirmed the dominance of ATRA+ATO over AIDA.

B 3.9 Subgroup analysis

Due to the low number of patients affected by APL, no subgroup analyses were conducted

B 3.10 Validation

B 3.10.1 Internal validation of cost-effectiveness analysis

Internal validation of the model concerned the technical accuracy of the model and intended to identify programming errors, data entry issues and logical inconsistencies in the model. Therefore, a variety of extensive tests was performed before the model was used for cost-effectiveness analysis. The aim of these tests was to demonstrate that the model was able to predict the results in a manner consistent with expectations.

Validation checks included all main aspects of the model: efficacy and safety of compared strategies, treatment schedules, treatment costs, resource use and mortality in the modelled population. Data, calculations and formulas were verified by a person not involved in the initial project.

This part of the validation included quality control conducted following the methodology proposed by the York Health Economics Consortium (YHEC). A summary of evidence on the internal validity of the model is reported in ***Table 3.27***.

Table 3.27. Summary of validation tests

#	Test (theoretical)	Expected effect	Test (conducted)	Observed effect
1	Set both treatment and comparator to same intervention	Costs and QALYs to be equal	Same Schedule, Efficacy, Safety and costs + Probability of relapse at 48 month set to 0 in ATRA + IDA (AIDA)	Costs and QALYs were equal in the sub-strategy ATRA+ATO / AIDA and in AIDA
2	Set all efficacy data equal for treatment and control and set disutility associated with treatment related adverse events to 0	Same QALY estimates for treatment and control	Same efficacy, safety (relapse at 48m set to 0 in AIDA) and treatment schedule ; Time horizon =1 year; Utilities=1 and disutilities=0	All health outcome were equal between strategy ATRA+ATO / AIDA and in AIDA
3	Set mortality rate to 0% at all ages	No deaths in model	- Death rate in Safety and Morality sheets equaled 0% at all ages; - Disease and treatment related mortality rates set to 0; - All mortality rates set to 0	- No background death; - No disease related death. Life expectancy = 65.66 in all arms; - No death, life expectancy = 145 (age of patients + 100 "Markov" years)
4	Set mortality rate to 100% at all ages	All patients dead at cycle 1, but still generate expected costs and QALYs	Mortality rates set to 100% at all ages	All patients die at cycle 1. Costs and QALYs are generated only in the 1st cycle
5	Set mortality rate to 100% at age 70	All patients dead after x years (starting age 70 - x) but still generate expected costs and QALYs	Set mortality rate to 100% at age 70	All patients die at 70 years or in the 326th cycle = beginning of the 26th year
6	Increase mortality rate	Reduced costs	- Increase disease related death inputs other than treatment related (1 by 1); - increase all disease related death probabilities; - Increase treatment (trt) related death.	- Disease related deaths increased in both arms; - increasing trt1 related death --> QALYs decreased, Costs increased, trt1 related death increased; - increasing trt2 related death --> QALYs increased, Costs decreased, trt2 related death increased.
7	Health state utilities same	Same QALYs for surviving	Same utility=0.9 in all state, disutilities=0	ratio QALY/LY equal utility value

	for all states	patients (life years and QALYs should have same ratio in both arms)		(0.9) in all arms
8	Health state utilities and adverse events all set to 0	Total QALYs = 0 for treatment and comparator.	- Utilities and AE duration = 0; - Utilities and disutilities set to 0	- Total QALYs = 0 for all treatments - Total QALYs = 0 for all treatments
9	Health state utilities for states all set to 1 and adverse events all set to 0	Total QALYs same as life years	Health state utilities set to 1 in all states, disutilities set to 0	Total QALYs same as life years in all arms
10	Unit costs of treatments set to 0	Total cost of treatment = 0	Unit costs of treatments set to 0	Total Cost of treatment =0
11	Doubled unit costs of treatment	Treatment costs doubled	Doubled treatment costs	Treatment costs doubled
12	Unit costs of background cost to 0	Total administration costs = 0	All costs = 0, except for treatment acquisition	All medical costs = 0 Supportive care and FU costs = 0
13	Doubled unit costs of Background cost	Total administration costs doubled	- All medical costs * 2 - AE costs * 2	- All medical costs doubled - Supportive care and FU costs doubled
14	Alter time horizon (TH)	Total costs and QALYs to increase/decrease in accordance with longer/shorter durations	Decrease TH to 1, 5 and 50 years	Total costs and QALYs to increase/decrease in accordance with longer/shorter durations. TH of 1 year -> LY's in accordance with mortality inputs
15	Altered transition probabilities	Varies by model	- Increase the duration in induction/consolidation phase; - probability of relapse following 1st line remission at 48 months equal to 0; - probability of relapse following 1st line remission at 24 months and at 48 months equal to 0, haematological and molecular remission rates equal to 1, switch from cardiac adverse events blocked	- The corresponding probabilities in transition matrices increased; - time spent in 2nd line reduced; - time spent in 2nd line equal to 0.
16	Discount rates set to 100%	Costs and QALYs should	Discount rates set to 100%	Long term costs (HSCT and

		be significantly reduced		palliative) and QALYs were significantly reduced
17	Discount rates set to 0%	Undiscounted and discounted results should be the same	Discount rates set to 0%	Undiscounted and discounted results are the same.

AE=Adverse event; ATO=arsenic trioxide; ATRA=all-trans retinoic acid; LY=life year; QALY=quality-adjusted life year

B 3.10.2 External validation

An external validation was conducted, comparing the outcomes from the model to those observed in clinical trials at different time points (24 and 50 months) for the following:

- DFS: regarding the model structure and programming, DFS could not be computed the same way as in the clinical trials (i.e. time from CHR to relapse). In the model, relapses were not counted separately and patients commencing second-line treatment may have switched from first line due to an AE or failed to reach molecular remission after the first-line induction course. Therefore, DFS was estimated in two different manners in the model:
  - proportion of patients in first remission health states (molecular remission and +2y remission).
  - proportion of patients in all remission health states (molecular remission, +2y remission, second line molecular remission and HSCT remission)
- OS estimated as the proportion of patients alive at a given time point in the model

Results of the external validation are presented in ***Table 3.28***. The absolute difference was computed as the difference between the values estimated in the model and those observed in the APL0406 trial for each arm, while the relative difference was calculated as the difference between both arms in the model compared with the RCT. The results showed that the model did not replicate exactly the results of the RCT, which it is normal, as the structure of the model reflected real-life clinical practices rather than the RCT protocol. Nevertheless, the results seemed coherent, with relatively low “absolute differences” as none exceeded ten percentage points. Furthermore, the differences can be explained as follows:

- Regarding DFS, the model slightly overestimated it in the ATRA+ATO arm, this could be due to the fact that mortality during the treatment phases was not considered in the model. On the contrary, in the AIDA arm DFS computed from first remission states was underestimated; this could be explained by the fact that in the model patients experiencing cardiac events (5% during the induction phase, see Table 2.1 in Appendix J) switched to second line, as did patients who failed to reach molecular remission after first-line consolidation. Considering DFS estimated based on all remission states, this was slightly overestimated in the model, albeit more so in the AIDA strategy. In this case, the difference between ATRA+ATO and AIDA was reduced in the model compared to the APL0406 trial.

- Regarding OS, the model tended to overestimate it, with the relative difference reduced in the model compared to the APL0406 trial and an absolute difference in favour of AIDA. As the same tendency was observed with DFS based on all remission states, this could demonstrate conservative trends of the model.

Other clinical outcomes of the model are compared to RCT results in ***Appendix J***.

Table 3.28. External validation results

Source of trial data	Time point	Trial		CEM		Absolute difference		Relative difference	
		ATRA +ATO	AIDA	ATRA+ ATO	AIDA	ATRA +ATO	AIDA	Trial	CEM
		DFS (following first remission states)-line treatment							
Lo-Coco[1]	24 months	97	90	99.59	82.61	2.59	-7.39	7.00	16.98
Platzbecker[3]	24 months	98.3	86.8	99.59	82.61	1.29	-4.19	11.50	16.98
Platzbecker[3]	50 months	97.3	82.6	98.34	76.48	1.04	-6.12	14.70	21.86
		DFS (all remission states)							
Lo-Coco[1]	24 months	97	90	99.62	92.99	2.62	2.99	7.00	6.63
Platzbecker[3]	24 months	98.3	86.8	99.62	92.99	1.32	6.19	11.50	6.63
Platzbecker[3]	50 months	97.3	82.6	98.79	90.83	1.49	8.23	14.70	7.96
		OS							
Lo-Coco[1]	24 months	99	91	99.63	96.70	0.63	5.7	8.00	2.93
Platzbecker[3]	24 months	99.2	94.8	99.63	96.70	0.43	1.9	4.40	2.93
Platzbecker[3]	50 months	99.2	92.6	99.09	93.40	-0.11	0.8	6.60	5.69

ATO=Arsenic trioxide; ATRA=all-trans retinoic acid; EM=Cost-effectiveness model; DFS=Disease-free survival; OS=Overall survival

B 3.11 Interpretation and conclusions of economic evidence

The use of ATO in relapsed/refractory APL is well-established, while its availability to the newly-diagnosed APL population treated within the NHS could completely change the approach to the treatment of these patients in England. Therefore, the model compared ATRA+ATO versus AIDA in first line, integrating second line treatment in a single analysis.

Cost-effectiveness of ATRA+ATO was modelled with AIDA as the comparator. This choice was driven by the fact the AIDA regimen is routinely used to treat newly-

diagnosed APL patients. As ATO was until recently only used in the relapsed/refractory APL setting, using this treatment first line leaves some uncertainty as to how patients may be treated following a relapse; however, with less than 2% of patients treated with ATRA+ATO in clinical trials relapsing[2, 3], this uncertainty will affect only very few patients, so the use of AIDA as a comparator in second line is unlikely to affect cost-effectiveness estimates substantially, especially since – although the model did not explicitly compare ATRA+ATO to AIDA in second line – results of scenario analyses showed that the cost-effectiveness of ATRA+ATO against AIDA was not driven by the benefits gained in second line. As described in ***Table 1.1***, we did not consider HSCT or best supportive care as direct comparators. This was because ATRA+ATO treatment usually precedes HSCT, albeit it may also offer effective treatment without subsequent transplantation. Regarding the use of best supportive care, patients are unlikely to receive it unless they fail other treatments. Considering that with first-line ATO use the number of patients who relapse will diminish so that the number of patients concerned would be very small, we decided not to overtly complicate the analysis by modelling individual patient trajectories and did not include best supportive care as a comparator.

ATRA+ATO proved to be cost saving against AIDA while also saving additional life years and QALYs. Sensitivity analyses showed that the cost-effectiveness results were robust. In the DSA, all variations in model inputs resulted in ATRA+ATO being dominant over AIDA. PSA showed that at a willingness-to-pay threshold of £30,000/QALY, the probability of ATRA+ATO being cost-effective was almost 94%.

In addition to PSA and DSA, further scenario analyses were explored to look at specific issues or assumptions:

- Cost-effectiveness in first line only: The model did not explicitly compare ATRA+ATO to AIDA in second line; however, results of scenario analyses excluding ATO as a second-line treatment showed that, although the ICUR for ATRA+ATO vs AIDA increased slightly, it remained in favours of ATRA+ATO.Thus, the overall ICUR was not driven by the benefits gained in the well-established second-line indication. In addition, the results of the model suggest that if ATRA+ATO was used as first-line therapy, the number of patients moving to second line would decrease substantially (see ***section B 1.3.3*** for a crude estimate), so that the use of second-line treatments (including ATO) would diminish.

- HSCT: the effect that utilisation of the two HSCT types ha on the ICUR was tested in a scenario analyses where the rate of HSCT was lower. This scenario resulted in an increased ICUR but ATRA+ATO remained the dominant strategy.
- Using utilities from a published cost-effectiveness model: As we were unable to identify APL-specific utilities, we used utilities from the US-based cost-effectiveness analysis by Tallman et al.[70]. This scenario led to an increase in incremental QALYs that favoured ATRA+ATO.

Given the small indication and limited data, the utility values had to be sourced from another leukaemia field (CLL) and adjusted to APL based on age differences. Sensitivity analyses and scenario analyses showed that this limitation did not have a substantial impact on the results. The use of ATO in first line was associated with more patients achieving long-term remission, which improved survival. Thus, the number of QALYs gained was driven primarily by the number of Lys gained, rather than by the use of any specific utility values.

Overall, our cost-effectiveness analysis supports the use of ATRA+ATO on the NHS as not only a clinically-effective but also a cost-effective treatment within its full licensed indication, that is for both newly-diagnosed patients with low- to intermediate-risk APL and those patients who have relapsed or refractory disease.

B.4 References

1 Lo-Coco F, et al. Retinoic acid and arsenic trioxide for acute promyelocytic leukemia. *N Engl J Med*. 2013;369(2):111-21.

2 Burnett AK, et al. Arsenic trioxide and all-trans retinoic acid treatment for acute promyelocytic leukaemia in all risk groups (AML17): results of a randomised, controlled, phase 3 trial. *Lancet Oncol*. 2015;16(13):1295-305.

3 Platzbecker U, et al. Improved Outcomes With Retinoic Acid and Arsenic Trioxide Compared With Retinoic Acid and Chemotherapy in Non-High-Risk Acute Promyelocytic Leukemia: Final Results of the Randomized Italian-German APL0406 Trial. *J Clin Oncol*. 2017;35(6):605-12.

4 Russell NH, et al. Long Term Follow up from the NCRI AML17 Trial of Attenuated Arsenic Trioxide and ATRA Therapy for Newly Diagnosed and Relapsed Acute Promyelocytic Leukaemia. *Blood*. 2016;128(22):897.

5 Platzbecker U, et al. German Intergroup recommendations on the diagnostic and therapeutic management of acute promyelocytic leukemia (APL). Version Final 3.0. [cited Oct, 2017]. Available from: https://www.kompetenznetz-leukaemie.de/sites/kompetenznetz-leukaemie/content/e53457/e55758/e57007/NapoleonRegister_TreatmentRecommendation_Final_V3.0_20150630.pdf.

6 Lengfelder E, et al. European recommendation for salvage therapy of relapsed acute promyelocytic leukemia (APL) including arsenic trioxide (ATO). [cited Oct, 2017]. Available from: https://www.leukemia-net.org/content/leukemias/aml/apl/apl_recommendations/e6479/infoboxContent6480/APLrelapserecommendation121207.pdf.

7 Lengfelder E, et al. Arsenic trioxide-based therapy of relapsed acute promyelocytic leukemia: registry results from the European LeukemiaNet. *Leukemia*. 2015;29(5):1084-91.

8 Sanz MA, et al. Management of acute promyelocytic leukemia: recommendations from an expert panel on behalf of the European LeukemiaNet. *Blood*. 2009;113(9):1875-91.

9 Fey MF, et al. Acute myeloblastic leukaemias in adult patients: ESMO Clinical Practice Guidelines for diagnosis, treatment and follow-up. *Ann Oncol*. 2013;24 Suppl 6vi138-43.

10 Trisenox (Arsenic trioxide). Summary of Product Characteristics. Teva B.V. 2017. Available from: http://www.ema.europa.eu/docs/en_GB/document_library/EPAR_-_Product_Information/human/000388/WC500042844.pdf.

11 Food and Drug Administration. Search Orphan Drug Designations and Approvals. Arsenic trioxide. [cited Oct, 2017]. Available from: https://www.accessdata.fda.gov/scripts/opdlisting/oopd/detailedIndex.cfm?cfgridkey=109697.

12 European Medicines Agency. EU/3/00/008 Orphan designation. [cited Oct, 2017]. Available from: http://www.ema.europa.eu/ema/index.jsp?curl=pages/medicines/human/orphans/2009/11/human_orphan_000009.jsp&mid=WC0b01ac058001d12b.

13 U.S National Library of Medicine. Genetics Home Reference. Acute promyelocytic leukemia. [cited Oct, 2017]. Available from: https://ghr.nlm.nih.gov/condition/acute-promyelocytic-leukemia.

14 Arber DA, et al. The 2016 revision to the World Health Organization classification of myeloid neoplasms and acute leukemia. *Blood*. 2016;127(20):2391-405.

15 American Cancer Society. How Is Acute Myeloid Leukemia Classified? [cited Oct, 2017]. Available from: https://www.cancer.org/cancer/acute-myeloid-leukemia/detection-diagnosis-staging/how-classified.html.

16 Sant M, et al. Incidence of hematologic malignancies in Europe by morphologic subtype: results of the HAEMACARE project. *Blood*. 2010;116(19):3724-34.

17 Visser O, et al. Incidence, survival and prevalence of myeloid malignancies in Europe. *Eur J Cancer*. 2012;48(17):3257-66.

18 Dores GM, et al. Acute leukemia incidence and patient survival among children and adults in the United States, 2001-2007. *Blood*. 2012;119(1):34-43.

19 Vickers M JG, Taylor P, et al. The incidence of acute promyelocytic leukemia appears constant over most of a human lifespan, implying only one rate limiting mutation. *Leukemia*. 2000;14(4):722-6.

20 Lehmann S, et al. Continuing high early death rate in acute promyelocytic leukemia: a population-based report from the Swedish Adult Acute Leukemia Registry. *Leukemia*. 2011;25(7):1128-34.

21 Rahme R, et al. Early death in acute promyelocytic leukemia (APL) in French centers: a multicenter study in 399 patients. *Leukemia*. 2014;28(12):2422-4.

22 Rashidi A, et al. Delay in the administration of all-trans retinoic acid and its effects on early mortality in acute promyelocytic leukemia: final results of a multicentric study in the United States. *Leuk Res*. 2014;38(9):1036-40.

23 Park JH, et al. Early death rate in acute promyelocytic leukemia remains high despite all-trans retinoic acid. *Blood*. 2011;118(5):1248-54.

24 Lo-Coco F, Cicconi L. History of acute promyelocytic leukemia: a tale of endless revolution. *Mediterr J Hematol Infect Dis*. 2011;3(1):e2011067.

25 Sanz MA, et al. Definition of relapse risk and role of nonanthracycline drags for consolidation in patients with acute promyelocytic leukemia: A joint study of the PETHEMA and GIMEMA cooperative groups. *Blood*. 2000;96(4):1247-53.

26 Lo-Coco F, et al. Targeted Therapy Alone for Acute Promyelocytic Leukemia. *N Engl J Med*. 2016;374(12):1197-8.

27 Döhner H, et al. Acute Myeloid Leukemia. *New Engl J Med*. 2015;373(12):1136-52.

28 Lengfelder E, et al. Onkopedia. Akute Promyelozyten Leukämie (APL). [cited Oct, 2017]. Available from: https://www.onkopedia.com/de/onkopedia/guidelines/akute-promyelozyten-leukaemie-apl/@@view/html/index.html. German.

29 Teva Pharmaceuticals Limited. APL Market Understanding & Role of Trisenox. Data on file. May, 2015.

30 Burnett AK, et al. Inclusion of chemotherapy in addition to anthracycline in the treatment of acute promyelocytic leukaemia does not improve outcomes: Results of the MRC AML15 trial. *Leukemia*. 2013;27(4):843-51.

31 Sanz MA, et al. Risk-adapted treatment of acute promyelocytic leukemia based on all-trans retinoic acid and anthracycline with addition of cytarabine in consolidation therapy for high-risk patients: further improvements in treatment outcome. *Blood*. 2010;115(25):5137-46.

32 Lo-Coco F, et al. Front-line treatment of acute promyelocytic leukemia with AIDA induction followed by risk-adapted consolidation for adults younger than 61 years: Results of the AIDA-2000 trial of the GIMEMA Group. *Blood*. 2010;116(17):3171-9.

33 Lengfelder E, et al. High dose ara-C in the treatment of newly diagnosed acute promyelocytic leukemia: long-term results of the German AMLCG. *Leukemia*. 2009;23(12):2248-58.

34 Ades L, et al. Treatment of newly diagnosed acute promyelocytic leukemia (APL): a comparison of French-Belgian-Swiss and PETHEMA results. *Blood*. 2008;111(3):1078-84.

35 European Medicines Agency. Committee for Orphan Medicinal Products (COMP). [cited Oct, 2017]. Available from: http://www.ema.europa.eu/ema/index.jsp?curl=pages/about_us/general/general_content_000263.jsp&mid=WC0b01ac0580028e30.

36 National Institute for Health and Care Excellence. NICE Citizens Council Report Ultra Orphan Drugs. [cited Oct, 2017]. Available from: https://www.nice.org.uk/Media/Default/Get-involved/Citizens-Council/Reports/CCReport04UltraOrphanDrugs.pdf.

37 Chamberlain C, et al. Equity of access to treatment on the Cancer Drugs Fund: A missed opportunity for cancer research? *J Cancer Policy*. 2015;525-30.

38 Sanz MA, et al. Arsenic trioxide in the treatment of acute promyelocytic leukemia. A review of current evidence. *Haematologica*. 2005;90(9):1231-5.

39 Breccia M, Lo-Coco F. Arsenic trioxide for management of acute promyelocytic leukemia: current evidence on its role in front-line therapy and recurrent disease. *Expert Opin Pharmacother*. 2012;13(7):1031-43.

40 Efficace F, et al. Randomized phase III trial of retinoic acid and arsenic trioxide versus retinoic acid and chemotherapy in patients with acute promyelocytic leukemia: health-related quality-of-life outcomes. *J Clin Oncol*. 2014;32(30):3406-12.

41 Platzbecker U, et al. Improved outcome with ATRA-arsenic trioxide compared to ATRA-chemotherapy in non-high risk acute promyelocytic leukemia-updated results of the Italian-German APL0406 trial on the extended final series. *Blood Conference: 56th Annual Meeting of the American Society of Hematology, ASH*. 2014;124(21).

42 Protocol APL0406 Intergroup Study AL WP GIMEMA - DSIL [cited November, 2017]. Available from: http://www.nejm.org/doi/suppl/10.1056/NEJMoa1300874/suppl_file/nejmoa1300874_protocol.pdf.

43 Raffoux E, et al. Combined treatment with arsenic trioxide and all-trans-retinoic acid in patients with relapsed acute promyelocytic leukemia. *J Clin Oncol*. 2003;21(12):2326-34.

44 Wang J, et al. Tamibarotene compared to all-trans retinoic acid (ATRA) as add-on to arsenic trioxide (ATO) in subjects with relapsed acute promyelocytic leukemia (APL). *Blood*. 2015;126 (23)220.

45 Soignet SL, et al. United States multicenter study of arsenic trioxide in relapsed acute promyelocytic leukemia. *J Clin Oncol*. 2001;19(18):3852-60.

46 Soignet SL, et al. Complete remission after treatment of acute promyelocytic leukemia with arsenic trioxide. *N Engl J Med*. 1998;339(19):1341-8.

47 Milligan DW, et al. Guidelines on the management of acute myeloid leukaemia in adults. *Br J Haematol*. 2006;135(4):450-74.

48 Cheson BD, 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.

49 Diverio D, et al. Early detection of relapse by prospective reverse transcriptase-polymerase chain reaction analysis of the PML/RARalpha fusion gene in patients with acute promyelocytic leukemia enrolled in the GIMEMA-AIEOP multicenter "AIDA" trial. GIMEMA-AIEOP Multicenter "AIDA" Trial. *Blood*. 1998;92(3):784-9.

50 Cheson BD, et al. Report of the National Cancer Institute-sponsored workshop on definitions of diagnosis and response in acute myeloid leukemia. *J Clin Oncol*. 1990;8(5):813-9.

51 Eclache V, et al. PML/RAR alpha rearrangement in acute promyelocytic leukaemia with t(1;17) elucidated using fluorescence in situ hybridization. *Br J Haematol*. 1997;98(2):440-3.

52 Fenaux P, et al. A randomized comparison of all transretinoic acid (ATRA) followed by chemotherapy and ATRA plus chemotherapy and the role of maintenance therapy in newly diagnosed acute promyelocytic leukemia. *Blood*. 1999;94(4):1192-200.

53 Thomas X, et al. Treatment of relapsing acute promyelocytic leukemia by all-trans retinoic acid therapy followed by timed sequential chemotherapy and stem cell transplantation. APL Study Group. Acute promyelocytic leukemia. *Leukemia*. 2000;14(6):1006-13.

54 Ravandi F, et al. Effective treatment of acute promyelocytic leukemia with all-trans-retinoic acid, arsenic trioxide, and gemtuzumab ozogamicin. *J Clin Oncol*. 2009;27(4):504-10.

55 Farrington CP, Manning G. Test statistics and sample size formulae for comparative binomial trials with null hypothesis of non-zero risk difference or non-unity relative risk. *Stat Med*. 1990;9(12):1447-54.

56 Cocks K, et al. Evidence-based guidelines for determination of sample size and interpretation of the European Organisation for the Research and Treatment of Cancer Quality of Life Questionnaire Core 30. *J Clin Oncol*. 2011;29(1):89-96.

57 Centre for Reviews and Dissemination UoY, . Systematic Reviews. CRD's guidance for undertaking reviews in health care. York: York Publishing Services Ltd; 2009.

58 EORTC Quality of Life. EORTC QLQ-C30. [cited Nov, 2017]. Available from: http://groups.eortc.be/qol/eortc-qlq-c30.

59 Young T, et al. Guidelines for assessing Quality of Life in EORTC clinical trials. [cited Nov, 2017]. Available from: http://groups.eortc.be/qol/sites/default/files/archives/clinical_trials__guidelines_qol.pdf.

60 Zigmond AS, Snaith RP. The hospital anxiety and depression scale. *Acta Psychiatr Scand*. 1983;67(6):361-70.

61 Lo-Coco F, et al. Treatment of acute promyelocytic leukemia. *N Engl J Med*. 2013;369(15):1472.

62 Ma Y, et al. All-Trans Retinoic Acid plus Arsenic Trioxide versus All-Trans Retinoic Acid plus Chemotherapy for Newly Diagnosed Acute Promyelocytic Leukemia: A Meta-Analysis. *PLoS One*. 2016;11(7):e0158760.

63 Wu F, et al. Bayesian network meta-analysis comparing five contemporary treatment strategies for newly diagnosed acute promyelocytic leukaemia. *Oncotarget*. 2016;7(30):47319-31.

64 Godley LA, Larson RA. Therapy-related myeloid leukemia. *Semin Oncol*. 2008;35(4):418-29.

65 Iland HJ, et al. Have all-trans retinoic acid and arsenic trioxide replaced all-trans retinoic acid and anthracyclines in APL as standard of care. *Best Pract Res Clin Haematol*. 2014;27(1):39-52.

66 Kremer LC, et al. Anthracycline-induced clinical heart failure in a cohort of 607 children: long-term follow-up study. *J Clin Oncol*. 2001;19(1):191-6.

67 Thomas X, et al. Anthracycline-related toxicity requiring cardiac transplantation in long-term disease-free survivors with acute promyelocytic leukemia. *Ann Hematol*. 2002;81(9):504-7.

68 ClinicalTrials.gov. Study for Patients With Newly Diagnosed, High-risk Acute Promyelocytic Leukemia (TUD-APOLLO-064). [cited Nov, 2017]. Available from: https://clinicaltrials.gov/ct2/show/NCT02688140.

69 Tallman M, et al. Cost-Effectiveness Analysis of Arsenic Trioxide for The Treatment of In Patients With Acute Promyelocytic Leukemia. *Value Health*. 2015;18(3):A202.

70 Tallman M, et al. Cost-Effectiveness Analysis of Treating Acute Promyelocytic Leukemia Patients with Arsenic Trioxide and Retinoic Acid in the United States. *Clin Lymphoma, Myeloma and Leuk*. 2015;15(12):771-7.

71 Lachaine J, et al. Economic evaluation of arsenic trioxide for treatment of newly diagnosed acute promyelocytic leukaemia in Canada. *Hematol Oncol*. 2015;33(4):229-38.

72 Kruse M, et al. Budgetary impact of treating acute promyelocytic leukemia patients with first-line arsenic trioxide and retinoic acid from an Italian payer perspective. *PLoS One*. 2015;10(8):e0134587.

73 Lachaine J, et al. Economic evaluation of arsenic trioxide compared to all-trans retinoic acid + conventional chemotherapy for treatment of relapsed acute promyelocytic leukemia in Canada. *Eur J Haematol*. 2015;95(3):218-29.

74 Schonsteiner S, et al. G-DRG 2016 based cost-effectiveness analysis of arsenic trioxide (ATO) and all-trans retinoid acid (ATRA) inpatients with acute promyelocytic leukemia treated within the APL0406 trial. *Ann Hematol*. 2017;96(1):S76.

75 Holter Chakrabarty JL, et al. Autologous is superior to allogeneic hematopoietic cell transplantation for acute promyelocytic leukemia in second complete remission. *Biol Blood Marrow Transplant*. 2014;20(7):1021-5.

76 Life tables for UK, 2013-2015. 2016, [cited November 2016]. Available from: https://www.ons.gov.uk/file?uri=/peoplepopulationandcommunity/birthsdeathsandmarriages/lifeexpectancies/datasets/nationallifetablesunitedkingdomreferencetables/current/nltuk1315reg.xls.

77 Woods B, et al. Bendamustine versus chlorambucil for the first-line treatment of chronic lymphocytic leukemia in England and Wales: a cost-utility analysis. *Value Health*. 2012;15(5):759-70.

78 Beusterien KM, et al. Population preference values for treatment outcomes in chronic lymphocytic leukaemia: a cross-sectional utility study. *Health Qual Life Outcomes*. 2010;850.

79 Szende AJ, B.; Cabasés, J. M. Self-reported population health : an international perspective based on EQ-5D. Springer Netherlands; 2014.

80 Breitscheidel L. Cost utility of allogeneic stem cell transplantation with matched unrelated donor versus treatment with imatinib for adult patients with newly diagnosed chronic myeloid leukaemia. *J Med Econ*. 2008;11(4):571-84.

81 Morton RL, et al. The cost-effectiveness of induction immunosuppression in kidney transplantation. *Nephrol Dial Transplant*. 2009;24(7):2258-69.

82 Cooperberg MR, et al. Primary treatments for clinically localised prostate cancer: a comprehensive lifetime cost-utility analysis. *BJU Int*. 2013;111(3):437-50.

83 Attard CL, et al. Cost-effectiveness of folfirinox for first-line treatment of metastatic pancreatic cancer. *Curr Oncol*. 2014;21(1):e41-51.

84 Stevenson SM, et al. Cost-effectiveness of neoadjuvant chemotherapy before radical cystectomy for muscle-invasive bladder cancer. *Urol Oncol*. 2014;32(8):1172-7.

85 Choi HW, et al. Cost-effectiveness of Xpert(R) MTB/RIF for diagnosing pulmonary tuberculosis in the United States. *Int J Tuberc Lung Dis*. 2013;17(10):1328-35.

86 Prica A, et al. Combined modality therapy versus chemotherapy alone as an induction regimen for primary central nervous system lymphoma: a cost-effectiveness analysis. *Neuro Oncol*. 2014;16(10):1384-91.

87 Nshimyumukiza L, et al. Dabigatran versus warfarin under standard or pharmacogenetic-guided management for the prevention of stroke and systemic thromboembolism in patients with atrial fibrillation: a cost/utility analysis using an analytic decision model. *Thromb J*. 2013;11(1):14.

88 British National Formulary. BMJ Group and Pharmaceutical Press. 2015-2016, [cited July 2016, January 2017]. Available from: https://www.evidence.nhs.uk/formulary/bnf/current/.

89 Douer D, Tallman MS. Arsenic trioxide: new clinical experience with an old medication in hematologic malignancies. *J Clin Oncol*. 2005;23(10):2396-410.

90 NHS trusts and NHS foundation trusts. National Schedule of Reference Costs. 2014-2015.

91 NHS Blood and Transplant. BLOOD AND COMPONENTS. [cited Dec, 2017]. Available from: http://hospital.blood.co.uk/media/27988/price_list_bc_nhs_2015-16.pdf.

92 Personal Social Services Research Unit (PSSRU). Unit Costs of Health and Social Care. Curtis L., Burns A.; 2015.

93 Leunis A, et al. The costs of initial treatment for patients with acute myeloid leukemia in the Netherlands. *Leuk Res*. 2013;37(3):245-50.

94 Marie Curie Cancer Care. Understanding the cost of end of life care in different settings. [cited Dec, 2017]. Available from: https://www.mariecurie.org.uk/globalassets/media/documents/commissioning-our-services/publications/understanding-cost-end-life-care-different-settingspdf.

95 Morgan AS, A. ; Wailoo, A. . The Risk and Costs of Febrile Neutropenia in Patients with Non-Small Cell Lung Cancer Treated with Docetaxel. Available from: http://www.nicedsu.org.uk/PDFs%20of%20reports/Erlotinib%20DSU%20final%20report1.pdf.

96 Soini E, et al. Cost-effectiveness of First-line Chronic Lymphocytic Leukemia Treatments When Full-dose Fludarabine Is Unsuitable. *Clin Ther*. 2016;38(4):889-904 e14.

97 Akhtar WC, Y. Saving the NHS one blood test at a time. BMJ Quality Improvement Reports. BMJ; 2014.

98 Calhoun EA, et al. Evaluating the total costs of chemotherapy-induced toxicity: results from a pilot study with ovarian cancer patients. *Oncologist*. 2001;6(5):441-5.

99 NICE. Chronic heart failure: costing report. London: NICE; 2010.

100 NICE. Intravenous fluid therapy in adults in hospital. NICE; 2013.

101 Saito AM, et al. Costs of allogeneic hematopoietic cell transplantation with high-dose regimens. *Biol Blood Marrow Transplant*. 2008;14(2):197-207.

102 Jones CAF, L. P.; Weimersheimer, P.; Zakai, N. A.; Sharf, M.; Mesa, O. A.; Peters, C.; di Carlo, A.; Norotsky, M. C. Estimating the Burden of Cost in Chronic Graft-versus-Host Disease: A Human Capital Approach. *JHEOR*. 2016;4(2):113-8.

103 Bridges S. DM, Fuller E., Knott C., Mindell J., Moody A., Ryley A., Scholes S., Seabury C., Wardle H., Whalley R. Chapter 10: Adult anthropometric measures, overweight and obesity. Health Survey for England 2012. Vol 1. Health and Social Care Information Centre; 2013.

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

Arsenic trioxide for treating acute promyelocytic leukaemia ID446

Dear Teva,

The Evidence Review Group, Kleijnen Systematic Reviews, and the technical team at NICE have looked at the submission received on 7 December 2017 from Teva. 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 **Wednesday 17 January 2018.** Your response and any supporting documents should be uploaded to NICE.

Two versions of your written response should be submitted; one with academic/commercial-in-confidence information clearly marked and one with this information removed.

Please underline all confidential information, and separately highlight information that is submitted as commercial in confidence in turquoise, and all information submitted as academic in confidence in yellow.

If you present data that are not already referenced in the main body of your submission and that are academic/commercial in confidence, please complete the attached checklist for confidential information.

Please do not embed documents (PDFs or spreadsheets) in your response because this may result in them being lost or unreadable.

If you have any queries on the technical issues raised in this letter, please contact Kirsty Pitt, Technical Lead (kirsty.pitt@nice.org.uk). Any procedural questions should be addressed to Stephanie Yates, Project Manager (stephanie.yates@nice.org.uk).

Yours sincerely

Alex Filby
Technical Adviser – Appraisals
Centre for Health Technology Evaluation
On behalf of:

NICE National Institute for Health and Care Excellence

Level 1A
City Tower
Manchester
M1 4BT
United Kingdom

+44 (0)300 323 0140

Frances Sutcliffe
Associate Director – Appraisals
Centre for Health Technology Evaluation

Encl. checklist for confidential information

Section A: Clarification on effectiveness data

Literature searching

A1. Please provide details of the MEDLINE and Embase date ranges searched (company submission B2 and Appendix D).

A2. Please clarify which database in the Cochrane Library was searched for randomised controlled trials (RCTs) e.g. Cochrane Database of Systematic Reviews.

A3. Please provide full details for the searches for conference abstracts (American Society of Clinical Oncology, American Society of Hematology and European Hematology Association), including the specific conference proceedings searched, the search strategies, search terms used and results.

A4. Please provide details of the host interface used to search the NHS Economic Evaluation Database, and which issue of this database was searched (company submission B3 and Appendix G).

A5. Please provide details of the targeted search conducted to retrieve adverse reaction unit costs and resource use studies (company submission B3.5.3).

Included and excluded studies

A6. **Priority question**: The company conducted a review of RCTs and a review of non-RCTs of any intervention and comparator for patients with acute promyelocytic leukaemia (APL). The flow charts on pages 25 and 26 of the submission indicate that 44 RCT and 36 non-RCT publications were included and these are listed in appendix D of the company submission. However, only 3 RCTs were used as primary sources (company submission summary, table A.5.1). Please provide reasons for not using the remaining 41 'included' RCT and 36 non-RCT publications as primary sources in the company submission. Please split the studies by population (first-line vs relapsed/refractory APL), arsenic trioxide (ATO) or comparator intervention and for ATO studies describe whether ATO use was in accordance with the license.

A7. Which studies were evaluated for inclusion in a possible network meta-analysis?

NICE National Institute for Health and Care Excellence

Level 1A
City Tower
Manchester
M1 4BT
United Kingdom

+44 (0)300 323 0140

A8. Why was a minimum sample size of 50 used for non-RCTs, given the rarity of the condition?

A9. **Priority question**: In the company submission, 2 reviews (company submission p23, references 38 and 39) were mentioned to demonstrate that ATO is effective in relapsed/refractory APL. However, the primary studies cited in these reviews are not included in the submission. Furthermore, page 26 refers to "one meta-analysis that were identified through the systematic literature review." Please list all the relevant systematic reviews identified (of newly diagnosed and relapsed/refractory patients), briefly list their included studies and results and explain why they were not considered relevant for the submission.

A10. **Priority question**: The company submission states that non-RCT evidence was not sought in relapsed/refractory APL because ATO is well-established and "has long been considered first-choice therapy for induction and consolidation in this setting". Please clarify which 70 non-RCTs were excluded because ATO was not used at first line (company submission, p26 figure 2.2).

A11. **Priority question**: No RCT evidence for the effectiveness of ATO in relapsed/refractory APL was provided in the company submission (the study by Raffoux *et al.* (2003) compared all-trans retinoic acid (ATRA) plus ATO with ATO; therefore it does not include a relevant comparator). If there is no RCT with a relevant comparator, please include all relevant non-RCTs of ATO and all relevant non-RCTs of the comparators listed in the scope.

A12. Please provide a reference to support the assertion that maintenance treatment is not used in the UK. "First-line therapy in APL generally consists of three consecutive treatment phases: induction, consolidation and maintenance, although maintenance is usually omitted in the UK clinical practice with the aim of minimising the risk of treatment-related myelodysplastic syndrome or acute myeloid leukaemia (tMDS/AML)." (company submission p15) Were any trials including maintenance treatment omitted from the review?

A13. **Priority question**: On page 79 of the company submission it is stated that "The estimated overall cumulative exposure to Teva Group products containing ATO was approximately 13,855 patients, with an estimated 363 patients exposed to ATO in 6 clinical trials sponsored by Teva Group." Which 6 studies are these and in which populations are they conducted? Please provide full data if they are relevant to the current decision problem.

The included trials: APL0406, AML17 and Raffoux *et al.*

A14. **Priority question**: How many patients from the UK were recruited to each of the three included trials?

NICE National Institute for Health and Care Excellence

Level 1A
City Tower
Manchester
M1 4BT
United Kingdom

+44 (0)300 323 0140

A15. **Priority question**: Please provide data on treatment-related deaths across the three trials, if available.

A16. **Priority question**: Please provide the full APL0406 Clinical Study Report (CSR). As mentioned in the APL0406 journal publication (LoCoco *et al.* 2013, New England Journal of Medicine) arsenic trioxide was donated by Teva Pharmaceutical Industries. Therefore, it might be possible to obtain the full CSR.

A17. In APL0406, how many patients went on to receive a haematopoietic stem cell transplant (HSCT) during or after the trial?

Ongoing research

A18. Are any further analyses planned or publications in process for any of the three included trials (APL0406, AML17 and Raffoux *et al.*)? When will these be available?

A19. Please confirm when the long-term quality of life analysis of the final APL0406 cohort will be reported (p61).

Section B: Clarification on cost-effectiveness data

Model structure

B1. The company's model includes 14 health states (62 including tunnel states). However, the existing economic evaluations identified through the systematic literature review only include 4-5 health states.

a. Please clarify why the inclusion of each of these health states was necessary.

b. Please discuss the impact of this more complex model on outcomes, considering that the other existing economic evaluations resulted in positive incremental costs and lower QALY gains.

B2. **Priority question:** The company submission states that patients on treatment could be at increased risk of mortality but this was not incorporated in the company base-case analysis (company submission, p98). In the base-case model, patients in induction and consolidation phases (almost 3 years) experience the same level of mortality as the general UK population, and only patients treated with AIDA (ATRA plus idarubicin) can also die from tMDS/AML in these phases. The impact of this assumption was not explored in scenario analysis.

a. Please justify the decision to exclude disease-related mortality from the induction and consolidation phases.

b. Please include a scenario analysis in which disease-related mortality is included in the induction and consolidation phases for both treatment groups.

NICE National Institute for Health and Care Excellence

Level 1A
City Tower
Manchester
M1 4BT
United Kingdom

+44 (0)300 323 0140

B3. **Priority question:** The company submission states that some adverse events could lead to a change in treatment (company submission, p98). However, in the base-case model, only cardiac events lead to a change in treatment. There is an error in the model, where, if other adverse events prompting a change in treatment are implemented, it results in negative numbers of patients in the Markov trace.

a. Please justify why only cardiac events can trigger a change in treatment in the model. Please comment on what other adverse events may prompt a change in treatment in practice, citing the relevant literature or expert opinion.

b. Please provide a version of the model, in which the selection of all or several adverse events prompting a change in treatment is correctly implemented (i.e. does not lead to negative numbers in the Markov trace).

c. Please explain why patients in the model who have a cardiac event in second-line treatment can only receive allogeneic HSCT, not autologous HSCT.

B4. The company submission states that *"During these first 2 years of remission, the probability of relapse was higher than in the following years. In case of a relapse, patients moved to the "Second-line induction" state; otherwise, they went through all tunnel states and after 24 Markov cycles they moved to the "+2y remission" state."* (Company submission, p98). This indicates that the probability of relapse decreases over time. Please justify the use of a 2 year cut-off for the change in relapse rate, citing relevant literature and expert opinion.

Intervention and comparator

B5. **Priority question:** In the company submission, ATRA+ATO is only assessed in first line, which is not in line with the scope issued by NICE. Although the submission states that the use of second-line treatment will decrease if ATO is used as first-line treatment (company submission, p98), it is plausible that the cost effectiveness of second-line treatment options will have an impact on the cost effectiveness of first-line treatments.

a. Please provide an analysis where ATRA+ATO is compared with AIDA in the second-line setting (the relapsed/refractory population). This could be implemented by removing second-line treatment and moving patients straight to HSCT (instead of second-line treatment).

b. Please provide an analysis with only ATO as the intervention (without ATRA) in the second-line setting.

c. Please provide an analysis with best supportive care as a comparator in the second-line setting.

Effectiveness

B6. **Priority question:** Please provide an overview of all transition probabilities that are used in the Excel model (i.e. in the Markov trace), including sources, description of

Level 1A
City Tower
Manchester
M1 4BT
United Kingdom

+44 (0)300 323 0140

any calculations performed and justification (for both the calculations and the source) for the parameters related to sections 3.3.3.1-3.3.3.5 and 3.3.4.

B7. **Priority question:** Please justify the extrapolation of the long-term relapse in the model for both patients in first remission (i.e. after 50 months) as well as patients in second remission. Moreover, please provide scenarios using alternative assumptions for extrapolating the long-term relapse probabilities.

B8. The probability of relapse for patients in second remission is retrieved from Tallman *et al.* 2015. Please provide details of how this probability was obtained, whether any alternative sources for this probability are available (for example the study by Raffoux *et al.*) provide justification for both the data source used and any calculations performed.

B9. The company assumed that patients in the model who did not undergo polymerase chain reaction (PCR) testing (and therefore whose molecular remission status could not be evaluated) were in remission. Please justify this assumption and explain its implications, given that it was also assumed that none of the patients who received ATRA + ATO as first line treatment underwent PCR testing in the model – was it assumed that these patients were in molecular remission?

Quality of life

B10. Some (not all) health state utility values are adjusted for both age and for the 'utility representing perfect health'. This method of adjustment results in counter-intuitive utility values, i.e. utility values for 45 year olds that are lower than utility values for 60 years olds with the same condition. Additionally, the 'allogeneic HSCT' and 'autologous HSCT' health state utility value obtained from Breitscheidel (2008) is adjusted by multiplying it by the 'second-line molecular remission' health state utility value.

a. Please justify why an age adjustment is necessary at all, given that the impact of this disease would far outweigh any age-related utility decrements.

b. Please justify why some utilities are adjusted for the 'utility representing perfect health'.

c. Please justify why the 'allogeneic HSCT' and 'autologous HSCT' health state utility value is adjusted by the 'second-line molecular remission' health state utility value.

d. Provide a scenario analysis without any of these utility value adjustments.

B11. Health state utility values in the cost-effectiveness analysis remain constant during the entire time horizon. However, when patients flow through the model, they also become older resulting in a decreased utility value. Please provide a scenario analysis implementing health state utility values that do not exceed general population utility values (e.g. by capping all health state utility values to the age-dependent general population utility values).

NICE National Institute for Health and Care Excellence

Level 1A
City Tower
Manchester
M1 4BT
United Kingdom

+44 (0)300 323 0140

B12. The choice of the sources for the different utility and disutility values are not always clearly described in the company submission. Additionally, it is unclear how studies informing the disutility values for the adverse events were identified.

a. Please describe how the studies informing the disutility values of adverse events were identified.

b. Please explain the choice of each source used to obtain health state utility values and disutilities for adverse events.

c. 'Hospitalisation', 'leukocytosis', 'differentiation syndrome' and 'QTc prolongation' disutility values are based on assumptions – please justify these assumptions as well as the disutility values. Please refer to relevant sources or expert opinion.

d. Please provide an overview of the health states in which patients are at risk of hospitalisation in the cost effectiveness model.

e. Please justify why acute and chronic graft versus host disease (GvHD) have the same disutility value

f. Please justify why a multiplicative framework is used to estimate the disutility for 'Acute GvHD'.

Adverse events

B13.

a. Please provide an overview of the adverse events implemented in the model, their frequency of occurrence, their duration and the sources on which this information is based. Please justify the choice of each source.

b. Please explain whether it is assumed that GvHD is reversible in the model and why?

Costs and resource use

B14. Please justify why PCR tests are not performed in the molecular remission health state for patients who received first-line treatment with ATRA+ATO. Please add a positive number of PCR tests into the model for the ATRA+ATO arm, if necessary.

B15. A systematic literature review was performed to collect appropriate costs and resource use data for England. Nevertheless, none of the 11 identified studies is used in the economic model because *"information captured was not compatible with that needed to populate the model, and in others by the fact NHS reference costs were preferentially used to ensure relevance to the current situation in England"* (CS page 124). Please provide more specific justification, for each resource use and cost item, why the chosen sources are the most appropriate.

NICE National Institute for Health and Care Excellence

Level 1A
City Tower
Manchester
M1 4BT
United Kingdom

+44 (0)300 323 0140

B16. The health state costs include costs related to follow-up appointments and PCR tests for the whole time horizon of the model

a. Please justify the inclusion of life-long costs related to follow-up appointments and PCR tests and provide a scenario with alternative assumptions related to this (for example, excluding these costs after a certain time in remission).

b. Please explain why costs related to monitoring of haematological response were not included in the model.

Validation and transparency

B17. **Priority question:** Please provide the experts' responses (for example meeting minutes) to all the questions asked in the presentation shown in Appendix M.

B18. Please provide a cross-validation of the assumptions, inputs and outputs in the cost-effectiveness model, with the cost-effectiveness analyses identified in the systematic literature review.

Scenario and sensitivity analyses

B19. **Priority question:** Please provide a model which includes an option to select the scenario analyses performed (both those in the original company submission as well as those performed in response to the clarification letter) by the company.

a. Please also provide a scenario analysis implementing maintenance treatment.

b. Please also provide a scenario analysis implementing a Gamma distribution for resource use (the Normal distribution that is currently used might result in negative numbers for resource use).

c. Please also provide a scenario analysis implementing the consolidation health state with 26 cycles (reflecting a maximum of 2 years in that health state), instead of the 24 cycles used. (The cycle length in the model is 4 weeks, meaning that one year contains 13 cycles. However, the 2-year consolidation health state consists of 24 cycles.)

B20. **Priority question:** At the end of the time horizon of 40 years chosen by the company, 45% of patients in the ATRA+ATO first and second line arm (48% of patients in the ATRA+ATO first line and AIDA second line) are still alive. The chosen time horizon is therefore not a lifetime horizon.

a. Please justify the long life expectancy of patients in the model (a proportion of patients are still alive at ages of 100+ years) and comment on the plausibility of this.

b. Please provide results with a lifetime time horizon that captures all relevant outcomes.

NICE National Institute for Health and Care Excellence

Level 1A
City Tower
Manchester
M1 4BT
United Kingdom

+44 (0)300 323 0140

B21. The CS states that *"For patients in first molecular remission, the probability of relapse at 48 months was assumed to be equal to that at 50 months" (page 105)*. Please clarify why these were assumed to be equal instead of converting the 50 month probability to a 48 month probability of relapse.

Section A: Clarification on effectiveness data

Literature searching

A1. Please provide details of the MEDLINE and Embase date ranges searched (company submission B2 and Appendix D).

Company response:

- Initial search: 1968–20.07.2016
- Updated search: 2016–10.10.2017

A2. Please clarify which database in the Cochrane Library was searched for randomised controlled trials (RCTs) e.g. Cochrane Database of Systematic Reviews.

Company response: It was the Cochrane Central Register of Controlled Trials, as stated in the submission.

A3. Please provide full details for the searches for conference abstracts (American Society of Clinical Oncology, American Society of Hematology and European Hematology Association), including the specific conference proceedings searched, the search strategies, search terms used and results.

Company response: The first search was performed on July 20th, 2016 and the update on October 10th, 2017. For the update, the search terms: “apl” and “acute promyelocytic leukemia” were used. Time restrictions applied when searching for ASCO and EHA abstracts were 2016–2017. For the ASH abstract database, we restricted the search to 2015–2017. These time restrictions were based on the fact the initial search did not capture all abstracts published in 2016 (ASCO, EHA) and 2015 (ASH). The search of the ASCO, ASH and EHA databases yielded 22, 157 and 118 conference abstracts, respectively. All of the retrieved abstracts were screened and relevant ones were included in our review.

A4. Please provide details of the host interface used to search the NHS Economic Evaluation Database, and which issue of this database was searched (company submission B3 and Appendix G).

Company response: The National Institute for Health Research (NIHR) Centre for Reviews and Dissemination (CRD) interface was used to search the CRD Database, CRD assessed economic evaluation (bibliographic) and CRD assessed economic evaluation (full abstract).

A5. Please provide details of the targeted search conducted to retrieve adverse reaction unit costs and resource use studies (company submission B3.5.3).

Company response: Please see the details provided in response to question B12.

Included and excluded studies

A6. **Priority question**: The company conducted a review of RCTs and a review of non-RCTs of any intervention and comparator for patients with acute promyelocytic leukaemia (APL). The flow charts on pages 25 and 26 of the submission indicate that 44 RCT and 36 non-RCT publications were included and these are listed in appendix D of the company submission. However, only 3 RCTs were used as primary sources (company submission summary, table A.5.1). Please provide reasons for not using the remaining 41 'included' RCT and 36 non-RCT publications as primary sources in the company submission. Please split the studies by population (first-line vs relapsed/refractory APL), arsenic trioxide (ATO) or comparator intervention and for ATO studies describe whether ATO use was in accordance with the license.

Company response: As we were hoping a network meta-analysis would be possible for the purpose of this submission, and the systematic search and review aimed to extract studies that would inform it. The search for both RCTs and non-RCTs was very broad with no restrictions applied regarding intervention or comparators. For non-RCTs the restriction to first-line ATO studies was applied only after the initial screening of abstracts and full-texts. Hence, the search returned a variety of publications testing a range of interventions in APL against a large number of comparators.

For the submission itself, we presented evidence from RCTs that used ATO alone or in combination with ATRA. We felt that with second-line ATO use being widespread in clinical practice, the submission should focus on strong RCT-based evidence supporting first-line ATO use, that is on the APL0406 and AML17 trials, providing enough details to allow NICE to make a well-informed decision on this recently-approved indication of ATO. Overall, we presented RCT evidence supporting the use of ATO in APL (first- or second-line). However, we excluded RCTs where ATO (with or without ATRA) was used in addition to chemotherapy (e.g. the US-based study by Powell, et al.[1]), since the approved indication for ATO does not include this type of combination. We also excluded RCTs conducted in China from the main submission, presenting them instead in Appendix L. The rationale for this was twofold, as described in the submission. First, APL treatment paradigms in China are likely to vary substantially from European ones. Second, Trisenox® is not marketed in China, so any studies from the country most likely used a different ATO formulation.

A7. Which studies were evaluated for inclusion in a possible network meta-analysis?

Company response: Two steps were proposed to assess the feasibility of a network meta-analysis:

1. Study and treatment characteristics:

- Line of treatment: First line
- Therapy phase: Induction, consolidation ± maintenance
- APL risk level: Low/intermediate

Thirty-five out of 44 studies were excluded from the NMA based on the aforementioned criteria. The remaining 9 studies were assessed in the second step, based on the outcomes presented.

2. Outcome characteristics

- Burnett, A.K., et al., Arsenic trioxide and all-trans retinoic acid treatment for acute promyelocytic leukaemia in all risk groups (AML17): Results of a randomised, controlled, phase 3 trial. The Lancet Oncology, 2015. 16(13): p. 1295-1305.
- Lo-Coco, F., et al., Retinoic acid and arsenic trioxide for acute promyelocytic leukemia. New England Journal of Medicine, 2013. 369(2): p. 111-121.
- Ades, L., et al., Long-term follow-up of European APL 2000 trial, evaluating the role of cytarabine combined with ATRA and Daunorubicin in the treatment of nonelderly APL patients. American Journal of Hematology, 2013. 88(7): p. 556-559.
- Powell, B.L., et al., Arsenic trioxide improves event-free and overall survival for adults with acute promyelocytic leukemia: North American Leukemia Intergroup Study C9710. Blood, 2010. 116(19): p. 3751-3757.
- Asou, N., et al., A randomized study with or without intensified maintenance chemotherapy in patients with acute promyelocytic leukemia who have become negative for PML-RARalpha transcript after consolidation therapy: The Japan Adult Leukemia Study Group (JALSG) APL97 study. Blood, 2007. 110(1): p. 59-66.
- Platzbecker, U., et al., Improved Outcomes With Retinoic Acid and Arsenic Trioxide Compared With Retinoic Acid and Chemotherapy in Non-High-Risk Acute Promyelocytic Leukemia: Final Results of the Randomized Italian-German APL0406 Trial. Journal of Clinical Oncology, 2016. 11: p. 11.
- Ades, L., et al., Is cytarabine useful in the treatment of acute promyelocytic leukemia? Results of a randomized trial from the European acute promyelocytic leukemia group. Journal of Clinical Oncology, 2006. 24(36): p. 5703-5710.
- Li, J., et al., Retinoic acid and arsenic trioxide with or without chemotherapy for acute promyelocytie leukemia with different risk stratifications: A interim analysis of China APL 2012 study. Blood, 2016. 128(22): p. 445.
- Russell, N.H., et al., Long term follow up from the NCRI AML17 trial of attenuated arsenic trioxide and ATRA therapy for newly diagnosed and relapsed acute promyelocytic leukaemia. Blood, 2016. 128(22): p. 897.

A8. Why was a minimum sample size of 50 used for non-RCTs, given the rarity of the condition?

Company response: Few RCTs in APL were identified, so that a sample size cut-off could not be justified. For non-RCTs, however, we restricted the sample size to 50 at minimum. While this cut-off appears arbitrary, it is worth noting that Central Limit Theorem cannot be applied to studies with fewer than 30 patients (per arm), limiting their statistical validity and the interpretation of their results. Thus, by selecting larger studies

we aimed to ensure that only the most informative and highest quality non-RCTs are analysed.

A9. **Priority question**: In the company submission, 2 reviews (company submission p23, references 38 and 39) were mentioned to demonstrate that ATO is effective in relapsed/refractory APL. However, the primary studies cited in these reviews are not included in the submission. Furthermore, page 26 refers to "one meta-analysis that were identified through the systematic literature review." Please list all the relevant systematic reviews identified (of newly diagnosed and relapsed/refractory patients), briefly list their included studies and results and explain why they were not considered relevant for the submission.

Company response: Eight systematic reviews were identified through the literature search; their characteristics are presented in *We* screened the reference lists of these reviews against the list of studies included in our review, to check if they had been identified during our search process. Studies that had not already been identified were screened using the inclusion/exclusion criteria described in the main submission and Appendix D. Of the studies cited in the 8 relevant reviews, 35 had previously been identified and assessed in our systematic review process. The remaining studies (over 40) had not been previously assessed. When reviewed for eligibility, most were excluded since they did not concern the relevant population or did not employ a study design of interest. Studies published in non-core journals in China were also excluded, due to their generally inferior quality to core journal publications.

Table 1. We screened the reference lists of these reviews against the list of studies included in our review, to check if they had been identified during our search process. Studies that had not already been identified were screened using the inclusion/exclusion criteria described in the main submission and Appendix D. Of the studies cited in the 8 relevant reviews, 35 had previously been identified and assessed in our systematic review process. The remaining studies (over 40) had not been previously assessed. When reviewed for eligibility, most were excluded since they did not concern the relevant population or did not employ a study design of interest. Studies published in non-core journals in China were also excluded, due to their generally inferior quality to core journal publications.

Table 1. Systematic reviews identified through the literature search

	Review
1.	Muchtar E, Vidal L, Ram R, Gafter-Gvili A, Shpilberg O, Raanani P. The role of maintenance therapy in acute promyelocytic leukemia in the first complete remission. Cochrane Database Syst Rev. 2013;(3):CD009594
2.	Fenfang Wu, Di Wu, Yong Ren, Chongyang Duan, Shangwu Chen, and Anlong Xu. Bayesian network meta-analysis comparing five contemporary treatment strategies for newly diagnosed acute promyelocytic leukaemia. Oncotarget. 2016; 7(30): 47319–47331
3.	Estcourt LJ, Desborough M, Brunskill SJ, Doree C, Hopewell S, Murphy MF, Stanworth SJ. Antifibrinolytics (lysine analogues) for the prevention of bleeding in people with haematological disorders. Cochrane Database Syst Rev. 2016;3:CD009733

4. Wang H, Chen XY, Wang BS, Rong ZX, Qi H, Chen HZ. The efficacy and safety of arsenic trioxide with or without all-trans retinoic acid for the treatment of acute promyelocytic leukemia: a meta-analysis. Leuk Res. 2011;35(9):1170-7

5. Kamimura T, Miyamoto T, Harada M, Akashi K. Advances in therapies for acute promyelocytic leukemia. Cancer Sci. 2011;102(11):1929-37.

6. Fenaux P, Chastang C, Chomienne C, Castaigne S, Sanz M, Link H, Löwenberg B, Fey M, Archim-Baud E, Degos L, et al. Treatment of newly diagnosed acute promyelocytic leukemia (APL) by all transretinoic acid (ATRA) combined with chemotherapy: The European experience. European APL Group. Leuk Lymphoma. 1995;16(5-6):431-7

7. Ades, L and Fenaux, P. Recent Results of the French Belgian Swiss Acute Promyelocytic Leukemia (APL) Group. Annals of Hematology. 2011; 90. S45-S48.

8. Huang J, Sun M, Wang Z, Zhang Q, Lou J, Cai Y, Chen W, Du X. Induction treatments for acute promyelocytic leukemia: a network meta-analysis. Oncotarget. 2016;7(44):71974-71986

A10. **Priority question**: The company submission states that non-RCT evidence was not sought in relapsed/refractory APL because ATO is well-established and “has long been considered first-choice therapy for induction and consolidation in this setting”. Please clarify which 70 non-RCTs were excluded because ATO was not used at first line (company submission, p26 figure 2.2).

Company response: The list of 70 studies excluded based on ATO not being used first-line is provided in *Table 2* below.

Table 2. List of non-RCTs excluded based on ATO not being used first-line.

1	Daniela G. Guolo F. Minetto P. Clavio M. Giannoni L. Coviello E. Pastori G. Rivoli G. Ballerini F. Colombo N. Grasso R. Miglino M. Lemoli R.M. Gobbi M. *High prognostic value of a 3-genes molecular score in non-high risk acute promyelocytic leukemia treated with AIDA 2000 protocol: A retrospective study from a regional register*. Blood. Conference: 57th Annual Meeting of the American Society of Hematology, ASH 2015 San Diego, CA United States. Conference Start: 20161203 Conference End: 20161206. Conference Publication: (var.pagings). 126 (23) (pp 2601), 2015. Date of Publication: 03 Dec 2015.
2	Hecht A. Nowak D. Nolte F. Nowak V. Oblaender J. Buechner T. Spiekermann K. Hofmann W.-K. Lengfelder E. *Different impact of expression levels of IGFBP2 and IGFBP7 on survival and relapse rate in acute promyelocytic leukemia.* Blood. Conference: 57th Annual Meeting of the American Society of Hematology, ASH 2015 San Diego, CA United States. Conference Start: 20161203 Conference End: 20161206. Conference Publication: (var.pagings). 126 (23) (pp 1382), 2015. Date of Publication: 03 Dec 2015.
3	Lu Y. Li F. Mu Q. Meng H. Qian W. Tong H. Mai W. Pei R. Yu M. Zhao X. Jin J. *The clinical efficacy of all-trans retinoic acid plus arsenic trioxide in 177 newly diagnosed acute promyelocytic leukemia patients.* [Chinese] Zhonghua xue ye xue za zhi = Zhonghua xueyexue zazhi. 36 (5) (pp 372-377), 2015. Date of Publication: 01 May 2015.
4	Rollig C. Schafer-Eckardt K. Hanel M. Kramer M. Schaich M. Thiede C. Oelschlagel U. Mohr B. Wagner T. Einsele H. Krause S.W. Bodenstein H. Martin S. Stuhlmann R. Ho A.D. Bornhauser M. Ehninger G. Schuler U. Platzbecker U. *Two cycles of risk-adapted consolidation therapy in patients with acute promyelocytic leukemia. Results from the SAL-AIDA2000 trial.* Annals of Hematology. 94 (4) (pp 557-563), 2015. Date of Publication: 2015.
5	Mitrovic M. Suvajdzic N. Elezovic I. Bogdanovic A. Djordjevic V. Miljic P. Djunic I. Gvozdenov M. Colovic N. Virijevic M. Lekovic D. Vidovic A. Tomin D. *Thrombotic events in acute promyelocytic leukemia.* Thrombosis Research. 135 (4) (pp 588-593), 2015. Date of Publication: 01 Apr 2015.
6	Gill H. Leung A. Tse E. Kwong Y.L. *Relapse characteristics and risk factors for central nervous system involvement in acute promyelocytic leukaemia in the oral arsenic trioxide era-a 13-year follow-up study*. Haematologica. Conference: 20th Congress of the European Hematology Association Vienna Austria. Conference Start: 20150611 Conference End: 20150614. Conference Publication: (var.pagings). 100 (pp 42), 2015. Date of Publication: 22 Jun 2015.

7	Albano F. Zagaria A. Anelli L. Orsini P. Minervini C.F. Impera L. Casieri P. Coccaro N. Tota G. Brunetti C. Minervini A. Pastore D. Carluccio P. Mestice A. Cellamare A. Specchia G. *Lymphoid enhancer binding factor-1 (LEF1) expression as a prognostic factor in adult acute promyelocytic leukemia.* Oncotarget. 5 (3) (pp 649-658), 2014. Date of Publication: 15 Feb 2014.
8	Lucena-Araujo A.R. Kim H.T. Jacomo R.H. Melo R.A. Bittencourt R. Pasquini R. Pagnano K. Fagundes E.M. de Lourdes Chauffaille M. Chiattone C.S. Lima A.S. Kwaan H.C. Gallagher R. Niemeyer C.M. Schrier S.L. Tallman M.S. Grimwade D. Ganser A. Berliner N. Ribeiro R.C. Lo-Coco F. Lowenberg B. Sanz M.A. Rego E.M. *Prognostic impact of KMT2E transcript levels on outcome of patients with acute promyelocytic leukaemia treated with all-trans retinoic acid and anthracycline-based chemotherapy: An International Consortium on Acute Promyelocytic Leukaemia study*. British Journal of Haematology. 166 (4) (pp 540-549), 2014. Date of Publication: August 2014.
9	Fujita H. Asou N. Iwanaga M. Hyo R. Nomura S. Kiyoi H. Okada M. Inaguma Y. Matsuda M. Yamauchi T. Ohtake S. Izumi T. Nakaseko C. Ishigatsubo Y. Shinagawa K. Takeshita A. Miyazaki Y. Ohnishi K. Miyawaki S. Naoe T. *Role of hematopoietic stem cell transplantation for relapsed acute promyelocytic leukemia: A retrospective analysis of JALSG-APL97.* Cancer Science. 104 (10) (pp 1339-1345), 2013. Date of Publication: October 2013.
10	Hecht A. Nowak D. Nowak V. Hanfstein B. Faldum A. Buchner T. Spiekermann K. Sauerland C. Lengfelder E. Hofmann W.K. Nolte F. *High expression of the Ets-related gene (ERG) is an independent prognostic marker for relapse-free survival in patients with acute promyelocytic leukemia*. Annals of hematology. 92 (4) (pp 443-449), 2013. Date of Publication: Apr 2013.
11	Grisariu S. Spectre G. Kalish Y. Gatt M.E. *Increased risk of central venous catheter-associated thrombosis in acute promyelocytic leukemia: A single-institution experience.* European Journal of Haematology. 90 (5) (pp 397-403), 2013. Date of Publication: May 2013.
12	Feng J. Liu J. Pang L. Wen J. Zhong F. Zhang Q. Meng Q. *Effects of pirarubicin in the consolidation chemotherapy of acute promyelocytic leukemia.* [Chinese] Chinese Journal of Clinical Oncology. 39 (12) (pp 861-863+870), 2012. Date of Publication: 30 Jun 2012.
13	Mcclellan J.S. Kohrt H.E. Coutre S. Gotlib J.R. Majeti R. Alizadeh A.A. Medeiros B.C. *Treatment advances have not improved the early death rate in acute promyelocytic leukemia*. Haematologica. 97 (1) (pp 133-136), 2012. Date of Publication: 01 Jan 2012.
14	Elliott M.A. Letendre L. Tefferi A. Hogan W.J. Hook C. Kaufmann S.H. Pruthi R.K. Pardanani A. Begna K.H. Ashrani A.A. Wolanskyj A.P. Al-Kali A. Litzow M.R. *Therapy-related acute promyelocytic leukemia: Observations relating to APL pathogenesis and therapy.* European Journal of Haematology. 88 (3) (pp 237-243), 2012. Date of Publication: March 2012.
15	Au W.-Y. Kumana C.R. Lee H.K.K. Lin S.-Y. Liu H. Yeung D.Y.M. Lau J.S.M. Kwong Y.-L. *Oral arsenic trioxide-based maintenance regimens for first complete remission of acute promyelocytic leukemia: A 10-year follow-up study*. Blood. 118 (25) (pp 6535-6543), 2011. Date of Publication: 15 Dec 2011.
16	Wiernik P.H. Sun Z. Gundacker H. Dewald G. Slovak M.L. Paietta E. Kim H.T. Appelbaum F.R. Cassileth P.A. Tallman M.S. *Prognostic implications of additional chromosome abnormalities among patients with de novo acute promyelocytic leukemia with t(15;17).* Medical Oncology. 29 (3) (pp 2095-2101), 2012. Date of Publication: September 2012.
17	Yuan Y.-H. Wu D.-P. Ouyang J. *Clinical study of elderly patients with acute promyelocytic leukemia*. [Chinese] Journal of Leukemia and Lymphoma. 19 (11) (pp 672-674), 2010. Date of Publication: November 2010.
18	Beitinjaneh A. Jang S. Roukoz H. Majhail N.S. *Prognostic significance of FLT3 internal tandem duplication and tyrosine kinase domain mutations in acute promyelocytic leukemia: A systematic review.* Leukemia Research. 34 (7) (pp 831-836), 2010. Date of Publication: July 2010.
19	Batzios C. Hayes L.A. He S.Z. Quach H. McQuilten Z.K. Wall M. Campbell L.J. *Secondary clonal cytogenetic abnormalities following successful treatment of acute promyelocytic leukemia.* American Journal of Hematology. 84 (11) (pp 715-719), 2009. Date of Publication: November 2009.
20	Cassinat B. de Botton S. Kelaidi C. Ades L. Zassadowski F. Guillemot I. Schlageter M.-H. Raffoux E. Harousseau J.-L. Legrand O. Escoffre-Barbe M. Reman O. Gardembas M. Himberlin C. Cahn J.Y. Guyotat D. Bouscary D. Parry A. Rousselot P. Baruchel A. Dombret H. Chevret S. Fenaux P. Chomienne C. *When can real-time quantitative RT-PCR effectively define molecular relapse in acute promyelocytic leukemia patients?* (Results of the French Belgian Swiss APL Group). Leukemia Research. 33 (9) (pp 1178-1182), 2009. Date of Publication: September 2009.
21	Liang J.Y. Wu D.P. Liu Y.J. Ma Q.F. Gong J.X. Zhu M.Q. Xue Y.Q. Chen Z.X. *The clinical and laboratory features of acute promyelocytic leukemia: an analysis of 513 cases.* [Chinese] Zhonghua nei ke za zhi [Chinese journal of internal medicine]. 47 (5) (pp 389-392), 2008. Date of Publication: May 2008.

22 Thomas X. Pigneux A. Raffoux E. Huguet F. Caillot D. Fenaux P. *Superiority of an arsenic trioxide-based regimen over a historic control combining all-trans retinoic acid plus intensive chemotherapy in the treatment of relapsed acute promyelocytic leukemia*. Haematologica. 91 (7) (pp 996-997), 2006. Date of Publication: July 2006.

23 Lee S. Kim Y.-J. Eom K.-S. Min C.-K. Kim H.-J. Cho S.-G. Lee J.-W. Min W.-S. Kim C.-C. *The significance of minimal residual disease kinetics in adults with newly diagnosed PML-RARalpha-positive acute promyelocytic leukemia: Results of a prospective trial.* Haematologica. 91 (5) (pp 671-674), 2006. Date of Publication: May 2006.

24 Callens C. Chevret S. Cayuela J.-M. Cassinat B. Raffoux E. de Botton S. Thomas X. Guerci A. Fegueux N. Pigneux A. Stoppa A.-M. Lamy T. Rigal-Huguet F. Vekhoff A. Meyer-Monard S. Ferrand A. Sanz M. Chomienne C. Fenaux P. Dombret H. Cailles Desablens Gardembas Hunault Ifrah Martin Corront Dor Sutton Pulik Lepeu Renoux Deconinck Ades Gardin Casassus Fenaux Pigneux Boiron Reiffers Abgrall Berthou Reman Leporrier Deveaux Salles de Revel Nedellec Plagne Legros Travade Audhuy Decaudin Dutel Pautas Cordonnier Reyes Caillot Guy Cahn Sotto Durand Solal-Celigny Tertian Morel Nelken de Botton Turlure Bordessoule Thomas Archimbaud Bastion Michallet Fiere Coiffier Philippe Stoppa Bouabdallah Vey Blaise Benothman Allard Christian Margueritte Fegueux Rossi Donadio Ojeda Henon Guerci Witz Harousseau Pesce Gratecos Cassuto Schoenwald Dreyfus Vilmer Marie Vekhoff *Prognostic implication of FLT3 and Ras gene mutations in patients with acute promyelocytic leukemia (APL): A retrospective study from the European APL Group*. Leukemia. 19 (7) (pp 1153-1160), 2005. Date of Publication: July 2005.

25 Abdulkadyrov K.M. Kuraleva V.V. Popova T.I. Kharchenko M.F. Balashova V.A. Perekatova T.N. Katsadze I.L. Martynkevich I.S. Balakova N.I. *Retrospective analysis of the morphological, cytochemical, cytogenetic, electron microscopic, and biochemical features of blast cells in acute promyelocytic leukemia.* [Russian] Klinicheskaia meditsina. 82 (8) (pp 51-56), 2004. Date of Publication: 2004.

26 Grimwade D. *The significance of minimal residual disease in patients with t(15;17). Best practice & research*. Clinical haematology. 15 (1) (pp 137-158), 2002. Date of Publication: Mar 2002.

27 Pulsoni A. Pagano L. Lo Coco F. Avvisati G. Mele L. Di Bona E. Invernizzi R. Leoni F. Marmont F. Mele A. Melillo L. Nosari A.M. Pogliani E.M. Vignetti M. Visani G. Zagonel V. Leone G. Mandelli F. *Clinicobiological features and outcome of acute promyelocytic leukemia occurring as a second tumor: The GIMEMA experience.* Blood. 100 (6) (pp 1972-1976), 2002. Date of Publication: 15 Sep 2002.

28 Jung J.I. Choi J.E. Hahn S.T. Min C.K. Kim C.C. Park S.H. *Radiologic features of all-trans-retinoic acid syndrome.* American Journal of Roentgenology. 178 (2) (pp 475-480), 2002. Date of Publication: 2002.

29 Sanz M.A. Lo Coco F. Martin G. Avvisati G. Rayon C. Barbui T. Diaz-Mediavilla J. Fioritoni G. Gonzalez J.D. Liso V. Esteve J. Ferrara F. Bolufer P. Bernasconi C. Gonzalez M. Rodeghiero F. Colomer D. Petti M.C. Ribera J.M. Mandelli F. *Definition of relapse risk and role of nonanthracycline drags for consolidation in patients with acute promyelocytic leukemia: A joint study of the PETHEMA and GIMEMA cooperative groups*. Blood. 96 (4) (pp 1247-1253), 2000. Date of Publication: 15 Aug 2000.

30 Slater S. Carter M. Howe K. Amess J. Rohatiner A.Z.S. Lister T.A. *Acute promyelocytic leukaemia: A retrospective analysis of patients treated at St. Bartholomew's Hospital 1969-1995.* Annals of Hematology. 78 (3) (pp 131-137), 1999. Date of Publication: March 1999.

31 Lee K.H. Kim D.H. Lee J.S. Suh C.W. Kim S.W. Kim S.B. Lee J.H. Doh B.S. Chi H.S. Lee M.S. *Acute promyelocytic leukemia is a distinct subset of acute myelocytic leukemia with unique clinicopathologic characteristics including longer duration of relapse free survival: experience in 13 cases*. Journal of Korean medical science. 9 (6) (pp 437-443), 1994. Date of Publication: Dec 1994.

32 Thomas X. Archimbaud E. Treille-Ritouet D. Fiere D. *Prognostic factors in acute promyelocytic leukemia: A retrospective study of 67 cases.* Leukemia and Lymphoma. 4 (4) (pp 249-256), 1991. Date of Publication: 1991.

33 Davey F.R. Davis R.B. MacCallum J.M. Nelson D.A. Mayer R.J. Ball E.D. Griffin J.D. Schiffer C.A. Bloomfield C.D. *Morphologic and cytochemical characteristics of acute promyelocytic leukemia*. American Journal of Hematology. 30 (4) (pp 221-227), 1989. Date of Publication: 1989.

34 Rashidi A Riley M, Goldin TA, Sayedian F Bayerl MG Aguilera NS Vos JA Goudar RK Fisher SI. *Delay in the administration of all-trans retinoic acid and its effects on early mortality in acute promyelocytic leukemia: final results of a multicentric study in the United States*. Leukemia Research. 38(9):1036-40, 2014 Sep.

35 Lou Y Suo S Tong Y Tong H Qian W Meng H Mai W Huang J Yu W Jin J *Outcomes and prognostic factors of first relapsed acute promyelocytic leukemia patients undergoing salvage therapy with intravenous arsenic trioxide and chemotherapy*. Annals of Hematology. 93(6):941-8, 2014 Jun.

36 Montesinos P Gonzalez JD Gonzalez J Rayon C de Lisa E Amigo ML Ossenkoppele GJ Penarrubia MJ Perez-Encinas M Bergua J Deben G Sayas MJ de la Serna J Ribera JM Bueno J Milone G Rivas C Brunet S Lowenberg B Sanz M *Therapy-related myeloid neoplasms in patients with acute promyelocytic leukemia treated with all-trans-retinoic Acid and anthracycline-based chemotherapy*. Journal of Clinical Oncology. 28(24):3872-9, 2010 Aug 20.

37 Chillon MC Santamaria C Garcia-Sanz R Balanzategui A Sarasquete ME Alcoceba M Marin L Caballero MD Vidriales MB Ramos F Bernal T Diaz-Mediavilla J Garcia de Coca A Penarrubia MJ Queizan JA Giraldo P San Miguel JF Gonzalez *M Long FLT3 internal tandem duplications and reduced PML-RARalpha expression at diagnosis characterize a high-risk subgroup of acute promyelocytic leukemia patients*. Haematologica. 95(5):745-51, 2010 May.

38 Su YC Dunn P Shih LY Kuo MC Chang H Wu JH Lin TL Wang PN Tang TC Hung *YS Retinoic acid syndrome in patients following the treatment of acute promyelocytic leukemia with all-trans retinoic acid.* Chang Gung Medical Journal. 32(5):535-42, 2009 Sep-Oct.

39 Sanz MA Montesinos P Vellenga E Rayon C de la Serna J Parody R Bergua JM Leon A Negri S Gonzalez M Rivas C Esteve J Milone G Gonzalez JD Amutio E Brunet S Garcia-Larana J Colomer D Calasanz MJ Chillon C Barragan E Bolufer P Lowenberg B *Risk-adapted treatment of acute promyelocytic leukemia with all-trans retinoic acid and anthracycline monochemotherapy: long-term outcome of the LPA 99 multicenter study by the PETHEMA Group.* Blood. 112(8):3130-4, 2008 Oct 15.

40 de Botton S Fawaz A Chevret S Dombret H Thomas X Sanz M Guerci A San Miguel J de la Serna J Stoppa AM Reman O Stamatoulas A Fey M Cahn JY Sotto JJ Bourhis JH Parry A Chomienne C Degos L Fenaux P *Autologous and allogeneic stem-cell transplantation as salvage treatment of acute promyelocytic leukemia initially treated with all-trans-retinoic acid: a retrospective analysis of the European acute promyelocytic leukemia group*. Journal of Clinical Oncology. 23(1):120-6, 2005 Jan 1.

41 Lobe I Rigal-Huguet F Vekhoff A Desablens B Bordessoule D Mounier C Ferrant A Sanz M Fey M Chomienne C Chevret S Degos L Fenaux P European APL group experience *Myelodysplastic syndrome after acute promyelocytic leukemia: the European APL group experience*. Leukemia. 17(8):1600-4, 2003 Aug.

42 Gallagher RE Yeap BY Bi W Livak KJ Beaubier N Rao S Bloomfield CD Appelbaum FR Tallman MS Slack JL Willman CL *Quantitative real-time RT-PCR analysis of PML-RAR alpha mRNA in acute promyelocytic leukemia: assessment of prognostic significance in adult patients from intergroup protocol 0129.* Blood. 101(7):2521-8, 2003 Apr 1.

43 Cassinat B Chevret S Zassadowski F Balitrand N Guillemot I Menot ML Degos L Fenaux P Chomienne C *European APL Group In vitro all-trans retinoic acid sensitivity of acute promyelocytic leukemia blasts: a novel indicator of poor patient outcome.* Blood. 98(9):2862-4, 2001 Nov 1.

44 Slack JL Arthur DC Lawrence D Mrozek K Mayer RJ Davey FR Tantravahi R Pettenati MJ Bigner S Carroll AJ Rao KW Schiffer CA Bloomfield CD *Secondary cytogenetic changes in acute promyelocytic leukemia--prognostic importance in patients treated with chemotherapy alone and association with the intron 3 breakpoint of the PML gene: a Cancer and Leukemia Group B study*. Journal of Clinical Oncology. 15(5):1786-95, 1997 May.

45 Miller WH Jr *Differentiation therapy of acute promyelocytic leukemia: clinical and molecular features*. [Review] [65 refs] Cancer Investigation. 14(2):142-50, 1996.

46 Bassan R Battista R Viero P d'Emilio A Buelli M Montaldi A Rambaldi A Tremul L Dini E Barbui T *Short-term treatment for adult hypergranular and microgranular acute promyelocytic leukemia.* Leukemia. 9(2):238-43, 1995 Feb.

47 Vahdat L Maslak P Miller WH Jr Eardley A Heller G Scheinberg DA Warrell RP Jr *Early mortality and the retinoic acid syndrome in acute promyelocytic leukemia: impact of leukocytosis, low-dose chemotherapy, PMN/RAR-alpha isoform, and CD13 expression in patients treated with all-trans retinoic acid.* [Review] [42 refs] Blood. 84(11):3843-9, 1994 Dec 1.

48 Ventura GJ Hester JP Dixon DO Khorana S Keating MJ *Analysis of risk factors for fatal hemorrhage during induction therapy of patients with acute promyelocytic leukemia.* Hematologic Pathology. 3(1):23-8, 1989.

49 Hoyle CF Swirsky DM Freedman L Hayhoe FG *Beneficial effect of heparin in the management of patients with APL.* British Journal of Haematology. 68(3):283-9, 1988 Mar.

50 Marty M Ganem G Fischer J Flandrin G Berger R Schaison G Degos L Boiron M *[Acute promyelocytic leukemia: retrospective study of 119 patients treated with daunorubicin*]. [French] Nouvelle Revue Francaise d Hematologie. 26(6):371-8, 1984.

51 Samyak Manandhar, Vivek Verma, Smith Giri, Ranjan Pathak, Vijaya Raj Bhatt; *Acute promyelocytic leukemia (APL) in pregnancy: A systematic review*. J Clin Oncol 33, 2015 (suppl; abstr e18017)

52 Massimo Breccia, Luca Mazzarella, Vincenzo Bagnardi, Davide Disalvatore, Giuseppina Loglisci, Giuseppe Cimino, Anna Maria Testi, Giuseppe Avvisati, Maria Concetta Petti, Clara Minotti, Roberto Latagliata, Robin Foà, Pier Giuseppe Pelicci, Francesco Lo-Coco *Increased BMI correlates with higher risk of disease relapse and differentiation syndrome in patients with acute promyelocytic leukemia treated with the AIDA protocols Blood 2012 119:49-54*

53 D Martinez-Cuadron, P Montesinos, G Ossenkoppele, AJ González-Huerta, O Salamero, A Holowiecka, S Brunet, P Fernandez, C Benavente, JM Ribera, M Perez Encinas, J De la Serna, J Esteve, V Rubio, J Gonzalez, L Escoda, ME Amutio, B Lowënberg, *Long-term outcome of elderly patients with acute promyelocytic leukemia treated with anthracycline monotherapy and ATRA-based pethema protocols.* Hematologica, 2016;101:s1 (Journal of the European Hematology Association, Abstract Book 2016)

54	S Kayser, J Krzykalla, K Norsworthy, M Elliott, D Grimwade, RK Hills, Z Racil, J Mayer, MR Baer, J Novak, P Zak, T Szotkowski, NH Russell, RB Walter, A Benner, A Krämer, BD Smith, AK Burnett, C Röllig, AD Ho, G Ehninger, MJ Levis, U Platzbecker. *Impact of first line arsenic trioxide and retinoic acid treatment on outcome in therapy-related acute promyelocytic leukemia patients.* Hematologica, 2016;101:s1 (Journal of the European Hematology Association, Abstract Book 2016)
55	M. Mitrovic N. Suvajdzic, A. Novkovic, I. Djunic, A. Vidovic, M. Virijevic, N. Colovic, I. Elezovic, D. Tomin. *Prognostic risk score for haemorrhagic early death in acute promyelocytic leukemia*. Hematologica,2015;100:s1 (Journal of the European Hematology Association, Abstract Book 2015)
56	A. Hecht, S. Doll, H. Altmann, D. Nowak, E. Lengfelder, C. Rollig, G. Ehninger, K. Spiekermann, W. Hiddemann, C. Weiss, W. K. Hofmann, F. Nolte and U. Platzbecker. *Validation of a Molecular Risk Score for Prognosis of Patients With Acute Promyelocytic Leukemia Treated With All-trans Retinoic Acid and Chemotherapy-containing Regimens Clinical Lymphoma,* Myeloma and Leukemia;
57	P. Montesinos and M. A. Sanz. *Secondary and therapy-related acute promyelocytic leukemia*. Experience of the PETHEMA registry Annals of Hematology; 96; S87-S89
58	V. R. Bhatt, V. Shostrom, S. Giri, K. Gundabolu, K. M. Monique Islam, F. R. Appelbaum and L. J. Maness. *Early mortality and overall survival of acute myeloid leukemia based on facility type*. American Journal of Hematology; 92; 8; 764-771
59	D. Martinez-Cuadron, P. Montesinos, G. Ossenkoppele, A. J. Gonzalez-Huerta, O. Salamero, A. Holowiecka, S. Brunet, P. Fernandez, C. Benavente, J. M. Ribera, M. Perez Encinas, J. De La Serna, J. Esteve, V. Rubio, J. Gonzalez, L. Escoda, M. E. Amutio, B. Lowenberg and M. A. Sanz. Long-term outcome of elderly patients with acute promyelocytic leukemia treated with anthracycline monotherapy and ATRA-based pethema protocols Haematologica; 101; 178
60	D. Guardo, F. Guolo, P. Minetto, E. Coviello, M. Gambella, G. Rivoli, N. Colombo, M. Miglino, F. Ballerini, M. Clavio, R. M. Lemoli and M. Gobbi. *A simple 3-gene molecular score retains high prognostic value in non-high risk acute promyelocytic leukemia treated with AIDA 2000 protocol: A retrospective study from a regional register*. Haematologica; 101; 377-378
61	J. Oak, N. Shahmarvand, M. Alcasid, M. Cascio, E. Goodman, B. Medeiros, D. A. Arber, J. Zehnder and R. Ohgami. *A study of disseminated intravascular coagulation in acute leukemia reveals markedly elevated d-dimer levels are a sensitive indicator of acute promyelocytic leukemia.* Laboratory Investigation; 97; 367A
62	I. D. F. Azevedo, M. G. Magalhaes, F. R. Souto, W. B. D. Neves, F. C. D. B. C. Melo, E. M. Rego and R. A. M. Melo. *Molecular and hematologic relapses in adult patients with acute promyelocytic leukemia: a cohort study.* Revista Brasileira de Hematologia e Hemoterapia; 39; 1; 46-51
63	M. Baljevic, B. Dumitriu, J. W. Lee, E. M. Paietta, P. H. Wiernik, J. Racevskis, C. Chen, E. M. Stein, R. E. Gallagher, J. M. Rowe, F. R. Appelbaum, B. L. Powell, R. A. Larson, S. E. Coutre, J. Lancet, M. R. Litzow, S. M. Luger, N. S. Young and M. S. Tallman. *Telomere length recovery: A strong predictor of overall survival in acute promyelocytic leukemia*. Acta Haematologica; 136; 4; 210-218
64	J. H. Yoon, H. J. Kim, D. H. Kwak, S. S. Park, Y. W. Jeon, S. E. Lee, B. S. Cho, K. S. Eom, Y. J. Kim, S. Lee, C. K. Min, S. G. Cho, D. W. Kim, J. W. Lee and W. S. Min. *High WT1 expression is an early predictor for relapse in patients with acute promyelocytic leukemia in first remission with negative PML-RARa after anthracycline-based chemotherapy: a single-center cohort study*. Journal of Hematology and Oncology; 10; 1; 1-4
65	M. Yanada, S. Yano, H. Kanamori, M. Gotoh, N. Emi, K. Watakabe, M. Kurokawa, A. Nishikawa, T. Mori, N. Tomita, M. Murata, H. Hashimoto, H. Henzan, Y. Kanda, M. Sawa, A. Kohno, Y. Atsuta, T. Ichinohe and A. Takami. *Autologous hematopoietic cell transplantation for acute promyelocytic leukemia in second complete remission: outcomes before and after the introduction of arsenic trioxide.* Leukemia and Lymphoma; 58; 5; 1061-1067
66	S. Giri, R. Pathak, M. G. Martin and V. R. Bhatt. *Survival of de novo and secondary acute promyelocytic leukemia: A propensity-matched analysis of the SEER database.* Leukemia and Lymphoma; 57; 2; 385-391
67	C. Ganzel, V. Mathews, K. Alimoghaddam, A. Ghavamzadeh, D. Kuk, S. Devlin, H. Wang, M. J. Zhang, D. Weisdorf, D. Douer, J. M. Rowe, E. Polge, J. Esteve, A. Nagler, M. Mohty and M. S. Tallman. *Autologous transplant remains the preferred therapy for relapsed APL in CR2.* Bone Marrow Transplantation; 51; 9; 1180-1183
68	M. S. Sarwar and U. Shaikh. *Respiratory failure and outcome in patients with acute promyelocytic leukemia, experience at a tertiary care centre in Pakistan*. Respirology; 21; 77
69	F. Khamruddin, A. A. Alex, K. Kathirvel, H. K. Babu, A. Abraham, A. Srivastava, B. George, P. Balasubramanian and V. Mathews. *Immunophenotype characterization of patients with acute promyelocytic leukemia: A singe centre experience Indian*. Journal of Hematology and Blood Transfusion; 32; S464-S465
70	D. Martinez-Cuadron, P. Montesinos, E. Vellenga, T. Bernal, O. Salamero, A. Holowiecka, S. Brunet, C. Gil, C. Benavente, J. M. Ribera, M. Perez-Encinas, J. De la Serna, J. Esteve, V. Rubio, J. Gonzalez-Campos, L. Escoda, M. E. Amutio, M. Arnan, J. Arias, S. Negri, B. Lowenberg and M. A. Sanz. *Long-term outcome of older patients with newly*

diagnosed de novo acute promyelocytic leukemia treated with ATRA plus anthracycline-based therapy. Leukemia; doi:10.1038/leu.2017.178

A11. **Priority question**: No RCT evidence for the effectiveness of ATO in relapsed/refractory APL was provided in the company submission (the study by Raffoux *et al.* (2003) compared all-trans retinoic acid (ATRA) plus ATO with ATO; therefore it does not include a relevant comparator). If there is no RCT with a relevant comparator, please include all relevant non-RCTs of ATO and all relevant non-RCTs of the comparators listed in the scope.

Company response: As presented in Appendix D, we performed a very wide literature search, with terms only concerning population and study design. No intervention or comparator restrictions were used during the initial screening of abstracts and full-texts. Consequently, we identified a variety of non-randomised studies in APL, and subsequently focused on those describing ATO use in the recently-approved first-line indication. However, the available non-randomised second-line studies of ATO are generally supportive of its use in this indication (see for instance the pivotal studies by Soignet et al.[2, 3] and a comprehensive European registry-based study by Lengfelder et al.[4]). This is also reflected in European clinical guidelines that unanimously recommend ATO-based salvage in patients relapsing after first-line treatment with ATRA and chemotherapy, and indeed in standard clinical practice. In terms of the NICE submission, Teva felt that the study by Raffoux et al. may provide more comprehensive evidence than some of the non-randomised studies identified, due to its randomised design and in-depth description of the treatment regimens that individual patients received. Among the studies on second-line ATO use that were initially identified by our literature search but later rejected as they did not focus on first-line indication, two deserve particular attention and are described below.

In a retrospective multicentre analysis, Thomas et al. reported the outcomes of 25 patients with relapsed APL treated with ATO for remission induction, and compared them with the outcomes of an earlier strategy, where patients were treated with ATRA and intensive timed sequential EMA chemotherapy (including etoposide, mitoxantrone and cytarabine)[5]. For both induction regimens, post-remission therapy included autologous or allogeneic HSCT, or maintenance treatment, although a smaller percentage of ATO-treated than ATRA-EMA-treated patients proceeded to receive a HSCT[5]. The outcomes appeared more favourable in the ATO-treated group than the ATRA-EMA group, with improved 2-year overall survival (77% vs 51%, respectively) and 2-year leukaemia-free survival (90% vs 47%, respectively)[5].

Transplantation in second remission may improve patient outcomes[4, 6] and a retrospective registry-based study from Japan showed that the annual number of autologous transplants among APL patients in second CR increased approximately 4-fold after ATO became commercially available in the country in late 2004[7]. The proportion of patients remaining relapse-free post-transplantation improved significantly after the introduction of ATO, corresponding to a significant improvement in 4-year relapse-free survival (75.0% pre-ATO vs 88.2% post-ATO, p=0.028) and a decrease in the cumulative incidence of post-transplant relapse (22.3% pre-ATO vs 8.5% post-ATO at 4 years from transplantation, p=0.008)[7]. Furthermore, this translated into a significant 4-year post-transplant OS improvement observed after the introduction of ATO (79.8%

pre-ATO vs 92.8% post-ATO, p=0.027)[7]. However, an important limitation of this study lied in the fact that the authors had no information on whether patients treated after the introduction of ATO actually received it[7], so that the study does not provide a direct ATO vs no ATO comparison, but only approximates it.

Overall, Teva feel that the use of ATO in the relapsed or refractory APL setting is already so well-established in routine clinical practice that it would be difficult to provide NICE with novel information based on the analysis of additional studies.

A12. Please provide a reference to support the assertion that maintenance treatment is not used in the UK. “First-line therapy in APL generally consists of three consecutive treatment phases: induction, consolidation and maintenance, although maintenance is usually omitted in the UK clinical practice with the aim of minimising the risk of treatment-related myelodysplastic syndrome or acute myeloid leukaemia (tMDS/AML).” (company submission p15) Were any trials including maintenance treatment omitted from the review?

Company response: Primary market research commissioned by Teva in 2015 suggested that APL treatment in the UK does not include maintenance therapy[8]. Only a small minority of the surveyed clinicians mentioned they prescribed maintenance treatment to high-risk patients at greater risk of relapse[8]. Earlier on, maintenance therapy was included in the UK-based AML15 trial (recruitment period: 2002–2007), which tested the AIDA regimen (inclusive of maintenance) against the MRC regimen that included more intensive chemotherapy but no maintenance treatment[9]. The trial reported quite a high 5-year cumulative incidence of tMDS/AML (6% in the AIDA arm vs 2% in the MRC arm, p=0.11)[9], and removal of maintenance in the AML17 trial aimed to address this issue. Indeed, the 4-year cumulative incidence of tMDS/AML was much lower in the AML17 trial (3% in the AIDA group vs 0% in the ATRA+ATO group, p=0.34)[10]. Consequently, at present, the standard treatment approach in the UK does not include maintenance therapy, as confirmed by Dr Dillon in a written statement attached to this document.

A13. **Priority question**: On page 79 of the company submission it is stated that “The estimated overall cumulative exposure to Teva Group products containing ATO was approximately 13,855 patients, with an estimated 363 patients exposed to ATO in 6 clinical trials sponsored by Teva Group.” Which 6 studies are these and in which populations are they conducted? Please provide full data if they are relevant to the current decision problem.

Company response: Exposure was estimated in the PERIODIC SAFETY UPDATE REPORT No. 750/10/16. The estimated cumulative clinical trials exposure to ATO in 6 clinical trials sponsored by Cephalon, Inc. (CTI 1073, CTI 1058, CTI 1061, ATO202, CTI 1064, C18477/3059/AM/USCA) and 5 clinical trials sponsored by Cell Therapeutics, Inc. (CTI1057, CTI1059, CTI1060, CTI1062, CTI1063) was approximately 363 patients (see *Table 3*). Cephalon, Inc. acquired Trisenox in 2005 from Cell Therapeutics, Inc. (CTI). The MAH (Teva) is aware of the fact that the cumulative number of patients exposed to arsenic trioxide in all clinical trials sponsored by MAHs (Cephalon, Inc. /CTI) prior the acquisition by Teva Group may be higher since, due to historical reasons, Teva’s access to much of the data regarding studies conducted with ATO was limited.

Table 3. Cumulative Patient Exposure from Marketing Authorisation Holder's Clinical Trials

Study	Title	Number of patients treated with arsenic trioxide
CTI1057	Phase II Study of Arsenic Trioxide in Patients with Multiple Myeloma	24
CTI 1058	Phase II Multicenter Study of Arsenic Trioxide in Patients With Myelodysplastic Syndromes	70
CTI1059	Phase II Study of Arsenic Tioxide in Neuroblastoma and Other Pediatric Solid Tumors	20
CTI1060	Phase II Clinical Trial of Arsenic Trioxide and Dexamethasone as Therapy for Relapsed or Refractory Multiple Myeloma	15
CTI 1061	Phase I/II Study of Arsenic Trioxide in Patients With Myelodysplastic Syndromes	115

The included trials: APL0406, AML17 and Raffoux et al.

A14. **Priority question**: How many patients from the UK were recruited to each of the three included trials?

Studies APL0406 and Raffoux et al. did not include any UK patients, to the best of our knowledge. The study by Raffoux et al. was conducted solely in France, while the APL0406 study was conducted jointly by GIMEMA and SAL–AMLSG in continental Europe. The AML17 trial, on the other hand, was primarily UK based. The initial publication reported on the outcomes of 235 patients who were randomised[10], but a more recent report from this study mention that a further 70 patients were treated with AIDA within AML17 after randomisation was closed[11]. Thus, in total, 305 patients from the UK were included in the AML17 trial.

A15. **Priority question**: Please provide data on treatment-related deaths across the three trials, if available.

Company response: Treatment-related deaths were not specifically reported in any of the three studies. However, some relevant information can be extracted from the publications and is summarised below.

Raffoux et al. Two patients in this study died during induction therapy[12]. One patient, who had previously suffered a CNS haemorrhage, died of septic shock with seizures[12]. The other patient died of an ATO-induced differentiation syndrome with

hyperleukocytosis, which failed to respond to treatment[12]. While the first death is difficult to attribute to study treatment, the second can likely be considered treatment-related, and the authors clearly attributed it to an ATO-induced syndrome[12].

Among the 16 patients in this study who achieved complete remission (CR), 5 patients died[12]. Of these deaths, 4 were caused by APL relapse and the remaining one was attributed to sepsis occurring after consolidation chemotherapy[12]. Haematological toxicity, such as neutropenia and thrombocytopenia, is common with cytotoxic chemotherapy and patients often recover over several weeks after treatment; thus, this death could reasonably be considered treatment-related. Overall, two deaths in this study could potentially be considered to be directly related to treatment – one with ATO (from differentiation syndrome) and the other with chemotherapy (from sepsis).

APL0406. While no induction deaths were observed in the ATRA + ATO group, in the AIDA group 4 patients died during induction therapy – 2 from differentiation syndrome, 1 from ischaemic stroke and 1 from bronchopneumonia[13]. Although the publication by Platzbecker, et al. did not directly report causal relationships between death and study treatment, differentiation syndrome is a common adverse effect of ATRA (and ATO), and, in these two cases, could be considered related to ATRA administration. Indeed, one of them was reported as a fatal serious adverse event (SAE), accompanied by respiratory failure – both the differentiation syndrome and respiratory failure were deemed related to treatment with idarubicin and ATRA[13]. In terms of the death from ischaemic stroke, this was reported in the publication as an SAE with a fatal outcome, and was deemed related to treatment with both ATRA and idarubicin[13]. The relationship between the death from bronchopneumonia and study treatment is difficult to evaluate based on the information available. There was a case of an acute respiratory distress syndrome reported as a fatal SAE related to treatment with ATRA and idarubicin[13]. It is, however, unclear from the publication if this fatal SAE corresponds to the other death from differentiation syndrome, or to a death from bronchopneumonia, whether during induction or in CR (see below).

Across both treatment groups, 6 patients died in CR[13]. The only patient in the ATRA + ATO group died of bronchopneumonia caused by infection with the H1N1 virus, which was reported as a fatal SAE unrelated to treatment with either ATRA or ATO[13]. The remaining 5 patients who died in CR were in the AIDA group: 2 deaths resulted from bronchopneumonia, and 1 each from haemorrhagic shock, pulmonary embolism and secondary myelodysplastic syndrome (MDS)[13]. Of these, the deaths from haemorrhagic shock and pulmonary embolism were reported as fatal SAEs and deemed not related to treatment[13]. The death from therapy-related secondary MDS can be seen as resulting from a delayed adverse effect of treatment, and therefore treatment-related. Regarding the deaths from bronchopneumonia, the publication by Platzbecker et al. listed two cases of bronchopneumonia as fatal SAEs[13]. Both were considered related to treatment – one of them with methotrexate and the other with ATRA[13]. While the former death clearly occurred in CR (methotrexate was part of maintenance treatment but not the earlier treatment phases), it is unclear if the second fatal bronchopneumonia reported as an SAE referred to the remaining death in CR from this cause, or to the induction death (see above).

AML17. Deaths in this study were not routinely assessed for relationship to study treatment[10] and limited information on causes of death is available from the publication. However, some information is available on the causes of deaths occurring up to day 60 (corresponding to the maximum duration of induction therapy)[10]. In this timeframe, 6 patients died in the ATRA + ATO group, compared with 11 in the AIDA group; this difference was not statistically significant (p=0.22)[10]. As very limited information on the causes of these deaths is available, their potential relationship to study treatment is difficult to discuss. In the ATRA + ATO group the causes of 3 deaths were cardiac events, with infection, renal failure, and multiple causes resulting in 1 death each[10]. ATO may cause QTc prolongation, so that further data on the deaths from cardiac events could be of interest. The causes of the 11 deaths in the AIDA group were: haemorrhage (3 deaths), infection (3 deaths), pulmonary causes (2 deaths), progressive disease (2 deaths) and renal cause (1 death)[10]. It is not clear if the deaths from pulmonary causes occurred in patients with ATRA-induced differentiation syndrome, so that we cannot speculate on their relationship with study treatment.

A16. **Priority question**: Please provide the full APL0406 Clinical Study Report (CSR). As mentioned in the APL0406 journal publication (LoCoco *et al.* 2013, New England Journal of Medicine) arsenic trioxide was donated by Teva Pharmaceutical Industries. Therefore, it might be possible to obtain the full CSR.

Company response: According to the Clinical Overview we submitted to the EMA, no clinical studies have been performed by Teva to support the use of ATO as a first-line treatment for APL. The indication is based primarily on published data from a pivotal clinical phase 3 study (APL0406) performed by the Italian GIMEMA group, the German-Austrian Acute Myeloid Leukemia Study Group, and Study Alliance Leukemia. Supportive data are provided by study AML17, performed by the AML Working Group of the UK National Cancer Research Institute. The APL0406 study, like the AML17 study, was an Investigator Sponsored Study and Teva only received the final publication from the investigator (Professor Lo-Coco). It is impossible for the company to receive additional data, including the CSR.

A17. In APL0406, how many patients went on to receive a haematopoietic stem cell transplant (HSCT) during or after the trial?

Company response: APL0406 was an Investigator Sponsored Study and Teva only received the final publication. However, according to the Professor Lo-Coco, the principal investigator for this trial, none of the patients received HSCT during the trial. Actually, HSCT for the treatment of APL was only considered at relapse, but patients relapsing went off study in APL0406.

Ongoing research

A18. Are any further analyses planned or publications in process for any of the three included trials (APL0406, AML17 and Raffoux *et al.*)? When will these be available?

Company response: APL0406 was an Investigator Sponsored Study and Teva only received the final publication. However, according to Professor Lo-Coco, a publication presenting the updated outcome of patients enrolled in the APL0406 trial at a 60-month

median follow-up is planned for 2018. Regarding the AML17 trial, we also enquired with Professor Russell, one of the investigators, and he confirmed that the updated analysis presented at ASH 2016[11, 14] is currently being prepared for publication. Teva believe it is very unlikely updated analyses from Raffoux et al. will be published, as the original study dates back to 2003.

A19. Please confirm when the long-term quality of life analysis of the final APL0406 cohort will be reported (p61).

Company response: APL0406 was an investigator-sponsored study, and publication of the long-term quality of life analysis will be based on a decision by the principal investigators, Prof. Efficace and Prof. Lo Coco. Teva is expecting the final publication of this analysis in 2019.

Section B: Clarification on cost-effectiveness data

Model structure

B1. The company's model includes 14 health states (62 including tunnel states). However, the existing economic evaluations identified through the systematic literature review only include 4-5 health states.

a. Please clarify why the inclusion of each of these health states was necessary.

Company response: The previous models described health states in a very schematic way that that does not adequately reflect the trajectory of APL patients. They only consider stable states and disease events, while we offer more granularity with treatment phases, molecular remission and HSCT. The granularity is necessary to better reflect the clinical trajectory of APL patients, and also to have a better understanding of which variables/inputs influence the model.

b. Please discuss the impact of this more complex model on outcomes, considering that the other existing economic evaluations resulted in positive incremental costs and lower QALY gains.

Company response: As the previously published models were developed for countries other than the UK, a direct comparison of outcomes with the submitted model is difficult. The ICER reported for the US was low (base-case ICERs of $4,512[15]), but drug prices in the US are usually much higher than in the UK, and indeed drug cost was the key cost driver in the analysis by Tallman et alC. The QALY gain reported by Tallman et al. (ATO+ATRA resulted in a QALY gain of 6.19 vs AIDA[15]) was higher than in our model, where ATRA+ATO resulted in a QALY gain of 2.63 vs. AIDA. Difference in costs might also be related to differences in model structure – as the US model only included disease events, HSCT as such was not modelled. One of the assumptions of that study was: "It was assumed that patients are hospitalized after a disease event post first-line treatment and any costs would be included in the DRG cost for second-line treatment." Detailed information was not available in the publication, so it is possible that the DRG for second line did not fully capture HSCT-related spending, thus underestimating the costs. By having a clear separation of health states, our model allows more detailed clinical and economic calculations.

B2. **Priority question:** The company submission states that patients on treatment could be at increased risk of mortality but this was not incorporated in the company base-case analysis (company submission, p98). In the base-case model, patients in induction and consolidation phases (almost 3 years) experience the same level of mortality as the general UK population, and only patients treated with AIDA (ATRA plus idarubicin) can also die from tMDS/AML in these phases. The impact of this assumption was not explored in scenario analysis.

a. Please justify the decision to exclude disease-related mortality from the induction and consolidation phases.

Company response: The mortality rate observed during treatment in both the APL0406 trial and the AML7 trial was numerically lower for ATRA+ ATO compared to AIDA[10, 13]. However, as the differences were not statistically significant, we decided not to include a disease-related mortality rate in the model, in order to keep it conservative.

b. Please include a scenario analysis in which disease-related mortality is included in the induction and consolidation phases for both treatment groups.

Company response: A scenario incorporating the mortality observed during treatment in the APL0406 clinical trial was implemented in version 3.17 of the model. Patients in the ATRA+ATO arm lived for 4.42 discounted years longer than in the AIDA arm, representing 3.8 incremental QALYs. Cost savings with ATRA+ATO were reduced to £21,099 vs AIDA. This could be mainly explained by lower savings from avoiding HSCT procedures (fewer AIDA patients were alive and underwent HSCT) and higher incremental treatment acquisition costs.

B3. **Priority question:** The company submission states that some adverse events could lead to a change in treatment (company submission, p98). However, in the base-case model, only cardiac events lead to a change in treatment. There is an error in the model, where, if other adverse events prompting a change in treatment are implemented, it results in negative numbers of patients in the Markov trace.

Company response: Thank you for flagging the issue. Indeed, if several AEs leading to discontinuation were modelled, this would result in negative numbers of patients in the Markov trace. Nevertheless, we would like to highlight that this error does not really affect the UK model submitted to NICE, as based on expert recommendation only a single type of AEs (i.e. cardiac) led to a treatment switch, so that the negative trace does not occur in the base case.

a. Please justify why only cardiac events can trigger a change in treatment in the model. Please comment on what other adverse events may prompt a change in treatment in practice, citing the relevant literature or expert opinion.

Company response: Due to the severity of the disease, experts stated that only cardiac events consistently lead to a change of treatment in clinical practice. While it is possible that other serious AEs may prompt a treatment switch, they were not frequent enough to either find adequate probabilities or have any impact on the end results.

b. Please provide a version of the model, in which the selection of all or several adverse events prompting a change in treatment is correctly implemented (i.e. does not lead to negative numbers in the Markov trace).

Company response: The model was corrected at version v3.13, which was used as the base for all requested scenarios. This means that all models provided with this clarification letter are corrected regarding this issue.

c. Please explain why patients in the model who have a cardiac event in second-line treatment can only receive allogeneic HSCT, not autologous HSCT.

Company response: According to an expert opinion from Dr Cicconi (Italy), if no remission is achieved in second line, patients receive allogeneic HSCT, while autologous HSCT is only considered for patients who have achieved molecular remission. Patients who experience a serious cardiac AE during treatment may likely do so before achieving molecular remission, so that in the model only allogeneic HSCT was possible for these patients. On the other hand, Dr Dillon, the clinical expert from the UK, explained that the approach to allogeneic HSCT in the UK differs somewhat from other countries in Europe. Indeed, fewer HSCTs are conducted in the UK and allogeneic HSCT is generally not recommended in APL. Nevertheless, we thought that the assumption in question was reasonable and conservative, and the situation concerned a low number of patients receiving AIDA in second line, (i.e. patients from the ATRA+ATO arm experiencing an early relapse [≤2 years from remission]).

B4. The company submission states that “During these first 2 years of remission, the probability of relapse was higher than in the following years. In case of a relapse, patients moved to the “Second-line induction” state; otherwise, they went through all tunnel states and after 24 Markov cycles they moved to the “+2y remission” state.” (Company submission, p98). This indicates that the probability of relapse decreases over time. Please justify the use of a 2 year cut-off for the change in relapse rate, citing relevant literature and expert opinion.

Company response: Selection of the 2-year cut off for the decline in relapse probability was based on clinical expert opinion from Prof. Lo-Coco and Dr Cicconi.

Intervention and comparator

B5. **Priority question:** In the company submission, ATRA+ATO is only assessed in first line, which is not in line with the scope issued by NICE. Although the submission states that the use of second-line treatment will decrease if ATO is used as first-line treatment (company submission, p98), it is plausible that the cost effectiveness of second-line treatment options will have an impact on the cost effectiveness of first-line treatments.

a. Please provide an analysis where ATRA+ATO is compared with AIDA in the second-line setting (the relapsed/refractory population). This could be implemented by removing second-line treatment and moving patients straight to HSCT (instead of second-line treatment).

Company response: A model assessing second-line treatment is provided with this clarification letter (Excel file ID446

Trisenox_NICE_CEM_2nd_Line_v1.1_BC_11012018.xlsm). It was based on the model with 26 tunnels for the first two years of remission. The health states representing first-line therapy were changed to second-line, and those representing second line were neutralised (no transitions to these states were possible). The analysis was conducted with the same data as in the base case company model and the results are presented in the 3 tables below (*Table 4–Table 6*). The analysis showed that ATRA+ATO was cost-effective versus AIDA in the second-line setting with an incremental cost-effectiveness ratio (ICER) of £16,733 per QALY gained. Salvage therapy with ATRA+ATO was £10,082 more costly than when conducted with AIDA, but produced 0.6 more QALYs and 0.7 more life years over the 40-year time horizon. Except for treatment acquisition and follow-up and monitoring costs, ATRA+ATO generated cost savings for all other cost items.

Table 4. Base-case incremental results (discounted)

Treatment	Total costs (£)	Total LYG	Total QALYs	Incremental costs (£)	Incremental LYG	Incremental QALYs	ICER versus baseline (£/QALY)	ICER incremental (£/QALY)
ATRA+ATO	£198,959	11.43	9.44	-	-	-	£16,733	£16,733
AIDA	£188,877	10.72	8.84	-£10,082	-0.70	-0.60	-	-

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

Table 5. Discounted disaggregated and total costs – base-case scenario

Cost category	ATRA+ATO	ATRA+IDA	ATRA+ATO vs. AIDA
Treatments	£29,688	£3,854	£25,834
Administration	£17,263	£20,012	-£2,749
Supportive care and antibiotics	£3,842	£4,309	-£467
Follow-up and monitoring	£14,832	£14,084	£748
Adverse Events	£3,811	£7,879	-£4,068
MDS	£0	£134	-£134
HSCT	£118,218	£125,691	-£7,473
Palliative care	£11,305	£12,915	-£1,610
Total	£198,959	£188,877	£10,082

ATO=arsenic trioxide; ATRA=all-trans retinoic acid; HSCT=haematopoietic stem cell transplantation; MDS=myelodysplastic syndrome

Table 6. Discounted health outcomes in the model

	ATRA+ATO	AIDA	ATRA+ATO vs. AIDA
Number of QALYs	9.44	8.84	0.60
Number of LYs	11.43	10.72	0.70
Second remission	24.35%	20.05%	4.30

Second long remission (> 2 years)	10.15%	5.59%	4.56
MDS	0.00%	0.83%	-0.83
Death	62.31%	64.72%	-2.41
APL related death	51.67%	54.74%	-3.07
Background death	10.64%	9.98%	0.66

ATO=arsenic trioxide; ATRA=all-trans retinoic acid; LY=life years; MDS=myelodysplastic syndrome; QALY=quality-adjusted life years

b. Please provide an analysis with only ATO as the intervention (without ATRA) in the second-line setting.

Company response: We were unable to identify suitable efficacy data for ATO alone other than those published by Raffoux et al., 2003[12]. The study, based on a low number of patients (20, 10 in each arm), did not show significant differences between ATO+ATRA and ATO alone, and, surprisingly, disease-free survival was better with ATO alone than with ATRA+ATO[12]. Conducting this scenario would lead to better cost-effectiveness results for ATO vs. AIDA, reducing treatment acquisition costs without changing the effectiveness results. Furthermore, according to all experts and especially to Dr Dillon (clinical expert for the UK), ATO alone is rarely used nowadays.

c. Please provide an analysis with best supportive care as a comparator in the second-line setting.

Company response: All experts strongly stated that, due to the severity of the disease, best supportive care is not a relevant comparator in the second-line setting, and that best supportive care is only a relevant alternative in 3rd or 4th line. However, given the very small number of affected patients, adding best supportive care as a comparator in 3rd or 4th line would have very little impact on the ICER. Thus, we decided not to include it in the model.

Effectiveness

Priority question: Please provide an overview of all transition probabilities that are used in the Excel model (i.e. in the Markov trace), including sources, description of any calculations performed and justification (for both the calculations and the source) for the parameters related to sections 3.3.3.1-3.3.3.5 and 3.3.4.

Company response: Please see details of the calculations in *Table 7*.

Table 7. Details of transition probabilities used in the model

From	To	Strategy	Value	Calculation	Source	Justification
First line - Induction cycle 1	First line - Consolidation cycle 1	ATRA+ ATO AIDA	0.45475 0.408712	$1 - \left[1 - 0.5 \times \left(1 - P(induction\ death) - P(induction\ switch) - P(induction\ MDS)\right)\right]^{(28/median\ time\ to\ remission)}$	Lo-Coco et al, 2013 Platzbecker et al., 2017	Lo-Coco et al., 2003 and Platzbecker et al., 2017 present the results of the APL0406 trial comparing ATRA+ATO vs. AIDA in newly-diagnosed patients with low/intermediate risk APL. The studied population correspond exactly to the population for which ATO is approved, and these papers provided the treatment schedule, dosage, efficacy and safety data of interest for the model. Calculation: 50% of the patients entered into remission after X days (median time to remission), expressed among patients still alive, who did not switch or did not experienced MDS during the induction. Then the probability was calculated for 28 days (cycle duration) according to an exponential distribution. The median time to remission was assumed to represent the average duration of the induction phase.
First line - Induction cycle 1	First line - Induction cycle 2	ATRA+ ATO AIDA	0.54525 0.561577	$1 - \left[P_{cycle}(induction\ to\ consolidation) + P_{cycle}(induction\ death) + P_{cycle}(induction\ switch) + P_{cycle}(induction\ MDS)\right]$		Patients staying in the health state are those who did not move to consolidation, die, experience MDS or switch.
First line - Induction cycle 1-2	Second line - Induction + 1 Consolidation	ATRA+ ATO AIDA	0 0.029711	$P(induction\ switch) = 1 - \prod_{i=1}^{N_{AE}} 1 - P(induction\ AE_i)$ $P_{cycle}(induction\ switch) = 1 - \left(1 - P(induction\ switch)\right)^{(28/median\ time\ to\ remission}$	Platzbecker et al., 2017	Platzbecker et al., 2017 provided the proportion of patients experiencing the adverse events of interest in both treatment phases. The probability of a switch was calculated as follows: B1. We calculated the probability that patients do not change treatment due to any of the adverse events that could prompt a switch in the model. This was done by multiplying the probabilities that each of the events does not occur, for all the events. B2. The product of probabilities described in step 1 was subtracted from 1, to render the probability of a

From	To	Strategy	Value	Calculation	Source	Justification
						switch. This probability was then expressed per cycle.
First line - Induction cycle 1-2	MDS	ATRA+ ATO AIDA	0 0	$P_{cycle}(induction\ MDS) = 1 - \left(1 - P(induction\ MDS)\right)^{(28/median\ time\ to\ remission)}$	Platzbecker et al., 2017	
First line - Induction cycle 1-2	APL-related death	ATRA+ ATO AIDA	0 0	$P_{cycle}(induction\ death) = 1 - \left(1 - P(induction\ death)\right)^{(28/median\ time\ to\ remission}$		Disease-related mortality was not considered in the base case model.
First line – Consolidation cycle$_i$	First line – Consolidation cycle$_{i+1}$	ATRA+ ATO AIDA	1 0.995109455	$1 - \left[P_{cycle}(consolidation\ death) + P_{cycle}(consolidation\ switch) + P_{cycle}(consolidation\ MDS)\right]$		The number of tunnels was determined according to the duration of consolidation treatment, computed as:$number\ of\ cycle \times (duration\ on\ treatment\ per\ cycle + duration\ off\ treatment\ per\ cycle)$. For ATRA+ATO the longer treatment duration between ATRA and ATO is considered.
First line – Consolidation until the second to last cycle	Second line - Induction + 1 Consolidation	ATRA+ ATO AIDA	0 0	$P_{cycle}(consolidation\ switch)$ $P_{cycle}(consolidation\ switch)$		The same methodology as for the probability of switch during induction was applied. Consolidation treatment duration (computed as explained above) was considered for expressing the probability per cycle (instead of the median time to remission).
First line – Consolidation last cycle	Second line - Induction + 1 Consolidation	ATRA+ ATO/AIDA AIDA	0.0155039 0.050845	$P_{cycle}(consolidation\ switch) + [CHR \times Eligibility \times (1 - CMR) + (1 - CHR)] \times \left[1 - \left[P_{cycle}(consolidation\ death) + P_{cycle}(consolidation\ switch) + P_{cycle}(consolidation\ MDS)\right]\right]$		In addition to patients switching in this cycle, patients who did not achieve haematological or molecular remission moved to second line. The proportion of these patients was expressed among patients still alive, who did not switch or experience MDS. Patients receiving ATRA+ATO in second line following first-line treatment with the same combination were not affected by this transition, as re-treatment with ATRA+ATO was only possible for patients relapsing after at least s2 years of remission.

From	To	Strategy	Value	Calculation	Source	Justification
First line – Consolidation	MDS	ATRA+ATO AIDA	0 0.004890545	$P_{cycle}(consolidation\ MDS) = 1 - \left(1 - P(consolidation\ MDS)\right)^{(28/consolidation\ duratio}$	Platzbecker et al., 2017	The per-cycle probability was computed as described above for the induction phase.
First line – Consolidation	APL related death	ATRA+ATO AIDA	0 0	$P_{cycle}(consolidation\ death) = 1 - \left(1 - P(consolidation\ death)\right)^{(28/consolidation\ durati}$		
First line – Consolidation last cycle	First line - Remission	ATRA+ATO/ATRA+ATO ATRA+ATO/AIDA AIDA	1 0.984496124 0.944264446	$[CHR \times Eligibility \times CMR + CHR \times (1 - eligibility)] \times \left[1 - \left[P_{cycle}(consolidation\ death) + P_{cycle}(consolidation\ switch) + P_{cycle}(consolidation\ MDS)\right]\right]$	Platzbecker et al., 2017	All patients who received ATRA+ATO in second line had been in remission more than 2 years before they relapsed. Patients were forced to enter remission in this sub-arm. In the other arms, patients entering remission were those who were in molecular remission and those in haematological remission but with unavailable PCR results. The probabilities were expressed among patients still alive, who did not switch or experience MDS.
First line – Remission cycle $cycle_i$	First line – Remission $cycle_{i+1}$	ATRA+ATO/ATRA+ATO ATRA+ATO/AIDA AIDA	1 0.999623373 0.996441435	$1 - P_{cycle}(relapse\ between\ 0\ and\ 24\ months)$ $P_{cycle}(relapse\ between\ 0\ and\ 24\ months) = 1 - \left(1 - P(relapse\ at\ 24\ months)\right)^{28/24\times30}$	Platzbecker et al., 2017	No patients relapsed before 24 months in this sub-arm and all patients moved through all tunnels. The per-cycle probability of relapse between 0 and 24 months post remission was computed from the cumulative probability of relapse at 24 months.
First line – Remission cycle $cycle_i$	Second line - Induction + 1	ATRA+ATO/ATRA+ATO	0			

From	To	Strategy	Value	Calculation	Source	Justification
	Consolidation	ATRA+ATO/AIDA AIDA	0.000376627 0.003558565	$P_{cycle}(relapse\ between\ 0\ and\ 24\ months)$ $= 1 - \left(1 - P(relapse\ at\ 24\ months)\right)^{28/24\times30}$	Platzbecker et al., 2017	Relapsing patients moved to second line.
First line – Remission +2 years	First line – Remission +2 years	ATRA+ATO/ATRA+ATO ATRA+ATO/AIDA AIDA	0.99958 0 0.997557613	$1 - P_{cycle}(relapse\ after\ 24\ months)$	Platzbecker et al., 2017	Patients stayed in this health state until they relapsed. All patients in the ATRA+ATO/AIDA sub-arm relapsed earlier than 24 months after achieving remission.
First line – Remission +2 years	Second line - Induction + 1 Consolidation	ATRA+ATO/ATRA+ATO ATRA+ATO/AIDA AIDA	0.00042 0 0.002442387	$P_{cycle}(relapse\ after\ 24\ months)$ $= 1$ $- \left(1 - P(relapse\ at\ 48\ months)\right.$ $\left. - P(relapse\ at\ 24\ months)\right)^{28/24\times30}$	Platzbecker et al., 2017	The probability of relapse after 24 months was computed as the difference between the cumulative probability of relapse at 48 months and the one at 24 months, representing the proportion of relapses occurring between 24 and 48 months. This proportion was expressed as a per-cycle probability.
Second line - Induction cycle 1	Second line – Consolidation cycle 1	ATRA+ATO/ATRA+ATO ATRA+ATO/AIDA AIDA	0.276365381 0.253170144 0.276365381	1 $- \left[1 - 0.5\right.$ $\times \left(1 - P(induction\ death)\right.$ $- P(induction\ switch)$ $\left. - P(induction\ MDS)\right)\left]^{\left(28/\text{median time to remissio}\right.}\right.$ $\text{consolidation cycle durat}$	Platzbecker et al., 2017	The probability that patients moved from the first cycle of induction to consolidation was computed the same ways as in first line.
Second line -	Second line -	ATRA+ATO/AT	0.723634619		Platzbecker et al., 2017	Patients staying in the health state are those who did not move to consolidation, die, experience MDS or switch.

From	To	Strategy	Value	Calculation	Source	Justification
Induction cycle 1	Induction cycle 2	RA+ATO ATRA+ATO/AIDA AIDA	0.717118681 0.723634619	$1 - [P_{cycle}(induction\ to\ consolidation) + P_{cycle}(induction\ death) + P_{cycle}(induction\ switch) + P_{cycle}(induction\ MDS)]$		
Second line - Induction cycle 2	Second line – Consolidation cycle 1	ATRA+ATO/ATRA+ATO ATRA+ATO/AIDA AIDA	1 1 1			According to clinical experts, patients always receive at least one consolidation cycle after second-line induction, before further treatment decisions are made. Thus, patients were forced to enter consolidation at the end of second-line induction.
Second line - Induction	Allogeneic HSCT	ATRA+ATO/ATRA+ATO ATRA+ATO/AIDA AIDA	0 0.029711175 0	$P_{cycle}(induction\ switch)$	Platzbecker et al., 2017	Patients experiencing an adverse event requiring a switch were assumed to undergo allogeneic HSCT.
Second line – Consolidation cycle 1	Second line – Consolidation cycle 2	ATRA+ATO/ATRA+ATO ATRA+ATO/AIDA AIDA	0 1 0			If the duration of a consolidation treatment cycle was greater than 28 days, patients were forced to follow a second tunnel of consolidation.
Second line – 1 cycle	Allogeneic HSCT	ATRA+ATO/ATRA+ATO	0.28240		Platzbecker et al., 2017 Russel et al., 2017	Patients undergoing allogeneic HSCT at the end of the first consolidation cycle included all patients who switched and a proportion of those who either 1) reached molecular

From	To	Strategy	Value	Calculation	Source	Justification
consolidation		ATRA+ATO/AIDA AIDA	0.364773159 0.28240	$P(Conso\ 1\ to\ Allo\ HSCT)$ $= P_{cycle}(induction\ switch) + CMR_{2nd\ line}$ $\times P(Allo\ HSCT \mid induc\ CMR)$ $+ \Big[1$ $- \Big(CMR_{2nd\ line} + P_{cycle}(consolidation\ death)$ $+ P_{cycle}(consolidation\ switch)$ $+ P_{cycle}(consolidation\ MDS)\Big)\Big]$ $\times P(Allo\ HSCT \mid induc\ no\ CMR)$		remission or 2) did not reach molecular remission, switch, die or experience MDS.
Second line – 1 cycle consolidation	Autologous HSCT	ATRA+ATO/ATRA+ATO ATRA+ATO/AIDA AIDA	0.27600 0.2415 0.27600	$P(Conso\ 1\ to\ Auto\ HSCT) = CMR_{2nd\ line}$ $\times P(auto\ HSCT \mid induc\ CMR)$ $+ \Big[1$ $- \Big(CMR_{2nd\ line} + P_{cycle}(consolidation\ death)$ $+ P_{cycle}(consolidation\ switch)$ $+ P_{cycle}(consolidation\ MDS)\Big)\Big]$ $\times P(Auto\ HSCT \mid induc\ no\ CMR)$	Platzbecker et al., 2017 Russel et al., 2017	Patients undergoing autologous HSCT at the end of the first consolidation cycle were a proportion of those who either 1) reached molecular remission or 2) did not reach molecular remission, switch, die or experience MDS. However, in line with advice from clinical experts, the proportion of patients who undergo autologous HSCT among those who are not in molecular remission was set to 0.
Second line – 1 cycle consolidation	Second line – Consolidation	ATRA+ATO/ATRA+ATO ATRA+ATO/AIDA AIDA	0.4416 0.3864 0.4416	1 $- \big[P(conso\ 1\ to\ Allo\ HSCT)$ $+ P(conso\ to\ Auto\ HSCT)$ $+ P_{cycle}(consolidation\ switch)$ $+ P_{cycle}(consolidation\ MDS)$ $+ P_{cycle}(consolidation\ death)\big]$		Patients continuing consolidation were those who did not undergo HSCT at the end of the first consolidation cycle, switch, die or experience MDS. Unless they switch treatment, these patients are forced to follow the full remaining course of consolidation.
Second line – Consolidation last cycle	Second line – Remission	ATRA+ATO/ATRA+ATO ATRA+ATO/AIDA AIDA	0.55200 0.5446732 0.55200	1 $- \big[P(conso\ 1\ to\ Allo\ HSCT)$ $+ P(conso\ to\ Auto\ HSCT)$ $+ P_{cycle}(consolidation\ switch)$ $+ P_{cycle}(consolidation\ MDS)$ $+ P_{cycle}(consolidation\ death)\big]$	Platzbecker et al., 2017 Russel et al., 2017	Patients moving to the second remission state are those who did not undergo HSCT, switch, experience MDS or die.

From	To	Strategy	Value	Calculation	Source	Justification
Second line – Consolidation last cycle	Allogeneic HSCT	ATRA+ATO/ATRA+ATO ATRA+ATO/AIDA AIDA	0.10300 0.10300 0.10300	$P(conso\ to\ Allo\ HSCT)$	Russel et al., 2017	At the end of the consolidation course, a certain proportion of patients moved to HSCT.
Second line – Consolidation last cycle	Autologous HSCT	ATRA+ATO/ATRA+ATO ATRA+ATO/AIDA AIDA	0.345 0.345 0.345	$P(conso\ to\ Auto\ HSCT)$	Russel et al., 2017	At the end of the consolidation course, a certain proportion of patients moved to HSCT.
Second line – Consolidation all cycles	MDS	ATRA+ATO/ATRA+ATO ATRA+ATO/AIDA AIDA	0 0.007326841 0	$P_{cycle}(consolidation\ MDS)$	Platzbecker et al., 2017	
Second line – Remission	Second line – Remission	ATRA+ATO/ATRA+ATO ATRA+ATO/AIDA AIDA	0.967461129 0.951495104 0.967461129	$1 - P_{cycle}(2nd\ remission\ to\ Auto\ HSCT) - P_{cycle}(2nd\ remission\ to\ Allo\ HSCT)$	Russel et al., 2017	
		ATRA+ATO/AT	0.01542		Russel et al., 2017	

From	To	Strategy	Value	Calculation	Source	Justification
Second line – Remission	Allogeneic HSCT	RA+ATO ATRA+ATO/AIDA AIDA	0.01873418 0.01542	$P_{cycle}(2nd\ remission\ to\ Allo\ HSCT)=1-\left(1-P(2nd\ remission\ to\ Allo\ HSCT)\right)^{28/time\ to\ 2nd\ rel}$		The proportion of patients undergoing HSCT from 2nd remission was adjusted per cycle according to the time to 2nd relapse.
Second line – Remission	Autologous HSCT	ATRA+ATO/ATRA+ATO ATRA+ATO/AIDA AIDA	0.01712 0.029770716 0.01712	$P_{cycle}(2nd\ remission\ to\ Auto\ HSCT)=1-\left(1-P(2nd\ remission\ to\ Auto\ HSCT)\right)^{28/time\ to\ 2nd}$	Russel et al., 2017	The proportion of patients undergoing HSCT from 2nd remission was adjusted per cycle according to the time to 2nd relapse.
Allogeneic HSCT	Allogeneic HSCT	All	0.932876645	$1-P_{cycle}(Allo\ HSCT\ death)$	Hosing et al., 2003	Patients who did not die during HSCT stayed in the health state for the duration of the follow-up.
Autologous HSCT	Autologous HSCT	All	0.97999461	$1-P_{cycle}(Auto\ HSCT\ death)$	Hosing et al., 2003	Patients who did not die during HSCT stayed in the health state for the duration of the follow-up.
Allogeneic HSCT	Allogeneic HSCT remission	All	0.673910088	$P(Allo\ HSCT\ CMR)\times\left(1-P_{cycle}(Allo\ HSCT\ death)\right)$	Holter Chakrabarty et al., 2013 Hosing et al., 2003	Patients reaching molecular remission after allogeneic HSCT who did not die.
Autologous HSCT	Autologous HSCT remission	All	0.96147	$P(Auto\ HSCT\ CMR)\times\left(1-P_{cycle}(Auto\ HSCT\ death)\right)$	Holter Chakrabarty et al., 2013 Hosing et al., 2003	Patients reaching molecular remission after autologous HSCT who did not die.
Allogeneic HSCT	Failure	All	0.25897	$\left(1-P(Allo\ HSCT\ CMR)\right)\times\left(1-P_{cycle}(Allo\ HSCT\ death)\right)$	Holter Chakrabarty et al., 2013 Hosing et al., 2003	Patients who neither reached molecular remission after allogeneic HSCT nor died.

From	To	Strategy	Value	Calculation	Source	Justification
Autologous HSCT	Failure	All	0.01852	$(1 - P(Auto\ HSCT\ CMR)) \times (1 - P_{cycle}(Auto\ HSCT\ death))$	Holter Chakrabarty et al., 2013 Hosing et al., 2003	Patients who neither reached molecular remission after autologous HSCT nor died.
Allogeneic HSCT	APL related death	All	0.06712	$P_{cycle}(Allo\ HSCT\ death) = 1 - (1 - P(Allo\ HSCT\ death))^{28/Allo\ HSCT\ FU\ duration}$	Hosing et al., 2003	
Autologous HSCT	APL related death	All	0.02001	$P_{cycle}(Auto\ HSCT\ death) = 1 - (1 - P(Auto\ HSCT\ death))^{28/Auto\ HSCT\ FU\ duration}$	Hosing et al., 2003	
Allogeneic HSCT remission	APL related death	All	0.00199	$P_{cycle}(death\ after\ HSCT) = 1 - (1 - P(fatal\ complications\ after\ HSCT))^{28/30}$	de Botton et al., 2005	Patients experiencing fatal complications during post-HSCT remission.
Autologous HSCT remission	APL related death	All	0.00199	$P_{cycle}(death\ after\ HSCT) = 1 - (1 - P(fatal\ complications\ after\ HSCT))^{28/30}$	de Botton et al., 2005	Patients experiencing fatal complications during post-HSCT remission.
Failure	APL related death	All	0.03117	$P_{cycle}(failure\ death) = 1 - (1 - (P(fatal\ complications\ after\ HSCT) + P(death\ after\ HSCT\ failure)))^{28/30}$	de Botton et al., 2005 Ramadan et al., 2012	In addition to the mortality following HSCT failure, those patients were affected by the mortality due to fatal complications from HSCT. Patients who did not die stayed in the failure health state.
MDS	APL related death	All	0.02685	$P_{cycle}(MDS\ death) = 1 - (1 - P(monthly\ MDS\ death))^{28/30}$	Ma et al., 2007	Patients who did not die stayed in the MDS health state.

B7. **Priority question:** Please justify the extrapolation of the long-term relapse in the model for both patients in first remission (i.e. after 50 months) as well as patients in second remission. Moreover, please provide scenarios using alternative assumptions for extrapolating the long-term relapse probabilities.

Company response: The assumption of a lifetime constant rate for a delayed relapse (from first long-term remission (+2 y) and second remission) is a limitation of the model which was validated with the experts. This assumption was used for simplification, as considering long-term relapses after 50 months would lead to a large number of tunnel states overcomplicating the model. The uncertainty around this assumption was studied in the DSA, with low values of cumulative relapse probabilities at 48 months set to be the same as the cumulative probabilities of relapse at 24 months (meaning no relapses occurred after 24 months, i.e. a plateau was reached for cumulative relapse incidence) and high values set to150% of the base case value. The results are presented in *Table 8* below.

Table 8. DSA results regarding relapse probability

	Incremental costs	Incremental QALYs	ICER
ATRA+ATO - Probability of relapse at 48 months – Low value: 0.009	- £46,300.71	3.02	Dominant
ATRA+ATO - Probability of relapse at 48 months – Low value: 0.029	- £18,537.76	2.27	Dominant
AIDA - Probability of relapse at 48 months – Low value: 0.082	£25,700.86	1.00	25,658.10
AIDA - Probability of relapse at 48 months – Low value: 0.209	- £66,545.5	3.78	Dominant

A scenario analysis where relapses after 24 months of first remission were not considered in both treatment arms was conducted. In this scenario, ATRA+ATO generated more costs (discounted incremental cost: £10,671) and the QALYs gained were reduced compared to the base case (1.40 vs. 2.62), leading to an ICER of £7,610 per QALY gained. Incremental results are presented in *Table 9* below.

Table 9. Incremental results of the scenario without relapses after 24 months of first remission

	Incremental costs		**Incremental effectiveness**		**Incremental ratio**	
ICERs	Not discounted	Discounted	Not discounted	Discounted	Not discounted	Discounted
Cost / QALY (ICUR)	-£1,793	£10,671	2.72	1.40	Dominant	**£7,610**
Cost / LY	-£1,793	£10,671	3.19	1.63	Dominant	**£6,529**

B8. The probability of relapse for patients in second remission is retrieved from Tallman *et al.* 2015. Please provide details of how this probability was obtained, whether any alternative sources for this probability are available (for example the study by Raffoux *et al.*) provide justification for both the data source used and any calculations performed.

Company response: There is a lack of data in relapsed/refractory APL, likely due to its rarity, and the best source we could identify was the publication by Tallman et al., which provided data that was validated by experts and peer-reviewed. A major limitation of using data from Raffoux et al. is the very low number of patients (10 patients per arm).

Regarding the question on how the probability reported by Tallman et al. was obtained, the publication specifies the following:[15] "The monthly estimates from the KM curves for EFS and OS were used to estimate the following monthly transition probabilities: 1) from first-line stable disease to second-line stable/first-line disease event, 2) from first-line stable disease to death, 3) from second-line stable/first-line disease event to second-line disease event, 4) from second-line stable/first-line disease event to death, 5) from second-line disease event to death. Calibration was conducted using Microsoft Excel Solver where the calibration process manipulated all of the transition probabilities so that the deviation between the observed clinical trial data and the predicted data (model-produced outcomes) was minimized.

Microsoft Excel Solver calculated the transition probabilities such that the sum of the percent absolute difference between EFS and OS in step 1 and step 2 was as small as possible. This process was repeated for each treatment arm in each stage (i.e., first- or second-line) of the analysis."

In the submitted model, we analysed the uncertainty around this parameter in the DSA and PSA. Despite a large variation used in the DSA (±50% of the base case values), the parameter did not appear to have an important impact on the results – it was never among the 20 parameters that influenced incremental results the most, and ATRA+ATO dominated AIDA in all tested scenarios.

B9. The company assumed that patients in the model who did not undergo polymerase chain reaction (PCR) testing (and therefore whose molecular remission status could not be evaluated) were in remission. Please justify this assumption and explain its implications, given that it was also assumed that none of the patients who received ATRA + ATO as first line treatment underwent PCR testing in the model – was it assumed that these patients were in molecular remission?

Company response: Based on the opinions from experts involved in the APL0406 clinical trial (Prof. Lo-Coco and Dr Cicconi), patients who did not undergo PCR testing followed the whole course of consolidation and were considered in remission until they relapsed. Our assumption is thus in line with clinical practice.

The comment "given that it was also assumed that none of the patients who received ATRA + ATO as first line treatment underwent PCR testing in the model" does not seem to be accurate. In line with the expert opinion from Dr Dillon, patients treated with ATRA+ATO in first line underwent 5 PCR tests over the course of health states

corresponding to treatment phases, while patients in the AIDA arm received only 4 PCR tests over the same time. However, when in remission, patients in the ATRA+ATO arm do not require PCR monitoring for signs of molecular relapse, and only clinical examinations are performed at 3-monthly follow up visits. The assumption that no PCR tests were performed in the first remission state in the ATRA+ATO arm is thus directly based on UK clinical expert opinion from Dr Dillon.

Quality of life

B10. Some (not all) health state utility values are adjusted for both age and for the 'utility representing perfect health'. This method of adjustment results in counter-intuitive utility values, i.e. utility values for 45 year olds that are lower than utility values for 60 years olds with the same condition. Additionally, the 'allogeneic HSCT' and 'autologous HSCT' health state utility value obtained from Breitscheidel (2008) is adjusted by multiplying it by the 'second-line molecular remission' health state utility value.

Company response: please note that utility values were extremely difficult to find in the literature; indeed, they were among the scarcest inputs. Since we did not identify any data specific to APL, we were forced to use data for other related conditions and make adjustments to make the values compatible. The relations between all utility values used in the model are logical (e.g. values for remission states are always higher than for on-treatment states) and all utilities compare appropriately to the utility in the general population. Please see further explanations for our assumptions below; we also provide a scenario with no age adjustments.

a. Please justify why an age adjustment is necessary at all, given that the impact of this disease would far outweigh any age-related utility decrements.

Company response: We have used the same method as in the study by Tallman et al. (2015) to adjust the utility values from another leukaemia field (chronic lymphocytic leukaemia, CLL) to APL based on age differences. We agree with the ERG that the impact of the disease is greater than the age-related utility decrement. However, we decided to consider the age impact nonetheless, as a 15-year difference in average age is quite significant.

It must be noted that all health state utilities are directly or indirectly adjusted for age, as shown in Table 3.4 of the main submission, except for the End of life state (Palliative care) and tMDS/AML. By indirect adjustment, we mean applying a disutility to the age-adjusted utility values, such as in the HSCT states. For palliative care, no adjustment was made, as this utility was based on patients with a malignancy in general, and not only on CLL patients. For tMDS/AML, the value was based on patients with a secondary malignancy following treatment for prostate cancer. Perhaps this value could also have been adjusted; however, the health state is an absorbing state with a very low utility (0.4) that would definitely outweigh the effect of any age differences.

b. Please justify why some utilities are adjusted for the 'utility representing perfect health'.

Company response: Some of the utility values were higher that the utility value of the general population and they needed to be adjusted. For instance, the utility value reported in Beusterien 2010 for remission was 0.91 (95%CI: 0.88, 0.93). This utility range is above the utility value for general population at age 60. Therefore, we had to perform a two-step adjustment, as described in the main submission. First, we adjusted the utility value to be in line with the utility of general population. Second, we performed an age adjustment to account for the age difference between CLL and APL populations. This is certainly what gave the impression of counter-intuitive values to the ERG, but the adjustment method does not led to "utility values for 45 year olds that are lower than utility values for 60 years olds with the same condition".

c. Please justify why the 'allogeneic HSCT' and 'autologous HSCT' health state utility value is adjusted by the 'second-line molecular remission' health state utility value.

Company response: As mentioned by the experts (Prof. Lo-Coco and Dr Cicconi), only patients who were in remission were eligible for HSCT, with the exception of patients experiencing a major adverse event (cardiac event), who moved to allogeneic HSCT. Given that the majority of HSCT patients were previously in the second remission state, we applied a disutility to the second remission state utilities.

d. Provide a scenario analysis without any of these utility value adjustments.

Company response: A scenario without any utility adjustment is implemented in the new version of the model provided (v3.17). We would like to insist on the fact that this scenario presents illogical and counter-intuitive ranking of utility values between health states (e.g. remission states with higher values than the general population, HSCT health states with higher values than remission health states). The incremental QALYs accumulated by patients treated with ATRA+ATO were slightly higher than in the base case, (2.96 discounted QALYs vs. 2.62). A more relevant scenario analysis regarding utility values was implemented with values used in the publication by Tallman et al. This scenario was presented in section B3.8.3.2 of the company submission.

B11. Health state utility values in the cost-effectiveness analysis remain constant during the entire time horizon. However, when patients flow through the model, they also become older resulting in a decreased utility value. Please provide a scenario analysis implementing health state utility values that do not exceed general population utility values (e.g. by capping all health state utility values to the age-dependent general population utility values).

Company response: We thank the ERG for flagging this critical point and we fully agree on the fact that this was a serious limitation of the model. Version 5.1 of the model (Excel file: ID446 Trisenox_NICE_CEM_v5.1_utility_capping_12012018.xlsm) accompanying this document with company responses implements this correction. For each health state, we compared 1) the utility of the general population according to the age of patients with 2) the utility value of the heath state defined in the original model submitted to NICE, and used the lower value of the two as the "new"utility for that health state. The correction resulted in a smaller number of incremental QALYs compared to the base case: 2.41 discounted incremental QALYs vs. 2.62. This had no critical impact on the ICER as ATRA+ATO still dominated AIDA.

B12. The choice of the sources for the different utility and disutility values are not always clearly described in the company submission. Additionally, it is unclear how studies informing the disutility values for the adverse events were identified.

a. Please describe how the studies informing the disutility values of adverse events were identified.

Company response: We conducted targeted literature searches with terms relevant to adverse events. For each event, we captured a number of potential articles. Then we selected the most plausible source by considering the following criteria:

- Articles that were more than 10 years old were excluded
- Articles that included utility values based on pure assumptions were excluded
- Articles that reported utility/ disutilities considering different states within the adverse event were removed, as our model only included one utility value
- Articles that provided disutilities too small or too large according to experts (Prof. Lo-Coco and Dr Cicconi) were excluded

b. Please explain the choice of each source used to obtain health state utility values and disutilities for adverse events.

Company response: We selected the reference based on the exclusion criteria listed under part a above. Since the search for disutilities was only a targeted search, we did not report the reasons for exclusion for each article. We extracted the utility/ disutility information from the remaining articles and, if needed, applied the adjustment factors. Finally we discussed the values identified in the literature with experts and, if there was more than one value available for a health state/ adverse event, we asked the expert to select the most relevant value.

c. 'Hospitalisation', 'leukocytosis', 'differentiation syndrome' and 'QTc prolongation' disutility values are based on assumptions – please justify these assumptions as well as the disutility values. Please refer to relevant sources or expert opinion.

Company response: We did not find any specific utility values for these AEs in literature, so we had to make different assumptions as listed below.

- For hospitalizations we selected the conservative value of -0.01, given that the number of hospital days was smaller for the ATO+ATRA strategy.
- For leucocytosis we selected the value of -0.08, as a value that appeared logical to experts in comparison with the disutilities of other AEs.
- For QTc prolongation, we received the following feedback from experts (Prof. Lo-Coco and Dr Cicconi): "Patients receive ECG daily until normalisation. Usually it normalises within 2–3 days from drug discontinuation and no major cardiac events have been registered. Among 2,900 cases treated by US-FDA approved ATO (for relapse) there have been no arrhythmia-related deaths." This lead to the

assumption of very small utility difference. We suggested the disutility of -0.01 which was approved by the experts.

- For differentiation syndrome, assigning a utility value was a challenge, as the syndrome comprises a range of conditions, such as dyspnea, pulmonary infiltrates, oedema, unexplained fever, weight gain, etc. We selected the value of -0.12 as it appeared logical based on experts' opinion.

We have ran a conservative scenario to test the worst cases for the ATRA + ATO strategy; i.e. high disutilities (double the base case values) for AEs that are not in favour of ATRA+ ATO (leucocytosis, hepatic toxicity, neurotoxicity, differentiation syndrome, QTC prolongation) and zero disutilities for AEs that are in favour of ATRA+ATO (thrombocytopenia, neutropenia, infection and cardiac event). The incremental discounted QALYs were hardly impacted: 2.57 vs. 2.62 in the base case.

d. Please provide an overview of the health states in which patients are at risk of hospitalisation in the cost effectiveness model.

Company response: The six health states in which patients may be hospitalised are:

- First line induction
- First line consolidation
- Second line induction +1 cycle consolidation
- Second line consolidation
- Allogeneic and autologous HSCT

e. Please justify why acute and chronic graft versus host disease (GvHD) have the same disutility value

Company response: The source we used for this disutility value did not distinguish acute and chronic GvHD, raising the assumption that chronic GvHD is a recurrence of the acute condition. This assumption seemed to be a reasonable and was validated by experts.

f. Please justify why a multiplicative framework is used to estimate the disutility for 'Acute GvHD'.

Company response: Thank you for flagging this. The idea was to take into consideration the fact that utilities where adjusted (re-scaled) in the source study (Breitscheidel L., 2008). Nevertheless, a better way to compute this disutility would be to simply compute the difference between the health state with the condition and the one without the condition, i.e. 0.769-0.836 = -0.067. A scenario using this value was implemented in the model and the results were unaffected, i.e. exactly the same number of QALYs was accumulated in each arm.

Adverse events

B13.

a. Please provide an overview of the adverse events implemented in the model, their frequency of occurrence, their duration and the sources on which this information is based. Please justify the choice of each source.

Company response: The rates of adverse event occurrence were retrieved from the publications on the APL0406 trial (Platzbecker et al.[13] and Lo-Coco et al.[16]). This was considered the most appropriate source, since patients included in the cost-effectiveness analysis presented the same characteristics and followed the same treatment schedule as in the APL0406 trial. Furthermore, efficacy data were also sourced from the same publications. Unfortunately, the duration of the adverse events was not provided in these publications. A targeted literature search was conducted in order to find the most appropriate data and the data collected were submitted to experts for validation.

	ATRA + ATO		AIDA		
	Induction	Consolidation	Induction	Consolidation	Source and justification
Thrombocytopenia (grade 3-4, >15 days)					
Probability of event	0.3800	0.1870	0.6200	0.8100	Platzbecker et al., 2017
Event duration (days)	20	25	20	25	Wolff et al., 1989. No data were found for both studied therapies in APL patients. The publication provided durations of thrombocytopenia during the induction and consolidation treatment phases for AML patients treated with ATRA+DNR+ARA-C. It was assumed that the average duration of this event for patients treated with ATRA+chemotherapy would be the same. To keep the assumption conservative, the same values were used for ATRA+ATO. The reasoning and values were validated by experts.
Neutropenia (grade 3-4, >15 days)					
Probability of event	0.3500	0.1270	0.6400	0.8130	Platzbecker et al., 2017
Event duration (days)	19	19	19	19	Fenaux et al., 1993. The data concerned patients treated with ATRA+DNR+ARA-C, and the values were derived as described above for neutoropenia.
Infection					
Probability of event	0.2300	0.0420	0.5500	0.1520	Platzbecker et al., 2017
Event duration (days)	17	17	17	17	Pneumonia – What happens (http://www.webmd.com/lung/tc/pneumonia-what-happens). Based on table A1 in Platzbecker et al. 2017, most infections are pneumonia; thus, the duration of pneumonia was used in the model. According to the brochure

					Pneumonia – What happens, pneumonia last between 2 and 3 weeks; we used an average of 17 days.
Leukocytosis					
	Probability of event	0.4730	0.0000	0.2410	0.0000	Lo-Coco et al., 2013
	Event duration (days)	14	14	14	14	Shoenfeld et al., 1981. “Leukocytosis reached maximal values within two weeks in most cases”. Expert: With the use of hydroxyurea the median time to normalisation of the leukocyte count is 10.5 days. The assumption that leucocytosis affects quality of life during the 14 days around the peak in both arms was validated by experts.
Hepatic toxicity					
	Probability of event	0.4000	0.0160	0.0300	0.0023	Platzbecker et al., 2017
	Event duration (days)	10	10	10	10	Zhu et al., 2013 observed hepatic toxicity lasting <2 weeks with ATRA+ATO. Based on this, we used a duration of 10 days and assumed the same duration for AIDA. This was validated by experts.
Neurotoxicity					
	Probability of event	0.0070	0.0503	0.0000	0.0000	Platzbecker et al., 2017
	Event duration (days)	365	365	365	365	Assumption based on Ratnaike, 2003. Acute poisoning from arsenic can lead to peripheral neuropathy which can last for up to 2 years. Neuropathy duration of 1 year was validated by experts. No neurotoxicity was observed with AIDA; nevertheless, the same duration was considered in SA when neurotoxicity is introduced in this arm.

Differentiation syndrome					
Probability of event	0.1940	0.0000	0.1600	0.0000	Platzbecker et al., 2017 (ATRA+ATO). Lo-Coco et al., 2013 (AIDA)
Event duration (days)	4	4	4	4	According to Breccia et al., 2008, resolution of retinoic acid syndrome occurred after a median time of four days in patients treated with AIDA. The same duration was assumed for ATRA+ATO. The assumption was validated by experts.
Cardiac events					
Probability of event	0.0000	0.0000	0.0370	0.0000	Platzbecker et al., 2017
Event duration (days)	1	1	1	1	Siu et al., 2006 stated that "The QTc prolongation were transient and lasted only 4 hours" in patients treated with ATRA+ATO. The same value was used for AIDA (conservative assumption).
QTc prolongation					
Probability of event	0.0850	0.0183	0.0070	0.0000	Platzbecker et al., 2017
Event duration (days)	0.5	0.5	0.5	0.5	
Myelodysplastic syndrome (MDS)					
Probability of event	0.0000	0.0000	0.0000	0.0146	Platzbecker et al., 2016

b. Please explain whether it is assumed that GvHD is reversible in the model and why?

Company response: As we considered both acute and chronic GvHD, the disease could be reversible in a proportion of patients. Nevertheless, to simplify the structure of the model and achieve calculation transparency (no VBA code in the Markov trace calculations), chronic GvHD was assumed to be a lifetime condition, and so not reversible. At each cycle, a proportion of the patients in the allogeneic HSCT remission state experienced chronic GvHD. Since only a small proportion of the cohort was affected by this chronic condition (0.032% in the ATRA+ATO arm, 0.135% in the AIDA arm), the experts stated that it was a reasonable assumption.

Costs and resource use

B14. Please justify why PCR tests are not performed in the molecular remission health state for patients who received first-line treatment with ATRA+ATO. Please add a positive number of PCR tests into the model for the ATRA+ATO arm, if necessary.

Company response: The UK clinical expert, Dr Dillon stated that patients in molecular remission after first-line ATRA+ATO treatment do not undergo PCR tests, but do have follow-up appointments three months. This assumption that assessing patients for clinical symptoms of relapse is sufficient is almost certainly due to the very low rate of relapse following first-line treatment with ATRA+ATO. A scenario where four PCR tests per year are performed during the first two years of remission in the ATRA+ATO arm (as in the AIDA arm) was analysed and is implemented in the last version of the model. This scenario led to a reduction of the discounted cost savings generated with the use of ATRA+ATO to £28,301 from £31,270 in the base case.

B15. A systematic literature review was performed to collect appropriate costs and resource use data for England. Nevertheless, none of the 11 identified studies is used in the economic model because *"information captured was not compatible with that needed to populate the model, and in others by the fact NHS reference costs were preferentially used to ensure relevance to the current situation in England"* (CS page 124). Please provide more specific justification, for each resource use and cost item, why the chosen sources are the most appropriate.

Company response: The systematic literature search for costs and resources identified 7 full-text publications and 4 abstracts; the full list is available in Appendix I. Three of the full-text studies were cost-effectiveness analyses by Lachaine et al.[17, 18] and Tallman et al.,[15] which focused on Canada and the US, respectively. Other studies provided data for the US[19], China[20, 21] and Japan[22]. Since the healthcare systems (and potentially treatment paradigms) in these countries are substantially different from the UK, these publications did not provide information suitable to inform our model. Another study by Kruse et al. (available as a publication[23] and an abstract[24]) concerned budgetary impact of ATO in Italy; however, little detail on inputs was provided in the publication. Furthermore, the authors stated that unit costs were based on local price lists and tariffs[23], suggesting they would not be readily applicable to the UK. Further, two abstracts based on the APL0406 trial[25, 26] provided some information on resource use and costs

associated with ATRA+ATO and AIDA, but the data was restricted to Germany and largely based on the local DRG system, making the results hardly applicable to the UK.

In light of the lack of appropriate data derived from a systematic approach, we aimed to use NHS reference costs and the PSSRU wherever possible, supplementing this with data from studies identified through a targeted search where necessary. Since the search was only targeted, we did not report the reasons for inclusion or exclusion for each article. However, we did aim to obtain plausible cost and resource use estimates, and tested the uncertainty around these values in deterministic and probabilistic sensitivity analyses. DSA showed that only the cost and number of bed days during 1st-line induction treatment affected the incremental costs (but not the ICER) substantially. Both of these inputs were derived from high-quality data. The number of bed days was based on the median time to remission reported in the APL0406 trial[13], and the cost per bed day on the relevant NHS reference cost.

B16. The health state costs include costs related to follow-up appointments and PCR tests for the whole time horizon of the model

a. Please justify the inclusion of life-long costs related to follow-up appointments and PCR tests and provide a scenario with alternative assumptions related to this (for example, excluding these costs after a certain time in remission).

Company response: The costs of follow-up are limited to 2 years for patients in first remission. Due to the structure of the model, it is not possible to limit follow-up costs for patients in remission following second-line treatment or a HSCT. The version of the model provided with this clarification letter proposes a scenario in which no follow-up visits and PCR tests are considered in second and post-HSCT remission. The discounted cost savings were reduced to -£26,510 vs. -£31,270 in the base case. Follow-up and monitoring costs were still lower with ATRA+ATO than with AIDA (respectively, £2,111 and £4,750).

Furthermore, please note that the “worst case” scenario presented in appendix J (section J 4.2) of the company submission does not consider any follow-up cost in either arm.

b. Please explain why costs related to monitoring of haematological response were not included in the model.

Company response: Costs related to monitoring haematological response were not considered in the model, as they were not mentioned by any of the experts. Most likely, this was because the benefits of treating at molecular relapse before progression into a haematological one are widely recognised, so the monitoring for relapse focuses more on molecular (PCR) testing. In comparison with the costs of bone marrow biopsy followed by PCR testing, the costs of haematological monitoring (blood counts, serum biochemistry, coagulations tests[27]) was considered negligible. Furthermore, not including these costs in the analysis is conservative, i.e. they would lead to equivalent monitoring cost for both first-line treatments and to higher costs following second-line treatment in the AIDA arm, since more patients in that arm relapse.

Validation and transparency

B17. Priority question: Please provide the experts' responses (for example meeting minutes) to all the questions asked in the presentation shown in Appendix M.

Company response: the minutes of the meeting are enclosed

B18. Please provide a cross-validation of the assumptions, inputs and outputs in the cost-effectiveness model, with the cost-effectiveness analyses identified in the systematic literature review.

Company response: As stated above, the previously published models only describe very schematic heath states that do not adequately reflect the trajectory of APL patients. Since we have developed a more comprehensive approach, it is not possible to perform an exact cross-validation of the assumptions, inputs and outputs. However, the scenario where we used utilities from Tallman et al. 2015 leads to an increased QALY gain compared to the base case. For key clinical inputs, we used the same study as Tallman et al (APL0406), the only difference being that, since the Tallman model was published in 2015, the authors used the initial report from Lo-Coco (2013)[16], while we used the updated report from Platzbecker (2017)[13].

Scenario and sensitivity analyses

B19. **Priority question:** Please provide a model which includes an option to select the scenario analyses performed (both those in the original company submission as well as those performed in response to the clarification letter) by the company.

Company response: Version v3.17 of the model accompanying this clarification letter includes a list box for scenario choice in the "Settings & Population" sheet. In addition to the base case, twelve scenarios are implemented. Due to their complexity, the following scenarios were not included in this version, but instead are provided in separated files:

- Second-line treatment with AIDA in both arms
- Second-line model
- Health state representing the first two years of remission comprising 26 tunnels
- Scenario where the utility value of health states cannot exceed the age-dependent utility value of the general population.
- PSA implementing a Gamma distribution for resource use.

a. Please also provide a scenario analysis implementing maintenance treatment.

Company response: A scenario is available in v3.17 of the model implementing two years of maintenance in the AIDA arm, with schedule and dosage as described in Lo-Coco et al, 2013[16]; i.e. "intramuscular or oral methotrexate (MTX) at a dose of 15 mg per square meter per week and oral 6-mercaptopurine (6-MP) at a dose of 50 mg per square meter per day, alternating with ATRA at a dose of 45 mg per square meter per day, for 15 days every 3 months for 2 years". We would like to insist on the fact that in clinical

practice maintenance is not recommended and used anymore in the UK. One should note that this scenario only impacts treatment acquisition costs in the AIDA arm and is therefore not conservative. Discounted incremental costs were increased to -£33,012 vs -£31,270 in the base case.

b. Please also provide a scenario analysis implementing a Gamma distribution for resource use (the Normal distribution that is currently used might result in negative numbers for resource use).

Company response: In the file ID446 Trisenox_NICE_CEM_v3.13_PSA_RU_gamma_11012018.xlsm, the PSA was run with a Gamma distribution for resource use instead of the "positive" Normal distribution initially used. The results of the PSA are presented below. Please note that in the initial PSA, the Normal distribution used could not produce negative numbers, as a new random number was generated each time a negative one was drawn. This approached aimed to avoid potential high values due to the long right tails of Gamma distributions. The Excel file contains a "PSA - Gamma vs. Normal+" presenting base-case values and average of the simulated numbers for the two distributions.

Results were not substantially impacted and stayed consistent with the previous PSA. The average incremental cost was -£31,668 vs. -£31,088 in the initial PSA. The cost-effectiveness acceptability curve demonstrated that the probability of ATRA+ATO being cost-effective was around 82.27% (vs. 81.33% in the initial PSA) at a £0 willingness-to-pay threshold, and increased further when this threshold increased, to reach 93.87% at £30,000 per QALY (same as in the initial PSA). Furthermore, ATRA+ATO dominated AIDA in 77.73% of the simulations (vs. 77.13% previously).

Table.10. Probabilistic sensitivity analysis results.

	Incremental costs (£)	Incremental QALYs	ICER (£/QALY)
Mean	-31,668	2.49378	Dominant
Std Deviation	36,212	2.08712	339,362
Median	-29,047	2.41434	9,571
Min	-192,363	-11.50735	10
Q 0.025	-111,497	-1.65964	303
Q 0.975	29,914	6.82671	182,293
Max	67,413	10.79954	5,152,754

Figure 1. Incremental cost-effectiveness plane for ATRA+ATO vs AIDA

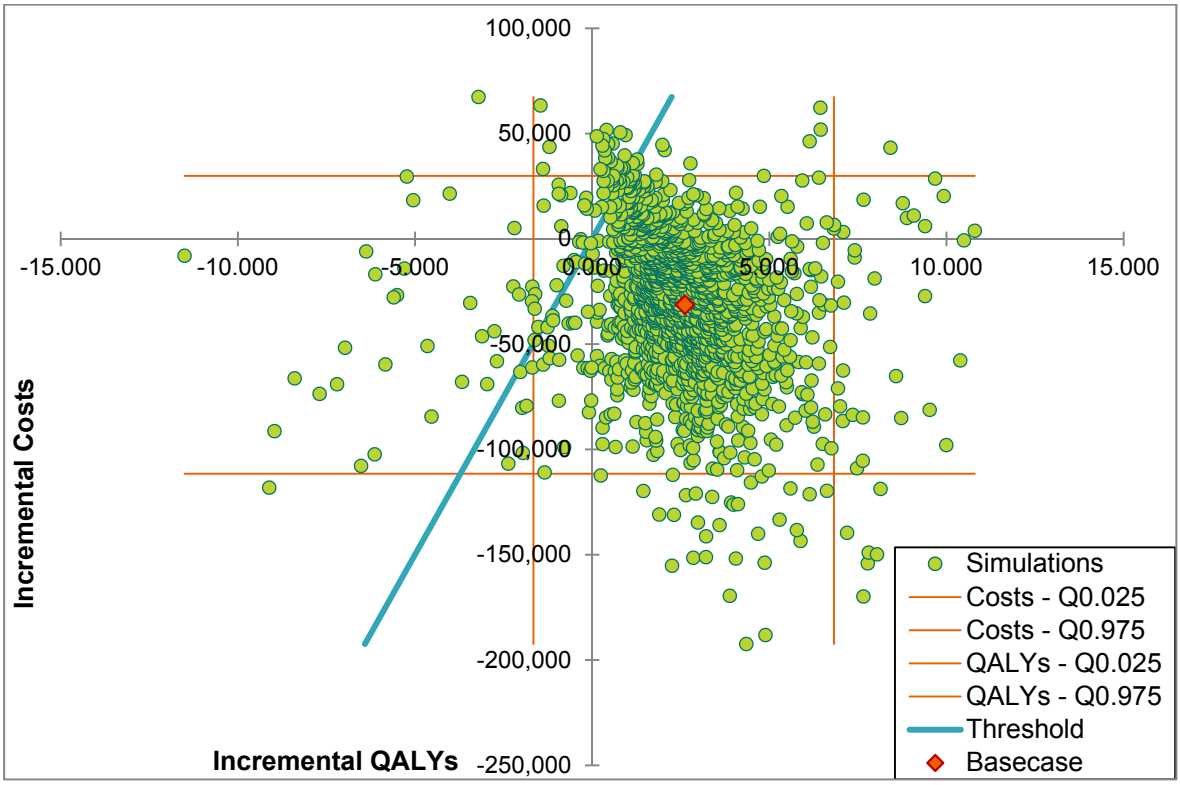
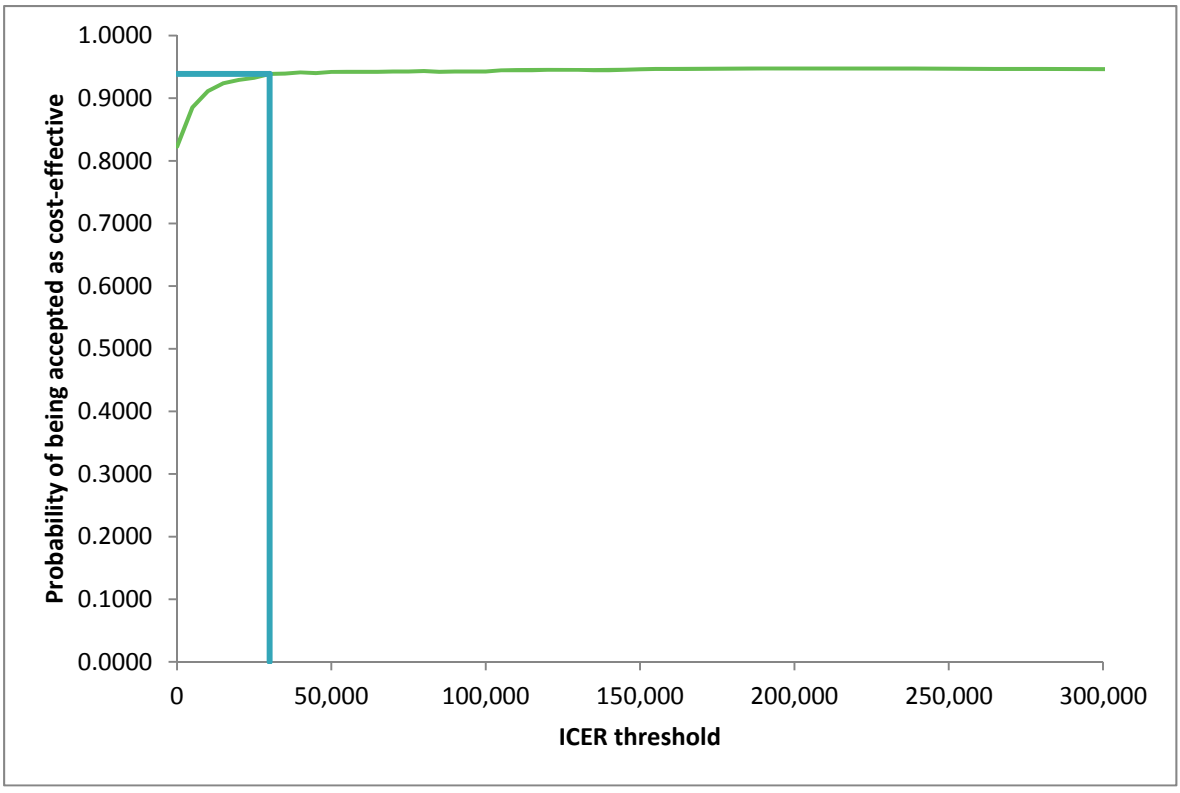


Figure 2. Cost-effectiveness acceptability curve for ATRA+ATO vs AIDA



c. Please also provide a scenario analysis implementing the consolidation health state with 26 cycles (reflecting a maximum of 2 years in that health state), instead of the 24

cycles used. (The cycle length in the model is 4 weeks, meaning that one year contains 13 cycles. However, the 2-year consolidation health state consists of 24 cycles.)

Company response: This scenario is implemented in the file ID446 Trisenox_NICE_CEM_v4.2_rem26_12012018.xlsm, in which the first two years of remission are represented with 26 tunnel states. The impact of the 2 missing tunnels was minor and the base case was conservative. The scenario with 26 tunnels provided slightly better results for ATRA+ATO: incremental costs were -£31,813 vs. -£31,270 in the base case and incremental QALYs were 2.63 vs. 2.62.

B20. **Priority question:** At the end of the time horizon of 40 years chosen by the company, 45% of patients in the ATRA+ATO first and second line arm (48% of patients in the ATRA+ATO first line and AIDA second line) are still alive. The chosen time horizon is therefore not a lifetime horizon.

a. Please justify the long life expectancy of patients in the model (a proportion of patients are still alive at ages of 100+ years) and comment on the plausibility of this.

Company response: The life expectancy of patients in the ATRA+ATO arm of the model is 80.85 years over the 100 years simulated; in the AIDA arm it is 73.33 years. The current life expectancy of a newborn is 79.2 years for a boy and 82.9 years for a girl, if mortality rates remain the same as they were in the UK in 2014–2016 throughout their lives[28]. As our hypothetical cohort was born before the current estimate, their life expectancy should be somewhat lower. Although life expectancy in the model is on the high end of the UK life expectancy, it is within the plausible range as we did not model disease-specific mortality in order to keep the model conservative (see the response to question B2).

Regarding the fact a proportion of patients in the model are still alive at 100+ years, this is also related to the life expectancy of APL patients in long-term remission being similar to the general population. Based on the UK mortality rates for 2014–2016, 40.9% of men and 53.8% of women in England and Wales can expect to live past the age of 85 (the maximum age modelled) and 1.2% of men and 2.7% of women past the age of 100[28]. Thus, the model is not implausible in that respect.

b. Please provide results with a lifetime time horizon that captures all relevant outcomes.

Company response: A scenario considering a time horizon of 56 years, meaning that patients were 100 years old at the end of the time horizon, was implemented in version v3.17. Discounted incremental savings increased to £32,922 (vs. £31,270 in the base case), incremental QALYs increased to 2.83 (vs. 2.62) and LYs to 3.25 (vs. 3.00). At the end of the time horizon, 0.96% of patients in the ATRA+ATO and 0.44% in the AIDA arm were still alive, which is slightly less than the current general population estimates. This result appears plausible, considering the modelled cohort was born before the current population life expectancy estimates presented in point a, and that some patients died in the course of the disease.

B21. The CS states that “For patients in first molecular remission, the probability of relapse at 48 months was assumed to be equal to that at 50 months” (page 105). Please clarify why these were assumed to be equal instead of converting the 50 month probability to a 48 month probability of relapse.

Company response: Many thanks for flagging this. This was a simple approximation used in our model. A scenario where these probabilities were adjusted to 48 months is implemented in the last version of the model. In this scenario, incremental costs were -£28,555, and 2.53 additional QALYs and 2.90 additional LYs were accumulated with ATRA+ATO than with AIDA. Consequently, ATRA+ATO was dominant vs AIDA.

References

1 Powell BL, et al. Arsenic trioxide improves event-free and overall survival for adults with acute promyelocytic leukemia: North American Leukemia Intergroup Study C9710. *Blood*. 2010;116(19):3751-7.
2 Soignet SL, et al. United States multicenter study of arsenic trioxide in relapsed acute promyelocytic leukemia. *J Clin Oncol*. 2001;19(18):3852-60.
3 Soignet SL, et al. Complete remission after treatment of acute promyelocytic leukemia with arsenic trioxide. *N Engl J Med*. 1998;339(19):1341-8.
4 Lengfelder E, et al. Arsenic trioxide-based therapy of relapsed acute promyelocytic leukemia: registry results from the European LeukemiaNet. *Leukemia*. 2015;29(5):1084-91.
5 Thomas X, et al. Superiority of an arsenic trioxide-based regimen over a historic control combining all-trans retinoic acid plus intensive chemotherapy in the treatment of relapsed acute promyelocytic leukemia. *Haematologica*. 2006;91(7):996-7.
6 Ganzel C, et al. Autologous transplant remains the preferred therapy for relapsed APL in CR2. *Bone Marrow Transplant*. 2016;51(9):1180-3.
7 Yanada M, et al. Autologous hematopoietic cell transplantation for acute promyelocytic leukemia in second complete remission: outcomes before and after the introduction of arsenic trioxide. *Leuk Lymphoma*. 2017;58(5):1061-7.
8 Teva Pharmaceuticals Limited. APL Market Understanding & Role of Trisenox. Data on file. May, 2015.
9 Burnett AK, et al. Inclusion of chemotherapy in addition to anthracycline in the treatment of acute promyelocytic leukaemia does not improve outcomes: Results of the MRC AML15 trial. *Leukemia*. 2013;27(4):843-51.
10 Burnett AK, et al. Arsenic trioxide and all-trans retinoic acid treatment for acute promyelocytic leukaemia in all risk groups (AML17): results of a randomised, controlled, phase 3 trial. *Lancet Oncol*. 2015;16(13):1295-305.
11 Russell N BA, Hills R, Betteridge S, Dennis M, Dillon R, Grimwade D,. National Cancer Reseaserch Institute,. The Feasibility of More Attenuated Dosage Schedules of ATO in Newly Diagnosed And Relapsed Acute PromyelocyAc Leukaemia. Data on file. 2017.
12 Raffoux E, et al. Combined treatment with arsenic trioxide and all-trans-retinoic acid in patients with relapsed acute promyelocytic leukemia. *J Clin Oncol*. 2003;21(12):2326-34.
13 Platzbecker U, et al. Improved Outcomes With Retinoic Acid and Arsenic Trioxide Compared With Retinoic Acid and Chemotherapy in Non-High-Risk Acute Promyelocytic Leukemia: Final Results of the Randomized Italian-German APL0406 Trial. *J Clin Oncol*. 2017;35(6):605-12.

14 Russell NH, et al. Long Term Follow up from the NCRI AML17 Trial of Attenuated Arsenic Trioxide and ATRA Therapy for Newly Diagnosed and Relapsed Acute Promyelocytic Leukaemia. *Blood*. 2016;128(22):897.
15 Tallman M, et al. Cost-Effectiveness Analysis of Treating Acute Promyelocytic Leukemia Patients with Arsenic Trioxide and Retinoic Acid in the United States. *Clinical Lymphoma, Myeloma and Leukemia*. 2015;15(12):771-7.
16 Lo-Coco F, et al. Retinoic acid and arsenic trioxide for acute promyelocytic leukemia. *N Engl J Med*. 2013;369(2):111-21.
17 Lachaine J, et al. Economic evaluation of arsenic trioxide compared to all-trans retinoic acid + conventional chemotherapy for treatment of relapsed acute promyelocytic leukemia in Canada. *Eur J Haematol*. 2015;95(3):218-29.
18 Lachaine J, et al. Economic evaluation of arsenic trioxide for treatment of newly diagnosed acute promyelocytic leukaemia in Canada. *Hematol Oncol*. 2015;33(4):229-38.
19 Eardley AM, et al. Morbidity and costs of remission induction therapy with all-trans retinoic acid compared with standard chemotherapy in acute promyelocytic leukemia. *Leukemia*. 1994;8(6):934-9.
20 Pei RZ, et al. [Clinical investigation of homoharringtonine in combination with all-transretinoic acid and arsenic trioxide for acute promyelocytic leukemia]. [Chinese]. *Zhonghua xue ye xue za zhi = Zhonghua xueyexue zazhi*. 2013;34(2):144-8.
21 Fan CS, et al. Budget impact analysis of oral tetra-arsenic tetra-sulfide formula in the treatment of acute promyelocytic leukemia. 2017.
22 A T. Significant reduction of medical costs by differentiation therapy with all-trans retinoic acid during remission induction of newly diagnosed patients with acute promyelocytic leukemia. 1995.
23 Kruse M, et al. Budgetary impact of treating acute promyelocytic leukemia patients with first-line arsenic trioxide and retinoic acid from an Italian payer perspective. *PLoS One*. 2015;10(8):e0134587.
24 Kruse MM, et al. Cost of treating acute promyelocytic leukemia (APL) patients in Italy. *Value in Health*. 2014;17 (7)A624.
25 Schlenk RF, et al. Cost-effectiveness analysis of arsenic trioxide in combination with all-trans retinoic acid in acute promyelocytic leukemia with pretreatment white blood counts <10g/l. *Blood Conference: 56th Annual Meeting of the American Society of Hematology, ASH*. 2014;124(21).
26 Schonsteiner S, et al. G-DRG 2016 based cost-effectiveness analysis of arsenic trioxide (ATO) and all-trans retinoid acid (ATRA) inpatients with acute promyelocytic leukemia treated within the APL0406 trial. *Annals of Hematology*. 2017;96(1):S76.
27 Platzbecker U, et al. German Intergroup recommendations on the diagnostic and therapeutic management of acute promyelocytic leukemia (APL). Version Final 3.0. [cited Oct, 2017]. Available from: https://www.kompetenznetz-leukaemie.de/sites/kompetenznetz-leukaemie/content/e53457/e55758/e57007/NapoleonRegister_TreatmentRecommendation_Final_V3.0_20150630.pdf.
28 Office for National Statistics. National life tables, UK: 2014 to 2016. [cited January 2018]. Available from: https://www.ons.gov.uk/peoplepopulationandcommunity/birthsdeathsandmarriages/lifeexpectancies/bulletins/nationallifetablesunitedkingdom/2014to2016.

NICE National Institute for Health and Care Excellence

Patient organisation submission

Arsenic trioxide for treating acute promyelocytic leukaemia [ID446]

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	Bloodwise
3. Job title or position	[redacted]
4a. Brief description of the organisation (including who funds it). How many members does it have?	Bloodwise's mission is to beat all blood cancers – stopping people from dying, improving the lives of everyone affected by blood cancer, and where possible preventing people getting blood cancer in the first place. We do this by funding world leading research, supporting all those affected by blood cancer, and campaigning for improvements in care and services. We are entirely funded by voluntary donations and have approximately 100 members of staff and 140 patient ambassadors plus many more volunteers and supporters.
4b. Do you have any direct or indirect links with, or funding from, the tobacco industry?	None
5. How did you gather information about the experiences of patients and carers to include in your submission?	We were unable to complete our submission for this appraisal before the earlier deadline and have been asked by the PIP team at NICE to make a submission at this stage with reference to the responses we received when preparing a recent submission to the All Wales Medicines Strategy Group (AWMSG) for their appraisal of the treatment - arsenic trioxide for the induction of remission, and consolidation in adult patients with newly diagnosed low-to-intermediate risk APL in combination with all trans retinoic acid (ATRA) characterised by the presence of the t(15;17) translocation and/or the presence of the promyelocytic leukaemia/retinoic-acid-receptor alpha (PML/RAR-alpha gene). In order to gather evidence for the AWMSG submission, we sent an email to our database of patient ambassadors asking them to contact us to share their experiences of APL and treatment with arsenic trioxide. We also consulted our medical advisory panel, an expert group of clinicians, to gain further

	insight into the condition and patients' experiences using this treatment from a clinical perspective. We heard from one patient who had the treatment as part of a clinical trial and also spoke to another patient who had not had the treatment but gave us great insight into life with the condition and had done his own research into the treatment so was fully informed on this. The clinician we spoke to was exceptionally knowledgeable about the treatment and had witnessed directly the effects on 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?	Patients with the condition told us that they felt exhausted and unwell constantly. The condition is very disruptive and the standard treatment (chemotherapy, usually idarubicin) very restrictive with one patient describing how he could not go out in public during treatment and had to give up work for 6 months.
Current treatment of the condition in the NHS	
7. What do patients or carers think of current treatments and care available on the NHS?	The standard treatment for APL is chemotherapy in combination with all-trans-retinoic acid (ATRA). The main disadvantages are that patients are exposed to the harsh side effects that accompany chemotherapy. Short term side effects include hair loss, severe nausea and vomiting, reduced white blood cell count leading to serious infection and long periods in hospital and inability to work during treatment. Longer term side effects include the risk of developing secondary cancers, effects on fertility in younger patients and increased risk of heart disease.
8. Is there an unmet need for patients with this condition?	The need for kinder treatments is not met by the highly toxic level of the treatment currently available. In addition there is a need for a lower chance of relapse as this risk has a significant impact on patients psychologically.

Advantages of the technology	
9. What do patients or carers think are the advantages of the technology?	Patients are treated with arsenic trioxide in combination with ATRA in 2 month treatment blocks. This involves attending hospital everyday for treatment in the first week, then twice a week for the next 3 weeks followed by a 4 week break. Patients are treated with idarubicin in one month blocks attending hospital most days in the first week followed by a 3 week break in treatment. However, although initial hospital attendance is higher for treatment with arsenic trioxide, the effect on patients' daily lives is much less significant than treatment with chemotherapy. Patients are usually able to continue working throughout treatment, whereas this is not possible for patients undergoing chemotherapy as their immune systems are so compromised. Arsenic trioxide also has minimal side effects compared with chemotherapy. The clinician and patients we consulted in preparing this submission highlighted one of the most significant advantages of arsenic trioxide as the exceptionally low relapse rate (close to zero). The relapse rate with idarubicin is in the region of 20-30% and those patients who relapse are often treated with arsenic trioxide at that point and so exposed to two lots of treatment. APL has a relatively high percentage of early deaths following diagnosis and more routine use of arsenic trioxide at an early stage could reduce this percentage.
Disadvantages of the technology	
10. What do patients or carers think are the disadvantages of the technology?	There are few disadvantages. The main side effect is neuropathy which is usually minimal. One patient who was randomised onto the APL trial in September 2011 describes how she suffered a major side effect from the arsenic trioxide, pseudo tumour cerebrai which caused serious headaches, vision changes and vomiting. She was advised by her treating consultant that this was due to her elevated white blood cell count and the reaction to rid her cells of leukaemia. She was monitored closely and had a combination of other drugs to counteract this and recovered after 3 weeks. The main other side effects she suffered were a sore mouth, cracked lips, mild tiredness and constipation. However, despite these

	effects, she describes being able to function well on the treatment. She has 9 months of treatment and has been in remission since then.
Patient population	
11. Are there any groups of patients who might benefit more or less from the technology than others? If so, please describe them and explain why.	
Equality	
12. Are there any potential equality issues that should be taken into account when considering this condition and the technology?	

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:

- Treatment with arsenic trioxide significantly reduces the risk of relapse so gives patients the reassurance of a good chance of survival. Whereas the risk of relapse with the current treatment for APL (chemotherapy with idarubicin) is much higher at 20-30%). This also helps patients psychologically as the psychological impact of fearing relapse once in remission cannot be underestimated.
- Arsenic trioxide has minimal side effects compared with the harsh side effect profile of chemotherapy.
- The treatment has less impact on day-to-day life than chemotherapy as although initial hospital attendance for treatment is higher with arsenic trioxide, patients are usually able to continue living their lives relatively normally, including working throughout treatment whereas patients cannot usually work while having chemotherapy due to their immune systems being so compromised.
- APL has a relatively high percentage of early deaths following diagnosis and more routine use of arsenic trioxide at an early stage could reduce this percentage.

NICE National Institute for Health and Care Excellence

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

Patient organisation submission

Arsenic trioxide for treating acute promyelocytic leukaemia [ID446]

Thank you for agreeing to give us your organisation's views on this technology and its possible use in the NHS.

You can provide a unique perspective on conditions and their treatment that is not typically available from other sources.

To help you give your views, please use this questionnaire with our guide for patient submissions.

You do not have to answer every question – they are prompts to guide you. The text boxes will expand as you type.

Information on completing this submission

- Please do not embed documents (such as a PDF) in a submission because this may lead to the information being mislaid or make the submission unreadable
- We are committed to meeting the requirements of copyright legislation. If you intend to include **journal articles** in your submission you must have copyright clearance for these articles. We can accept journal articles in NICE Docs.
- Your response should not be longer than 10 pages.

About you	
1.Your name	[redacted]

2. Name of organisation	Leukaemia Care
3. Job title or position	[redacted]
4a. Brief description of the organisation (including who funds it). How many members does it have?	Leukaemia Care is a national blood cancer charity, first registered with the Charity Commission in 1969. We work to ensure that everybody has the right information, advice and support. Our key services include: Freephone helpline, Nurse Advisor, LiveChat, Nationwide Support Groups, Conferences, Campaigning and Advocacy, Buddy Support and Patient Booklets. Over 85% of our funding comes from our own fundraising activities and those of our volunteers. This includes a wide range of activities – such as legacies, community events, marathons, recycling campaigns etc. Leukaemia CARE also receives funding from a wide range of pharmaceutical companies, but in total those funds equate to approximately 15% of our total annual income. Any funds received from the pharmaceutical industry are in accordance with the ABPI Code of Practice and the Leukaemia Care Code of Practice, our voluntary commitment that governs how we work with, and accept funding from, the pharmaceutical industry: www.leukaemiacare.org.uk/resources/code-of-practice
4b. Do you have any direct or indirect links with, or funding from, the tobacco industry?	N/A
5. How did you gather information about the	Information primarily gathered through Leukaemia Care patient experience survey – 'Living with Leukaemia' (www.leukaemiacare.org.uk/living-with-leukaemia). The survey was run from September to

experiences of patients and carers to include in your submission?	December 2016, as a follow up to NHS England's annual Cancer Patient Experience Survey (NCPES). The Leukaemia Care survey involved 85 questions and had responses from 2519 blood cancer patients, including 63 acute promyelocytic leukaemia (APL) patients. The results of this survey have been used to inform our submission. Additionally, we have gathered information through our helpline, support groups, communication with our membership and one to one discussion with patients. We also work closely with other patient groups and share expertise.
Living with the condition	
6. What is it like to live with the condition? What do carers experience when caring for someone with the condition?	Acute promyelocytic leukaemia (APL) is a rare subtype of acute myeloid leukaemia, with around 170 new cases of APL in the UK each year. APL is a rapidly progressing condition, with 62% of patients experiencing symptoms for less than a month before visiting a GP. Common symptoms prior to diagnosis include bruising or bleeding (75%), fatigue (57%), feeling weak or breathless (46%), headaches (22%), fever/night sweats (21%). Being diagnosed with an aggressive blood cancer like APL can be difficult, both practically and emotionally. Being told you have cancer can be very upsetting. It can be especially difficult with acute leukaemia as you often get ill suddenly, and have to start treatment quickly (63% on the same day as diagnosis and an additional 33% within a week). There is usually very little time to take in information and start to cope with it. As a result, 43% of APL patients report being depressed or anxious more often since diagnosis. The emotional impact does not affect the patient in isolation. A diagnosis can place huge emotional strain on families, many of whom may also be affected.

	The most common symptoms encountered by patients since their diagnosis include fatigue (62%), feeling weak or breathless (38%), bone/joint pain (38%), sleeping problems (32%), bruising and bleeding (29%). APL also has a practical impact, with 45% of APL patients experiencing pain as a direct result of their condition. Additionally, 35% of APL patients have difficulty moving around and 35% of ALL patients have difficulty performing some of their daily routines, such as cooking or cleaning. Of those in work or education before their diagnosis, 51% have been impacted (27% reduced hours, 24% no longer able to work or continue education). Consequently, 38% of patients reported a negative financial impact as a result of having APL (increased costs or reduced income).
7. What do patients or carers think of current treatments and care available on the NHS?	APL is highly curable with ATRA and anthracycline-based chemotherapy. However, these can be associated with high levels of toxicity, which may be difficult to tolerate for less fit patients. In our survey, 56% of APL patients reported being hospitalised as a result of side effects.
8. Is there an unmet need for patients with this condition?	

9. What do patients or carers think are the advantages of the technology?	In the newly diagnosed setting, arsenic trioxide (in combination with ATRA) offers an alternative to anthracyclines, with improved event free survival (97% after 2 years v 86%). In the second line setting ATO (in combination with ATRA) offers high CR rates (87%) in previously treated patients. Arsenic trioxide also offers an alternative for patients who cannot tolerate anthracycline-based chemotherapy. In our survey, intravenous infusion was rated as a preferred method of treatment administration by APL patients. It was selected by 54% of patients (compared with Oral Tablet 43%; Oral Suspension 16%; Subcutaneous Injection 13% and Intramuscular Injection 10%).
10. What do patients or carers think are the disadvantages of the technology?	Common side effects include: blood levels (high glucose, low magnesium & potassium); cardiovascular events (e.g. tachycardia); diarrhoea, vomiting, nausea; differentiation syndrome; breathlessness, dizziness, headaches; fatigue; fever; increased levels of liver enzymes; itching/rash; muscle pain; pins and needles; swelling.

11. Are there any groups of patients who might benefit more or less from the technology than others? If so, please describe them and explain why.	
12. Are there any potential equality issues that should be taken into account when considering this condition and the technology?	

13. Are there any other issues that you would like the committee to consider?	

Key messages

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

- Acute promyelocytic leukaemia (APL) is a rare subtype of acute myeloid leukaemia, with around 170 new cases of APL in the UK each year.
- APL patients experience a range of symptoms, as well as both a practical (pain, mobility) and financial impact. The most common symptoms reported by patients since their diagnosis include fatigue (62%), feeling weak or breathless (38%), bone/joint pain (38%), sleeping problems (32%), bruising and bleeding (29%).
- Being diagnosed with APL also has an emotional impact, with 43% of APL patients report being depressed or anxious more often since diagnosis.
- Arsenic trioxide offers a curative, tolerable alternative to anthracycline-based chemotherapy with improved event free survival (first-line) and high CR rates (in previously treated patients).
- Infusions – viewed as an acceptable method of treatment administration by APL 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

Arsenic trioxide for treating acute promyelocytic leukaemia [ID446]

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/British Society for Haematology**

3. Job title or position	[redacted]
4. Are you (please tick all that apply):	x☐ an employee or representative of a healthcare professional organisation that represents clinicians? x☐ a specialist in the treatment of people with this condition? x☐ a specialist in the clinical evidence base for this condition or technology? ☐ other (please specify):
5a. Brief description of the organisation (including who funds it).	
5b. Do you have any direct or indirect links with, or funding from, the tobacco industry?	**No**
The aim of treatment for this condition	
6. What is the main aim of treatment? (For example, to stop progression, to improve mobility, to cure the condition,	Arsenic trioxide in combination with all-trans-retinoic acid (ATRA) for induction of remission and consolidation in adult patients with newly diagnosed standard risk acute promyelocytic leukaemia (APL). This is curative therapy with less toxicity than current chemotherapy approaches The use of ATO+ATRA has transformed the treatment of relapsed disease with over 90% of patients achieving a complete remission and the survival of patients who have had a relapse after chemotherapy now approaches 90%. ATO has been used at relapse for 10 years and is regarded as a standard of care in the UK and is routinely commissioned.

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.)	Achievement of molecular complete remission is the first clinically important response
8. In your view, is there an unmet need for patients and healthcare professionals in this condition?	Yes. ATO allows a chemotherapy free approach for the treatment of standard risk APL with a cure rate of at least 90%
What is the expected place of the technology in current practice?	
9. How is the condition currently treated in the NHS?	Acute Promyelocytic Leukaemia (APL) is an uncommon haematological malignancy, the number of new cases per year in the United Kingdom is 150-200. . Since the mid-1990's the therapy of choice for newly diagnosed patients has consisted of chemotherapy combined with ATRA (all-trans retinoic acid). One form of this treatment that is commonly used is called AIDA (ATRA plus IDArubicin) and has been standard therapy for APL for the last 10 years. With this approach the 2-year Kaplan-Meier estimates of overall survival were improved to about 80%. Patients who achieve a remission then routinely undergo regular bone marrow monitoring for 3 years. If patients relapse then the standard treatment that has been routinely

	commissioned by the NHS for 10 years is ATO. Some relapsed patients may then go on to either allogeneic or autologous BMT.
• Are any clinical guidelines used in the treatment of the condition, and if so, which?	Yes the AML19 trial gives advice on the management of APL. The European guidelines were written and published in 2009 by an international panel of experts on behalf of the European Leukaemia Net. To quote "given the high anti leukaemic efficacy of ATO in relapsed patients and its relatively favourable toxicity profile this agent is presently regarded as the best treatment option in the setting of relapse of APL". I am aware that hese guideines are being updated to recommend ATO as frontline therapy for standard risk APL
• 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 in general. APL treatment is a medical emergency
• What impact would the technology have on the current pathway of care?	It would remove the requirement to treat standard risk APL patients with chemotherapy and protracted molecular monitoring
10. Will the technology be used (or is it already used) in the same way as current care in NHS clinical practice?	ATO is licensed in Europe (since 2002) for the treatment of relapsed Acute Promyelocytic Leukaemia and in combination has become the standard treatment for relapsed patients. The use of ATO+ATRA has transformed the treatment of relapsed disease with over 90% of patients achieving a complete remission and the survival of patients who have had a relapse now approaches 90%. Over the last 10 years some newly diagnosed APL patients in the UK have also been treated with ATO+ATRA if there were reasons where they could not withstand the toxicity associated with chemotherapy. This includes older patients with

	APL or those whose leukaemia has developed as a consequence of previous exposure to chemotherapy (so called secondary APL) and in whom it was not possible to give AIDA,
• How does healthcare resource use differ between the technology and current care?	Intensive anthracycline-based chemotherapy is the current standard treatment for low-intermediate risk acute promyelocytic leukaemia (APL) patients. Patients receive an AIDA (idarubicin + all-trans-retinoic acid [ATRA]) induction course followed by consolidation with three alternating cycles of anthracycline-based chemotherapy plus ATRA. This is in accordance with the NCRI AML17 and 19 protocols. Induction: ATRA 45 mg/m²/day administered starting on day 1. ATRA treatment continued until haematologic complete remission (CR) and for a maximum of 60 days. Idarubicin, 12 mg/m² on days 2, 4, 6 and 8 by short (20 minute) intravenous infusion. First consolidation cycle: Idarubicin, 5 mg/m2/day by short (20 minute) intravenous infusion on days 1, 2, 3, 4. ATRA 45 mg/m2/day, given from day 1 to day 15. Second consolidation cycle: Mitoxantrone (MTZ), 10 mg/m2/day as 30 minute intravenous infusion on days 1, 2, 3, 4, and 5. ATRA 45 mg/m2/d, given from day 1 to day 15. Third consolidation cycle: Idarubicin, 12 mg/m2/day as short (20 minute) intravenous infusion only on day 1. ATRA 45 mg/m2/d, given from day 1 to day 15. Minimal residual disease assessment is essential to guide the need for further therapy. Maintenance may be used up to 2 years (not commonly used in the UK): 6-mercaptopurine (6-MP) 50 mg/m2/day, methotrexate (MTX) 15 mg/m2/week plus ATRA 45 mg/m2 for 15 days, every 3 months (during these 15 days 6-MP and MTX are discontinued).2
• In what clinical setting should the technology be used? (For example, primary or secondary care, specialist clinics.)	Secondary/tertiary care

• What investment is needed to introduce the technology? (For example, for facilities, equipment, or training.)	I think very little is required. Haematologists already have significant experience of using ATO from clinical trials and in the relapsed setting.
11. Do you expect the technology to provide clinically meaningful benefits compared with current care?	There is a current unmet need for an alternative to chemotherapy-based treatment that can cure APL with less toxicity. Early mortality in APL due to haemorrhagic complications is a substantial problem, affecting up to 30% of patients. Although all-trans-retinoic acid (ATRA) + anthracycline-based chemotherapy can achieve complete remission in 94–99% of standard risk APL patients up to 20% of patients subsequently relapse, necessitating second-line treatment. Deaths during remission can also result from chemotherapy-related toxicity, including excessive myelosuppression. Also late cardiotoxicity is a cause of morbidity and mortality as is treatment-related myelodysplastic syndrome (2-5% of patients).
• Do you expect the technology to increase length of life more than current care?	Yes. Overall survival has been demonstrated to be improved. The final analysis of the APL0406 study was published recently reporting the updated and extended follow up on a series of 276 patients. Of 263 patients evaluable for response to induction, 127 (100%) of 127 patients and 132 (97%) of 136 patients achieved complete response (CR) in the ATRA-ATO and ATRA-CHT arms, respectively (p=0.12). After a median follow-up of 40.6 months, the event-free survival, cumulative incidence of relapse, and overall survival at 50 months for patients in the ATRA-ATO vs ATRA-CHT arms were 97.3% vs 80%, 1.9% vs 13.9%, and 99.2% vs 92.6%, respectively (p<0.001, p=.0013, and p=0.0073, respectively). A recent Phase III, multicentre trial (AML17) has also recently been conducted by members of the United Kingdom National Cancer Research Institute (NCRI) Acute Myeloid Leukaemia Working Group comparing the ATRA-ATO treatment regimen with the chemotherapy-based regimen (ATRA and idarubicin) in both low-to-intermediate risk (white-cell count <10x109/L) and high-risk patients (white-cell count ≥10x109/L) with APL. 235 patients were enrolled and randomly assigned to ATRA and idarubicin (n=119) or ATRA and ATO (n=116), including 57 high-risk patients.The combination of ATRA and ATO achieved a high cure rate (CR: 94% vs 89%. p=0.18) and less relapse (4-year cumulative molecular relapse 0% vs 27%; p<0.0001) when compared with ATRA and idarubicin in

	both low- and high- risk patients with APL. Confirmed molecular negativity was 91% in the ATRA and ATO group, compared with 88% in the ATRA-idarubicin arm. 19 (95%) of the 20 patients who relapsed on the ATRA and idarubicin arm received ATO salvage therapy.
• Do you expect the technology to increase health-related quality of life more than current care?	Yes the avoidance of chemotherapy will reduce early toxicity, requirement for hospitalisation etc. Deaths during remission can also result from chemotherapy-related toxicity, including excessive myelosuppression, late cardiotoxicity and treatment-related myelodysplastic syndrome (2-5% of patients). These can be avoided.
12. Are there any groups of people for whom the technology would be more or less effective (or appropriate) than the general population?	We think it should be used within the licenced indication for standard risk patients as first line therapy. Also for high risk patients relapsing after initial therapy with AIDA chemotherapy (either molecular or haematological relapse). However some newly diagnosed high risk APL patients (WCC>10 x 109/L) are candidates for ATO upfront if there were reasons why they could not withstand the toxicity associated with chemotherapy. This includes patients with secondary APL occurring as the result of previous chemotherapy exposure and elderly frail patients.
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	The use of ATO plus ATRA would allow for omitting the use of any chemotherapy in the frontline management of APL in patients with low-intermediate risk APL (white-cell count ≤10x109/L). The ability to cure patients with this aggressive form of cancer and reduce the early death rate without chemotherapy could help to avoid unnecessary complications associated with chemotherapeutic agents such as severe haematologic toxicity, a sizeable risk of toxic death and development of therapy-related malignancies.

treatments needed, additional clinical requirements, factors affecting patient acceptability or ease of use or additional tests or monitoring needed.)	The therapy once remission has been achieved is given as a Daycase with a significant reduction in antibiotic use, blood product use and hospital in patient stay Furthermore molecular monitoring may not be required for patients treated with ATO once molecular remission has been achieved as the relapse rate is extremely low. This is in contrast with the AIDA schedule, where intensive minimal residual disease monitoring is required for 3 years, incurring an additional cost.
14. Will any rules (informal or formal) be used to start or stop treatment with the technology? Do these include any additional testing?	PCR monitoring should be continued until the patient has achieved a molecular CR (usually after the second block of therapy). This methodology is available to treating centres
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?	Yes. Reduction in the risk of relapse and hence of BMT in second remission

16. Do you consider the technology to be innovative in its potential to make a significant and substantial impact on health-related benefits and how might it improve the way that current need is met?	Yes I do. For APL patients this is completely transforming technology
• Is the technology a 'step-change' in the management of the condition?	The use of upfront ATO (with ATRA) has transformed this hyper-acute, once rapidly fatal disease into the most highly curable acute leukaemia in adults without the requirement for administering chemotherapy. It is the opinion of the NCRI AML Working Group that ATO is the treatment of choice for standard risk APL patients at diagnosis and for high risk patients who have relapsed after initial treatment with AIDA chemotherapy
• Does the use of the technology address any particular unmet need of the patient population?	Yes. First line therapy with ATO is associated with a very low risk of relapse in APL. This compares with current chemotherapy where the relapse risk is at least 20% ATO is also suitable for patients who have secondary APL due to previous chemotherapy exposure or older frail patients not suitable for intensive chemotherapy. At present, in order to gain access to ATO, an independent funding request demonstrating exceptionality needs to be completed for each individual

	patient placing a considerable administrative burden on physicians and delaying access to treatment in what is essentially an emergency situation.
17. How do any side effects or adverse effects of the technology affect the management of the condition and the patient's quality of life?	The major SE is hepatotoxicity but this can be managed by temporary dose reduction
Sources of evidence	
18. Do the clinical trials on the technology reflect current UK clinical practice?	The NCRI AML17 trial was carried out at over 100 UK centres. The trial used an alternative attenuated dosing schedule (less frequent administration) and this schedule is also widely used in the relapsed setting in the UK and is being widely used in Europe. For example it has been adopted as the preferred regimen by the Nordic AML group. The advantage is that identical results can be used as with conventional dosing but with reduced drug acquisition costs and reduced day case attendances
• If not, how could the results be extrapolated to the UK setting?	Not applicable
• What, in your view, are the most important	

outcomes, and were they measured in the trials?	The rate of molecular remission, haematological complete remission; overall survival rate; rate of cumulative incidence of relapse; event-free survival Yes they were measured in both trials: Lo-Coco F et al. Retinoic acid and arsenic trioxide for acute promyelocytic leukaemia. NEJM 2013; 369: 111-121. Burnett A et al. Arsenic trioxide and all-trans retinoic acid treatment for acute promyelocytic leukaemia in all risk groups (AML17): results of a randomised, controlled, phase 3 trial. Lancet Oncol 2015; 16: 1295-1305.
• If surrogate outcome measures were used, do they adequately predict long-term clinical outcomes?	Surrogate measures were not used
• Are there any adverse effects that were not apparent in clinical trials but have come to light subsequently?	No. There are no long term adverse effects reported
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?	I am aware that the German Napoleon registry is collecting real world data on first line ATO for standard risk APL although not published data was presented at the recent 7th International APL Symposium which concluded that the early experience of first line therapy was in keeping with the results of the RCTs
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.

- ATO is an alternative to chemotherapy-based treatment that can cure acute promyelocytic leukaemia (APL) with less toxicity
- ATO is associated with a reduced risk of relapse compared to standard chemotherapy
- Survival is at least equivalent, if not superior to the standard chemotherapy-based regimen (AIDA)
- Once remission is achieved there is no requirement for molecular monitoring as is needed with chemotherapy
- Avoidance of late cardiotoxicity and treatment-related myelodysplastic syndrome

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

Arsenic trioxide for treating acute promyelocytic leukaemia [ID446]

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	**NCRI-ACP-RCP**

3. Job title or position	[redacted]
4. Are you (please tick all that apply):	
5a. Brief description of the organisation (including who funds it).	
5b. Do you have any direct or indirect links with, or funding from, the tobacco industry?	**No**
The aim of treatment for this condition	
6. What is the main aim of treatment? (For example, to stop progression, to improve mobility, to cure the condition, or prevent progression or disability.)	Arsenic trioxide in combination with all-trans-retinoic acid (ATRA) for induction of remission and consolidation in adult patients with newly diagnosed standard risk acute promyelocytic leukaemia (APL). This is curative therapy with less toxicity than current chemotherapy approaches The use of ATO+ATRA has transformed the treatment of relapsed disease with over 90% of patients achieving a complete remission and the survival of patients who have had a relapse after chemotherapy now approaches 90%. ATO has been used at relapse for 10 years and is regarded as a standard of care in the UK and is routinely commissioned.
7. What do you consider a	Achievement of molecular complete remission is the first clinically important response

clinically significant treatment response? (For example, a reduction in tumour size by x cm, or a reduction in disease activity by a certain amount.)	
8. In your view, is there an unmet need for patients and healthcare professionals in this condition?	Yes. ATO allows a chemotherapy free approach for the treatment of standard risk APL with a cure rate of at least 90%
What is the expected place of the technology in current practice?	
9. How is the condition currently treated in the NHS?	Acute Promyelocytic Leukaemia (APL) is an uncommon haematological malignancy, the number of new cases per year in the United Kingdom is 150-200. Since the mid-1990's the therapy of choice for newly diagnosed patients has consisted of chemotherapy combined with ATRA (all-trans retinoic acid). One form of this treatment that is commonly used is called AIDA (ATRA plus IDArubicin) and has been standard therapy for APL for the last 10 years. With this approach the 2-year Kaplan-Meier estimates of overall survival were improved to about 80%. Patients who achieve a remission then routinely undergo regular bone marrow monitoring for 3 years. If patients relapse then the standard treatment that has been routinely commissioned by the NHS for 10 years is ATO. Some relapsed patients may then go on to either allogeneic or autologous BMT.
• Are any clinical guidelines used in the treatment of the	Yes the AML19 trial gives advice on the management of APL. The European guidelines were written and published in 2009 by an international panel of experts on behalf of the European Leukaemia Net. To quote 'given the high anti leukaemic efficacy of ATO in relapsed patients and its relatively favourable toxicity

condition, and if so, which?	profile this agent is presently regarded as the best treatment option in the setting of relapse of APL'. I am aware that these guidelines are being updated to recommend ATO as frontline therapy for standard risk APL
• 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 in general. APL treatment is a medical emergency
• What impact would the technology have on the current pathway of care?	It would remove the requirement to treat standard risk APL patients with chemotherapy and protracted molecular monitoring
10. Will the technology be used (or is it already used) in the same way as current care in NHS clinical practice?	ATO is licensed in Europe (since 2002) for the treatment of relapsed Acute Promyelocytic Leukaemia and in combination has become the standard treatment for relapsed patients. The use of ATO+ATRA has transformed the treatment of relapsed disease with over 90% of patients achieving a complete remission and the survival of patients who have had a relapse now approaches 90%. Over the last 10 years some newly diagnosed APL patients in the UK have also been treated with ATO+ATRA if there were reasons where they could not withstand the toxicity associated with chemotherapy. This includes older patients with APL or those whose leukaemia has developed as a consequence of previous exposure to chemotherapy (so called secondary APL) and in whom it was not possible to give AIDA,
• How does healthcare resource use differ between the technology	Intensive anthracycline-based chemotherapy is the current standard treatment for low-intermediate risk acute promyelocytic leukaemia (APL) patients. Patients receive an AIDA (idarubicin + all-trans-retinoic acid [ATRA]) induction course followed by consolidation with three alternating cycles of anthracycline-based chemotherapy plus ATRA. This is in accordance with the NCRI AML17 and 19

and current care?	protocols. Induction: ATRA 45 mg/m²/day administered starting on day 1. ATRA treatment continued until haematologic complete remission (CR) and for a maximum of 60 days. Idarubicin, 12 mg/m² on days 2, 4, 6 and 8 by short (20 minute) intravenous infusion. First consolidation cycle: Idarubicin, 5 mg/m2/day by short (20 minute) intravenous infusion on days 1, 2, 3, 4. ATRA 45 mg/m2/day, given from day 1 to day 15. Second consolidation cycle: Mitoxantrone (MTZ), 10 mg/m2/day as 30 minute intravenous infusion on days 1, 2, 3, 4, and 5. ATRA 45 mg/m2/d, given from day 1 to day 15. Third consolidation cycle: Idarubicin, 12 mg/m2/day as short (20 minute) intravenous infusion only on day 1. ATRA 45 mg/m2/d, given from day 1 to day 15. Minimal residual disease assessment is essential to guide the need for further therapy. Maintenance may be used up to 2 years (not commonly used in the UK): 6-mercaptopurine (6-MP) 50 mg/m2/day, methotrexate (MTX) 15 mg/m2/week plus ATRA 45 mg/m2 for 15 days, every 3 months (during these 15 days 6-MP and MTX are discontinued).2
• 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 that very little is required. Haematologists already have significant experience of using ATO from clinical trials and in the relapsed setting.
11. Do you expect the	There is a current unmet need for an alternative to chemotherapy-based treatment that can cure APL with less toxicity. Early mortality in APL due to haemorrhagic complications is a substantial problem, affecting up

technology to provide clinically meaningful benefits compared with current care?	to 30% of patients. Although all-trans-retinoic acid (ATRA) + anthracycline-based chemotherapy can achieve complete remission in 94–99% of standard risk APL patients up to 20% of patients subsequently relapse, necessitating second-line treatment. Deaths during remission can also result from chemotherapy-related toxicity, including excessive myelosuppression. Also late cardiotoxicity is a cause of morbidity and mortality as is treatment-related myelodysplastic syndrome (2-5% of patients).
• Do you expect the technology to increase length of life more than current care?	Yes. Overall survival has been demonstrated to be improved. The final analysis of the APL0406 study was published recently reporting the updated and extended follow up on a series of 276 patients. Of 263 patients evaluable for response to induction, 127 (100%) of 127 patients and 132 (97%) of 136 patients achieved complete response (CR) in the ATRA-ATO and ATRA-CHT arms, respectively (p=0.12). After a median follow-up of 40.6 months, the event-free survival, cumulative incidence of relapse, and overall survival at 50 months for patients in the ATRA-ATO vs ATRA-CHT arms were 97.3% vs 80%, 1.9% vs 13.9%, and 99.2% vs 92.6%, respectively (p<0.001, p=.0013, and p=0.0073, respectively). A recent Phase III, multicentre trial (AML17) has also recently been conducted by members of the United Kingdom National Cancer Research Institute (NCRI) Acute Myeloid Leukaemia Working Group comparing the ATRA-ATO treatment regimen with the chemotherapy-based regimen (ATRA and idarubicin) in both low-to-intermediate risk (white-cell count <10x109/L) and high-risk patients (white-cell count ≥10x109/L) with APL. 235 patients were enrolled and randomly assigned to ATRA and idarubicin (n=119) or ATRA and ATO (n=116), including 57 high-risk patients.The combination of ATRA and ATO achieved a high cure rate (CR: 94% vs 89%. p=0.18) and less relapse (4-year cumulative molecular relapse 0% vs 27%; p<0.0001) when compared with ATRA and idarubicin in both low- and high- risk patients with APL. Confirmed molecular negativity was 91% in the ATRA and ATO group, compared with 88% in the ATRA-idarubicin arm. 19 (95%) of the 20 patients who relapsed on the ATRA and idarubicin arm received ATO salvage therapy.
• Do you expect the technology to increase health-related quality of life more than current	Yes the avoidance of chemotherapy will reduce early toxicity, requirement for hospitalisation etc. Deaths during remission can also result from chemotherapy-related toxicity, including excessive myelosuppression, late cardiotoxicity and treatment-related myelodysplastic syndrome (2-5% of patients). These can be avoided.

care?	
12. Are there any groups of people for whom the technology would be more or less effective (or appropriate) than the general population?	We think it should be used within the licenced indication for standard risk patients as first line therapy. Also for high risk patients relapsing after initial therapy with AIDA chemotherapy (either molecular or haematological relapse). However some newly diagnosed high risk APL patients (WCC>10 x 109/L) are candidates for ATO upfront if there were reasons why they could not withstand the toxicity associated with chemotherapy. This includes patients with secondary APL occurring as the result of previous chemotherapy exposure and elderly frail patients.
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.)	The use of ATO plus ATRA would allow for omitting the use of any chemotherapy in the frontline management of APL in patients with low-intermediate risk APL (white-cell count ≤10x109/L). The ability to cure patients with this aggressive form of cancer and reduce the early death rate without chemotherapy could help to avoid unnecessary complications associated with chemotherapeutic agents such as severe haematologic toxicity, a sizeable risk of toxic death and development of therapy-related malignancies. The therapy once remission has been achieved is given as a Daycase with a significant reduction in antibiotic use, blood product use and hospital in patient stay Furthermore molecular monitoring may not be required for patients treated with ATO once molecular remission has been achieved as the relapse rate is extremely low. This is in contrast with the AIDA schedule, where intensive minimal residual disease monitoring is required for 3 years, incurring an additional cost.

14. Will any rules (informal or formal) be used to start or stop treatment with the technology? Do these include any additional testing?	PCR monitoring should be continued until the patient has achieved a molecular CR (usually after the second block of therapy). This methodology is available to treating centres
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?	Yes. Reduction in the risk of relapse and hence of BMT in second remission
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	Yes. For APL patients this is completely transforming technology

need is met?	
• Is the technology a 'step-change' in the management of the condition?	The use of upfront ATO (with ATRA) has transformed this hyper-acute, once rapidly fatal disease into the most highly curable acute leukaemia in adults without the requirement for administering chemotherapy. It is the opinion of the NCRI AML Working Group that ATO is the treatment of choice for standard risk APL patients at diagnosis and for high risk patients who have relapsed after initial treatment with AIDA chemotherapy
• Does the use of the technology address any particular unmet need of the patient population?	Yes. First line therapy with ATO is associated with a very low risk of relapse in APL. This compares with current chemotherapy where the relapse risk is at least 20% ATO is also suitable for patients who have secondary APL due to previous chemotherapy exposure or older frail patients not suitable for intensive chemotherapy. At present, in order to gain access to ATO, an independent funding request demonstrating exceptionality needs to be completed for each individual patient placing a considerable administrative burden on physicians and delaying access to treatment in what is essentially an emergency situation.
17. How do any side effects or adverse effects of the technology affect the management of the condition and the patient's quality of life?	The major SE is hepatotoxicity but this can be managed by temporary dose reduction

Sources of evidence	
18. Do the clinical trials on the technology reflect current UK clinical practice?	The NCRI AML17 trial was carried out at over 100 UK centres. The trial used an alternative attenuated dosing schedule (less frequent administration) and this schedule is also widely used in the relapsed setting in the UK and is being widely used in Europe. For example it has been adopted as the preferred regimen by the Nordic AML group. The advantage is that identical results can be used as with conventional dosing but with reduced drug acquisition costs and reduced day case attendances
• If not, how could the results be extrapolated to the UK setting?	Not applicable
• What, in your view, are the most important outcomes, and were they measured in the trials?	The rate of molecular remission, haematological complete remission; overall survival rate; rate of cumulative incidence of relapse; event-free survival Yes they were measured in both trials: Lo-Coco F et al. Retinoic acid and arsenic trioxide for acute promyelocytic leukaemia. NEJM 2013; 369: 111-121. Burnett A et al. Arsenic trioxide and all-trans retinoic acid treatment for acute promyelocytic leukaemia in all risk groups (AML17): results of a randomised, controlled, phase 3 trial. Lancet Oncol 2015; 16: 1295-1305.
• If surrogate outcome measures were used, do they adequately predict	Surrogate measures were not used

long-term clinical outcomes?	
• Are there any adverse effects that were not apparent in clinical trials but have come to light subsequently?	No. There are no long term adverse effects reported
19. Are you aware of any relevant evidence that might not be found by a systematic review of the trial evidence?	No
20. Are you aware of any new evidence for the comparator treatment(s) since the publication of NICE technology appraisal guidance [TAXXX]?	No
21. How do data on real-world experience compare with the trial data?	Our experts note that the German Napoleon registry is collecting real world data on first line ATO for standard risk APL although not published data was presented at the recent 7th International APL Symposium which concluded that the early experience of first line therapy was in keeping with the results of the RCTs

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. • ATO is an alternative to chemotherapy-based treatment that can cure acute promyelocytic leukaemia (APL) with less toxicity • ATO is associated with a reduced risk of relapse compared to standard chemotherapy • Survival is at least equivalent, if not superior to the standard chemotherapy-based regimen (AIDA) • Once remission is achieved there is no requirement for molecular monitoring as is needed with chemotherapy • Avoidance of late cardiotoxicity and treatment-related myelodysplastic syndrome

Thank you for your time.

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

NHS England submission on the NICE appraisal of arsenic trioxide for treating acute promyelocytic leukaemia

1. NHS England notes that the Summary of Product Characteristics (SPC) for arsenic trioxide (ATO) states that it is indicated for induction of remission, and consolidation in adult patients with disease characterised by the presence of the t(15;17) translocation and/or the presence of the Pro-Myelocytic Leukaemia/Retinoic-Acid-Receptor-alpha (PML/RAR-alpha) gene in the following two clinical settings:
   • Newly diagnosed low-to-intermediate risk acute promyelocytic leukaemia (APL) (white blood cell count, ≤ 10 x 103/µl) in combination with all-trans-retinoic acid (ATRA)
   • Relapsed/refractory acute promyelocytic leukaemia (Previous treatment should have included a retinoid and chemotherapy).
2. NHS England would agree that ATO is commissioned routinely in the relapsed/ refractory setting and that it is rarely used as a single agent in practice. However, whilst commissioned by NHS England, the use in combination with ATRA would be considered off-label in this relapsed/ refractory indication. Additionally and when considering the relapsed/ refractory ATO indication, it is assumed that, if ATO is approved in the newly diagnosed setting, previous treatment will not have included chemotherapy i.e. use would be off label. Therefore, when matching to current UK practice in the use of ATO in the relapsed/ refractory setting in combination and to potential re-treatment following a chemotherapy-free previous treatment, both of these scenarios should be considered off-label. NHS England would suggest that consideration needs to be given as to whether such off-label uses can be recommended by NICE in the relapsed/ refractory setting.
3. NHS England agrees that the evidence suggests that there is a very low risk of relapse with ATO + ATRA in the newly diagnosed setting and that this means the use of second-line treatment would markedly decrease. NHS England further understands the rationale behind the clinical expert opinion that patients who remained in remission for ≥2 years following first-line ATO + ATRA treatment would be re-treated with ATO + ATRA upon relapse. However, NHS England would note that no evidence was presented by the manufacturer to support the use of re-treatment with ATO + ATRA nor the 2-year cut-off.
4. NHS England observes that the dosing of ATO that is licensed and used in the APL0406 trial is different to that used in the UK-based AML17 trial, which many UK clinicians may be used to using in practice. If NICE recommends ATO to the NHS for the treatment of acute promyelocytic leukaemia, NHS England would wish to commission its use at the licensed dose. This may become an issue for the design of future UK-based clinical trials.
5. NHS England also notes that the AML17 trial included newly-diagnosed patients with high-risk APL. If ATO is recommended by NICE within its marketing authorisation,

NHS England will wish to (at least initially) commission its use in the licensed risk groups only (low-to-intermediate). Any commissioning in the high risk group would require a NHS England commissioning policy.

6. NHS England understands that maintenance treatment is usually omitted in UK practice and that maintenance treatment with ATO is not in line with the SPC. If ATO is recommended by NICE within its marketing authorisation, NHS England will not commission its use as maintenance treatment.
7. The license for ATO is limited to adults. Acute promyelocytic leukaemia is seen in patients aged less than 18 years and there is no biological reason why any NICE recommendation as to the clinical and cost effectiveness of ATO for either the newly diagnosed or the relapsed/refractory populations it has considered would not be valid in paediatric and teenager populations. In this situation, NHS England would ensure the funding of ATO within baseline commissioning to extend to relevant patients under the age of 18 years.

[redacted]

[redacted]

March 2018

NICE National Institute for Health and Care Excellence

Clinical expert statement

Arsenic trioxide for treating acute promyelocytic leukaemia [ID446]

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 submission

- Please do not embed documents (such as a PDF) in a submission because this may lead to the information being mislaid or make the submission unreadable
- We are committed to meeting the requirements of copyright legislation. If you intend to include **journal articles** in your submission you must have copyright clearance for these articles. We can accept journal articles in NICE Docs.
- Your response should not be longer than 13 pages.

About you	
1. Your name	**Nigel Russell**
2. Name of organisation	

3. Job title or position	**Professor of Haematology, Nottingham University Hospital**
4. Are you (please tick all that apply):	**x**☐ an employee or representative of a healthcare professional organisation that represents clinicians? **x**☐ a specialist in the treatment of people with this condition? **x**☐ a specialist in the clinical evidence base for this condition or technology? ☐ other (please specify):
5a. Brief description of the organisation (including who funds it).	
5b. Do you have any direct or indirect links with, or funding from, the tobacco industry?	**No**
The aim of treatment for this condition	
6. What is the main aim of treatment? (For example, to stop progression, to improve mobility, to cure the condition,	Arsenic trioxide in combination with all-trans-retinoic acid (ATRA) for induction of remission and consolidation in adult patients with newly diagnosed standard risk acute promyelocytic leukaemia (APL). This is curative therapy with less toxicity than current chemotherapy approaches The use of ATO+ATRA has transformed the treatment of relapsed disease with over 90% of patients achieving a complete remission and the survival of patients who have had a relapse after chemotherapy now approaches 90%. ATO has been used at relapse for 10 years and is regarded as a standard of care in the UK and is routinely commissioned.

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.)	Achievement of molecular complete remission is the first clinically important response
8. In your view, is there an unmet need for patients and healthcare professionals in this condition?	Yes. ATO allows a chemotherapy free approach for the treatment of standard risk APL with a cure rate of at least 90%
What is the expected place of the technology in current practice?	
9. How is the condition currently treated in the NHS?	Acute Promyelocytic Leukaemia (APL) is an uncommon haematological malignancy, the number of new cases per year in the United Kingdom is 150-200. Since the mid-1990's the therapy of choice for newly diagnosed patients has consisted of chemotherapy combined with ATRA (all-trans retinoic acid). One form of this treatment that is commonly used is called AIDA (ATRA plus IDArubicin) and has been standard therapy for APL for the last 10 years. With this approach the 2-year Kaplan-Meier estimates of overall survival were improved to about 80%. Patients who achieve a remission then routinely undergo regular bone marrow monitoring for 3 years. If patients relapse then the standard treatment that has been routinely

	commissioned by the NHS for 10 years is ATO. Some relapsed patients may then go on to either allogeneic or autologous BMT.
• Are any clinical guidelines used in the treatment of the condition, and if so, which?	Yes the AML19 trial gives advice on the management of APL. The European guidelines were written and published in 2009 by an international panel of experts on behalf of the European Leukaemia Net. To quote "given the high anti leukaemic efficacy of ATO in relapsed patients and its relatively favourable toxicity profile this agent is presently regarded as the best treatment option in the setting of relapse of APL". I am aware that hese guideines are being updated to recommend ATO as frontline therapy for standard risk APL
• 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 in general. APL treatment is a medical emergency
• What impact would the technology have on the current pathway of care?	It would remove the requirement to treat standard risk APL patients with chemotherapy and protracted molecular monitoring
10. Will the technology be used (or is it already used) in the same way as current care in NHS clinical practice?	ATO is licensed in Europe (since 2002) for the treatment of relapsed Acute Promyelocytic Leukaemia and in combination has become the standard treatment for relapsed patients. The use of ATO+ATRA has transformed the treatment of relapsed disease with over 90% of patients achieving a complete remission and the survival of patients who have had a relapse now approaches 90%. Over the last 10 years some newly diagnosed APL patients in the UK have also been treated with ATO+ATRA if there were reasons where they could not withstand the toxicity associated with chemotherapy. This includes older patients with

	APL or those whose leukaemia has developed as a consequence of previous exposure to chemotherapy (so called secondary APL) and in whom it was not possible to give AIDA,
• How does healthcare resource use differ between the technology and current care?	Intensive anthracycline-based chemotherapy is the current standard treatment for low-intermediate risk acute promyelocytic leukaemia (APL) patients. Patients receive an AIDA (idarubicin + all-trans-retinoic acid [ATRA]) induction course followed by consolidation with three alternating cycles of anthracycline-based chemotherapy plus ATRA. This is in accordance with the NCRI AML17 and 19 protocols. Induction: ATRA 45 mg/m²/day administered starting on day 1. ATRA treatment continued until haematologic complete remission (CR) and for a maximum of 60 days. Idarubicin, 12 mg/m² on days 2, 4, 6 and 8 by short (20 minute) intravenous infusion. First consolidation cycle: Idarubicin, 5 mg/m2/day by short (20 minute) intravenous infusion on days 1, 2, 3, 4. ATRA 45 mg/m2/day, given from day 1 to day 15. Second consolidation cycle: Mitoxantrone (MTZ), 10 mg/m2/day as 30 minute intravenous infusion on days 1, 2, 3, 4, and 5. ATRA 45 mg/m2/d, given from day 1 to day 15. Third consolidation cycle: Idarubicin, 12 mg/m2/day as short (20 minute) intravenous infusion only on day 1. ATRA 45 mg/m2/d, given from day 1 to day 15. Minimal residual disease assessment is essential to guide the need for further therapy. Maintenance may be used up to 2 years (not commonly used in the UK): 6-mercaptopurine (6-MP) 50 mg/m2/day, methotrexate (MTX) 15 mg/m2/week plus ATRA 45 mg/m2 for 15 days, every 3 months (during these 15 days 6-MP and MTX are discontinued).2
• In what clinical setting should the technology be used? (For example, primary or secondary care, specialist clinics.)	Secondary/tertiary care

• What investment is needed to introduce the technology? (For example, for facilities, equipment, or training.)	I think very little is required. Haematologists already have significant experience of using ATO from clinical trials and in the relapsed setting.
11. Do you expect the technology to provide clinically meaningful benefits compared with current care?	There is a current unmet need for an alternative to chemotherapy-based treatment that can cure APL with less toxicity. Early mortality in APL due to haemorrhagic complications is a substantial problem, affecting up to 30% of patients. Although all-trans-retinoic acid (ATRA) + anthracycline-based chemotherapy can achieve complete remission in 94–99% of standard risk APL patients up to 20% of patients subsequently relapse, necessitating second-line treatment. Deaths during remission can also result from chemotherapy-related toxicity, including excessive myelosuppression. Also late cardiotoxicity is a cause of morbidity and mortality as is treatment-related myelodysplastic syndrome (2-5% of patients).
• Do you expect the technology to increase length of life more than current care?	Yes. Overall survival has been demonstrated to be improved. The final analysis of the APL0406 study was published recently reporting the updated and extended follow up on a series of 276 patients. Of 263 patients evaluable for response to induction, 127 (100%) of 127 patients and 132 (97%) of 136 patients achieved complete response (CR) in the ATRA-ATO and ATRA-CHT arms, respectively (p=0.12). After a median follow-up of 40.6 months, the event-free survival, cumulative incidence of relapse, and overall survival at 50 months for patients in the ATRA-ATO vs ATRA-CHT arms were 97.3% vs 80%, 1.9% vs 13.9%, and 99.2% vs 92.6%, respectively (p<0.001, p=.0013, and p=0.0073, respectively). A recent Phase III, multicentre trial (AML17) has also recently been conducted by members of the United Kingdom National Cancer Research Institute (NCRI) Acute Myeloid Leukaemia Working Group comparing the ATRA-ATO treatment regimen with the chemotherapy-based regimen (ATRA and idarubicin) in both low-to-intermediate risk (white-cell count <10x109/L) and high-risk patients (white-cell count ≥10x109/L) with APL. 235 patients were enrolled and randomly assigned to ATRA and idarubicin (n=119) or ATRA and ATO (n=116), including 57 high-risk patients.The combination of ATRA and ATO achieved a high cure rate (CR: 94% vs 89%. p=0.18) and less relapse (4-year cumulative molecular relapse 0% vs 27%; p<0.0001) when compared with ATRA and idarubicin in

	both low- and high- risk patients with APL. Confirmed molecular negativity was 91% in the ATRA and ATO group, compared with 88% in the ATRA-idarubicin arm. 19 (95%) of the 20 patients who relapsed on the ATRA and idarubicin arm received ATO salvage therapy.
• Do you expect the technology to increase health-related quality of life more than current care?	Yes the avoidance of chemotherapy will reduce early toxicity, requirement for hospitalisation etc. Deaths during remission can also result from chemotherapy-related toxicity, including excessive myelosuppression, late cardiotoxicity and treatment-related myelodysplastic syndrome (2-5% of patients). These can be avoided.
12. Are there any groups of people for whom the technology would be more or less effective (or appropriate) than the general population?	We think it should be used within the licenced indication for standard risk patients as first line therapy. Also for high risk patients relapsing after initial therapy with AIDA chemotherapy (either molecular or haematological relapse). However some newly diagnosed high risk APL patients (WCC>10 x 109/L) are candidates for ATO upfront if there were reasons why they could not withstand the toxicity associated with chemotherapy. This includes patients with secondary APL occurring as the result of previous chemotherapy exposure and elderly frail patients.
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	The use of ATO plus ATRA would allow for omitting the use of any chemotherapy in the frontline management of APL in patients with low-intermediate risk APL (white-cell count ≤10x109/L). The ability to cure patients with this aggressive form of cancer and reduce the early death rate without chemotherapy could help to avoid unnecessary complications associated with chemotherapeutic agents such as severe haematologic toxicity, a sizeable risk of toxic death and development of therapy-related malignancies.

treatments needed, additional clinical requirements, factors affecting patient acceptability or ease of use or additional tests or monitoring needed.)	The therapy once remission has been achieved is given as a Daycase with a significant reduction in antibiotic use, blood product use and hospital in patient stay Furthermore molecular monitoring may not be required for patients treated with ATO once molecular remission has been achieved as the relapse rate is extremely low. This is in contrast with the AIDA schedule, where intensive minimal residual disease monitoring is required for 3 years, incurring an additional cost.
14. Will any rules (informal or formal) be used to start or stop treatment with the technology? Do these include any additional testing?	PCR monitoring should be continued until the patient has achieved a molecular CR (usually after the second block of therapy). This methodology is available to treating centres
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?	Yes. Reduction in the risk of relapse and hence of BMT in second remission

16. Do you consider the technology to be innovative in its potential to make a significant and substantial impact on health-related benefits and how might it improve the way that current need is met?	Yes I do. For APL patients this is completely transforming technology
• Is the technology a 'step-change' in the management of the condition?	The use of upfront ATO (with ATRA) has transformed this hyper-acute, once rapidly fatal disease into the most highly curable acute leukaemia in adults without the requirement for administering chemotherapy. It is the opinion of the NCRI AML Working Group that ATO is the treatment of choice for standard risk APL patients at diagnosis and for high risk patients who have relapsed after initial treatment with AIDA chemotherapy
• Does the use of the technology address any particular unmet need of the patient population?	Yes. First line therapy with ATO is associated with a very low risk of relapse in APL. This compares with current chemotherapy where the relapse risk is at least 20% ATO is also suitable for patients who have secondary APL due to previous chemotherapy exposure or older frail patients not suitable for intensive chemotherapy. At present, in order to gain access to ATO, an independent funding request demonstrating exceptionality needs to be completed for each individual

	patient placing a considerable administrative burden on physicians and delaying access to treatment in what is essentially an emergency situation.
17. How do any side effects or adverse effects of the technology affect the management of the condition and the patient's quality of life?	The major SE is hepatotoxicity but this can be managed by temporary dose reduction
Sources of evidence	
18. Do the clinical trials on the technology reflect current UK clinical practice?	The NCRI AML17 trial was carried out at over 100 UK centres. The trial used an alternative attenuated dosing schedule (less frequent administration) and this schedule is also widely used in the relapsed setting in the UK and is being widely used in Europe. For example it has been adopted as the preferred regimen by the Nordic AML group. The advantage is that identical results can be used as with conventional dosing but with reduced drug acquisition costs and reduced day case attendances
• If not, how could the results be extrapolated to the UK setting?	Not applicable
• What, in your view, are the most important	

outcomes, and were they measured in the trials?	The rate of molecular remission, haematological complete remission; overall survival rate; rate of cumulative incidence of relapse; event-free survival Yes they were measured in both trials: Lo-Coco F et al. Retinoic acid and arsenic trioxide for acute promyelocytic leukaemia. NEJM 2013; 369: 111-121. Burnett A et al. Arsenic trioxide and all-trans retinoic acid treatment for acute promyelocytic leukaemia in all risk groups (AML17): results of a randomised, controlled, phase 3 trial. Lancet Oncol 2015; 16: 1295-1305.
• If surrogate outcome measures were used, do they adequately predict long-term clinical outcomes?	Surrogate measures were not used
• Are there any adverse effects that were not apparent in clinical trials but have come to light subsequently?	No. There are no long term adverse effects reported
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?	I am aware that the German Napoleon registry is collecting real world data on first line ATO for standard risk APL although not published data was presented at the recent 7th International APL Symposium which concluded that the early experience of first line therapy was in keeping with the results of the RCTs
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.	
Topic-specific questions	

23. The company's submission assumes that if arsenic trioxide is used as a first-line treatment for acute promyelocytic leukaemia, the use of second-line treatment would decrease. Is this an appropriate assumption?	Yes as relapse is so rare after first line ATO
24. The company's submission has not included stem cell transplant as a comparator for arsenic trioxide in second-line treatment. Is this appropriate? How is stem cell transplant used in the treatment pathway?	ATO is standard salvage therapy for relapsed patients with the aim of achieving molecular CR. Once that has been achieved SCT can be considered as an option to consolidate the remission but we have found that many patients do as well with completing a full course of ATO and do not need SCT

Key messages
24. In up to 5 bullet points, please summarise the key messages of your submission. • ATO is an alternative to chemotherapy-based treatment that can cure acute promyelocytic leukaemia (APL) with less toxicity • ATO is associated with a reduced risk of relapse compared to standard chemotherapy • Survival is at least equivalent, if not superior to the standard chemotherapy-based regimen (AIDA) • Once remission is achieved there is no requirement for molecular monitoring as is needed with chemotherapy • Avoidance of late cardiotoxicity and treatment-related myelodysplastic syndrome

Thank you for your time.

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

in collaboration with:

ERASMUS UNIVERSITEIT ROTTERDAM
INSTITUTE OF HEALTH POLICY & MANAGEMENT

Maastricht University

Arsenic trioxide for treating acute promyelocytic leukaemia

Produced by Kleijnen Systematic Reviews Ltd. in collaboration with Erasmus University Rotterdam (EUR) and Maastricht University

Authors
Rob Riemsma, Reviews Manager, KSR Ltd, UK
Bram Ramaekers, Health Economist, Maastricht UMC
Sabine Grimm, Health Economist, Maastricht UMC
Debra Fayter, Systematic Reviewer, KSR Ltd, UK
Sohan Deshpande, Systematic Reviewer, KSR Ltd, UK
Piet Portegijs, Systematic Reviewer, KSR Ltd, UK
Nigel Armstrong, Health Economist, KSR Ltd, UK
Willem Witlox, Health Economist, Maastricht UMC
Xavier Pouwels, Health Economist, Maastricht UMC
Steven Duffy, Information Specialist, KSR Ltd, UK
Gill Worthy, Statistician, KSR Ltd, UK
Manuela Joore, Health Economist, Maastricht UMC
Jos Kleijnen, Director, KSR Ltd, Professor of Systematic Reviews in Health Care, Maastricht University

Correspondence to
Rob Riemsma, Kleijnen Systematic Reviews
Unit 6, Escrick Business Park
Riccall Road, Escrick
York, UK
YO19 6FD

Date completed 21/02/2018

Source of funding: This report was commissioned by the NIHR HTA Programme as project number STA 17/56/11.

Declared competing interests of the authors
None.

Acknowledgements
None

Commercial in confidence (CiC) data are highlighted in blue throughout the report.
Academic in confidence (AiC) data are highlighted in yellow throughout the report.

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:
Riemsma R, Ramaekers B, Grimm S, Fayter D, Deshpande S, Portegijs, P, Armstrong N, Witlox W, Pouwels X, Duffy S, Worthy G, Joore M, Kleijnen J. Arsenic trioxide for treating acute promyelocytic leukaemia; A Single Technology Assessment. York: Kleijnen Systematic Reviews Ltd, 2018.

Contributions of authors
Rob Riemsma acted as project lead and systematic reviewer on this assessment, critiqued the clinical effectiveness methods and evidence and contributed to the writing of the report. Bram Ramaekers, acted as health economic project lead, critiqued the company's economic evaluation and contributed to the writing of the report. Sabine Grimm, Willem Witlox, Xavier Pouwels and Nigel Armstrong acted as health economists on this assessment, critiqued the company's economic evaluation and contributed to the writing of the report. Debra Fayter, Sohan Deshpande and Piet Portegijs acted as systematic reviewers, critiqued the clinical effectiveness methods and evidence and contributed to the writing of the report. Gill Worthy acted as statistician, critiqued the analyses in the company's submission and contributed to the writing of the report. Steven Duffy critiqued the search methods in the submission and contributed to the writing of the report. Manuela Joore critiqued the manufacturer's economic evaluation, contributed to the writing of the report and provided general guidance. Jos Kleijnen critiqued the company's definition of the decision problem and their description of the underlying health problem and current service provision, contributed to the writing of the report and supervised the project.

Abbreviations

6-MP	6-mercaptopurine
AATO	All-trans retinoic acid (ATRA) plus Arsenic trioxide (ATO)
AE	Adverse Events
AHSCT	Autologous haematopoietic stem cell transplantation
AML	Acute myeloid leukaemia
Amsa	Amsacrine
APL	Acute promyelocytic leukaemia
Ara-C	Cytarabine
ASH	American Society of Hematology (annual meeting)
ATO	Arsenic trioxide
ATRA	All-trans retinoic acid
BI	Budget impact
BIC	Bayesian information criterion
BNF	British National Formulary
BSC	Best supportive care
CDF	Cancer Drugs Fund
CE	Cost Effectiveness
CEA	Cost effectiveness Analysis
CEAC	Cost effectiveness Acceptability Curve
CHMP	Committee for Medicinal Products for Human Use
CHR	Complete Haematological Remission;
CI	Confidence Interval
CLL	Chronic lymphocytic leukaemia
CML	Chronic myeloid leukaemia
CMR	Complete Molecular Remission;
CNS	Central nervous system
CR	Complete response/remission
CRD	Centre for Reviews and Dissemination
CrI	Credible interval
CS	Company Submission
CSR	Clinical study report
CTCAE	Common Terminology Criteria for Adverse Events (NCI)
DFS	Disease-free survival
DNR	Daunorubicin
DT	Decision tree
DXM	Dexamethasone;
ECG	Electrocardiogram
ED	Early death
EFS	Event-free survival
EMA	European Medicines Agency
EORTC	European Organisation for Research and Treatment of Cancer
EPAR	European public assessment report
EQ-5D	European Quality of Life-5 Dimensions
ERG	Evidence Review Group
EUR	Erasmus University Rotterdam
FDA	Food and Drug Administration
FU	Follow-up
GO	Gemtuzumab ozogamicin
GvHD	Graft-versus-host disease

HADS	Hospital Anxiety and Depression Scale
HIV	Human immunodeficiency virus
HR	Hazard ratio
HRQL	Health-related Quality of Life
HSCT	Hematopoietic stem cell transplantation
HTA	Health Technology Assessment
ICER	Incremental cost-effectiveness ratio
ICUR	Incremental cost-utility ratio
IDA	Idarubicin
ITT	Intention to Treat
IV	Intravenous
KM	Kaplan–Meier
KSR	Kleijnen Systematic Reviews
L-VEF	Left-ventricular ejection fraction
LYG	Life Year Gained
MDS	Myelodysplastic syndrome
MeSH	Medical Subject Headings
MTC	Mixed Treatment Comparison
MTX	Methotrexate
MTZ	Mitoxantrone
NA	Not applicable
NCI	National Cancer Institute
NHS	National Health Services
NHSBT	National Health Service Blood and Transplant
NICE	National Institute for Health and Care Excellence
NIHR	National Institute for Health Research
NR	Not Reported
NS	Not significant
OS	Overall survival
PCR	Polymerase chain reaction
PRESS	Peer Review of Electronic Search Strategies
PSA	Probabilistic Sensitivity Analyses
PSS	Personal Social Services
QALY(s)	Quality-adjusted Life Year(s)
QLQ-C30	Quality of Life Questionnaire – Core 30
QoL	Quality of life
RFS	Relapse-free/Recurrence-free survival
RT-PCR	Reverse transcription polymerase chain reaction
RCT	Randomised Controlled Trial
SAE	Serious Adverse Events
SCT	Stem cell transplant
SD	Standard deviation
SEM	Standard error of the mean
SmPC	Summary of product characteristics
STA	Single Technology Appraisal
tMDS-AML	Therapy-related myelodysplastic syndrome or acute myeloid leukaemia
UK	United Kingdom
UMC	University Medical Centre
WBC	White blood cell
WHO	World Health Organisation

Table of Contents

Table of Tables

Table of Figures

1. SUMMARY

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

The NICE scope describes the decision problem as the clinical and cost effectiveness of arsenic trioxide (ATO) (with or without all-trans retinoic acid (ATRA)) within its marketing authorisation for adults with untreated low-to-intermediate risk acute promyelocytic leukaemia or relapsed/refractory acute promyelocytic leukaemia (APL).

The population in the submission is in line with the scope. Two main trials were included in the submission for patients with newly diagnosed APL (APL0406 and AML17). APL0406 took place in Italy and Germany whereas AML17 had trial centres in the UK, Denmark and New Zealand. The clinical expert from the company advised that "in the UK patients are treated following the AML17 protocol." However, AML17 also included patients at high risk who do not form part of the scope of this submission. In addition, in AML17 the dosing and regimens for ATO in the intervention arm (ATRA plus ATO (AATO)) were not in accordance with the licence; whilst the dosing and regimens in APL0406 were in accordance with the licence. As NICE can only issue guidance for interventions in accordance with the UK licence indication, APL0406 seems the most appropriate trial. However, AML17 might be a better reflection of UK practice.

The comparators listed in the NICE scope are: AIDA regimen (ATRA in combination with idarubicin), haematopoietic stem cell transplantation (HSCT) (for people with relapsed or refractory APL) and best supportive care (for people with relapsed or refractory APL). For first line treatment, AIDA was the comparator considered in the company submission (CS), both in the APL0406 trial and in the economic analysis. For adults with relapsed/refractory APL the company presented one randomised controlled trial (RCT) that included two arms: AATO versus ATO. Therefore, no comparative evidence for ATO in relation to any of the relevant comparators listed in the scope has been presented in the CS. Best supportive care and HSCT were not considered as comparators for people with relapsed or refractory APL in the submission.

1.2 Summary of clinical effectiveness evidence submitted by the company

The company presented evidence from three RCTs: Two of these were trials in newly diagnosed APL (APL0406 and AML17) and the third was a study in patients with relapsed APL (Raffoux, et al. 2003).

Newly diagnosed APL

Both trials in newly diagnosed APL (APL0406 and AML17) compared AATO (all-trans retinoic acid (ATRA) + ATO) with AIDA (ATRA + idarubicin). APL0406 included 266 patients with low-to-intermediate risk APL aged 18 to 71 years; while AML17 included 235 patients APL of any risk group, aged 16 or over (no upper age limit). APL0406 took place in Italy and Germany whereas AML17 had trial centres in the UK, Denmark and New Zealand. The dosing and regimens for the intervention arm (AATO) in AML17 were not in accordance with the licence; whilst the dosing and regimens for the intervention arm (AATO) in APL0406 were in accordance with the licence.

Results from APL0406 showed that AATO significantly improved overall survival (OS) at 50 months compared with AIDA (99.2% vs 92.6% respectively, p=0.007). The primary endpoint of this trial was event-free survival (EFS) at two years in an initial cohort of 156 patients (97% with AATO vs 86% with AIDA, p<0.001 for non-inferiority, p=0.02 for superiority). EFS was significantly better in the AATO group across all subsequent analyses to reach 97.3% at 50 months in the full cohort of 266 patients, compared with 80.0% in the AIDA group (p<0.001). The primary source of the observed EFS benefit was a reduction in the number of relapses with AATO – at 50 months, the cumulative incidence

of relapse was 1.9% in the AATO group compared with 13.9% in the AIDA group (p=0.0013). In terms of adverse events, corrected QT interval (QTc) prolongation was more common in the AATO group in the induction phase of treatment (8.5% vs 0.7%); as was grade 3 to 4 hepatic toxicity (40% vs 3%). However, there were no significant differences between groups in numbers of patients with moderate to severe differentiation syndrome in induction. During all treatment phases there were 19 instances of neurotoxicity with AATO and 0 with AIDA. In the AATO group patients experienced fewer haematological adverse events including fever and infection episodes and fewer grade 3 to 4 neutropenia and thrombocytopenia lasting over 15 days.

Results from AML17 showed an EFS benefit of AATO over AIDA (four-year EFS of 91% vs 70%, p=0.002), particularly in low-risk patients (four-year EFS was 92% in the AATO group [n=86] vs 71% in the AIDA group [n=92], p=0.008). The four-year cumulative incidence of haematological relapse was 18% in the AIDA arm and 1% in the AATO arm (p=0.0007). In this trial, patients were closely monitored for molecular relapse and many were treated before progression into a full haematological relapse, so that the cumulative incidence of molecular relapse at four years was 27% in the AIDA group and 0% in the AATO group (p<0.0001).

Relapsed or refractory APL

The study by Raffoux et al. (2003) compared AATO with ATO, which is not a relevant comparison according to the NICE scope. OS was similar between the AATO and ATO study arms. Across both groups, the estimated two-year OS was 59% (95% CI: 35%–77%). EFS was not reported in this study.

EMA approval of ATO in patients with relapsed or refractory APL was based on two single-arm studies conducted in the US, with no additional European studies supporting the EMA approval in this indication. However, these two studies were not included in the company submission.

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

The company conducted systematic reviews of the evidence for arsenic trioxide and its comparators in newly diagnosed and relapsed/refractory patients as per the NICE scope. The submission and response to clarification provided sufficient details for the ERG to appraise the literature searches. A good range of databases were searched, and additional searches of conference proceedings were conducted. Searches were carried out in accordance with the NICE guide to the methods of technology appraisal sections 5.2.2 and 5.2.4.

Of the two trials presented as evidence for untreated APL (APL0406 and AML17) only one (APL0406) is in accordance with the licence. We have thus prioritised an assessment of this trial in our report and presented AML17 as supporting evidence only. There are further differences between the trials which are outlined in this report. A full assessment of the quality of APL0406 by the company and by the ERG is hampered by the fact that only published information is available for assessment as the trial was not conducted by Teva. Overall the trial appears to have been well conducted. It is important to note that there are no UK patients in APL0406. The committee will need to consider the importance of this issue given that the population (low and intermediate risk), the intervention and the comparator are relevant to the UK setting. The effectiveness data show that relevant patient outcomes are improved. The safety data show that patients will need to be carefully selected and informed of the particular risks of the chosen regimen. Knowledge of long-term toxicity of AATO for newly diagnosed patients awaits a post-authorisation long term safety cohort study.

The company presented one trial in relapsed/refractory patients. The trial by Raffoux et al. (2003) compared AATO with ATO, which is not a relevant comparison according to the NICE scope. We have

not reported in detail on this small trial. In view of this lack of relevant evidence, the ERG considers that non-RCTs could have been included in the submission for the relapsed/refractory population. The committee will need to consider whether it is necessary to explore the evidence further given the company's view that "the use of ATO in the relapsed or refractory APL setting is already so well-established in routine clinical practice that it would be difficult to provide NICE with novel information based on the analysis of additional studies."

No trials of ATO alone were presented for those with relapsed/refractory disease. The committee will need to decide if they are in agreement with the company that ATO alone is rarely used in UK practice. It should also be noted that no trials in the CS compared ATO regimes with hematopoietic stem cell transplantation or with best supportive care as specified in the NICE scope.

1.4 Summary of cost effectiveness evidence submitted by the company

The company conducted systematic literature reviews (SLRs) to identify relevant cost effectiveness studies, health-related quality of life studies, resources and costs studies. Although the SLR identified cost effectiveness analyses (CEAs) in the literature, the company decided to develop a de novo model. The model structure proposed by the company however diverges from the one used in the CEAs identified in the SLR. The company justified this by stating that the existing economic evaluations did "not adequately reflect the trajectory of APL patients" and hence developed a more complex model structure to "offer more granularity with treatment phases, molecular remission and HSCT" and better reflect the clinical trajectory of APL patients. The model structure developed by the company considered different treatment phases: first line, second line, hematopoietic stem cell transplantation (HSCT) (including both alloHSCT and autoHSCT) and other phases (i.e. treatment-related myelodysplastic syndrome or acute myeloid leukaemia (tMDS/AML) and death).

The model adopts the perspective of the NHS and Personal and Social Services (PSS) in England and Wales. The model time horizon is 40 years, at the end of which a significant proportion of patients in the model are still alive (>40% of patients in the ATRA+ATO first line and AIDA second line arm). The model cycle length is four weeks to capture the treatment schedule and a half-cycle correction is applied. All costs and health gains were discounted at a rate of 3.5% per year.

The company only assessed the cost effectiveness of ATRA+ATO (AATO) in the newly diagnosed low-to-intermediate risk APL population, i.e. in first line treatment. The cost effectiveness of AATO in the relapsed/refractory APL population was not assessed.

AATO was modelled with up to two cycles (of four weeks) of induction therapy followed by eight cycles (of four weeks) of consolidation therapy. The only comparator, first line AIDA, was implemented with up to two cycles (of four weeks) of induction therapy followed by three cycles (of four weeks) of consolidation therapy. For both AATO and AIDA, maintenance treatment was not modelled and the justification provided by the company was that it is usually omitted in UK clinical practice with the aim of minimising the risk of tMDS/AML.

The transition probabilities from the first line phase of the model were informed by the APL0406 trial. The transitions from second line states and the HSCT states were only sparsely described.

Both the APL0406 and the AML17 trials used the EORTC QLQ-C30 instrument, and not the EQ-5D, to measure health-related quality of life (HRQoL) outcomes. Hence, utility values were obtained from the literature. However, no study reporting utility values based on the EQ-5D for APL patients was identified in the literature. Instead, utilities obtained in other diseases (e.g. chronic lymphocytic leukaemia and acute myeloid leukaemia) were used as a proxy for APL utilities. Additionally, the

company performed multiple adjustments to these utilities with the intention to make them more relevant for the modelled population.

The cost categories included in the model were treatment acquisition costs, medical costs (treatment administration, supportive care, monitoring and follow-up, HSCT, palliative care), and costs of managing adverse events. Drug costs were based on the British National Formulary (BNF) while NHS reference costs, BNF and PSSRU were mainly used for the medical costs. NHS reference costs were used to inform the costs of managing adverse events; alternatively, published literature was used.

In the company base-case (probabilistic) AATO was less expensive (£31,088 saved) and more effective (2.546 QALYs gained) than AIDA and thus the dominating strategy for newly diagnosed low-to-intermediate risk APL (i.e. the first line population). The probability of AATO being cost effective at a willingness-to-pay (WTP) of £30,000 per QALY was 94%. AATO remained dominant in most of the sensitivity and scenario analyses conducted by the company.

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

The cost effectiveness searches in the company submission were reported in enough detail for the ERG to appraise them. Separate searches were conducted to identify cost effectiveness studies, health-related quality of life data, and cost and healthcare resource use data.

The ERG considers that, although it is more complex than published cost effectiveness studies, the model structure is appropriate to reflect this condition and treatment pathway. The main ERG concerns regarding the model structure relate to inconsistencies between treatments regarding the modelling of patients that cannot be evaluated for molecular remission, an error in the number of tunnels used to represent the two year molecular remission health state, the absence of disease-related mortality from on treatment health states and the applicability of alloHSCT to the UK clinical setting. These issues were considered in the additional analyses performed by the ERG.

The model time horizon of 40 years results in a significant proportion of patients alive at the end of the model time horizon. Hence, the time horizon was extended to 56 years in the ERG base-case.

AATO was only assessed in the newly diagnosed APL population (first line). Although, in its clarification response, the company provided an analysis in the relapsed/refractory population (second line), the company's description of this analysis did not provide clarity over precisely how this analysis was performed. The ERG therefore implemented their own scenario by removing the first line health states and using the second line transition probabilities to reflect the relapsed/refractory population.

Inconsistent with the scope, the company did not consider ATO stand-alone nor best supportive care (BSC) as comparators in the second line setting. However, the ERG believed the justifications from the company to exclude these comparators, highlighting (based on expert opinion) that ATO alone and BSC would only rarely be used in UK clinical practice in the second line setting, to be reasonable.

The ERG had multiple concerns related to the estimation of treatment effectiveness. This included multiple reference/calculation errors, the overestimation of cardiac events and thus patients switching to second line induction for AIDA, assumptions and calculation errors related to the relapse probabilities and not considering treatment switching due to reversible arrhythmia in the model. Additionally, the evidence to inform transitions from second line health states was weak and it was frequently not transparently reported how the transition probabilities were obtained. Similarly, most of the evidence sources to inform transition probabilities from the HSCT health states are not described in the CS (neither are the transition probabilities reported). The lack of detailed description and

justification is worrying, given treatment effectiveness (including implicit assumptions made and selection of evidence sources to obtain transition probabilities) is often an influential part of the cost effectiveness model. This includes assumptions regarding the extrapolation of treatment effectiveness which is not extensively discussed in the CS. These issues were considered in the additional analyses performed by the ERG.

The ERG agrees with the company that utility values for APL patients elicited through the EQ-5D are probably not available in the literature. However, the ERG is concerned with the validity of the utility values for the following reasons: the selection process of the utility values and the assumptions underlying disutilities associated with adverse events were unclear, the non-adherence with the NICE reference case, and the lack of justification supporting the adjustments made by the company. The ERG preferred not to use the company's adjustment in its base-case analysis, and instead used the unadjusted health state utilities. Additionally, in order to prevent health state utility values exceeding the general population utility values (over time), the ERG decided to cap the health state utility values in the model using the general population utility values.

The main concerns regarding resource use and costs in the model relate to the lack of justification regarding some of the sources used. The ERG asked the company to provide more specific justification for each resource use and cost item. The company responded that they aimed to use NHS reference costs and the PSSRU wherever possible, supplementing this with data from studies identified through a targeted search where necessary. However, the company did not provide further justification and details about the included targeted sources, and the ERG was therefore unable to assess whether these sources were the best available evidence to inform resource use and costs estimates.

Considering the validity of the cost effectiveness results presented by the company, the ERG perceives the expected life expectancy outcome of the model to be relatively long. This is likely linked to the lack of disease-related mortality in the model during the first line and second line health states (only general population mortality is considered) as well as assumptions concerning (the extrapolation of) treatment benefits. The undiscounted life years (LYs) and QALYs for AATO, estimated in the model, are 33.22 and 27.91 respectively. When extending the model time horizon to 56 years, to represent a life time horizon, which is consistent with the NICE reference case, these increase to 35.83 and 30.12 respectively. The ERG is uncertain whether these outcomes have face validity. Particularly given that in the general UK population, the LY and QALYs estimated for patients aged 45 (with 48.7% being male) are 37.62 and 29.62 respectively.

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

1.6.1 Strengths

Overall, the company submission searches were well presented and reproducible. Searches were carried out on a range of databases and supplementary resources. The clinical evidence for untreated patients is based on a randomised controlled trial which is relevant to the population in this appraisal.

Strengths related to the economic evaluation include the granularity the model structure provides in comparison with other CEAs identified in the SLR. However, related to this, the (lack of) data to inform post first line transition probabilities can be regarded as a limitation. Additionally, the lack of (EQ5D) utility values for the APL population is a concern. Nevertheless, it is reassuring that AATO for the first line population remained dominant in the ERG base-case, and that the worst-case scenario produced by the ERG resulted in an (deterministic) ICER of £21,622 per QALY gained.

1.6.2 Weaknesses and areas of uncertainty

The ERG was concerned about the overall quality of the searches conducted, as there were numerous inconsistencies, inaccuracies and redundancy throughout. It is possible that relevant evidence may have been missed. However, the main weakness of the submission is that only one trial is directly relevant to the appraisal (APL0406) which provides data on an untreated population only. The trial does not have any UK patients. The company presented one trial in relapsed/refractory patients. However, the trial did not present a relevant comparison according to the NICE scope. The committee will need to consider whether it is necessary to explore further the evidence for relapsed/refractory patients or whether it is sufficiently well-established in routine clinical practice.

Although decision uncertainty in the economic evaluation is relatively low, suggested research priorities regarding the cost effectiveness might be focused on obtaining health state utility values for the APL population as well as transition probabilities from and to the HSCT health states reflective of UK clinical practice.

1.7 Summary of exploratory and sensitivity analyses undertaken by the ERG

In the company base-case (probabilistic) AATO is less expensive (£31,088 saved) and more effective (2.546 QALYs gained) than AIDA and thus the dominating strategy for newly diagnosed low-to-intermediate risk APL (i.e. the first line population). AATO remained dominant in most of the sensitivity and scenario analyses conducted by the company. The ERG has incorporated various adjustments to the company base-case. This resulted in the (deterministic) ERG base-case, wherein AATO remained dominant. Moreover, the ERG produced a worst-case scenario (combination of some of the scenario analyses explored by the ERG), to acknowledge the uncertainties discussed by the ERG in this report. This resulted in an ICER of £21,622 per QALY gained (deterministic). The ERG was unable to perform probabilistic analysis for its base-case. However, the ERG does not consider this to be a major issue as AATO is likely to remain dominant if the ERG would be able to produce probabilistic results for its base-case.

In conclusion, despite the ERG's criticism of the economic model and several highlighted uncertainties, it is reassuring that AATO for the first line population remained dominant in the ERG base-case, and that the worst-case scenario produced by the ERG resulted in an ICER of £21,622. However, as indicated by the subgroup analysis performed by the ERG, the cost effectiveness of AATO for the second line might be substantially different (estimated ICER of £31,184 per QALY gained).

2. BACKGROUND

In this report the ERG provides a review of the evidence submitted by Teva in support of arsenic trioxide, trade name TRISENOX®, for the treatment of patients with acute promyelocytic leukaemia (APL). In this section we outline and critique the company's description of the underlying health problem and the overview of current service provision. The information is taken from Chapter 3 of the company submission (CS) with sections referenced as appropriate.

2.1 Critique of company's description of underlying health problem.

The underlying health problem of this appraisal is APL which is a distinct subtype of acute myeloid leukaemia (AML). The company describes APL as a rare disease which is "caused by a translocation between chromosomes 15 and 17, abbreviated as t(15;17), fusing the PML gene with the RARA gene, which results in formation of the PML-RARα fusion protein".[1]

According to the NICE scope, "there were 2,590 diagnoses of acute myeloid leukaemia and 2,127 deaths in England in 2014. Around 10% of AML cases are APL.[2] The CS states that the exact incidence estimates vary across reports for example. Sant et al. analysed 2000 to 2002 data from 44 cancer registries across Europe and reported an overall annual crude incidence rate of 0.14 per 100,000[3]; Visser et al. analysed 1995 and 2002 data from 64 European cancer registries reporting a crude annual incidence rate of 0.11 per 100,000 people;[4] Dores et al. conducted a study based on the US Surveillance, Epidemiology and End Results (SEER) Program registry reported that age-adjusted incidence of APL was 0.27 per 100,000 person-years.[5]

The CS states that the "age distribution is a key difference between APL and most other AML types, which are diagnosed at a median age exceeding 60 years".[1] Hence, APL is likely to pose a considerable societal burden, affecting people of working age.[1]

The CS states that APL can progress rapidly with very poor survival prognosis. The company mentions a retrospective analysis that reported about 10–29% of patients die within 30 days of hospital admission or diagnosis. The majority of these deaths are due to haemorrhage (CNS or pulmonary) (31 to 55%) because of high risk of coagulopathy in APL patients.[6]

The CS refers to relapse risk stratification which is used to determine the most appropriate treatment options for APL patients.[1] The CS states that "assessment of relapse risk in APL is primarily based on white blood cell (WBC) count at presentation, with patients whose WBC count exceeds 10×10^9/L generally predicted to have a higher risk of relapse. Risk stratification was developed through a joint analysis of two multicentre trials (AIDA0493 and LPA96)".[7] The relapse risk categories are: low-risk (WBC $\leq10\times10^9$/L and platelet count $>40\times10^9$/L), intermediate-risk (WBC and platelet counts $\leq10\times10^9$/L and $\leq40\times10^9$/L, respectively) and high-risk (WBC count $>10\times10^9$/L).[7] In this submission, the population under consideration is adults with untreated low-to-intermediate risk and relapsed/refractory APL.

ERG comment:

- The company provides a good overview of the underlying health problem. The ERG checked the references provided to support the statements in the company submission. In general, these were found to be appropriate.

2.2 Critique of company's overview of current service provision

The company correctly reports that there is no relevant technology appraisal guidance on APL published in the UK to date and that ATO has never been assessed by the NICE. The CS mentions that, in the past

patients with newly-diagnosed APL were commonly treated with the standard chemotherapy-based treatment approach, AIDA which is a combination of all-trans retinoic acid (ATRA) and idarubicin.[1] In 2015, Teva conducted primary market research in seven European countries including Austria, France, Germany, Italy, Spain, Switzerland and the UK to understand the current treatment patterns in APL.[1] The CS states that "patients newly-diagnosed with APL in the UK are commonly treated according to clinical trial protocols (MRC AML trials), as they are recommended to enrol in ongoing trials upon diagnosis".[1]

Arsenic trioxide has a UK marketing authorisation for induction and consolidation in adult patients with: newly diagnosed low-to-intermediate risk APL (white blood cell count, ≤ 10 x 10^3/µl) in combination with ATRA and relapsed/refractory APL (previous treatment should have included a retinoid and chemotherapy).[2]

The CS states "first-line therapy in APL generally consists of three consecutive treatment phases: induction, consolidation and maintenance, although maintenance is usually omitted in the UK clinical practice with the aim of minimising the risk of treatment-related myelodysplastic syndrome or acute myeloid leukaemia (tMDS/AML)"[1] The ATO-based first-line treatment regimen was not explicitly recommended for wider use by the 2013 guidelines from the European Society for Medical Oncology (ESMO)[8] and the 2009 European LeukemiaNet guidelines.[9] However, the two German guidelines Deutsche Gesellschaft für Hämatologie und Medizinische Onkologie (DGHO)[10] and the German Intergroup[11] have listed the AATO (ATRA+ATO) combination as an option for treating newly-diagnosed low-to-intermediate-risk patients.

The CS states that "patients with relapsed or refractory APL may receive a HSCT to consolidate second remission" if considered at risk of additional relapses.[1]However, patients who are not transplant candidates may receive additional ATO cycles.[1]

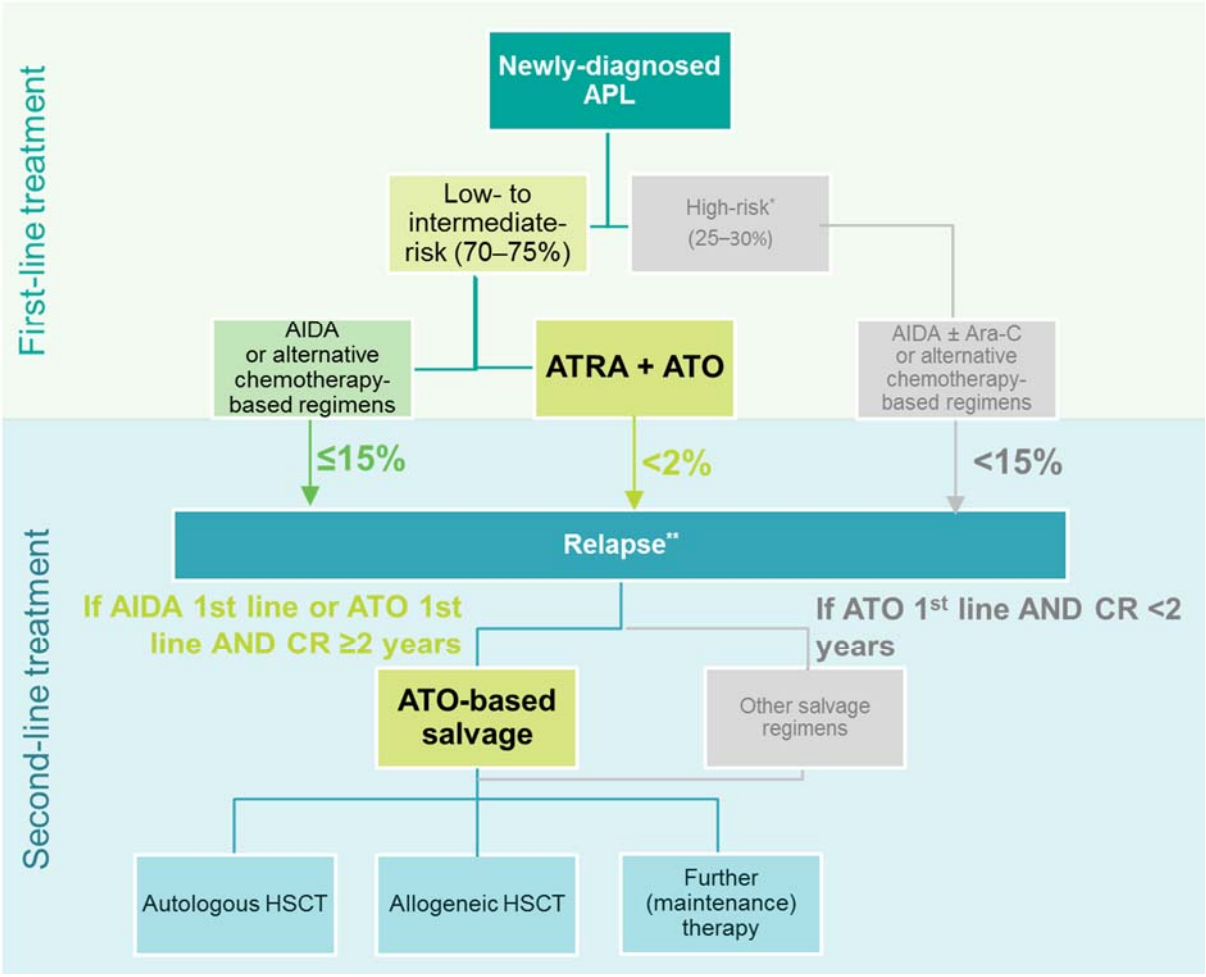
The CS states that "according to expert opinion, patients in the UK are treated as soon as molecular relapse is detected and before the patient progresses into a haematological relapse."[1] Once a relapse is confirmed, the choice of second line treatment depends on the type of first line therapy the patient has received for e.g. UK patients could switch from AIDA to AATO (ATRA+ATO) and from AATO to AIDA.[1]

The CS highlights the fact that "there is a lack of well-established paradigms or guidelines for second-line treatment following AATO administration in first line, and the field is constantly evolving with growing experience of first-line ATO use".[1] Hence, based on expert opinion the economic analysis in this submission included mixed re-treatment/switch approach which assumes that the "patients who remained in remission for 2 years or longer following first-line AATO treatment were re-treated with AATO upon relapse, while patients who achieved only a short (<2 years) remission after first-line treatment with AATO were treated with AIDA".[1]

The CS states that in "UK patients, second remissions are often consolidated with a HSCT". Also, according to the clinical expert "allogeneic HSCT is generally used in patients who enter haematological remission following second-line treatment but fail to achieve molecular remission; in patients who achieve a second molecular remission, allogeneic HSCT is rarely considered due its associated risks."[1] Further, the company's clinical expert also suggested that the "patients salvaged with ATO do not necessarily need transplantation, while those salvaged with chemotherapy generally do."[1]

Figure 2.1 shows the perceived role of ATO in the treatment of both newly diagnosed and relapsed/refractory patients with APL

Figure 2.1: Simplified treatment pathway in APL showing the licensed indications for ATO



Source: Section B1.3.2.2 of the CS
AIDA = ATRA in combination with idarubicin; APL = acute promyelocytic leukaemia; ATO = arsenic trioxide; ATRA = all-trans retinoic acid; CR = Complete remission; HSCT = haematopoietic stem cell transplantation

ERG comment:

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

3. CRITIQUE OF COMPANY'S DEFINITION OF DECISION PROBLEM

Table 3.1: Statement of the decision problem (as presented by the company)

	Final scope issued by NICE	Decision problem addressed in the company submission and rationale	ERG comments
Population	Adults with: • untreated low-to-intermediate risk acute promyelocytic leukaemia • relapsed/refractory acute promyelocytic leukaemia (APL)	Adults with: • untreated low-to-intermediate risk acute promyelocytic leukaemia • relapsed/refractory acute promyelocytic leukaemia (APL) characterised by the presence of the t(15;17) translocation and/or the presence of the promyelocytic leukaemia/ retinoic-acid-receptor-alpha (PML/RAR-alpha) gene.	In line with the scope.
Intervention	ATO (with or without ATRA)	• First line treatment: ATO combined with ATRA; both administered according to the APL0406[12] protocol. AML17[13] protocol was studied as a scenario. • Second line treatment: ATO administered according to the SmPC + ATRA administered according to the APL0406[12] protocol (as in first line). The AML17 protocol[13] was studied in a scenario analysis. Rationale: In line with both the pivotal APL0406 trial[12, 14] and the AML17 trial[13], ATO is authorised for use in newly-diagnosed patients in combination with ATRA. No treatment combinations are specified for use in relapsed/refractory patients, although in the AML17 trial treatment with ATRA+ATO (administered as in first line) was used in patients who relapsed.[15] Based on clinical expert opinion, it appears ATO alone (without ATRA) is now rarely used in the relapsed/refractory setting. Thus, for both first- and second-line treatment, only the ATRA+ATO combination was considered in the economic analysis.	The company presented evidence for AATO (ATRA+ATO) only. No evidence was presented for ATO alone.
Comparator(s)	• AIDA regimen (ATRA in combination with idarubicin) • haematopoietic stem cell transplantation (HSCT)	• Following a relapse, the choice of therapy strongly depends on prior treatments the patient has received. It is therefore difficult to separate first- and second-line indications of ATO, as they're closely linked. To optimally reflect the treatment pathway of APL patients in the UK, Teva has decided to submit a single model which evaluates the cost-	For first line treatment, one trial is presented comparing AATO versus AIDA (the APL0406 trial[12, 14]). A second trial was presented (the AML17 trial);[13]

	Final scope issued by NICE	Decision problem addressed in the company submission and rationale	ERG comments
	(people with relapsed or refractory APL) • best supportive care (people with relapsed or refractory APL)	effectiveness of ATO (+ATRA) in newly-diagnosed patients (first line indication) with second line treatments included, rather than presenting a separate cost effectiveness evaluation of ATO as a second line treatment. • For first line treatment, AIDA was the comparator considered in both the pivotal APL0406 trial[12, 14] and in the economic analysis • For the second-line part of the model, we considered a situation where ATO was available first-line and some of the patents who received ATO first line switched to AIDA in second line, so that AIDA was retained as the comparator. Rationale: • In the second line indication, HSCT was not considered as a direct comparator, since administration of ATRA+ATO usually precedes transplantation rather than replaces it. Upon relapse, ATRA+ATO can be used to induce remission, which, if possible, would be consolidated with HSCT.[11, 16] Although additional ATO (+ ATRA) cycles may be used in patients who do not undergo a transplant,[16, 17] ATO-based maintenance treatment is not included in the licensed administration schedule, and was therefore not considered in the economic analysis. Furthermore, other maintenance treatment options are also available to APL patients who do not undergo transplantation,[16] and it would be difficult to include all of them without overtly complicating the analysis. We therefore took a simplified approach of not modelling second line maintenance treatment, especially given that the number of patients concerned would be very small. • Best supportive care was not considered as a direct comparator in the second line indication. Following ATO-based treatment of first APL relapse, Lengfelder et al. reported 3-year EFS of ≥45%,[17] suggesting that attempting curative treatment may be most appropriate in patients with relapsed/refractory APL. Given the severity of APL, best supportive care can be seen as a palliative approach, and thus expected to be used where the disease is refractory to all other treatments, including ATO in second (or subsequent) treatment lines. Thus, it is unlikely that best	however, in this trial ATO was not administered according to its licensed indication). For second line treatment, a trial is presented comparing AATO vs ATO (Raffoux et al. 2003[18]). As ATO is part of both arms, this trial is not informative for the effectiveness of ATO in second line. No evidence is presented for HSCT and best supportive care. The company did not consider the relapsed/refractory APL population neither did they consider BSC nor ATO alone in the health economic sections of the CS. This is discussed in more detail in sections 5.2.4 and 5.2.3 of this report.

	Final scope issued by NICE	Decision problem addressed in the company submission and rationale	ERG comments
		supportive care will be considered an alternative to ATO or AIDA (see below) for treatment of relapsed APL. It is, however, worth noting that the economic analysis does take into account best supportive care – upon failure of second line treatment, patients in the model progressed to an end-of-life state, where they received palliative care. • The choice of second line treatment is largely determined by the first line therapy that the patient has received, and ATO (usually + ATRA) is the standard treatment for APL relapses after first line treatment containing ATRA and an anthracycline (e.g. AIDA). However, the choice of optimal salvage treatment in patients who relapse following first line ATO use is less clear. This is largely due to the absence of established guidelines, as many treatment guidelines in APL (e.g. from the European LeukemiaNet[9] and ESMO[8]) precede the approval of ATO for first-line use. In the economic analysis, treatment of relapses following first line ATO use was therefore based on clinical expert opinion. It was assumed that patients who remained in remission for ≥2 years following first line ATRA+ATO treatment were re-treated with ATRA+ATO upon relapse. However, patients who achieved only a short (<2 years) remission after first line treatment with ATRA+ATO, were assumed to be treated with AIDA upon relapse. Thus, AIDA was considered as a comparator also in the relapsed/refractory APL setting.	
Outcomes	• Overall survival (OS) • Progression-free survival (PFS) • Response rates (bone marrow remission) • Adverse effects of treatment • Health-related quality of life (HRQoL)	• OS • Event-free survival (EFS) • Complete haematological and molecular remission rates • Cumulative incidence of relapse (CIR) • Disease-free survival (DFS) or relapse-free survival (RFS) • Adverse effects of treatment • HRQoL Rationale: • PFS was not an endpoint in the pivotal APL0406 trial[12, 14] or in the AML17 trial,[13] and is thus not presented. Instead, the manufacturer	EFS was used instead of PFS. In the main trial, APL0406, EFS was assessed at 2 years after diagnosis, with treatment failure defined as any of the following: 1) no achievement of hematologic complete remission (CR) after induction; 2) no achievement of molecular CR after three consolidation courses; 3) molecular relapse;

	Final scope issued by NICE	Decision problem addressed in the company submission and rationale	ERG comments
		presented data on EFS – the primary endpoint of the APL0406 trial.[12, 14] It is, however, worth noting that in the APL0406 trial patients failing treatment were those who did not achieve remission, relapsed, or died, which is similar to what would be considered treatment failure when analysing PFS. In the AML17 trial, an additional event of treatment-related myelodysplastic syndrome or acute myeloid leukaemia (tMDS/AML) was also included in the EFS analysis; however, only a single patient in this study developed tAML,[13] so that inclusion of this event in EFS evaluation could be considered to have little effect on the overall result. In conclusion, although EFS rather than PFS is presented, the two outcomes are similar, so this does not represent a major deviation from the scope. • In addition to the outcomes listed in the Final Scope, the manufacturer will also present data on cumulative incidence of relapse and DFS (or RFS), if available. Given the curative intent of APL treatment, these endpoints are of particular importance, as they provide information on the proportion of patients who remain disease-free.	4) haematological relapse, or 5) death. EFS in this case is similar to PFS.

Source: Table 1.1, Section B.1.1 of the CS

AATO = ATRA+ATO; AIDA = ATRA in combination with idarubicin; AML = acute myeloid leukaemia; APL = acute promyelocytic leukaemia; ATO = arsenic trioxide; ATRA = all-trans retinoic acid; CIR = cumulative incidence of relapse; CR = Complete remission; CS = company submission; DFS = disease-free survival; EFS = event-free survival; ERG = Evidence Review Group; HRQoL = health-related quality of life; HSCT = haematopoietic stem cell transplantation; OS = overall survival; PFS = progression-free survival; PML = promyelocytic leukaemia; RAR-alpha = retinoic-acid-receptor-alpha; RFS = relapse-free survival; SmPC = Summary of product characteristics; tMDS = treatment-related myelodysplastic syndrome

3.1 Population

The population defined in the scope is adults with untreated low-to-intermediate risk acute promyelocytic leukaemia (APL) and adults with relapsed/refractory APL. The population in the submission is in line with the scope.

Two main trials were included in the submission for patients with newly diagnosed APL (APL0406 and AML17). APL0406 took place in Italy and Germany whereas AML17 had trial centres in the UK, Denmark and New Zealand. The clinical expert from the company advised that "in the UK patients are treated following the AML17 protocol."[19] However, AML17 also included patients at high risk who do not form part of the scope of this submission and the dosing and regimens for the intervention arm (AATO) in AML17 were not in accordance with the licence; while the dosing and regimens for the intervention arm (AATO) in APL0406 were in accordance with the licence.

APL0406 seems the most appropriate study as NICE can only issue guidance for interventions in accordance with the UK licence indication. However, AML17 might be a better reflection of UK practice.

3.2 Intervention

The intervention (ATO with or without ATRA) is in line with the scope. Regulatory approval by the EMA for the treatment of relapsed or refractory patients was granted in 2002. In November 2016 it was approved in the EU for the treatment of newly-diagnosed patients with low-to-intermediate risk APL.[20]

ATO is indicated for induction of remission, and consolidation in adult patients with:

- Newly diagnosed low-to-intermediate risk acute promyelocytic leukaemia (APL) (white blood cell count, $\leq 10 \times 10^3/\mu l$) in combination with all-trans-retinoic acid (ATRA)
- Relapsed/refractory APL (Previous treatment should have included a retinoid and chemotherapy) characterised by the presence of the t(15;17) translocation and/or the presence of the Pro-Myelocytic Leukaemia/Retinoic-Acid-Receptor-alpha (PML/RAR-alpha) gene.

ATO must be administered under the supervision of a physician who is experienced in the management of acute leukaemias and special monitoring procedures apply (see CS, Table 1.2, pages 9-11).

ATO is indicated for the treatment of APL characterised by the presence of the t(15;17) translocation and/or the presence of the Promyelocytic Leukaemia/Retinoic-Acid-Receptor-alpha (PML/RAR-alpha) gene. This translocation accounts for up to 98% of APL cases; however, other translocations involving the *RARA* gene have also been identified in APL.[21] It is widely accepted that the diagnosis of APL (as opposed to other types of AML) should be confirmed through molecular testing for PML-RARA. Although the pivotal APL0406 trial accepted a number of methods through which genetic confirmation of APL diagnosis could be established,[12] the diagnostic tests that appear most feasible for routine use are polymerase chain reaction (PCR) and fluorescent *in situ* hybridisation (FISH).[1]

APL patients also undergo repeated bone marrow biopsies and the collected material is PCR-tested for the presence of PML-RARA, which allows the treating clinician to establish how the patient responds to treatment (i.e. if molecular remission has been achieved or if minimal residual disease can be detected), and to monitor the patient for molecular relapse (i.e. the reappearance of PML-RARA in the bone marrow), which allows second line treatment to be administered early, before the patient progresses into a full haematological relapse that may be life-threatening. The frequency of monitoring depends on treatment choice.

3.3 Comparators

The description of the comparators in the NICE scope is as follows:

- AIDA regimen (ATRA in combination with idarubicin)
- haematopoietic stem cell transplantation (HSCT) (people with relapsed or refractory APL)
- best supportive care (people with relapsed or refractory APL)

For first line treatment, AIDA was the comparator considered in the CS, both in the APL0406 trial[12, 14] and in the economic analysis.

For adults with relapsed/refractory APL the company presented one randomised controlled trial (RCT) that included two arms: AATO versus ATO. Therefore, no evidence for ATO in relation to any of the relevant comparators listed in the scope has been presented in the CS. The company justifies this by stating: "This was motivated by the well-established and widespread use of ATO in relapsed/refractory APL, and the fact it has long been considered first-choice therapy for induction and consolidation in this setting." (CS, section B2.2.1, page 23).

Best supportive care and HSCT were not considered as comparators for people with relapsed or refractory APL in the CS.

3.4 Outcomes

The NICE final scope lists the following outcome measures:

- overall survival (OS)
- progression free survival (PFS)
- response rates (bone marrow remission)
- adverse effects of treatment (AE)
- health-related quality of life (HRQoL)

These outcomes are reported in the CS with one exception: PFS; instead event-free survival (EFS) was used. In the APL0406 trial, EFS was assessed at two years after diagnosis, with treatment failure defined as any of the following: 1) no achievement of haematologic CR after induction; 2) no achievement of molecular CR after three consolidation courses; 3) molecular relapse; 4) haematological relapse, or 5) death. EFS is similar to PFS in this instance.

3.5 Other relevant factors

The company states that "Given the high rates of overall survival achieved with APL treatments, ATO is unlikely to meet the end-of-life criteria." (CS, Page 91). The ERG agrees, this STA does not meet the end-of-life criteria.

There is no Patient Access Scheme (PAS) application.

The company states that: "making ATO available on the NHS is likely to allow a greater number of elderly patients to be treated, which may be an important step towards addressing the topical issue of under-treatment among elderly oncology patients" (CS, B1.4, page 20). In addition, the company mentions Jehovah's Witness patients as a potential equality concern. No further equity or equality issues were mentioned in the CS.

4. CLINICAL EFFECTIVENESS

4.1 Critique of the methods of review(s)

The company conducted two systematic reviews to identify evidence on the clinical effectiveness and safety of ATO and other treatments for adults with APL. One review focused on randomised controlled trial (RCT) evidence and the other on non-RCT evidence. The non-RCT evidence was intended to inform the use of ATO as first line treatment only. This section critiques the methods of the reviews including searching, inclusion criteria, data extraction, quality assessment and evidence synthesis.

4.1.1 Searches

The company submission stated that in order to address the decision problem two separate searches were conducted in July 2016 which were then updated in October 2017. One search was designed specifically to identify RCTs, whilst a second search was conducted to identify non-RCTs "in order to provide the widest possible range of data."[1] Search strategies were reported in detail in Appendix D of the company submission for the following databases: MEDLINE, MEDLINE in-Process, Embase, and Cochrane Central Register of Controlled Trials (CENTRAL). The host provider was reported for MEDLINE and Embase, but not for CENTRAL. The date the searches were conducted was provided, though the date span of the databases searched was not. In response to the ERG clarification letter the company provided the database date of inception, and the date the searches were conducted, but not the date span. Searches utilised study design filters based on the Cochrane Highly Sensitive Search Strategy for identifying randomised trials (although this was not explicitly reported).[22] It is not clear where the study design filters were derived from for the non-RCTs searches. Searches of the trials register ClinicalTrials.gov were also conducted.

Additional searches of the following conference proceedings were reported in the main text of the company submission (section B.2.2.1) for 2011-2017: American Society for Clinical Oncology (ASCO), American Society of Hematology (ASH) and European Hematology Association (EHA). However, no details of the conference proceedings search strategies, date of searches or results were provided in Appendix D. Details of the conference proceedings searches were provided in response to the ERG clarification letter: search terms used, dates of the conferences searched, and number of abstracts retrieved.

ERG comment:

- Relevant studies could have been missed due to sub-optimal use of proximity operators, truncation and synonyms in search strategies. The eligibility criteria provided in Table 2.1 of the company submission included systematic reviews and meta-analyses, but no attempt to search for these study designs was made.
- The search strategy provided in Appendix D of the CS reported a simultaneous search of MEDLINE and Embase using the Ovid interface without including both MeSH and EMTREE subject headings. Search filters were used for the wrong databases and safety data may have been missed because study design terms used to search for non-RCTs were possibly too restrictive to capture all safety data.
- It is possible that potentially relevant studies were excluded from the final search results because the method used to limit the MEDLINE and Embase searches to human studies was incorrect. See Appendix 1 for further details.

4.1.2 Inclusion criteria

As stated above, the company conducted two systematic reviews to identify evidence on the clinical effectiveness and safety of ATO and other treatments for adults with APL. One review focused on RCT evidence and the other on non-RCT evidence. The non-RCT evidence was intended to inform the use of ATO as first line treatment only. The eligibility criteria used in the search strategy for RCTs and non-RCTs are presented in Table 4.1. The CS stated that two independent reviewers screened the studies identified through the searches, in order to determine the eligibility of each study. The two lists of selected references were then compared and all disagreements were solved by discussion, or if persistent, by a third reviewer.

Table 4.1: Eligibility criteria used in the review search strategy

	Review of RCTs	Review of non-RCTs
Population	*Inclusion Criteria* Adult participants with APL, aged ≥16 years, of both genders *Exclusion criteria* • Paediatric-only population • High-risk newly diagnosed APL • Significant cardiac comorbidities • Significant pulmonary comorbidities • Active non-APL malignancy • Pregnant women • Women who were breastfeeding during the time of the study	*Inclusion Criteria* Adult participants with APL, aged ≥16 years, of both genders *Exclusion criteria* Paediatric-only population (aged ≤ 15 years)
Interventions and Comparators	*Inclusion Criteria:* Any intervention *Exclusion criteria:* None	
Outcomes	*Inclusion Criteria* • Adverse events • OS • EFS • DFS or RFS • Cumulative incidence of relapse • Response rates (complete haematological and molecular remission rates) *Exclusion criteria:* None	
Study design	*Inclusion Criteria* RCTs, Phase II/III studies, systematic literature reviews of RCTs, or meta-analysis *Exclusion criteria* Opinion, editorial letter	*Inclusion Criteria* • Observational study • Cohort study • Prospective study (non-RCT) • Patient registry • Cross sectional study • Case-control study • Cases series including ≥ 6 cases *Exclusion criteria* • Opinion, editorial letter • RCTs • Case reports • Case series with ≤5 cases

	Review of RCTs	Review of non-RCTs
Other	*Inclusion Criteria:* None *Exclusion criteria* • Old conference abstracts: conference abstracts published prior to 2014 were excluded. • No full text available online. • Chinese articles published in non-core journals were excluded.	*Inclusion Criteria:* None *Exclusion criteria* • Old conference abstracts: conference abstracts published prior to 2014 were excluded • Studies including a population of <50 patients • Studies that did not include ATO in first line
Source: CS, Tables 2.1 and 2.2 DFS = EFS = event-free survival; OS = overall survival; RCT = randomised controlled trials		

ERG comment:

- Two reviewers were involved in the selection of studies for the reviews which helps to minimise bias.
- The ERG queried the exclusion of non-RCT "studies that did not include ATO in first line". The company provided a list of the 70 non-RCT studies excluded on this basis and stated they were "generally supportive of its use in this indication".[19] Further details of the company's response is provided in section 4.2.1. The ERG considers that non-RCTs could have been included for the relapsed/refractory population particularly as no directly relevant RCT evidence is presented (see section 4.2.1). The committee will need to consider whether it is necessary to explore the evidence further given the company's view that "the use of ATO in the relapsed or refractory APL setting is already so well-established in routine clinical practice that it would be difficult to provide NICE with novel information based on the analysis of additional studies."[19]
- The company further stated in the CS that "Chinese articles published in non-core journals were excluded, due to their frequently poor quality. Furthermore, Trisenox® is not marketed in China, so Chinese studies may be expected to report on the use of other ATO formulations. Nonetheless, relevant Chinese articles that met the inclusion criteria are summarised in Appendix L."[1] The ERG examined the Chinese RCTs that met the inclusion criteria and believes that the company emphasised the most relevant RCTs at first line in a UK setting. The Chinese trials used different treatment regimens when compared to APL0406, the main relevant trial in the CS. Therefore, it was reasonable to exclude them from more detailed analysis.

4.1.3 Critique of data extraction

The CS stated that one reviewer extracted relevant data from included studies and the results were reviewed by a senior manager for quality control.

ERG comment: Data extraction appears to have been conducted appropriately.

4.1.4 Quality assessment

The CS did not explicitly state that two reviewers were involved in assessment of trial quality. However, given that study selection and data extraction included two reviewers it is assumed that this process also included two reviewers to minimise risk of bias. Quality was assessed using a tool adapted from the Centre for Reviews and Dissemination's (CRD's) guidance for undertaking reviews in health care.[23] Elements assessed were randomisation, allocation concealment, comparability of groups, blinding of care providers, patients and outcome assessors and drop out, selective reporting of outcomes and use of intention to treat analysis and appropriate methods for dealing with missing data. The three main trials

(APL0406, AML17 and Raffoux et al.) were quality assessed using published papers as the company was not involved in the trials.

ERG comment: Study quality appears to have been assessed appropriately. Results of the company's quality assessment and the ERG's assessment of APL0406 are presented in section 4.2.4. We have not presented an assessment of AML17 as the intervention was not delivered according to the licence and therefore not of direct relevance to the decision problem. Neither have we assessed Raffoux et al. as this trial was not considered as meeting the NICE scope.

4.1.5 Evidence synthesis

The authors stated that as the trials included in the review used different comparators a network meta-analysis (NMA) would be most appropriate. However, after evaluation of the included studies, the authors concluded that an NMA was not feasible for any of the outcomes. Studies that were comparable in terms of time point and outcome had no mutual comparator for inclusion in a network.

ERG comment:

- The two trials identified for newly diagnosed patients had different dosing and regimens for the intervention arm and as only one of these trials was in accordance with the licence and therefore of direct relevance to the decision problem (APL0406) it would not be possible to conduct a meta-analysis in this population. Additionally, AML17 included 57 high risk patients, a population which is not part of the NICE scope, although subgroup analysis was conducted by risk.
- For patients with relapsed/refractory disease one trial only was identified, which was not relevant for the decision problem as ATO was included in both treatment arms; therefore, a meta-analysis could not be performed in this population either.

4.2 Critique of trials of the technology of interest, their analysis and interpretation (and any standard meta-analyses of these)

4.2.1 Overview of the evidence in the submission

Newly diagnosed patients

Two main trials were included in the submission for patients with newly diagnosed APL (APL0406 and AML17).[13, 14] Both of these compared AATO (ATRA+ATO) to AIDA. Both trials focused on adults (age 18 in APL0406 and age 16 in AML17). Both were RCTs and both were open label. APL0406 took place in Italy and Germany whereas AML17 had trial centres in the UK, Denmark and New Zealand. AML17 also included patients at high risk who do not form part of the scope of this submission. Both trials presented a final analysis of patient outcomes at 53 months. However primary outcomes differed. APL0406 assessed event-free survival (EFS) at two years after diagnosis whilst AML17 assessed quality of life outcomes. Both considered a range of secondary outcomes including overall survival.

Relapsed/refractory APL

The only trial presented in the CS relating to relapsed/refractory patients was Raffoux et al.[18] All patients had been previously treated with ATRA and anthracycline-based chemotherapy. The trial compared AATO with ATO alone which is not a relevant comparison according to the NICE scope. The trial had just 20 patients and a median follow up of 21 months. No trials compared ATO regimes with hematopoietic stem cell transplantation or with best supportive care as specified in the NICE scope.

No relevant comparative trials of ATO alone were presented for either newly diagnosed patients or those with relapsed/refractory disease.

An overview of the three main trials in the CS is presented in Table 4.2.

Table 4.2: Overview of RCTs in the submission

Trial name	APL0406	AML17	Raffoux et al (2003)
Population	Patients with newly-diagnosed, low to intermediate risk APL aged 18 to 71 years	Patients with newly-diagnosed APL, of any risk group aged ≥ 16 years	Patients with APL in first or subsequent relapse, aged ≥ 12 years. All previously treated with ATRA and anthracycline-based chemotherapy.
Intervention	AATO	AATO	AATO
Comparator	AIDA	AIDA	ATO alone
Outcomes	Primary: EFS at 2 years after diagnosis Secondary: • Rate of haematological CR after induction • Rate of molecular CR after 3 consolidation cycles • Probability of OS • Cumulative incidence of relapse • Toxic effects • QoL	Primary: QoL (EORTC QLQ-C30 and HADS) Secondary: • OS • RFS • EFS • Incidence of relapse (morphological and molecular)	Primary: 2 week reduction in time to haematological CR Secondary: • Safety • Molecular response • OS • DFS
Trial design and duration	Prospective, randomised, open-label, phase III non-inferiority trial	Randomised, controlled, phase III open-label trial	Randomised study
Median follow up	Initial cohort: 34.4 months (updated analysis 53 months) Final cohort: 40.6 months	30.5 months (53.4 months in updated analysis)	21 months
Location	40 centres in Italy and 27 in Germany	81 hospitals in the UK, Denmark and New Zealand	Details not reported. Patients were referred onto the study from 17 hospitals in France.
Number of participants	156 in initial cohort and 266 in final cohort	235 randomised patients	20

Source: Tables 2.3, 2.4 and 2.5 of the CS
AATO = ATRA+ATO; AIDA = ATRA + idarubicin; APL = Acute promyelocytic leukaemia; ATO = arsenic trioxide; ATRA = All-trans retinoic acid; CR = complete remission; DFS = disease-free survival; EFS = event-free survival; EORTC = European Organisation for Research and Treatment of Cancer; HADS = Hospital Anxiety and Depression Scale; OS = overall survival; QoL = quality of life; RCT = randomised controlled trial; RFS = relapse-free survival

The target sample size for the APL0406 trial was 162 patients at which point randomisation and enrolment were closed. This represented the initial cohort of patients. However, it was found, on preliminary analysis, that compliance with quality of life assessment was suboptimal. In order to ascertain the effects of arsenic-based treatment on quality of life the protocol was amended to increase the sample size to 276 patients (a final cohort). It is important to realise that the initial cohort of patients are included in the final cohort. Numbers available for analysis in the initial cohort were 156 and 266 in the final cohort.

Newly diagnosed patients taking part in AML17 did not receive ATO at its licensed indication. Additionally, gemtuzumab ozogamicin (GO) was an optional treatment in high-risk patients randomised to AATO and seven low-to-intermediate risk patients in this study received GO to counteract rising white blood cell (WBC) counts. Treatments given in AML17 are shown in Table 4.3 and for APL0406, which used ATO according to its licensed indication, in Table 4.4.

Table 4.3: Overview of treatments in AML17

Intervention	AATO (ATRA+ATO)	AIDA
Induction	Oral ATRA (45 mg/m^2/day until CR or for up to 60 days) + IV ATO (0.3 mg/kg on days 1–5 and 0.25 mg/kg twice-weekly in weeks 2–8) Gemtuzumab ozogamicin (6 mg/m^2 single IV infusion within days 1–4).[1]	Oral ATRA (45 mg/m^2/day until CR or up to 60 days) IV idarubicin (12 mg/m^2/day for a total of 4 doses)
Consolidation	Cycles 1–3: oral ATRA (45 mg/m^2/day for 15 days, two weeks on, two weeks off) + IV ATO (0.3 mg/kg on days 1–5 and 0.25 mg/kg twice-weekly in weeks 2–4) Cycle 4: oral ATRA (45 mg/m^2/day for 15 days) + IV ATO (0.3 mg/kg on days 1–5 and 0.25 mg/kg twice-weekly in weeks 2–4)	1st cycle: oral ATRA (45 mg/m^2/day for 15 days) + IV idarubicin (5 mg/m^2/day for a total of 4 doses) 2nd cycle: oral ATRA (45 mg/m^2/day for 15 days) + IV mitoxantrone (10 mg/m2/day for a total of 4 days) 3rd cycle: oral ATRA (45 mg/m^2/day for 15 days) + IV idarubicin (12 mg/m^2/day for 1 dose)
Maintenance	No maintenance phase	No maintenance phase

Source: Table 2.4 of the CS

1) Gemtuzumab ozogamicin (GO) was an optional treatment in high-risk patients randomised to AATO. Of 30 high-risk patients in this group, 28 (93%) received GO, with the remaining two patients given an anthracycline instead. Additionally, seven low- to intermediate-risk patients in this study received GO to counteract rising WBC counts.

AATO = ATRA+ATO; AIDA = ATRA + idarubicin; ATO = arsenic trioxide; ATRA = All-trans retinoic acid; CR = complete remission; IV = intravenous

Table 4.4: Overview of treatments in APL0406

Intervention	AATO (ATRA+ATO)	AIDA
Induction	Oral ATRA (45 mg/m^2/day) + IV ATO (0.15 mg/kg/day) Both continued until CR or up to 60 days	Oral ATRA (45 mg/m^2/day until CR or up to <60 days) + IV idarubicin (12 mg/m^2/day for a total of 4 doses)
Consolidation	Cycles 1–3: oral ATRA (45 mg/m^2/day for 15 days, two weeks on, two weeks off) + IV ATO (0.15 mg/kg/day 5 days per week, four weeks on, four weeks off) Cycle 4: oral ATRA (45 mg/m^2/day for 15 days) + IV ATO (0.15 mg/kg/day 5 days per week for four weeks)	1st cycle: oral ATRA (45 mg/m^2/day for 15 days) + IV idarubicin (5 mg/m^2/day for a total of 4 doses) 2nd cycle: oral ATRA (45 mg/m^2/day for 15 days) + IV mitoxantrone (10 mg/m^2/day for a total of 5 days) 3rd cycle: oral ATRA (45 mg/m^2/day for 15 days) + IV idarubicin (12 mg/m^2/day for 1 dose)
Maintenance	No maintenance	Oral ATRA (45 mg/m^2/day for 15 days every 3 months for 2 years, for a total of 6 courses) alternating with intramuscular or oral methotrexate (15 mg/m^2/week) + oral 6-MP (50 mg/m^2/day) for a total of 7 courses
Source: Table 2.3 of the CS AATO = ATRA+ATO; AIDA = ATRA + idarubicin; ATO = arsenic trioxide; ATRA = All-trans retinoic acid; CR = complete remission; IV = intravenous		

ERG comment:

Newly diagnosed patients

- The most important point to note is that only one directly relevant RCT is presented in the submission (APL0406). Newly diagnosed patients taking part in AML17 (a mainly UK-based trial) did not receive ATO at its licensed indication. For this reason, the remainder of this report focuses on APL0406 in newly diagnosed patients which was the main trial used in economic modelling. AML17 is briefly described under section 4.2.7 'Supporting evidence'.
- The population in APL0406 is relevant to the scope as it includes adults with low-to-intermediate risk APL.
- The intervention and comparator in APL0406 are relevant to the scope of this appraisal. ATO is delivered at its licensed indication.
- The outcomes in APL0406 included in the scope of this appraisal are assessed. Event-free survival is assessed rather than progression-free survival but these outcomes are similarly defined in APL0406.
- APL0406 is randomised, open-label, non-inferiority trial. The fact that the trial is open-label means that care providers, participants and outcome assessors are not blind to treatment allocation so in this respect bias can be introduced. A quality assessment of this trial is found in section 4.2.4.
- APL0406 follows an initial cohort of 156 patients up to a median of 53 months. The final cohort including all 266 patients is followed up to a median of 40.6 months.
- APL0406 is a multicentre trial with centres in Italy and Germany. There are no UK patients. The committee will need to consider the importance of this issue given that the treatment and comparator are relevant to the UK setting. The trial does not include a maintenance phase for

ATO + AIDA. The company clarified that "Primary market research commissioned by Teva in 2015 suggested that APL treatment in the UK does not include maintenance therapy."[19] Furthermore, the ERG notes that licensing for ATO does not specifically include a maintenance phase.[20]

- The evidence for the efficacy and safety of ATO + AIDA in patients with low-to-intermediate risk APL is based on 266 patients from the APL0406 trial.

Patients with relapsed/refractory disease

- The only trial presented in the CS relating to relapsed/refractory patients was Raffoux et al which included 20 patients (10 in each arm). The trial compared AATO with ATO alone which is not a relevant comparator according to the NICE scope (both arms include ATO). Therefore, no relevant evidence in patients with relapsed/refractory disease was presented in the CS for relapsed/refractory patients.
- In the clarification letter the company was invited to include all relevant non-RCTs of ATO if no RCTs were available for this patient group. The company had excluded non-RCT studies which did not address first line patients in the CS. In response the company stated "the available non-randomised second-line studies of ATO are generally supportive of its use in this indication"[19] and added "Among the studies on second-line ATO use that were initially identified by our literature search but later rejected as they did not focus on first-line indication, two deserve particular attention….".[19] They described a retrospective analysis of 25 patients with relapsed APL treated with ATO for remission induction [24] and a retrospective registry-based study from Japan showing that the annual number of autologous transplants among APL patients in second complete response (CR) increased approximately four-fold after ATO became commercially available in the country in late 2004; however, it was not clear how many patients in this study had actually used ATO.[25] It was unclear why these two particular studies had been chosen and whether other evidence supporting or refuting the use of ATO was available.
- In response to clarification the company stated "Overall, Teva feel that the use of ATO in the relapsed or refractory APL setting is already so well-established in routine clinical practice that it would be difficult to provide NICE with novel information based on the analysis of additional studies."[19] The committee will need to decide if this is acceptable particularly given the low numbers of patients expected to be treated at this stage. The company estimates that if ATO-based treatment were provided at first line the number of patients to be treated for relapsed disease would be approximately 10 to 16 patients in England.[1]
- No trials of ATO alone were presented for those with relapsed/refractory disease. The company stated that "We were unable to identify suitable efficacy data for ATO alone other than those published by Raffoux et al."[19] and added "Furthermore, according to all experts and especially to Dr Dillon (clinical expert for the UK), ATO alone is rarely used nowadays".[19] The committee will need to decide if they are in agreement with this perspective.
- It should also be noted that no trials in the CS compared ATO regimes with hematopoietic stem cell transplantation or with best supportive care as specified in the NICE scope.

4.2.2 Statistical analysis of APL0406

APL0406 was designed as a non-inferiority trial aiming to show that AATO was non-inferior to AIDA. This was interpreted as the experimental (AATO) arm being at most 5% inferior to the control (AIDA) arm in terms of the percentage of patients who were alive and failure-free at two years (EFS at two years).

Expected two-year EFS was 85% in the AIDA arm, based on the AIDA-2000 trial,[26] and 95% in the AATO arm, based on a previous non-randomised study.[27] The trialists calculated that 73 patients per treatment arm (146 in total) would be required based on a non-inferiority limit of 5%. This was increased to 162 to allow 10% loss to follow-up. The trial reached its target accrual in September 2010, at which point randomisation and enrolment were closed. However, based on a preliminary analysis of available quality of life data, the trial protocol was amended to increase the target accrual for the final cohort to 276 patients (57 additional patients per arm) to reach optimal quality of life (QoL) compliance.

Non-inferiority was assessed by estimating the two-sided 95% confidence interval for the between-group difference in crude rates of two-year EFS and was confirmed if the lower bound was ≥-5%. The trialists conducted a sensitivity analysis that addressed all relevant scenarios for the patients who could not be evaluated, assuming poor outcome for all patients, favourable outcome for all patients, or poor outcome for patients in the AATO group and favourable outcome for those in the AIDA group.

All efficacy analyses in the APL0406 trial were stated to be based on the 'intention-to-treat (ITT)'principle, comparing groups according to the randomly assigned treatment. This was defined as all patients who received at least one dose of assigned therapy following randomisation (n=156 in the initial cohort, n=266 in the final cohort). A per-protocol non-inferiority analysis was also carried out for the primary efficacy endpoint (EFS at two years). The per-protocol analysis set included 229 patients with sufficient follow up (>24 months).

EFS was assessed by comparing Kaplan–Meier curves, taking into account time to treatment failure and loss to follow-up. Survival distributions (EFS, OS and DFS) were estimated with the use of the Kaplan–Meier product-limit estimator and compared between groups using a log-rank test. Cumulative incidence of relapse was compared between groups using the non-parametric Gray K-sample test. Differences in percentages and other categorical variables (response rates, toxicity) were compared using Fisher's exact test or a chi-squared test. Continuous variables were compared using Mann-Whitney and Kruskal-Wallis tests. All tests were two-sided.

HRQoL was a secondary end point of the APL0406 trial. The European Organisation for Research and Treatment of Cancer Quality of Life Questionnaire–Core 30 was used to assess HRQoL at end of induction and after consolidation therapy. All analyses were based on those 156 patients (the initial cohort) who received at least one dose of treatment, with groups defined according to randomly assigned treatment. Primary analysis was performed, estimating mean HRQoL score over time and differences between treatment arms using a linear mixed model.[28]

ERG comment:

- Although APL0406 was designed as a non-inferiority trial, trialists were able to demonstrate the superiority of AATO at least on certain outcomes.
- Analyses appeared to have been conducted appropriately. However, it should be noted that an ITT analysis should normally be conducted on all patients randomised to an intervention whether or not any treatment was received. In this case the analysis of the final cohort in regard to EFS was conducted for 263 of 266 randomised.
- According to the CS, the APL0406 trial protocol was amended to increase the target accrual for the final cohort to 276 patients (57 additional patients per arm) to reach optimal quality of life (QoL) compliance. However, all QoL analyses were based on the initial cohort of 156 patients who received at least one dose of treatment.

4.2.3 Participants in APL0406

Table 4.5 shows the inclusion and exclusion criteria for the APL0406 trial.

Table 4.5: Participant inclusion and exclusion criteria in APL0406

Inclusion criteria	Exclusion criteria
• Age 18–71 years • Newly-diagnosed APL • Low- to intermediate-risk APL (WBC count at diagnosis $\leq 10\times10^9$/L) • Genetic confirmation of diagnosis required after initial enrolment* • WHO performance status score ≤2 • Creatinine level ≤3.0 mg/dL (≤265 μmol/L) • Bilirubin level ≤3.0 mg/dL (≤51 μmol/L)	• Age <18 and ≥71 • WBC count at diagnosis $>10\times10^9$/L • Other active malignancy at time of study entry • Lack of diagnostic confirmation at genetic level • Significant arrhythmias, ECG abnormalities** or neuropathy • Cardiac contraindications for intensive chemotherapy (L-VEF <50%) • Uncontrolled, life-threatening infections • Severe uncontrolled pulmonary or cardiac disease • Pregnancy*** or breastfeeding • Concomitant severe psychiatric disorder • HIV positivity • Use of other investigational drugs at the time of enrolment or within 30 days before study entry

Source; Table 2.6 of the CS

*Confirmation of diagnosis at genetic level was required for patient eligibility. However, to avoid delay in treatment initiation, patients were randomised on the basis of morphologic diagnosis only, before the results of genetic tests were available. APL diagnosis was genetically confirmed by one or more of the following methods: 1) detection of the PML–RARA fusion gene by RT-PCR, 2) demonstration of the t(15;17) translocation by conventional karyotyping or FISH, 3) evidence of a microspeckled PML pattern by indirect immunofluorescence assay

** Including: 1) congenital long QT syndrome, 2) history or presence of significant ventricular or atrial tachyarrhythmia, 3) clinically significant resting bradycardia (<50 beats per minute), 4) QTc >450 ms on screening EKG, 5) Right bundle branch block plus left anterior hemiblock, bifascicular block

*** Women who were either pregnant or breast feeding, or of child-bearing potential were excluded, defined as all women physiologically capable of becoming pregnant, unless they meet one of the following definitions: amenorrhea; post-surgical bilateral oophorectomy with or without hysterectomy; using a highly effective method of birth control (defined as those which result in a failure rate less than 1% per year) when used consistently and correctly, such as implants, injectables, oral contraceptives, IUDs, sexual abstinence or vasectomized partner.

APL = acute promyelocytic leukaemia; ECG = electrocardiogram; HIV = human immunodeficiency virus; L-VEF = left-ventricular ejection fraction;WBC = white blood count; WHO = world health organisation

The APL0406 trial included 266 patients with genetically confirmed newly diagnosed, low-to-intermediate risk APL. Table 4.6 shows the characteristics of the patients in the APL0406 trial. These include the initial cohort of 156 patients as results were presented for this group in addition to the final cohort. Details of patient characteristics are limited as the company did not conduct the trial and relied on published information for these data.

Table 4.6: Patient characteristics in APL0406

	APL0406 initial cohort		APL0406 final cohort	
Treatment arm	AATO (n = 77)	AIDA (n = 79)	AATO (n = 129)	AIDA (n = 137)
Male gender; n (%)	40 (52)	36 (46)	60 (46.5)	70 (51.1)
Age, years; median (range)	44.6 (19.1 to 70.2)	46.6 (18.7 to 70.2)	46.6 (18.8 to 70.2)	46.6 (18.0 to 70.3)
WBC count, x 10^9/L; median (range)	1.49 (0.32 to 10.00)	1.60 (0.30 to 9.61)	1.4 (0.3 to 10.0)	1.5 (0.3 to 9.6)
Platelet count, x 10^9/L; median (range)	31 (3 to 224)	27 (3 to 236)	36.5 (3 to 224)	31.5 (3 to 236)
Low risk, n (%)	33 (43)	27 (34)	57 (45.2)	55 (41.3)
Intermediate risk, n (%)	44 (57)	52 (66)	69 (54.7)	78 (58.6)
High risk, n (%)	NA	NA	NA	NA
Source: Table 2.7 of the CS[1, 12, 14] AATO = ATRA+ATO; AIDA = ATRA + idarubicin; ATO = arsenic trioxide; ATRA = All-trans retinoic acid; NA = not applicable; WBC = white blood cell				

The median age of participants in APL0406 was 46.6 years in both arms of the trial with ages ranging from 18 to 70 years. Just under half of the participants in APL0406 are male. Approximately 42% had low risk disease with the remainder having intermediate risk.

ERG comment: The ERG asked if the company had access to the clinical study report (CSR) for APL0406 but the company stated that as it was an investigator-sponsored study it was impossible for Teva to obtain additional data including the CSR. From the information available and using the AML17 trial as a proxy for UK practice, the ERG concludes that the patients appear to reflect those seen in UK clinical practice.

4.2.4 Quality assessment of APL0406

Quality was assessed in the CS using a tool adapted from CRD's guidance for undertaking reviews in health care.[23] Elements assessed were randomisation, allocation concealment, comparability of groups, blinding of care providers, patients and outcome assessors and drop out, selective reporting of outcomes and use of intention to treat analysis and appropriate methods for dealing with missing data. The company assessed the APL0406 trial using published papers as they were not involved in the trials. The ERG has also assessed the trial using the published papers. Results are shown in Table 4.7.

Table 4.7: Quality assessment of APL0406

Quality dimension	CS evaluation[1]	ERG evaluation[1]	ERG comment
Was randomisation carried out appropriately?	Not clear	Not clear	No information although the protocol states that 'central randomisation' was to be used.
Was the concealment of treatment allocation adequate?	Not clear	Not clear	No information
Were the groups similar at the outset of the study in terms of prognostic factors?	Yes	Yes	
Were the care providers, participants and outcome assessors blind to treatment allocation?	No	No	This was an open label trial.
Were there any unexpected imbalances in drop-outs between groups?	No	No	
Is there any evidence to suggest that the authors measured more outcomes than they reported?	Not clear	Not clear	Hospitalisation days were listed in the protocol but these do not appear to have been reported.
Did the analysis include an ITT analysis? If so, was this appropriate and were appropriate methods used to account for missing data?	Not clear[2]	No	Analysis is best described as 'modified ITT' as patients were required to have received at least one dose of assigned therapy after randomisation.

Source: Table 2.12 of the CS
1.Based on Platzbecker et al 2017[14] and Lo Coco et al 2013[12]
2. The ITT population was described as including all patients who received at least one dose of assigned therapy after randomisation, i.e. 266 and 156 patients in the final and initial cohorts, respectively. However, the ITT analysis for the primary endpoint actually included 263 and 150 patients, respectively, and the available information is insufficient to conclude if this analysis was appropriate and if appropriate methods were used to account for missing data.
ITT = intention to treat

ERG comment:

- It was not possible for the ERG to fully assess the quality of the trial without access to the full CSR. We agree with the company that issues relating to randomisation, allocation concealment and assessment of outcomes are unclear based on published information**.**
- The fact that the trial is open-label means that care providers, participants and outcome assessors are not blind to treatment allocation so in this respect bias can be introduced.
- Analysis was not strictly based on intention-to-treat as only patients who had received at least one dose of assigned therapy after randomisation were included in the analysis.

4.2.5 Results of APL0406

The main results for APL0406 are given in Table 4.8. The primary endpoint (EFS at two years for the initial cohort) showed that more patients were event-free at two years with AATO (97%) compared to AIDA (86%) ($p < 0.001$ for non-inferiority; $p = 0.02$ for superiority). Based on the final cohort of 129 patients receiving AATO and 137 receiving the AIDA regimen, AATO was found to be superior to AIDA. Significantly more patients were event-free ($p <0.001$) at two years with AATO (98.3%) compared to AIDA (86.8%) and at 50 months (97.3% vs. 80.0%).

Based on the final cohort, overall survival was significantly better ($p = 0.007$) in the AATO group (99.2% vs. 94.8%, $p = 0.007$) at two years and at 50 months (99.2% vs. 92.6% (87.9 to 97.5).

There were no statistically significant differences in the proportions of patients with haematological complete response after induction (100% vs 97%, $p = 0.12$) or in molecular complete response rate after third consolidation cycle, (100% vs 98.3%, p = NR) in the final cohort.

Quality of life results from the APL0406 trial are available only for the initial patient cohort (156 patients) assessed at the end of induction and following the third consolidation course. However, no pre-treatment baseline assessment was performed. Of 150 patients eligible for HRQoL assessment at the end of induction, 115 returned HRQoL forms (77%). After the third consolidation cycle 119 of 142 eligible patients (84%) returned forms. Compliance rates did not differ significantly between the two treatment arms. Measured on the EORTC QLQ-C30 (version 3), a significant overall difference between treatment arms was only detected for fatigue ($p=0.022$). The company stated that comparison of scores at individual time points showed that AATO was associated with significantly lower fatigue severity after induction but not after the third consolidation course. A long-term QoL analysis in the final APL0406 patient cohort remains to be reported (see section 4.2.8).

Table 4.8: APL0406: key clinical efficacy results

Endpoint and time frame	Initial cohort			Final cohort		
	AATO (n = 77)	**AIDA (n = 79)**	**P value**	**AATO (n = 129)**	**AIDA (n = 137)**	**P value**
EFS at 2 years, % (95% CI)	97 (NR)	86 (NR)	< 0.001 for non-inferiority; 0.02 for superiority	98.3 (95.9 to 100)	86.8 (81.1 to 92.8)	< 0.001
EFS at 50 months, % (95% CI)	96 (92 to 100)	81 (73 to 91)	0.003	97.3 (94.3 to 100)	80.0 (72.9 to 88.0)	< 0.001
OS at 2 years, % (95% CI)	99 (96 to 100)	91 (85 to 97)	0.020	99.2 (97.7 to 100)	94.8 (91.1 to 98.6)	0.007
OS at 50 months, % (95% CI)	99 (96 to 100)	88 (81 to 96)	0.006	99.2 (97.7 to 100)	92.6 (87.9 to 97.5)	0.007
DFS at 2 years, % (95% CI)	97 (94 to 100)	90 (84 to 97)	0.110	98.3 (95.9 to 100)	89.4 (84.1 to 95.0)	< 0.001
DFS at 50 months, % (95% CI)	NA	NA	NA	97.3 (94.3 to 100)	82.6 (75.6 to 90.3)	< 0.001
Haematological CR rate after induction, %	100	95	0.120	100	97.0	0.120
Molecular CR rate after 3rd consolidation cycle, n (%)	75 (100)	70 (100)	NR	115 (100)	117 (98.3)	NR
CIR at 2 years, % (95% CI)	1 (0 to 4)	6 (0 to 11)	0.240	0.9 (0 to 2.7)	8.2 (3.3 to 13.2)	0.0013
CIR at 50 months, % (95% CI)				1.9 (0.0 to 4.5)	13.9 (7.1 to 20.6)	0.0013

Source: Table 2.13 of the CS
AATO = ATRA+ATO; AIDA = ATRA + idarubicin; ATO = arsenic trioxide; ATRA = All-trans retinoic acid; CIR = Cumulative incidence of relapse; CR = complete remission; DFS = disease-free survival; EFS = event-free survival; OS = overall survival

4.2.6 Safety results of APL0406

The CS noted that all adverse events (AEs), adverse drug reactions (ADRs), serious adverse events (SAEs) and serious unexpected adverse reactions were recorded during the treatment period in the APL0406 study. No long-term safety data were collected. The company stated that "All patients in the APL0406 study received differentiation syndrome prophylaxis with prednisone (0.5 mg/kg/day) from day 1 until the end of induction treatment…. At the earliest manifestations of suspected differentiation syndrome (e.g., unexplained respiratory distress) temporary discontinuation of ATRA and/or ATO treatment and prompt administration of dexamethasone was recommended."[1] Adverse events in the final patient cohort of APL0406 are given in Table 4.9.

Table 4.9: Adverse events in the final patient cohort of APL0406

Adverse event	Time frame	AATO (n = 129)	AIDA (n = 137)	P value
Induction-specific adverse events, n (%)				
Patients with moderate to severe differentiation syndrome	During induction	21 (17)	17 (13)	0.38
Leukocytosis*	During induction	56 (43)	NR	NR
Haematological adverse events				
Patients with grade 3–4 neutropenia lasting >15 days, n (%)	During induction	61 (35)	109 (64)	< 0.001
	1st consolidation cycle	8 (16)	40 (67)	< 0.001
	2nd consolidation cycle	7 (7)	90 (92)	< 0.001
	3rd consolidation cycle	5 (15)	28 (85)	< 0.001
Patients with grade 3–4 thrombocytopenia lasting >15 days, n (%)	During induction	74 (38)	120 (62)	< 0.001
	1st consolidation cycle	6 (26)	17 (74)	< 0.001
	2nd consolidation cycle	6 (7)	77	< 0.001
	3rd consolidation cycle	8 (23)	16 (76)	< 0.001
FUO and infection episodes, n (%)	During induction	30 (23)	75 (55)	< 0.001
	1st consolidation cycle	10 (8)	8 (6)	0.540
	2nd consolidation cycle	4 (3)	46 (38)	< 0.001
	3rd consolidation cycle	2 (1.6)	2 (1.7)	1.000
Non-haematological adverse events				
Patients with QTc prolongation**, n (%)	During induction	11 (8.5)	1 (0.7)	0.002
	1st consolidation cycle	3 (2)	0	0.110
	2nd consolidation cycle	3 (2)	0	0.110
	3rd consolidation cycle	2 (1.5)	0	0.230
Patients with grade 3–4 hepatic toxicity, n (%)	During induction	51 (40)	4 (3)	< 0.001
	1st consolidation cycle	5 (4)	1 (0.7)	0.110
	2nd consolidation cycle	1 (0.8)	0	0.490
	3rd consolidation cycle	0	0	NA
Patients with grade 3–4 gastrointestinal toxicity, n (%)	During induction	3 (2)	25 (18.2)	< 0.001
	1st consolidation cycle	0	1 (0.8)	1.000
	2nd consolidation cycle	0	6 (4.9)	0.03

Adverse event	Time frame	AATO (n = 129)	AIDA (n = 137)	P value
	3rd consolidation cycle	0	0	1.000
Patients with grade 3–4 cardiac function abnormalities, n (%)	During induction	0	5 (3.7)	0.060
	1st consolidation cycle	0	0	NA
	2nd consolidation cycle	0	0	NA
	3rd consolidation cycle	0	0	NA
Neurotoxicity (all grades), n (%)	During induction	1 (0.7)	0	0.480
	1st consolidation cycle	5 (4.2)	0	0.020
	2nd consolidation cycle	6 (5)	0	0.010
	3rd consolidation cycle	7 (5.9)	0	0.006
Hypercholesterolemia, n (%)	During induction	14 (10)	12 (8.7)	0.550
	1st consolidation cycle	19 (16)	12 (9.6)	0.130
	2nd consolidation cycle	19 (16)	12 (9.7)	0.140
	3rd consolidation cycle	16 (14)	11 (9.0)	0.270
Hypertriglyceridemia, n (%)	During induction	29 (22)	29 (22)	0.760
	1st consolidation cycle	22 (18.4)	19 (15.2)	0.490
	2nd consolidation cycle	17 (14.4)	10 (8)	0.120
	3rd consolidation cycle	16 (14)	13 (11)	0.500

Source: Table 2.18 of the CS
* Leukocytosis was defined as WBC count >10 × 10^9/L
** Defined as QTc increased to >450 msec in males and >460 msec in females
AATO = ATRA+ATO; AIDA = ATRA + idarubicin; ATO = arsenic trioxide; ATRA = All-trans retinoic acid; FUO = fever of unknown origin

From Table 4.9 it can be seen that there were no significant differences between groups in numbers of patients with moderate to severe differentiation syndrome in the induction phase. However, in the AATO group there was a high incidence (43%) of leukocytosis during induction.

In the AATO group patients experienced fewer haematological adverse events including fever and infection episodes and grade 3 to 4 neutropaenia and thrombocytopaenia lasting over 15 days.

In terms of non-haematological adverse events, AATO was more favourable than AIDA for grade 3-4 gastrointestinal toxicity. However, a greater number of patients experienced QTc prolongation with AATO. This was particularly the case in the induction phase (8.5% vs 0.7%). A greater number of patients experienced grade 3 to 4 hepatic toxicity, again particularly in the induction phase (40% vs. 3%). In almost all patients, this toxicity was reversible and manageable with temporary drug interruption and dose adjustments as per protocol recommendations.[14] There were no instances of neurotoxicity with AIDA but 19 events were noted with AATO. Rates of hypercholesterolemia and hypertriglyceridemia were similar across groups.

ERG comment:

- Safety information on the AATO combination at the licensed dose for the first line treatment of APL is currently limited to one trial in which 129 patients have been exposed. Furthermore, the EMA commented that, "due to the potential synergistic toxicity of ATRA and ATO (i.e. on

hepatotoxicity), no direct extrapolation of safety data observed with single-agent ATO is considered adequate".[20]

- Knowledge of long-term toxicity of AATO is very limited. It is drawn to the attention of the committee that the EMA has recommended that the company conduct a post-authorisation long term safety cohort study. This is designed to explore further the long-term safety of AATO in newly diagnosed low-to-intermediate risk APL patients in a real-world clinical practice setting.
- The ERG draws to the attention of the committee the increase in rates of hepatotoxicity particularly during the induction phase. The EMA noted that this might be due to a possible synergistic toxic effect of ATRA and ATO. However, they noted that the observed hepatic damage was reversible with suspension of ATO and/or ATRA, and that no additional safety measures beyond a warning on the SmPC were necessary.[20]
- Patients will need to be carefully informed of the particular risks of the treatment regimen chosen.
- The company was asked to clarify a statement from the CS. They stated that "The estimated overall cumulative exposure to Teva Group products containing ATO was approximately 13,855 patients, with an estimated 363 patients exposed to ATO in 6 clinical trials sponsored by Teva Group." The ERG asked which six trials were being described and whether full data could be provided if relevant to the current decision problem. The company responded that the estimated cumulative clinical trials exposure to ATO in six clinical trials sponsored by Cephalon, Inc. (CTI 1073, CTI 1058, CTI 1061, ATO202, CTI 1064, C18477/3059/AM/ USCA) and 5 clinical trials sponsored by Cell Therapeutics, Inc. (CTI1057, CTI1059, CTI1060, CTI1062, CTI1063) was approximately 363 patients. The company (Teva) stated that they were aware of the fact that the cumulative number of patients exposed to arsenic trioxide in all clinical trials prior to the acquisition by Teva Group may be higher since, due to historical reasons, Teva's access to much of the data regarding studies conducted with ATO was limited.

Serious adverse events (SAEs) occurring in APL0406 are displayed in Table 4.10. Overall, 95 SAEs were reported in 65 patients: 43 SAEs in the AATO group and 52 in the ATRA + chemotherapy group.

Table 4.10: Serious adverse events in APL0406

System organ class **Preferred term, n (%)**	**AATO** **(n = 129)**	**AIDA** **(n = 137)**
Blood and lymphatic system disorders	**0**	**10 (7.3)**
Febrile neutropaenia	0	8 (5.8)
Bone marrow failure	0	1 (0.7)
Neutropaenia	0	1 (0.7)
Cardiac disorders	**3 (2.3)**	**7 (5.1)**
Pericarditis	1 (0.8)	2 (1.5)
Acute myocardial infarction	1 (0.8)	1 (0.7)
Cardiac failure	0	1 (0.7)
Ejection fraction decreased	0	1 (0.7)
Myocardial ischaemia	0	1 (0.7)
Syncope	1 (0.8)	0
Tachyarrhythmia	0	1 (0.7)
Eye disorders	**1 (0.8)**	**0**

System organ class **Preferred term, n (%)**	**AATO** **(n = 129)**	**AIDA** **(n = 137)**
Diplopia	1 (0.8)	0
Gastrointestinal disorders	**1 (0.8)**	**5 (3.6)**
Anal haemorrhage	0	1 (0.7)
Diarrhoea	0	1 (0.7)
Dyspepsia	1 (0.8)	0
Emesis	0	1 (0.7)
Inguinal hernia	0	1 (0.7)
Pancreatitis acute	0	1 (0.7)
General disorders	**1 (0.8)**	**5 (3.6)**
Mucusal inflammation	0	2 (1.5)
Pyrexia	1 (0.8)	2 (1.5)
Fever in aplasia	0	1 (0.7)
Hepatic disorders	**4 (3.1)**	**0**
Hepatotoxicity	1 (0.8)	0
Hypertransaminasemia	1 (0.8)	0
Hepatic failure	1 (0.8)	0
Cholelithiasis	1 (0.8)	0
Injury, poisoning and procedural complications	**1 (0.8)**	**1 (0.7)**
Maternal exposures before pregnancy	1 (0.8)	1 (0.7)
Infections and infestations	**6 (4.7)**	**10 (7.3)**
Pneumonia	2 (1.6)	2 (1.5)
Bronchopneumonia	0	2 (1.5)
Catheter site infection	2 (1.6)	0
Infection	0	2 (1.5)
Sepsis	0	2 (1.5)
Febrile infection	1 (0.8)	0
Herpes zoster	1 (0.8)	0
Bacteraemia	0	1 (0.7)
Urinary tract infection	0	1 (0.7)
Investigations	**7 (5.4)**	**2 (1.5)**
Electrocardiogram QT prolonged	2 (1.6)	0
ALT increased	2 (1.6)	0
AST increased	1 (0.8)	0
Hepatic enzyme increased	1 (0.8)	0
C-reactive protein increased	1 (0.8)	0
Hyperglycaemia	0	1 (0.7)
Transaminases increased	0	1 (0.7)

System organ class Preferred term, n (%)	AATO (n = 129)	AIDA (n = 137)
Nervous system	**4 (3.1)**	**1 (0.7)**
Cerebrovascular accident	1 (0.8)	0
Cerebral haemorrhage	1 (0.8)	0
Depression	1 (0.8)	0
Hydrocephalus	1 (0.8)	0
Ischaemic stroke	0	1 (0.7)
Psychiatric disorders	**1 (0.8)**	**0**
Confusional state	1 (0.8)	0
Reproductive system and breast disorders	**1 (0.8)**	**0**
Endometriosis	1 (0.8)	0
Respiratory, thoracic and mediastinal disorders	**10 (7.8)**	**7 (5.1)**
Retinoic acid syndrome	1 (0.8)	3 (2.2)
Respiratory failure	2 (1.6)	2 (1.5)
APL differentiation syndrome	3 (2.3)	0
Dyspnoea	3 (2.3)	0
Acute respiratory distress syndrome	0	1 (0.7)
Pneumonia	1 (0.8)	0
Pulmonary embolism	0	1 (0.7)
Skin and subcutaneous tissue disorders	**1 (0.8)**	**0**
Leucocytoclastic vasculitis	1 (0.8)	0
Vascular disorders	**1 (0.8)**	**4 (2.9)**
Extradural haematoma	0	1 (0.7)
Intracranial aneurysm	1 (0.8)	0
Pulmonary embolism	0	1 (0.7)
Shock haemorrhagic	0	1 (0.7)
Thrombosis	0	1 (0.7)

Source: EMA assessment report[20]
AATO = ATRA+ATO; AIDA = ATRA + idarubicin; APL = Acute promyelocytic leukaemia; ATO = arsenic trioxide; ATRA = All-trans retinoic acid

ERG comment: The ERG asked if information was available on treatment-related deaths in the included trials. The company responded that treatment-related deaths were not specifically reported. However, they provided additional information based on the trial publications. They stated that in APL0406, whilst no induction deaths were observed in the AATO group, in the AIDA group four patients died during induction therapy – two from differentiation syndrome, one from ischaemic stroke and one from bronchopneumonial causes. The company stated that as differentiation syndrome is a common adverse effect of ATRA (and ATO) these two cases, could be considered related to ATRA administration. In terms of the death from ischaemic stroke, they stated that this was reported in the publication as an SAE with a fatal outcome, and was deemed related to treatment with both ATRA and idarubicin. They further stated that the relationship between the death from bronchopneumonia and

study treatment was difficult to evaluate based on the information available. The company also stated that across both treatment groups, six patients died in complete remission (CR). One patient in the AATO group died of bronchopneumonia caused by infection with the H1N1 virus, (reported as unrelated to treatment with either ATRA or ATO). The remaining five patients who died in CR were in the AIDA group: bronchopneumonia (two, both considered related to treatment), and one each from haemorrhagic shock (unrelated), pulmonary embolism (unrelated) and secondary myelodysplastic syndrome (MDS) (reported as treatment-related).[19]

4.2.7 Supporting evidence

As the AML17 trial was conducted largely in UK patients, it is useful to compare its characteristics with those of APL0406. One difference is that while the APL0406 trial enrolled patients aged 18 to 71, the AML17 trial was open to patients aged 16 or over, with no upper age limit. Median age was similar (age 47) as were the proportion of male patients (51%). 235 patients were included in the trial of whom 57 were at high risk and 178 were at low risk. Patients in the low risk category in AML17 had a similar WBC (up to 9.9)[13] to those in APL0406.

It has already been discussed that the intervention in AML17 was not according to the licensed dose. This trial had less frequent arsenic dosing and higher dosage compared to APL0406. Additionally, gemtuzumab ozogamicin (GO) was an optional treatment in high-risk patients randomised to AATO and seven low-to-intermediate risk patients in this study received GO to counteract rising WBC counts. In contrast to APL0406, no prophylaxis for differentiation syndrome was recommended in the AML17 trial.

Primary outcomes also differed. APL0406 assessed event-free survival (EFS) at two years after diagnosis whilst AML17 assessed quality of life. The AML17 trial was due to recruit 300 patients, allowing more than 80% power to detect a difference of 6 to 7 points (out of 100) on the global health scale of the EORTC QLQ-C30 questionnaire based on data the AML15 trial.[29] However, AML17 closed randomisation after recruiting 235 eligible patients as no further drug supply was available. As a result, the trial had 80% power to detect a difference of 7.5 points on the global health scale of the EORTC QLQ-C30 questionnaire. The Hospital Anxiety and Depression Scale (HADS) was also used to assess quality of life. Patients enrolled in the AML17 trial returned a total of 671 completed QoL forms (156 at baseline, 137 at three months, 139 at six months, 136 at 12 months and 103 at 24 months). The company reported that no statistically significant difference was detected in the primary outcome of global functioning (effect size = 2.17 (95% CI: 2.79 to 7.12)). The company reported that, based on the power calculation, the confidence intervals ruled out a minimally clinically important disadvantage of six points for AATO compared with AIDA. For other measures, including fatigue, which was significantly better with AATO than AIDA in the APL0406 trial, benefits of AATO were of modest size and results not statistically significant. Small but statistically significant benefits of AATO over AIDA were seen for cognitive functioning (effect size = 5.95 (95% CI: 0.26 to 11.63)) and role functioning (effect size = 6.74 (95% CI: 0.26 to 13.21)). The remainder of the results are given in Table 4.11.

Table 4.11: AML17 key clinical efficacy results

Endpoint and time frame	AATO (n = 116)	AIDA (n = 119*)	HR or OR (95% CI)	P value
Haematological CR, NR, n (%)	109 (94)	106 (89)	OR 0.54 (0.21 to 1.34)	0.180
Molecular CR, NR n (%)	106 (91)	105 (88)	OR 0.71 (0.31 to 1.65)	0.430
OS at 4 years, % (95% CI)	93 (86 to 96)	89 (81 to 93)	HR 0.60 (0.26 to 1.42)	0.250
Early mortality at 30 days, % (95% CI)	4 (2 to 10)	6 (3 to 12)	HR 0.72 (0.23 to 2.31)	0.560
Early mortality at 60 days, % (95% CI)	5 (2 to 11)	9 (5 to 16)	HR 0.55 (0.21 to 1.43)	0.220
EFS at 4 years, % (95% CI)	91 (84 to 95)	70 (56 to 80)	HR 0.35 (0.18 to 0.68)	0.002
Haematological RFS at 4 years, % (95% CI)	97 (90 to 99)	78 (63 to 88)	HR 0.24 (0.09 to 0.60)	0.004
Molecular RFS at 4 years, % (95% CI)	98 (91 to 99)	70 (62 to 83)	HR 0.17 (0.08 to 0.39)	< 0.001
Cumulative incidence of death in remission at 4 years, % (95% CI)	2 (1 to 9)	1 (0.2 to 8)	HR 1.72 (0.18 to 16.6)	0.640
Cumulative incidence of haematological relapse at 4 years, % (95% CI)	1 (0.1 to 7)	18 (10 to 34)	HR 0.16 (0.06 to 0.46)	< 0.001
Cumulative incidence of molecular relapse at 4 years, % (95% CI)	0	27 (18 to 45)	HR 0.12 (0.05 to 0.30)	< 0.001
Cumulative incidence of tMDS -AML at 4 years, % (95% CI)	0	3 (0.4 to 17)	HR 0.15 (0.003 to 7.48)	0.340

Source: Table 2.16 of the CS

*No data were available for survival or relapse for two patients in the ATRA + chemotherapy group (1 low risk, 1 high risk)

AATO = ATRA+ATO; AIDA = ATRA + idarubicin; ATO = arsenic trioxide; ATRA = All-trans retinoic acid; CI = confidence interval; EFS = event-free survival; HR = hazard ratio; NR = not reported; OR = odds ratio; OS = overall survival; RFS = recurrence-free survival; tMDS-AML = treatment-related acute myeloid leukaemia or myelodysplastic syndrome

As for APL0406, EFS was superior in the AATO group (HR 0.35 (0.18 to 0.68). Both haematological and molecular RFS were superior in the AATO group. However, outcomes relating to early mortality and overall survival were not significantly different between treatment groups. This was in contrast to APL0406 where overall survival was superior for AATO.

ERG comment:

- The AML17 trial provides supporting evidence only for this submission as the intervention in AML17 was not according to the licensed dose.
- The patients in AML17 are predominantly from the UK so represent a population relevant to clinical practice. However, the trial includes high risk patients who are not relevant to this submission.

- A comparison between the results of APL0406 and AML17 is difficult due to differences in population, intervention and other factors such as provision of prophylaxis for differentiation syndrome in APL0406.
- AML17 provides additional evidence of the efficacy of the AATO regime for selected clinical outcomes.
- It is noted that the primary outcome of quality of life was not found to be superior for AATO but it is possible that the trial was underpowered to investigate this.

4.2.8 Ongoing trials

The company mentioned in the CS that a post-authorisation safety study (PASS) is in the process of approval by the EMA. This study will start in 2018 and run for five years to evaluate long-term safety in APL patients treated at first line with AATO.

The ERG asked if any further analyses were planned or publications in process regarding the trials in the CS and if any details were available on the quality of life assessment of the final cohort of APL0406. The company responded that 'APL0406 was an Investigator Sponsored Study and Teva only received the final publication. However, according to Professor Lo-Coco, a publication presenting the updated outcome of patients enrolled in the APL0406 trial at a 60-month median follow-up is planned for 2018.'[19] and "the long-term quality of life analysis will be based on a decision by the principal investigators, Prof. Efficace and Prof. Lo Coco. Teva is expecting the final publication of this analysis in 2019."[19]

ERG comment:

- The ERG is satisfied that none of the ongoing trials could have been used to inform the submission.
- The ERG notes that efficacy and safety of AATO for the treatment of patients at first line (beyond 50 months assessed in the trial) is unknown.
- Ongoing research will highlight longer-term efficacy, safety and quality of life issues.

4.3 Critique of trials identified and included in the indirect comparison and/or multiple treatment comparison

Not applicable.

4.4 Critique of the indirect comparison and/or multiple treatment comparison

Not applicable.

4.5 Additional work on clinical effectiveness undertaken by the ERG

No further work on clinical effectiveness was undertaken by the ERG.

4.6 Conclusions of the clinical effectiveness section

The CS included systematic reviews of the evidence for arsenic trioxide and its comparators in newly diagnosed and relapsed/refractory patients as per the NICE scope. The company presented evidence from three RCTs: Two of these were trials in newly diagnosed APL (APL0406 and AML17) and the third was a study in patients with relapsed APL (Raffoux, et al. 2003).

Untreated APL

Both trials in newly diagnosed APL (APL0406 and AML17) compared AATO (all-trans retinoic acid (ATRA) + ATO) with AIDA (ATRA + idarubicin). APL0406 included 266 patients with newly-diagnosed, low-to-intermediate risk APL aged 18 to 71 years; while AML17 included 235 patients with newly-diagnosed APL of any risk group, aged 16 or over (no upper age limit). APL0406 took place in Italy and Germany whereas AML17 had trial centres in the UK, Denmark and New Zealand. The dosing and regimens for the intervention arm (AATO) in AML17 were not in accordance with the licence; while the dosing and regimens for the intervention arm (AATO) in APL0406 were in accordance with the licence. There were further differences between populations (inclusion of high risk patients in AML17) and outcomes. It is not, therefore, appropriate to pool the results. In this report we focused primarily on APL0406 as this was according to the licence. The trial was not conducted by TEVA so both the company and the ERG relied on published information for details. This meant that specific issues regarding trial quality were not always clear to the ERG. There are no UK patients in APL0406. The committee will need to consider the importance of this given that the treatment and comparator are relevant to the UK setting. The trial does not include a maintenance phase.

Efficacy results from APL0406 showed that AATO significantly improved overall survival (OS) at 50 months compared with AIDA (99.2% vs 92.6% respectively, p=0.007). The primary endpoint of this trial was event-free survival (EFS) at two years in the initial cohort of 156 patients (97% with AATO vs 86% with AIDA, p<0.001 for non-inferiority, p=0.02 for superiority). EFS was significantly better in the AATO group across all subsequent analyses to reach 97.3% at 50 months in the full cohort of 266 patients, compared with 80.0% in the AIDA group (p<0.001). The primary source of the observed EFS benefit was a reduction in the number of relapses with AATO – at 50 months, the cumulative incidence of relapse was 1.9% in the AATO group compared with 13.9% in the AIDA group (p=0.0013). Efficacy results from AML17 were generally supportive.

Safety information on the AATO combination at the licensed dose for the first line treatment of APL is limited to 129 patients exposed to AATO in APL0406. In this trial in the induction phase there were no significant differences between groups in numbers of patients with moderate to severe differentiation syndrome but in the AATO group there was a high incidence (43%) of leukocytosis. In the AATO group patients experienced fewer haematological adverse events including fever and infection episodes and grade 3 to 4 neutropaenia and thrombocytopaenia lasting over 15 days. AATO was also more favourable than AIDA for grade 3-4 gastrointestinal toxicity. However, a greater number of patients experienced QTc prolongation with AATO particularly during induction (11% vs 0.7%). A greater number of patients experienced grade 3 to 4 hepatic toxicity, again particularly in induction (40% vs. 3%). In almost all patients, this toxicity was reversible and manageable with temporary drug interruption and dose adjustments as per protocol recommendations.[14] There were no instances of neurotoxicity with AIDA but 19 instances were noted with AATO. Patients will need to be carefully selected and informed of the particular risks of the chosen regimen. Knowledge of long-term toxicity of AATO for this group of patients is limited. It is drawn to the attention of the committee that the EMA has recommended a post-authorisation long term safety cohort study to explore this.

Relapsed or refractory APL

The CS presented one study in relapsed/refractory patients. The study by Raffoux et al. (2003) compared AATO with ATO, which is not a relevant comparison according to the NICE scope. OS was similar between the AATO and ATO study arms. Across both groups, the estimated two-year OS was 59% (95% CI: 35%–77%). EFS was not reported in this study.

The ERG considers that non-RCTs could have been included in the CS for the relapsed/refractory population particularly as no directly relevant RCT evidence is presented. The committee will need to consider whether it is necessary to explore the evidence further given the company's view that "the use of ATO in the relapsed or refractory APL setting is already so well-established in routine clinical practice that it would be difficult to provide NICE with novel information based on the analysis of additional studies."[19]

No trials of ATO alone were presented for those with relapsed/refractory disease. The committee will need to decide if they are in agreement with the company that ATO alone is rarely used in UK practice. It should also be noted that no trials in the CS compared ATO regimes with hematopoietic stem cell transplantation or with best supportive care as specified in the NICE scope.

5. COST EFFECTIVENESS

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

Three SLRs were performed with the objectives to identify and select relevant 1) cost effectiveness analysis (CEA) studies in APL (CS Appendix G); 2) utility studies identify in APL (CS Appendix H); 3) costs and healthcare resource use studies in APL (CS Appendix I).

5.1.1 Searches performed for cost effectiveness section

The following paragraphs contain summaries and critiques of all searches related to cost effectiveness presented in the company submission.

Searches for cost effectiveness analysis review

A SLR was conducted to identify cost effectiveness evaluations. No details of the search methods used were provided in the main company submission (section B.3). Full details of the search strategies were reported in Appendix G for MEDLINE, Embase and the NHS Economic Evaluation Database (NHS EED). The host provider for MEDLINE and Embase was reported, but not for NHS EED. The company response to the ERG clarification letter confirmed that the Centre for Reviews and Dissemination (CRD) interface was used to search NHS EED. The date searches were conducted was provided, but not the database date range searched. Initial searches were conducted in July 2016, and an update search was conducted in October 2017.

Measurement and valuation of health effects

A SLR was conducted to identify health-related quality of life studies. No details of the search methods used were provided in the main company submission, section B 3.4.7. Full details of the search strategies were reported in Appendix H, although this was not indicated in the main company submission. MEDLINE, Embase and the NHS EED were searched. The host provider for MEDLINE and Embase was reported, but the host provider used to search NHS EED was not reported. The company response to the ERG clarification letter confirmed that the CRD interface was used to search NHS EED. The date searches were conducted was provided, but not the database date range searched. Initial searches were conducted in July 2016, and an update search was conducted in October 2017.

Cost and healthcare resource identification, measurement and valuation

A SLR was conducted to identify costs and resource use data for England. Details of the search methods used were not provided in the main company submission, section B 5. Full details of the search strategies were reported in Appendix I for MEDLINE, Embase and NHS EED. Searches were conducted in July 2016, and an update search was conducted in October 2017. The company submission reported that targeted searches were conducted to identify adverse event costs (per occurrence) if the required data were not available in the National Schedule of Reference Costs, 2014-15: these targeted searches were not provided. The company described how these data were identified via targeted searches in their response to the ERG clarification letter.

ERG comment: As per the clinical effectiveness search comments above (4.1.1), better use of adjacency, truncation and synonyms would have increased the sensitivity of the searches. Studies may have been missed due to inappropriate use of subject headings and search filters. Additionally, it is possible that potentially relevant studies were excluded from the final search results because the method used to limit the MEDLINE and Embase searches to human studies was incorrect. See Appendix 1 for further details.

5.1.2 Inclusion/exclusion criteria used in the study selection

Screening of publications by title and abstract was performed; followed by full publication review. Eligibility criteria for the review are presented in Table 5.1.

Table 5.1: Eligibility criteria for the systematic literature reviews

Eligibility domain	Inclusion criteria	Exclusion criteria
Population	Adult APL population	Children-only population (≤15 years)
Intervention(s)	Any intervention	-
Comparator(s)[a]	Any intervention	
Outcomes(s) 1 (Published economic evaluations)	Model structure (health states & transitions, decision tree), model specifications	
Outcomes(s) 2 (Utility studies)	Any relevant health utility data	
Outcomes(s) 3 (Cost/resource use studies)	Any relevant cost and resource use information	
Study design 1 (Cost effectiveness analysis studies)	Health economic evaluation, any methodology	Opinion, editorial letter
Study design 2 (Utility studies)	Any kind of study including utility data (utility elicitation studies or models referring to utility data)	
Study design 3 (Cost/resource use studies)	Any study including models, analysis of insurance databases or medical records, cross-sectional surveys, chart reviews or prospective observational studies	

ERG comment: The in- and exclusion criteria presented in Table 5.1 seem appropriate for the objective of this review. However, after considering the PRISMA charts, it appeared that additional exclusion criteria were applied. This included "Full text not available" and "Old study (>2 years)". As a result, some relevant studies might have been missed. Additionally, given the company eventually informed the model partly based on primary sources focusing on other populations than APL, extending the population for the SLR (beyond the APL population only) might have been informative.

5.1.3 Included/excluded studies in the cost effectiveness review

The searches related to CEA, utility and cost studies resulted in 145, 273 and 280 hits respectively (after removing duplicates) for screening. Eventually this resulted in six included publications for the review of CEA (of which two were abstracts), two for the utility review and 11 (of which four were abstracts) included publications for the review of cost studies respectively.

ERG comment: It is noticeable that publications were excluded based on "Full text not available" and "Old study (>2 years)", this was applicable to four, one and nine studies for the CEA, utility and cost studies SLRs respectively. As stated above, some relevant studies might therefore have been missed.

5.1.4 Conclusions of the cost effectiveness review

The cost effectiveness searches in the company submission were all documented and reproducible. However, there were a number of inconsistencies and inaccuracies, and some redundancy. The MEDLINE and Embase search strategies used an inappropriate 'animals' limit, and it is possible that relevant evidence may have been missed as a consequence

Considering the CEA SLR, the company concluded that in general, all of the included studies considered the cost effectiveness of AATO or ATO alone, compared to the combination of ATRA and chemotherapy. In all cases, the number of QALYs was higher in the groups receiving ATO than in the comparator groups. Mean total costs of AATO or ATO alone were higher than the costs of comparator treatments. Incremental cost effectiveness ratios (ICERs) differed between studies, which could be related to a number of methodological factors, including the fact that the studies concerned different countries.

Considering the utility SLR, the company stated that both included studies (which were also included in the CEA SLR) presented utilities which were based on conditions other than APL (i.e. chronic lymphocytic leukaemia and AML). Hence it was concluded that no utility values that were specific to APL could be identified from the SLR.

Considering the cost and resource use SLR, the company did not use the 11 identified studies in the economic model. This was justified by stating that the information captured was not compatible with that needed to populate the model, and in others by the fact NHS reference costs were preferentially used to ensure relevance to the current situation in England.

5.2 Summary and critique of company's submitted economic evaluation by the ERG

Table 5.2: Summary of the company's economic evaluation (with signposts to CS)

	Approach	Source/Justification	Signpost (location in CS)
Model	A Markov cohort model with 14 health states	To fully capture the impact of making ATO available to UK APL patients treated within the NHS	Chapter B 3.2
States and events	Health states include: - First line treatment induction, consolidation, < 2yrs and >2yrs remission health states - Second line treatment induction, consolidation and remission health states - HSCT (allogeneic or autologous) and post-HSCT remission and End of Life health states - tMDS/AML - Death	To capture the course of the disease, based on expert opinion.	Chapter B 3.2.2

	Approach	Source/Justification	Signpost (location in CS)
Comparators	AIDA	AIDA was the only comparator in the pivotal APL0406 trial, and the primary comparator in the AML17 trial. BSC was not used as a comparator because it only applies to the second line setting.	Chapter B 3.2.3
Population	Adult patients with newly diagnosed low-to-intermediate risk APL. Cost effectiveness was not assessed in the refractory / relapsed patients.	It is expected that clinical practice will shift towards the use of ATO as standard of care in newly-diagnosed patients.	Chapter B 3.2.1
Treatment effectiveness	Different sources are used to inform the different treatment effectiveness parameters. First line treatment effectiveness estimates (including probabilities of remission at different time points and probabilities of relapse) are derived from the APL0406 trial.[30] Second line treatment effectiveness estimates are derived from Raffoux et al (2003)[18] for probabilities of remission and treatment failure, and Tallman (2015)[31] for probabilities of relapse; and Russel et al (2017)[32] and Platzbecker et al (2016)[14] for probabilities of alloHSCT, autoHSCT. Post HSCT transitions are informed by Hosing et al (2003),[33] Ramadan et al (2012)[34] and de Botton et al (2005)[35] for mortality risk and by Holter Chakrabarty et al (2013)[36] for probability of molecular remission. The probability of death in the tMDS/AML state is informed by Ma et al (2007).[37] In addition to the RCT by Raffoux et al.,[18] the efficacy data in the second line part of the model were informed by clinical expert opinions and a previous cost effectiveness	The outcomes related to the use of ATRA+ATO and AIDA in newly-diagnosed APL were mainly estimated based on the head-to-head APL0406 clinical trial. A scenario analysis was also conducted with the treatment schedule and outcomes from the AML17 clinical trial. However, no head-to-head data versus AIDA were available for second line treatment.	Chapter B 3.3

	Approach	Source/Justification	Signpost (location in CS)
	model developed for ATRA+ATO in the US.[31]		
Adverse events	Several treatment-induced adverse events were considered in the model in terms of costs and patients QoL. Some could prompt treatment switch or discontinuation.	No justification for the selection of AEs was provided.	Chapter B 3.4.8
Health related QoL	Utilities are based on Tallman et al (2015),[31] which provided utility values in a chronic lymphocytic leukaemia population, and based on Lachaine et al (2015),[38] which reported utilities from acute myeloid leukaemia patients.	No studies reporting utility values specific to APL were identified. Previous cost-effectiveness studies in APL used proxy utilities for other conditions that the authors considered to be associated with utilities analogous to APL.	Chapter B 3.4.9
Resource utilisation and costs	Resource use and costs accounted for in the model are treatment acquisition costs, medical costs (treatment administration, supportive care, monitoring and follow-up, HSCT, palliative care), and management of adverse events costs. These were informed using NHS reference costs, the BNF, the PSSRU and publications of relevant trials.	The information captured in the 11 studies identified through the SLR was not compatible with that needed to populate the model, or not relevant to the setting in England.	Chapter B 3.5
Discount rates	Discount of 3.5% for utilities and costs	As per NICE reference case	Table 3.13
Sub groups	Not applicable		
Sensitivity analysis	Both DSA and PSA were performed as well as scenario analyses	As per NICE reference case	Chapter B 3.8

Source: CS[1]

AE = adverse events; AIDA=chemotherapy combined with all-trans retinoic acid; AML=acute myeloid leukaemia; APL=acute promyelocytic leukaemia; ATO=arsenic trioxide; ATRA=all-trans retinoic acid; BNF=British National Formulary; BSC=best supportive care; CS=company submission; DSA=deterministic sensitivity analysis; HSCT=haematopoietic stem cell transplant; PSA = probabilistic sensitivity analysis; PSSRU=Personal Social Services Research Unit; SLR=systematic literature review; tMDS/AML=treatment-related myelodysplastic syndrome or acute myeloid leukaemia.

5.2.1 NICE reference case checklist (TABLE ONLY)

Table 5.3: Summary of the company's economic evaluation (with signposts to CS): NICE reference case checklist

Elements of the economic evaluation	Reference Case	Included in submission	Comment on whether *de novo* evaluation meets requirements of NICE reference case
Population	As per NICE scope[2]	Partly	Cost effectiveness is not assessed in the refractory/relapsed (second line) setting.
Comparator(s)	Therapies routinely used in the National Health Service (NHS), including technologies regarded as current best practice	Partly	For the second line setting, as per the NICE scope, BSC should be considered as a comparator.
Type of economic evaluation	Cost effectiveness analysis	Yes	
Perspective on costs	NHS and Personal Social Services (PSS)	Yes	
Perspective on outcomes	All health effects on individuals	Yes	
Time horizon	Sufficient to capture differences in costs and outcomes	Partly	Time horizon of 40 years, used in the base-case, does not capture all relevant costs and effects
Synthesis of evidence in outcomes	Systematic review	Yes	
Measure of health effects	Quality adjusted life years (QALYs)	Yes	
Source of data for measurement HRQoL	Described using a standardised and validated instrument	Yes	
Source of preference data for valuation of changes in HRQoL	Time-trade off or standard gamble	No	HRQoL data used in the model are from studies in AML and CLL patients. Utility values were derived from cost effectiveness publications, not from the original studies.
Discount rate	An annual rate of 3.5% on both costs and health effects	Yes	
Equity weighting	An additional QALY has the same weight regardless of the other characteristics of the individuals receiving the health benefit	Yes	

Elements of the economic evaluation	Reference Case	Included in submission	Comment on whether *de novo* evaluation meets requirements of NICE reference case
Sensitivity analysis	Probabilistic modelling	Yes	
Source: CS[1] AML=acute myeloid leukaemia; CLL=chronic lymphocytic leukaemia; HRQoL=Health-related quality of life; NHS=National Health Service; PSS=Personal Social Services; QALY=quality-adjusted life year			

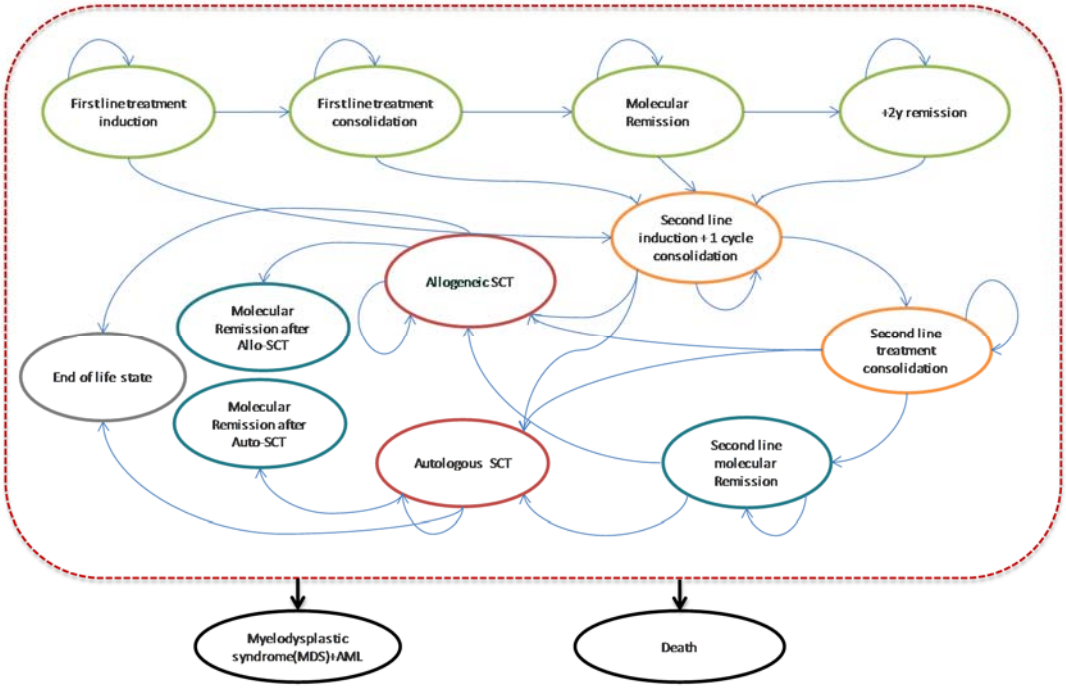
5.2.2 Model structure

The company developed a de novo Markov model comprising of 14 health states (or 69 when separately considering tunnel states, as indicated by the number of columns used in the Markov trace). No justification was provided for the chosen modelling approach, although the other economic evaluations identified in the SLR for CEA studies also used a Markov model structure. The number of health states deviated significantly from the four to five health states employed in the published economic evaluations, and this deviation was not justified.

The 14 health states of the model are shown in Figure 5.1. It should be noted that in Figure 5.1, curved arrows representing that patients can remain in health state, are missing for “End of life”, “Molecular remission after SCT”, “tMDS/AML”, “Death”. Furthermore, the arrow between “tMDS/AML” and “Death” is missing.

It is noteworthy that the length of time spent in some of these tunnel states depends on the treatment arm. Furthermore, the ATRA+ATO in first line treatment arm is implemented using two different subsequent treatment strategies: subsequent ATRA+ATO (AATO) in a large proportion of patients and subsequent AIDA in a small proportion of patients. Expert opinion suggested that the choice of subsequent treatment would depend on the duration of remission the patient has experienced. The first line AATO and second line AATO (subsequently referred to as AATO+AATO) strategy would be adopted in patients that had achieved two or more years of remission. The first line AATO, second line AIDA (AATO+AIDA) strategy is adopted in patients with less than two years remission. However, in the model, the proportion of patients experiencing two or more years of remission are implemented a priori, that is in different Markov traces.

Figure 5.1: Markov model structure with 14 health states



SCT= stem cell transplant; MDS+AML= myelodysplastic syndrome and acute myeloid leukaemia; y=year
Source: CS model file[1]

First line treatment health states

On treatment health states

There are two first line treatment health states in which patients are on treatment: the "first line induction" (during which patients are hospitalised) and the "first line consolidation" health states. All patients in the model are allocated to either AATO or AIDA first line treatment. For AATO "first line induction" consists of a maximum of two cycles of four weeks in the model and "first line consolidation" consists of a maximum of eight cycles of four weeks. In contrast, the maximum number of cycles in the "first line consolidation" state for the AIDA treatment arm is only three model cycles. As a note, the model implementation differs from the company's description in the CS, in which a maximum of three model cycles of induction and 10 cycles for consolidation phases was stated (although this might be technically possible in the economic model, first line treatment is restricted to fewer cycles).

From the "first line induction" state, patients can move to the "first line consolidation" state based on the median time necessary to achieve complete haematological remission before the maximum of two cycles and they can move to second line treatment if a cardiac event occurs. In contrast, patients would remain in the "consolidation" phase until the maximum of eight cycles, unless a cardiac event prompted treatment switch or they experienced tMDS/AML. At the end of the "first line consolidation" phase, patients that experienced treatment failure are moved to second line treatment.

Remission health states

There are two first line treatment remission health states: "first line molecular remission" and "+2y remission". In case of molecular remission after the "first line consolidation" phase, or, according to the company, if the patient could not be evaluated for remission (i.e. in the absence of evidence for treatment failure), the patient moves to the molecular remission health state. However, the latter is true

only for the AATO+AATO strategy in the model. There is an inconsistency in that, in the AIDA+AATO and AATO+AIDA strategies in the model, patients that could not be evaluated with PCR would be evaluated based on haematological response instead of being assumed to move to the molecular remission health state, and only in the case of haematological response would they move to the molecular remission health state.

Patients can remain in the "first line molecular remission" health state for a maximum of two years (24 model cycles, which is closer to 22 months due to the model cycle length of 28 days) and then move to the "+2y remission" health state. In the "first line molecular remission" health state, the probability of relapse is increased compared with the health state of "+2y remission". In case of a relapse in either one of these two remission health states, the patient moves to second line treatment. If there is no relapse, the patient remains in the "+2y remission" health state until death.

Second line treatment health states

Patients arrive in the second line induction phase in three cases: a) because they had experienced a cardiac event in first line induction or consolidation phases, b) because treatment failed after completion of the first line consolidation phase (40 weeks for AATO or 20 weeks for AIDA), or c) because of relapse when the patient had achieved molecular remission.

On treatment health states

There are two second line treatment health states in which patients are on treatment: the "second line induction + 1 cycle consolidation" and the "second line consolidation" health states. The "second line induction + 1 cycle of consolidation" health state consists of two model cycles of induction (mirroring the first line induction health state) and two model cycles of consolidation for AATO or one cycle of consolidation for AIDA in second line. Patients can move to the second part of this "second line induction + 1 cycle consolidation" health state (the consolidation cycle) if remission is achieved at one model cycle of induction therapy, to ensure that patients would always follow at least one cycle of consolidation. Patients in the AATO+AIDA strategy can transit from the "second line induction + 1 cycle consolidation" health state to the "tMDS/AML" health state. If, at the end of the "second line induction + 1 cycle consolidation" consolidation cycles, complete molecular remission is achieved, patients can continue consolidation treatment, or move to allogeneic or autologous HSCT. If complete molecular remission is not achieved, patients would undergo allogeneic HSCT (if they do not transit to "tMDS/AML"). Patients who experience a cardiac event discontinue second line treatment and undergo allogeneic HSCT.

The "second line treatment consolidation" health state comprises a maximum of six or two model cycles for AATO and for AIDA in second line respectively, at the end of which most patients undergo HSCT (allogeneic or autologous). Patients in the AATO+AIDA strategy can transit from the "second line treatment consolidation" health state to the "tMDS/AML" health state. Patients that cannot receive HSCT move to the "second line molecular remission" health state with no maintenance treatment.

Remission health state

There is only one second line remission health state: the "second line molecular remission" health state. Patients can stay in the second molecular remission health state until death, although there is a risk of relapse, which would prompt allogeneic HSCT. Patients can also move to allogeneic or autologous HSCT without experiencing a relapse.

Haematopoietic stem cell transplant related health states

The HSCT health states are populated with: a) patients that did not achieve molecular remission after the "second line induction + 1 cycle of consolidation" health state (only allogeneic HSCT), b) a proportion of patients that achieved complete molecular remission after the "second line induction + 1 cycle of consolidation" health state (autologous or allogeneic HSCT), c) patients that had a cardiac event in the "second line induction + 1 cycle of consolidation" or the "second line treatment consolidation" health states (only allogeneic HSCT), d) patients experiencing a relapse after achieving second molecular remission (only allogeneic HSCT), and e) proportions of patients in the "second line molecular remission" health states that did not experience a relapse (autologous or allogeneic HSCT).

Haematopoietic stem cell transplant health states

There are two HSCT health states: "alloHSCT" and "autoHSCT". The "alloHSCT" health state consists of six model cycles and reflects patients' hospitalisation and monitoring. Patients in this health state are at increased risk of acute graft versus host disease (GvHD) and at an increased risk of mortality compared with the general population. The "autoHSCT" health state consists of three model cycles and is also associated with an increased risk of mortality compared with that of the general population. Patients from both "alloHSCT" and "autoHSCT" health states would move to the respective remission after HSCT health states if the transplant was successful, or to the "End of life" health state if the transplant was not successful at the end of the respective tunnels.

Post HSCT health states

There are three post HSCT health states: "molecular remission after alloHSCT", "molecular remission after autoHSCT", and "End of life" (also called "Failure" in the model file). The "molecular remission after alloHSCT" is associated with increased costs, lower health-related quality of life, an increased mortality risk (compared with the general population) and a risk of developing chronic GvHD. The "molecular remission after autoHSCT" is associated an increased risk of mortality (compared with the general population), but with lower costs and better quality of life compared with the "molecular remission after alloHSCT", and no risk of chronic GvHD was applied. The "End of life" state is associated with low quality of life, high costs caused by extensive palliative care and a higher mortality risk than the molecular remission after HSCT health states.

Treatment-related MDS/AML and death health states

There are two other health states in the model: "tMDS/AML" and "death". Patients treated with AIDA in first or second line can experience tMDS/AML during the first line treatment consolidation phases. Patients stay there until they die, and mortality risk is increased compared to that of the general population.

Patients can die at any time in the model due to general population background mortality. Patients have an additional mortality risk when they are in the "tMDS/AML", the "alloHSCT" and "autoHSCT" health states, the "End of Life", the "molecular remission after alloHSCT" and the "molecular remission after autoHSCT" health states. The increased risk of mortality during induction and consolidation phases (due to bleeding and infection) was not modelled. Patients in the model therefore do not experience increased mortality in first and second line treatment induction, consolidation or remission health states.

ERG comment: The ERG's concerns relate to (a) a model structure that diverges from existing economic models in this therapeutic area, (b) inconsistent modelling of patients that cannot be evaluated for molecular remission, (c) of adverse events it was assumed that only cardiac events could prompt

treatment switch, (d) an error in the number of tunnels used to represent the two year molecular remission health state, (e) the absence of disease-related mortality from on treatment health states and (f) the applicability of alloHSCT to the UK clinical setting.

(a) A model structure that diverges from the one used in other economic evaluations in this condition. In response to the clarification question B1, the company justified the more complex model structure by stating that the existing economic evaluations did "not adequately reflect the trajectory of APL patients".[19] According to the company, the aim in this economic evaluation was to "offer more granularity with treatment phases, molecular remission and HSCT" to better reflect the clinical trajectory of APL patients. The company also explored the potential impact of their adopted model structure compared with the simpler model structures by comparing model outcomes and found, with the caveat that the settings in the model were different and a straight comparison is therefore not possible, that the inclusion of HSCT in this model and differences in drug costs across the models may account for differences in estimated costs between the models. The ERG considers that the model structure is appropriate to reflect this condition and treatment pathway.

(b) There is an inconsistency in what happens in the model when patients could not be evaluated for molecular remission. Patients in the AATO+AATO strategy would be assumed to be in molecular remission, while patients in the AATO+AIDA and AIDA+AATO strategies that could not be evaluated with PCR would be evaluated based on haematological response, and only if this was given patients were assumed to be in molecular remission (instead of assuming that all patients, regardless of haematological response, are in molecular remission). This was not justified and the ERG prefers to implement this in a consistent manner across treatment strategies. This is further explored in the treatment effectiveness section (Section 5.2.6) of this report.

(c) It was assumed in the model that among adverse events, only cardiac events could prompt a treatment switch. In response to the clarification question B3, the company stated that this was based on expert opinion. The company stated that "while it is possible that other serious AEs may prompt a treatment switch, they were not frequent enough to either find adequate probabilities or have any impact on the end results."[19] The minutes of the company's expert consultation, however, revealed that in a small proportion of patients reversible arrhythmia would also cause treatment switch. The ERG therefore has explored this in scenario analysis.

(d) The "first line molecular remission" health state is a tunnel state consisting of 24 model cycles. However, the company intended this to represent two years in remission. Due to the cycle length of four weeks, the appropriate number of cycles to reflect two years would be 26 cycles. In response to clarification question B19.c, the company acknowledged that these health states were missing and implemented 26 model cycles for this health state in a scenario, which resulted in slightly more favourable model outcomes for AATO, but with "minor impact".[19] The ERG implemented the 26 cycles in its base-case.

(e) No disease-related mortality was modelled during on treatment and remission phases. The company excluded disease-related mortality from the on treatment health states. In response to the clarification question B2, the company justified this modelling choice by stating that "the mortality rate observed during treatment in both the APL0406 trial and the AML17 trial was numerically lower for ATRA+ ATO compared to AIDA" and that the difference was not statistically significant.[19] The company explored the impact of adding disease-related mortality to the on treatment health states of their model and found that incremental QALYs increased and costs savings with AATO decreased. The ERG considers that the disease-related mortality risk is likely to be larger than the general population mortality risk in the treatment induction phase, which is an assumption consistent with the evidence

shown in the AML17 study. This additional mortality risk, is therefore implemented, during treatment induction, in exploratory analyses performed by the ERG.

(f) In the model, patients can undergo either autologous or allogeneic HSCT. However, it is questionable whether this is reflective of UK practice. For instance, patients who have a cardiac event in second line treatment can only receive allogeneic HSCT, not autologous HSCT. In response to clarification question B3.c, the company stated that this assumption was based on expert opinion that only patients with molecular remission are considered for autologous HSCT in the model and patients with a cardiac event would likely experience this before molecular remission. The company also explained that according to a UK expert "fewer HSCTs are conducted in the UK and allogeneic HSCT is generally not recommended in APL". The meeting minutes of the company's expert consultation support this. To reflect the uncertainty over the use of alloHSCT in the UK clinical practice, the ERG adopted a scenario in which only autoHSCT is performed.

5.2.3 Population

Arsenic trioxide (ATO), as per its marketing authorisation, is indicated for the treatment of:

- newly diagnosed low-to-intermediate risk APL (white blood cell count $\leq 10 \times 10^3/\mu l$) in combination with all-trans retinoic acid (ATRA) *(also referred to as first line treatment)*
- relapsed/refractory APL (previous treatment should have included a retinoid and chemotherapy) *(also referred to as second line treatment)*

The company only assesses the cost effectiveness of ATRA+ATO (AATO) in the newly diagnosed low-to-intermediate risk APL population, i.e. in first line treatment. The indication in the relapsed/refractory APL population is not assessed.

In the model, patients have an average age of 45 years, an average weight of 81kg and an average height of 169 cm. In total, 48.7% were assumed to be male. See Table 5.4 for the baseline characteristics of patients from the main evidence sources considered in the model.

ERG comment: AATO was only assessed in the newly diagnosed population (first line). In response to clarification question B5.a, the company provided an analysis in the relapsed/refractory population (second line) in which "the health states representing first line therapy were changed to second line, and those representing second line were neutralised (no transitions to these states were possible)".[19] The company further stated that "the analysis showed that ATRA+ATO was cost-effective versus AIDA in the second line setting with an incremental cost-effectiveness ratio (ICER) of £16,733 per QALY gained." The company's description of this analysis did not provide clarity over how this analysis was exactly performed. There was also a lack of clarity as to where the transition probabilities in the model were sourced from, and whether they reflected first line or second line treatment. The ERG therefore implemented their own scenario using the second line transition probabilities to reflect the relapsed/refractory population (second line).

Table 5.4: Key baseline patient characteristics in the APL0406 and AML17 trial

Study population	**APL0406 initial cohort**		**APL0406 final cohort**		**AML17**	
Treatment arm	**ATRA+ATO** (n=77)	**AIDA** (n=79)	**ATRA+ATO** (n=129)	**AIDA** (n=137)	**ATRA+ATO** (n=116)	**AIDA** (n=119)
Age, years; median (range)	44.6 (19.1–70.2)	46.6 (18.7–70.2)	46.6 (18.8–70.2)	46.6 (18.0–70.3)	47 (16–75)	47 (16–77)
Male gender; n (%)	50 (52%)	36 (46%)	60.0 (46.5%)	70.0 (51.1%)	60 (52%)	60 (50%)
WBC count, $\times 10^9$/L; median (range)	1.49 (0.32–10.00)	1.60 (0.30–9.61)	1.4 (0.3–10.0)	1.5 (0.3–9.6)	3.0 (0.4–100.9)	2.2 (0.4–78.2)
Platelet count, $\times 10^9$/L; median (range)	31 (3–224)	27 (3–236)	36.5 (3–224)	31.5 (3–236)	Not reported	Not reported
Low risk, n (%)	33 (43%)	27 (34%)	57.0 (45.2%)	55.0 (41.3%)	86 (74%)	92 (77%)
Intermediate risk, n (%)	44 (57%)	52 (66%)	69 (54.7%)	78 (58.6%)	Not reported	Not reported
High risk, n (%)	Not applicable	Not applicable	Not applicable	Not applicable	30 (26%)	27 (23%)

ATRA=All-trans retinoic acid; ATO=Arsenic trioxide; WBC=White blood cell

5.2.4 Interventions and comparators

First line therapy in APL generally consists of three consecutive treatment phases: induction, consolidation and maintenance.[1] However, maintenance treatment was not modelled and the justification provided by the company was that it is usually omitted in UK clinical practice with the aim of minimising the risk of treatment-related myelodysplastic syndrome or acute myeloid leukaemia (tMDS/AML) (CS Section B.1.3.2.1).

First and second line treatment with AATO was modelled with up to two cycles (of four weeks) of induction therapy followed by eight cycles (of four weeks) of consolidation therapy. Treatment protocols were in line with the APL0406 study.[12]

The only comparator used in the model was AIDA in first line. AIDA was implemented with up to two cycles (of four weeks) of induction therapy followed by three cycles (of four weeks) of consolidation therapy.

After first line treatment, subsequent treatment is prompted by relapse or by a cardiac event in the model. As per its marketing authorisation, ATO does not have to be administered in combination with ATRA in this second line population. However, the company only provided the analysis with ATRA+ATO in combination, in line with expert opinion stating that ATO alone would only rarely be used in the relapsed/refractory population.

Furthermore, there is a lack of clarity as to whether re-treatment with AATO can occur when a patient had relapsed. Informed by expert opinion, the company assumes in the economic model that patients who remained in remission for two years or longer following first line treatment with AATO would be re-treated with AATO upon relapse. Patients who achieved only a short (<2 years) remission after first line AATO treatment would be treated with AIDA. Of course, all patients treated with AIDA in first line, would switch to AATO in second line after relapse or cardiac event, independent of how long the period of remission was. In the model, the proportion of patients achieving two or more years of remission (and therefore assumed to be treated with AATO+AATO instead of with AATO+AIDA) are implemented a priori, that is in different Markov traces.

A comparison of ATO in the second line setting was not performed. If it had been implemented, according to the scope, Best Supportive Care should be considered a comparator in the second line setting. In response to the request for clarification, the company did perform an analysis of ATO at second line. However, Best Supportive Care was not implemented in the model as a comparator.

ERG comment: The ERG's concerns relate to (a) the lack of maintenance treatment in the company's model, (b) the absence of an analysis with ATO stand-alone in second line, and (c) the absence of BSC as a comparator in the second line setting.

(a) The company did not consider maintenance treatment in their model. In response to clarification question A.12, the company stated that their earlier insight based on market research, that in the UK maintenance therapy would only be provided in rare cases, was confirmed by a UK expert. It should be noted though that the rationale for maintenance therapy being rarely used in the UK is based on the AML15 study showing a higher incidence of tMDS/AML with maintenance therapy than the AML17. The incidences in both of these studies are for patients treated with chemotherapy regimens. Since the incidence of tMDS/AML is not a concern with AATO treatment, this does not justify not including maintenance treatment with AATO. However, maintenance therapy with AATO is not in line with the SmPC and therefore the ERG considers it as appropriate that maintenance therapy was not considered in the model.

(b) Only ATRA+ATO is modelled in second line, not ATO stand-alone. In response to clarification question B5.b, the company stated that there was no other evidence for second line treatment than the Raffoux et al study, which "did not show significant differences between ATO+ATRA and ATO alone, and, surprisingly, disease-free survival was better with ATO alone than with ATRA+ATO. Conducting this scenario would lead to better cost-effectiveness results for ATO vs. AIDA, reducing treatment acquisition costs without changing the effectiveness results."[19] The ERG was satisfied with this justification, especially given that experts stated that ATO alone would only rarely be used in UK clinical practice.

(c) BSC in second line was not included as a comparator in the model. In response to clarification question B5.c, the company stated that "all experts strongly stated that, due to the severity of the disease, best supportive care is not a relevant comparator in the second line setting, and that best supportive care is only a relevant alternative in 3rd or 4th line."[19] Furthermore, "given the very small number of affected patients, adding best supportive care as a comparator in 3rd or 4th line would have very little impact on the ICER."[19] The ERG was satisfied with this justification.

5.2.5 Perspective, time horizon and discounting

The model adopts the perspective of the NHS and Personal and Social Services (PSS) in England and Wales. The model time horizon is 40 years, at the end of which a significant proportion of patients in the model is still alive (> 40% of patients in the ATRA+ATO first line and AIDA second line arm). The model cycle length is 4 weeks to capture the treatment schedule and a half-cycle correction is applied. All costs and health gains were discounted at a rate of 3.5% per year.

ERG comment: The time horizon was too short to capture all relevant costs and outcomes. The company, in response to clarification question B20,[19] provided a scenario analysis with an extended time horizon of 56 years, which increased both cost savings and incremental QALYs. The ERG remains concerned about the long life expectancy of patients in the model and thinks that this calls the validity of the model into question (see section 5.2.12 for more details).

5.2.6 Treatment effectiveness and extrapolation

The treatment effectiveness section was structured according to different phases: first line, second line, haematopoietic stem cell transplant and other phases (i.e. tMDS/AML or death).

Treatment effectiveness of the AATO strategy was estimated by separately estimating Markov traces (as well as costs and QALYs) for AATO+AATO and AATO+AIDA. Subsequently, a weighted average was calculated with weights of 98% and 2% for AATO+AATO and AATO+AIDA respectively. No justification was provided for these weights.

First line health states

Transition from the first line health states were informed based on evidence from the APL0406 trial[12, 14] in combination with expert opinion (to inform assumptions related to the estimation of these transition probabilities).

From the "first line treatment induction" health state, patients can transit to "first line treatment consolidation" depending on the median time to CR (32 versus 35 days for AATO and AIDA respectively) and the occurrence of adverse events (i.e. cardiac event) requiring a treatment switch (probabilities of 0.0% and 3.0% for AATO and AIDA respectively). In case of adverse events requiring a treatment switch (i.e. cardiac events), patients transit to the "second line treatment induction phase +

1 cycle consolidation" health state. In case of no cardiac event after two cycles (i.e. 56 days), patients transit to "first line treatment consolidation".

From the "first line treatment consolidation" health state, patients can transit to the first line "molecular remission" health state if they have remission and do not transit to the "tMDS/AML" health state. For the Markov traces for AATO+AIDA and AIDA+AATO, the probability of remission is determined by the haematological response rate (98.4% versus 96.4% for AATO and AIDA respectively) for patients not evaluable with PCR (9.4% versus 9.8% for AATO and AIDA respectively). For patients that are evaluable with PCR, the above mentioned haematological response rate is multiplied by the molecular remission rate (100.0% versus 98.3% for AATO and AIDA respectively). Moreover, for the Markov trace for AATO+AATO, it was assumed that all surviving patients would be in remission.

Once the patients are in the first line remission health states (i.e. "molecular remission" and "+2y remission" health states), patients can transit to the "second line treatment induction phase + 1 cycle consolidation" health state based on the probability of relapse which was different for the first two years after remission and thereafter.

The transition probabilities from the first line health states (retrieved from the model) are presented in Table 5.5. This excludes general population mortality that is subsequently applied to the transition probabilities presented in this table (no additional mortality is assumed).

ERG comment: The ERG's concerns relate to (a) the overestimation of cardiac events and thus patients switching to second line induction for AIDA; (b) the calculation of patients transiting to first line consolidation early for AIDA; (c) calculations and assumptions regarding the remission probability; (d and e) assumptions and calculation concerning the relapse probabilities and; (f) not considering treatment switching due to reversible arrhythmia in the model. The transition probabilities that are adjusted in the ERG base-case are presented between square brackets in Table 5.5.

(a) The proportion of patients switching to second line induction due to experiencing a cardiac event during first line induction AIDA treatment is 4.6% in the economic model (based on the 3.0% probability per cycle) while only 3.7% (five out of 136) experience grade 3-4 cardiac events in the APL0406 trial. This overestimation (also reflected in Table 2.2 of CS Appendix J) was induced by the company using the median time to complete haematological remission (i.e. 35 days for AIDA) in order to convert this 3.7% to a cycle probability (of 3.0%). The ERG corrected this overestimation in its base-case by converting the 3.7% to a cycle probability of 2.4% using the average duration patients actually remained in the first line induction phase (in the model) for AIDA treatment (i.e. 44 days). This resulted in 3.7% of the patients switching to second line induction due to experiencing a cardiac event during first line induction AIDA treatment (consistent with the APL0406 trial).

(b) The proportion of patients transiting to first line consolidation early (i.e. after one cycle of induction) from first line induction AIDA treatment was calculated by using the median time to complete haematological remission of 35 days. However, for the calculation to convert this to a transition probability, the company assumed that the probability after 35 days (the median time) was 48.2% and not 50.0%. Presumably this was done to reflect patients switching treatment due to cardiac events. However, the ERG believes this is incorrect as the patients switching treatment due to cardiac events are already considered using separate transition probabilities. Hence this is corrected in the ERG base-case (implicitly assuming that patients switching treatment due to cardiac events have no 'early' response).

(c) In the CS it is assumed for AATO+AATO, that all surviving patients would transit to remission. This was done despite available evidence from the APL0406 trial to inform this parameter in the model. Given the lack of appropriate justification in the CS, the ERG preferred to inform the remission probability for AATO+AATO based on evidence from the APL0406 trial, consistent with what was done for AATO+AIDA and AIDA+AATO. Related to this, the probability of transiting to remission for patients that are evaluable with PCR was informed by the molecular remission rate in the ERG base-case for all strategies (removing the additional multiplication with the haematological response rate).

(d) The a priori division of first line AATO patients in the model into a group which will experience remission for ≥2 years (receiving AATO also in second line if necessary, AATO+AATO group) and another group that will not achieve this (receiving AIDA in second line, AATO+AIDA group) causes problems in the use of the probabilities of relapse in both parts of the model. For patients in the remission health states, the company assumed no relapse during the first 24 cycles for AATO+AATO, while for AATO+AIDA, the company assumed no relapse after 24 months. Although the company did not justify this approach, the ERG presumes that the company based this on the assumption that patients who remained in remission for ≥2 years following first line AATO would receive AATO+AATO, else patients would receive AATO+AIDA. However, irrespective of the exact rationale, the company should have used conditional probabilities of relapse when adopting this approach. The group of patients that are assumed to relapse during the first 24 months (AATO+AIDA) likely have a larger probability during this period than the average probability of all patients observed in the APL0406 trial (i.e. probability of relapse conditional on a relapse within two years after transiting to the remission health state). The company had intended that none of the patients should remain in remission for more than two years in this group of patients, which is not the case in the economic model. Similarly, for the group of patients that relapse after the first 24 months, the post 24 months relapse probability is likely higher than in the first 24 months (i.e. probability of relapse conditional on having no relapse within two years after transiting to the remission health state). By using unconditional relapse probabilities, the company likely underestimates the relapse probability for AATO. Given that the ERG did not have access to these conditional relapse probabilities, the average relapse probabilities were applied for AATO in the ERG base-case (i.e. 0.038% per cycle for the first two years and 0.042% per cycle thereafter) instead of the company's approach assuming 0% probabilities of relapse during and after the first 24 months for AATO+AATO and AATO+AIDA respectively.

(e) The 48 month relapse probabilities (used in the model to inform relapse more than two years after remission), were assumed to be equal to the 50 month relapse probabilities reported in the APL0406 trial publication. This was corrected in the ERG base-case by converting the 50 month probabilities to 48 month probabilities.

(f) The company submitted minutes considering the expert meeting the company organised to validate the CEA model. Although these minutes state that for AATO approximately 2% of patients experience reversible arrhythmia and therefore switch treatment, this was not incorporated in the model. Therefore, the ERG explored this scenario (using 2% cardiac events during the induction phase).

Table 5.5: Transition probabilities (per cycle) from the first line health states (excluding background mortality)

AATO+AATO									
TO **FROM**		**First line – Induction**	**First line – Consolidation**		**First line – Remission**			**Second line – Induction[b]**	**tMDS/AML**
	Tunnel #	**2**	**1**	**2-8**	**1**	**2-24**	**25+**	**1**	**-**
First line – Induction	**1**	54.5%	45.5%					0.0%	
	2		100.0%					0.0%	
First line – Consolidation	**1-7**			100.0%					0.0%
	8				100.0% [99.9%]			0.0% [0.1%]	0.0%
First line – Remission	**1-24**					100.0% [100.0%[a]]		0.0% [0.0%[a]]	
	25+						100.0%[a]	0.0%[a]	
AATO+AIDA									
TO **FROM**		**First line – Induction**	**First line – Consolidation**		**First line – Remission**			**Second line – Induction[b]**	**tMDS/AML**
	Tunnel #	**2**	**1**	**2-8**	**1**	**2-24**	**25+**	**1**	**-**
First line – Induction	**1**	54.5%	45.5%					0.0%	
	2		100.0%					0.0%	
First line – Consolidation	**1-7**			100.0%					0.0%
	8				98.4% [99.9%]			1.6% [0.1%]	0.0%
First line – Remission	**1-24**					100.0%[a]		0.0%[a]	
	25+						100.0% [100.0%[a]]	0.0% [0.0%[a]]	
AIDA+AATO									
TO **FROM**		**First line – Induction**	**First line – Consolidation**		**First line – Remission**			**Second line – Induction[b]**	**tMDS/AML**
	Tunnel #	**2**	**1**	**2-3**	**1**	**2-24**	**25+**	**1**	**-**
First line – Induction	**1**	56.2% [55.1%]	40.9% [42.6%]					3.0% [2.4%]	
	2		97.0%					3.0%	
First line – Consolidation	**1-2**			99.5%					0.5%
	3				94.4% [97.6%]			5.1% [1.9%]	0.5%
First line – Remission	**1-24**					99.6%		0.4%	
	25+						99.8%	0.2%	

Note the probabilities between square brackets are used in the ERG base-case

[a]This transition probability rounded to 100.0% or 0.0% (i.e. it would be less than 100.0% or more than 0.0% if more decimals would be displayed)

[b]This refers to the "second line treatment induction phase + 1 cycle consolidation" health state

Second line health states

From the "second line treatment induction phase + 1 cycle consolidation" health state patients can transit to second line consolidation (either the consolidation tunnel state within this health state if patients are still in the induction phase, or to the "second line treatment consolidation" health state), to "alloHSCT", "autoHSCT" or "tMDS/AML". The transition from second line treatment induction (first cycle) to second line treatment consolidation is based on the average induction duration[12, 39] plus the duration of one consolidation cycle (in total this resulted in an assumed median duration of 60 and 63 days for second line AATO and AIDA respectively). During the second line treatment induction phase for AIDA it is possible to transit to "alloHSCT" (not possible during second line treatment induction phase for AATO) if patients experience an adverse event requiring a treatment switch (i.e. cardiac event, identical probability as for first line AIDA treatment). From the last tunnel of the "second line treatment induction phase + 1 cycle consolidation" health state (i.e. during the second or first consolidation cycle for second line AATO and AIDA respectively) patients can transit to both HSCT health states or continue to the "second line treatment consolidation". These transitions were conditional on the haematological complete response rates of 80%[18] and 70% (expert opinion) for AATO and AIDA respectively. Due to lack of data, the company assumed that the molecular remission rate was identical to haematological complete response. Subsequently, it was assumed, that all patients without remission, that did not transit to "tMDS/AML", would transit to "alloHSCT". For the patients with remission, 10.3% and 34.5% would transit to the "alloHSCT" and "autoHSCT" health states respectively. These probabilities of transiting to the HSCT health states were based on data on file, without providing detailed information on this source.[32] The remaining patients transit to the "second line treatment consolidation" health state.

From the "second line treatment consolidation" health state patients can transit to "alloHSCT", "autoHSCT", "tMDS/AML" or "second line molecular remission". In the initial "second line treatment consolidation" tunnel states, patients can only transit to the next tunnel state or to "tMDS/AML". After the last consolidation cycle (maximum total of eight and three cycles of consolidation treatment for AATO and AIDA respectively), again 10.3% and 34.5% (based on data on file[32]) transit to the "alloHSCT" and "autoHSCT" health states respectively. The remainder, that did not transit to "tMDS/AML", transits to "second line molecular remission", implicitly assuming a 100% molecular remission rate.

The cycle probability of transiting to "tMDS/AML" from the second line AIDA treatment consolidation tunnel states was slightly higher than for the first line AIDA treatment consolidation tunnel states (0.5% versus 0.7% respectively) due to a difference in the assumed maximum duration of the AIDA consolidation phase (84 versus 56 days respectively).

From "second line molecular remission" health state, patients can transit to "alloHSCT" or "autoHSCT". It is assumed that all patients that relapse (monthly probability of 1.1%[31]) would transit to "alloHSCT". For patients without relapse, the previously mentioned probabilities of transiting to the HSCT health states (i.e. 10.3% and 34.5%, based on data on file,[32] transit to the "alloHSCT" and "autoHSCT" respectively) are adjusted using the median time to relapse (of 24.5 and 14.0 months based on first line data from the APL0406 trial[14]) after second line treatment with AATO and AIDA respectively.

The transition probabilities from the second line health states (retrieved from the model) are presented in Table 5.6. This excludes general population mortality that is subsequently applied to the transition probabilities presented in this table (no additional mortality is assumed).

ERG comment: The ERG's concerns relate to (a) evidence and transparency of descriptions for transitions from second line health states and; errors concerning the transition probability (b) from second line induction to consolidation for AIDA, (c) from second line AIDA consolidation to "tMDS/AML", as well as (d) from "second line molecular remission" to the HSCT states. The transition probabilities that are adjusted in the ERG base-case are presented between square brackets in Table 5.6.

(a) The evidence presented to inform transitions from second line health states was weak and it was frequently not transparently reported how the transition probabilities were obtained. For instance, evidence was obtained from expert opinion, based on 10 patients that received AATO[18] and first line data from the APL0406 trial was used. Moreover, it was unclear how transition probabilities were obtained from Tallman et al[31] (secondary data source that describes a calibration process to obtain transition probabilities that could not be reproduced by the ERG) and Russell et al[32] (data on file). This hampered the assessment of these transition probabilities.

(b) The ERG found a reference error in cell 'Calc - TransMat ATRA+ATO'!AT141 (cell refers to J10 while it should be J96, in the economic model initially submitted by the company). This affects the transition from second line induction to consolidation for AIDA.

(c) For calculating the probability of transiting to "tMDS/AML" during second line AIDA consolidation treatment, a maximum consolidation duration of 56 days (i.e. two cycles) is assumed while the probability is applied during 84 days (three cycles). This is corrected in the ERG base-case.

(d) The transition from "second line molecular remission" to the HSCT states is adjusted using the median time to relapse following second line remission. The rationale of this adjustment is unclear to the ERG. However, the uncorrected transition probabilities to the HSCT states seem relatively high (given these are applied each cycle). Hence, the ERG explored using the uncorrected transition probabilities to the HSCT states as well as setting these to 0.0% in scenario analyses.

Table 5.6: Transition probabilities (per cycle) from the second line health states

Second line AATO (re)treatment										
TO **FROM**		**Second line – Induction + 1 Consolidation[a]**			**Second line – Consolidation**		**Second line – Remission**	**AlloHSCT**	**AutoHSCT**	**tMDS/AML**
	Tunnel #	**2**	**3[b]**	**4[c]**	**1**	**2-6**	**-**	**1**	**1**	**-**
Second line – Induction + 1 Consolidation	**1**	72.4%	27.6%					0.0%		
	2		100.0%					0.0%		
	3[b]			100.0%						
	4[c]				44.2%			28.2%	27.6%	0.0%
Second line – Consolidation	**1-5**					100.0%				0.0%
	6						55.2%	10.3%	34.5%	0.0%
Remission	**-**						96.7%	1.5%	1.7%	
Second line AIDA treatment (after first line AATO)										
TO **FROM**		**Second line – Induction + 1 Consolidation**			**Second line – Consolidation**		**Second line – Remission**	**AlloHSCT**	**AutoHSCT**	**tMDS/AML**
	Tunnel #	**2**	**3[b]**	**4[c]**	**1**	**2**	**-**	**1**	**1**	**-**
Second line – Induction + 1 Consolidation	**1**	69.4% [71.1%]	27.6% [26.5%]					3.0% [2.4%]		
	2		97.0% [97.6%]					3.0% [2.4%]		
	3[b]				38.6% [38.6%]			36.5% [36.7%]	24.2%	0.7% [0.5%]
	4[c]									
Second line – Consolidation	**1**					99.3% [99.5%]				0.7% [0.5%]
	2						54.5% [54.7%]	10.3%	34.5%	0.7% [0.5%]
Remission	**-**						95.1%	1.9%	3.0%	

Note the probabilities between square brackets are used in the ERG base-case
[a]For AATO (re)treatment this health state contains 2 consolidation cycles (tunnel states 3 and 4)
[b]This tunnel state represents the first consolidation cycle
[c]This tunnel state represents the second consolidation cycle (not used for second line AIDA treatment)

Haematopoietic stem cell transplant and other health states

This section considers the haematopoietic stem cell transplant as well as "tMDS/AML" and "death" health states in the model.

From the "alloHSCT" health state patients can transit to "molecular remission after alloHSCT", "end of life" (health state for patients that failed on HSCT) and "death". In the initial "alloHSCT" tunnel states, patients can only transit to the next tunnel state or to "death". After the last "alloHSCT" tunnel state (i.e. after six tunnel states; 168 days), patients can transit to "molecular remission after alloHSCT", "end of life" and "death". The death probability is identical (i.e. 6.7% per cycle[33]) as in the previous "alloHSCT" tunnel states. For the surviving patients it is assumed that 72.2%[36] would have a molecular remission and hence transit to "molecular remission after alloHSCT" while the remaining patients transit to the "end of life" health state.

Similar to the transitions from alloHSCT, from the "autoHSCT" health state patients can transit to "molecular remission after autoHSCT", "end of life" and "death". In the initial "autoHSCT" tunnel states, patients can only transit to the next tunnel state or to "death". After the last "autoHSCT" tunnel state (i.e. after three tunnel states; 84 days), patients can transit to "molecular remission after autoHSCT", "end of life" and "death". The death probability is identical (i.e. 2.0% per cycle[33]) to the previous "alloHSCT" tunnel states. For the surviving patients it is assumed that 98.1%[36] would have a molecular remission and hence transit to "molecular remission after alloHSCT" while the remaining patients transit to the "end of life" health state.

From the "molecular remission after alloHSCT" and "molecular remission after autoHSCT" health states patients can only transit to "death". The transition probabilities are 0.2%[35] per cycle for both health states. Similarly, from the "end of life" and "tMDS/AML" health states, patients can only transit to "death" as well. The transition probabilities were 3.1%[34, 35] and 2.7%[37] per cycle for the "end of life" and "tMDS/AML" health states respectively.

The transition probabilities from the abovementioned health states (retrieved from the model) are presented in Table 5.7. This excludes general population mortality that is subsequently applied to the transition probabilities presented in this table.

ERG comment: The ERG's concerns relate to a general lack of (a) detailed descriptions and justification for calculations, assumptions and selected sources and; (b) elaborate consideration in the CS of (implicit) assumptions regarding extrapolation.

(a) Most of the evidence sources mentioned above are not described in the CS (neither are the transition probabilities). Although this is most prominent for the transition probabilities described in the preceding section, the lack of detailed description and justification in the CS is applicable to the majority of the transition probabilities described in the treatment effectiveness section of the ERG report. This is worrying, given treatment effectiveness (including implicit assumptions made and selection of evidence sources to obtain transition probabilities) is often an influential part of the cost effectiveness model. Although the company's response to clarification question B6 was helpful, justification for calculations, assumptions and selected sources remain largely unclear to the ERG.

(b) The extrapolation of treatment effectiveness is a limitation of the model (as also indicated by the company in response to clarification question B7) that is not extensively discussed in the CS. This includes for instance the extrapolation of relative treatment effectiveness, i.e. implicitly assuming that treatment benefits are maintained for the entire time horizon. In the first line for example, the relapse transition probability is assumed to be constant from two years after remission until the end of the time

horizon (higher for AIDA than for AATO). Alternative assumptions regarding the extrapolation could be influential as illustrated by the analyses performed by the company in response to clarification question B7. In this scenario, assuming no relapse after 24 months in the first line remission health state, AATO did not remain dominant as it became more expensive than AIDA with an ICER of £7,610 per QALY gained. Acknowledging this uncertainty, the ERG added a scenario analyses assuming an equal relapse probability two years after first line remission for AATO and AIDA.

Table 5.7: Transition probabilities (per cycle) from the haematopoietic stem cell transplant or tMDS/AML health states

TO FROM		AlloHSCT	AlloHSCT – Remission	AutoHSCT	AutoHSCT – Remission	End of life	tMDS/AML	Death[a]
	Tunnel #	**2-6**	**-**	**2-3**	**-**	**-**	**-**	**-**
AlloHSCT	**1-5**	93.3%						6.7%
	6		67.4%			25.9%		6.7%
AlloHSCT – Remission	**-**		99.8%					0.2%
AutoHSCT	**1-2**			98.0%				2.0%
	3				96.1%	1.9%		2.0%
AutoHSCT – Remission	**-**				99.8%			0.2%
End of life	**-**					96.9%		3.1%
tMDS/AML	**-**						97.3%	2.7%

[a]Disease related mortality

5.2.7 Adverse events

The following treatment-induced adverse events were considered in the model:

- Thrombocytopenia (grade 3–4, duration >15 days)
- Neutropenia (grade 3–4, duration >15 days)
- Infection
- Leukocytosis
- Hepatic toxicity
- Neurotoxicity
- Differentiation syndrome
- Cardiac events
- QTc prolongation

Except for cardiac events, the adverse events listed above did not lead to a change of treatment, but impacted only the costs and patients' QoL. The duration of each adverse event was used to compute the QALYs lost due to the QoL impairment in patients experiencing the event.

In addition to the adverse events listed above, tMDS or AML was incorporated in the model structure (see section 5.2.2).

ERG comment: Similar to the treatment effectiveness parameters, the description of adverse events (AEs) in the CS lacked transparency. The AE probabilities were not mentioned in the CS. See Table 5.8 for an overview that was retrieved from the model submitted by the company (see section 5.2.6 for the probability of therapy-induced MDS or AML). Although the company's response to clarification question B6 was helpful, justifications for the selected sources are largely unclear to the ERG. This was particularly the case for the sources selected to inform the duration of AEs, including multiple sources that are published >25 years ago. Moreover, the selection of these specific AEs is unclear to the ERG; this includes that is was unclear why reversible arrhythmia was not considered in the model (as discussed in section 5.2.6).

Table 5.8: Adverse events used in the economic model (for both first line and second line AATO/AIDA treatment)

	AATO		AIDA	
	Induction	**Consolidation**	**Induction**	**Consolidation**
Thrombocytopenia (grade 3-4, >15 days)				
Probability	0.380[14]	0.187[14]	0.620[14]	0.810[14]
Duration (days)	20[40]	25[40]	20[40]	25[40]
Neutropenia (grade 3-4, >15 days)				
Probability	0.350[14]	0.127[14]	0.640[14]	0.813[14]
Duration (days)	19[41]	19[41]	19[41]	19[41]
Infection				
Probability	0.230[14]	0.042[14]	0.550[14]	0.152[14]
Duration (days)	17[42]	17[42]	17[42]	17[42]
Leukocytosis				
Probability	0.473[12]	0.000[12]	0.241[12]	0.000[12]
Duration (days)	14[43]	14[43]	14[43]	14[43]
Hepatic toxicity (grade 3-4)				
Probability	0.400[14]	0.016[14]	0.030[14]	0.002[14]
Duration (days)	10[44]	10[44]	10[44]	10[44]
Neurotoxicity (all grades)				

Probability	0.007[14]	0.050[14]	0.000[14]	0.000[14]
Duration (days)	365[45]	365[45]	365[45]	365[45]
Differentiation syndrome				
Probability	0.194[12]	0.000[12]	0.160[12]	0.000[12]
Duration (days)	4[46]	4[46]	4[46]	4[46]
Cardiac events (grade 3-4)				
Probability	0.000[14]	0.000[14]	0.037[14]	0.000[14]
Duration (days)	1[47]	1[47]	1[47]	1[47]
QTc prolongation				
Probability	0.085[14]	0.018[14]	0.007[14]	0.000[14]
Duration (days)	0.5[48]	0.5[48]	0.5[48]	0.5[48]

5.2.8 Health-related quality of life

Health state utility values

Both the APL0406 and the AML17 trials used the EORTC QLQ-C30 instrument, and not the EQ-5D, to measure HRQoL outcomes. Therefore, the company performed a SLR to identify relevant quality of life studies for the current decision problem which yielded two CEA studies focussing on APL patients.[31, 38] The study by Tallman et al. 2015[31] used utility values from chronic lymphocytic leukaemia (CLL) patients which they adjusted for age and country (adjustment method not described in the CS). Lachaine et al. 2015[38] used utility values from AML patients. The CS did not report the primary sources informing the utility values in these CEA studies. No study reporting utility values based on the EQ-5D for APL patients was identified in the literature (see section 5.1 for more details regarding the SLR).

Utility values used in the model were obtained from the study by Woods et al. 2012,[49] which reported utility values from CLL patients. The company states that this study was selected because it "presented utility values for similar health states to those in our model, reflecting the treatment pathway".[1] Woods et al. 2012[49] was presumably identified through the study by Tallman et al. 2015.[31] Additionally, the company used Beusterien et al. 2010[50] (also considering CLL patients) because it was referred to in Woods et al., 2012.[49] Szende et al. 2014[51] provided general population utility values for the current assessment. The CS did not describe how the utility values were obtained in these studies (Woods et al., 2012[49], Beusterien et al. 2010[50] and Szende et al. 2014[51]) and why these sources were deemed to be the most appropriate for the current decision problem.

The company adjusted the utility values obtained from the literature, with the intention to make them more relevant for the modelled population. Two adjustments were made: 1) an adjustment for age and; 2) an adjustment for the utility representing perfect health:

1. The age adjustment consisted of multiplying with the ratio of the utility in the general population having the same age as the modelled population (i.e. 45 years old) to the utility value in the general population with the same age as the population in which the utility value was obtained (e.g. 60 years old).[1] The UK general utility values for patients aged 45 and 60 are 0.849 and 0.804 respectively, which resulted in a factor of 1.056 (=0.849/0.804) for the age adjustment.
2. The adjustment for the utility representing perfect health consisted of multiplying the (age-adjusted) utility values by the utility value in the UK general population with the same age as the population in which the utility was elicited. This adjustment for a 60 year old patient population would then be 0.804 (i.e. UK general population utility value of a 60 year old person).

When applying both adjustments to a utility value 0.910 obtained from the literature for a 60 year old patient population, these adjustments would result in a utility value of 0.773 (= 0.910×1.056

(=0.849/0.804) × 0.804) for a 45 year old patient in the modelled population. These two adjustments combined effectively equal the multiplication by 0.849 (i.e. the general population utility value of the 45 year old modelled population). It should be noted that the company did not apply these adjustments consistently on all health state utility values obtained from the literature (see Table 5.9 for details regarding the application of the adjustments). No evidence was provided to support the need to adjust the original utility values and no justification supported the inconsistencies in adjustments between health states.

The company applied another adjustment for patients receiving second line treatment (second line induction and consolidation phase). Here, the second line treatment utility was assumed to be 91% of the first line treatment utility value. The 91% was calculated by dividing 0.71 by 0.78, which represented utility values of Stable CLL during second and first line treatment respectively (utility values presumably obtained from Woods et al., 2012[49]). No evidence was provided in the CS to justify this adjustment.

Finally, an adjustment was made to obtain the utility values for the "Allogeneic HSCT" and "Autologous HSCT" health states: the utility value of the "CML after HSCT without GvHD" (i.e. 0.979) was multiplied by the utility in the "second line molecular remission" state (i.e. 0.702), which resulted in 0.687 (Table 5.9). The primary source for the "CML after HSCT without GvHD" utility value and the rationale for adjusting it were not provided.

Patients in the long-term remission health states were assumed to have a utility value equal to the general population at the age of 45. Table 5.9 presents the utility values from the original source, the different adjustments and the utility values used in the cost effectiveness model.

Table 5.9: Overview of the health state utility values used in the model

Health state	Mean utility value in cost effectiveness model (95% CI)	Age adjustment[a]	Adjustment for general population utility value[a]	Original utility value (95% CI)	Adjustment for second line treatment	Disutility for hospitalisation
First line induction treatment	0.74 (0.71, 0.77)	1.06	None	0.70[b] (0.67, 0.73)	NA	-0.01[g]
First line consolidation treatment	0.74 (0.71, 0.77)	1.06	None	0.70[b] (0.67, 0.73)	NA	No
First molecular remission	0.77 (0.75, 0.79)	1.06	0.81	0.91[c] (0.88, 0.93)	NA	No
First long-term remission (>2 years)	0.85 (NR)	NA	NA	0.85[a] (NR)	NA	No
Second line induction + 1 cycle consolidation	0.67 (0.64, 0.70)	1.06	None	0.70[b] (0.67, 0.73)	0.91[c]	-0.01[g]
Second line treatment consolidation	0.70 (NR)	None	None	0.77[c,h] (NR)	0.91[c]	-0.01[g]
Second molecular remission	0.85 (NR)	NA	NA	0.85[a] (NR)	None	No
Allogeneic HSCT*	0.69 (NR)	None	None	0.98[d] (NR)	None	-0.01[g]
Autologous HSCT*	0.69 (NR)	None	None	0.98[d] (NR)	None	-0.01[g]
Allogeneic HSCT molecular remission	0.85 (NR)	None	None	0.85[a] (NR)	None	No
Autologous HSCT molecular remission	0.85 (NR)	None	None	0.85[a] (NR)	None	No
End of life state (Palliative care)	0.40[f] (NR)	None	None	0.40[e] (NR)	None	No
tMDS/AML	0.40[g] (NR)	None	None	0.40[f] (NR)	None	No

Source: Adapted from CS, Table 3.4

* The utility weight in 'CML after HSCT without GvHD' (i.e. 0.979) was adjusted by the utility in the "second line molecular remission" state (i.e. 0.702): 0.979×0.702= 0.687.

[a] Age-adjustments, adjustments for general population utility values, and general population utility values were based on Szende et al., 2014[51]

[b] Obtained from Woods et al., 2012[49]; [c] Obtained from Beusterien et al., 2010[50]; [d] Obtained from Breitscheidel L., 2008[52]; [e] Obtained from Morton et al., 2009[53]; [f] Obtained from Cooperberg et al., 2013[54]; [g] Assumption

[h] In the ERG base-case, this utility value is assumed to be equal to the "Second line induction + 1 cycle consolidation" health state utility value (i.e. 0.70)

Abbreviations: CI, confidence interval; HSCT, hematopoietic stem cell transplantation; NA, not applicable; NR, not reported; tMDS/AML, treatment-related myelodysplastic syndrome or acute myeloid leukaemia.

Adverse events

Disutilities for adverse events were included in the treatment induction and consolidation health states of the cost effectiveness model. Adverse events were assumed not to occur in the remission health states. The duration (see Section 5.2.7 for more details) and disutilities associated with adverse events were obtained from the literature. Disutilities associated with adverse events were calculated for each cycle and were obtained by multiplying the proportion of patients experiencing the adverse event by the duration of the adverse event and the disutilities associated with the adverse events. The proportion of patients experiencing each adverse event was not reported in the CS (Section 5.2.7). The CS did not report how the sources informing the disutilities were identified and did not justify why these sources were the most appropriate.

Besides disutilities specifically applied to the treatment induction and consolidation health states, patients experienced a disutility for hospitalisation (i.e. -0.01) during first and second line treatment (both induction and consolidation) as well as during HSCT treatment. Additionally, patients in the "Allogeneic HSCT" and the "Allogeneic HSCT molecular remission" health states were at risk of experiencing a disutility for graft versus host disease (GvHD). The proportion of patients experiencing GvHD, the duration of GvHD, and the duration of hospitalisation in the above-mentioned health states were not reported in the CS. The CS also emphasised that patients could experience acute and chronic GvHD but did not describe how these were differentiated in the cost effectiveness model. Table 5.10 provides an overview of the duration and disutility values associated with adverse events.

Table 5.10: Overview of adverse events duration and associated disutilities

State	Mean utility value	Reference	Justification (comment)	Mean duration	Reference	Justification (comment)
Hospitalisation	-0.01	Assumption				
Thrombocytopenia (grade 3-4, >15 days)	-0.18	Attard et al., 2014[55]		Induc.: 20 days Cons.: 25 days	Wolff et al., 1989[40]	
Neutropenia (grade 3-4, >15 days)	-0.18	Attard et al., 2014[55]		19 days	Fenaux et al., 1993[41]	Assumed to be the same as for ATRA+DNR+ARA-C
Infection	-0.15	Stevenson et al., 2014[56]	Based on table A1 in Platzbecker 2016[3], most infections are pneumonia. Disutility of pneumonia was considered.	17 days	Pneumonia – What happens[42]	Based on table A1 in Platzbecker et al. 2016[3], most infections are pneumonia, thus duration of pneumonia (2-3 weeks) was considered.
Leukocytosis	-0.08	Assumption		14 days	Shoenfeld et al., 1981[43]	
Hepatic toxicity	-0.2	Choi et al., 2013[57]		10 days	Zhu et al., 2013[44]	“Less than two weeks”
Neurotoxicity	-0.21	Prica et al., 2014[58]		365 days	Assumption based on Ratnaike, 2003[45]	"Acute poisoning from arsenic can lead to peripheral neuropathy which can last for max 2 years"
Differentiation syndrome	-0.12	Assumption		4 days	Breccia et al., 2008[46]	Assumed to be same as AIDA
Cardiac events	-0.16	Nshimyumukiza et al., 2013[59]	· Myocardial infarction (MI)	1 day	Mathews et al., 2002[47]	Assumed to be same as ATRA+ATO
QTc prolongation	-0.001	Assumption		0.5	Siu et al., 2006[48]	Assumed to be same as ATRA+ATO
Acute GvHD	-0.08	Breitscheidel L., 2008[52]	· Mean utility weight after HSCT without GvHD, re-scaled: 0.836	NR		

State	Mean utility value	Reference	Justification (comment)	Mean duration	Reference	Justification (comment)
			· Mean utility weight after HSCT with GvHD, re-scaled: 0.769	NR		
			· Disutility of GvHD: 1-(0.769/0.836)= 0.080	NR		
			· Applied for the duration of the monitoring phase for the proportion of patients experiencing acute GvHD	NR		
Chronic GvHD	-0.08	Breitscheidel L., 2008[52]	Assumed to be the same as acute GvHD	NR		
			Applied for a lifetime for the proportion of patients experiencing chronic GvHD	NR		

Source: Adapted from CS, Table 3.4

Abbreviations: AIDA, all-trans retinoic acid and idarubicin; Ara-C, cytarabine; ATO, arsenic trioxide; ATRA, all-trans retinoic acid; Cons., consolidation; DNR, daunorubicin; GvHD, graft versus host disease; Induc., induction; NR, not reported.

ERG comment: The ERG agrees with the company that utility values for APL patients elicited through the EQ-5D are likely not to be available in the literature. However, the ERG is concerned by the validity of the utility values for the following reasons: (a) the selection process of the utility values and the assumptions underlying disutilities associated with adverse events were unclear, (b) the non-adherence with the NICE reference case, and (c) the lack of justification supporting the adjustments made by the company.

(a) In the CS, the company refers to Woods et al. 2012, Beusterien et al. 2010 and Szende et al. 2014 as main sources informing health state utility values. Breitscheidel 2008[52], Morton et al. 2009[53], and Cooperberg et al. 2013[54] were also used to inform the utility values of the "End of life state (Palliative care)", "tMDS/AML", and "Allogeneic HSCT" and "Autologous HSCT" health states, respectively (Table 5.9). Moreover, additional sources and/or assumptions were used to inform the disutility values associated with adverse events. However, the CS did not describe how all these studies were identified and justifications or evidence supporting the assumptions made by the company were largely lacking. Therefore, the ERG requested that the company clarify the choice of different sources and assumptions made in the CS. The company responded that a targeted search was performed to identify health state utility values and disutility values associated with adverse events. The selection of utility and disutility values and the assumptions supporting the disutility values associated with adverse events had then been discussed with experts. No detail was provided on the targeted literature search (e.g. key words used to identify studies and databases which were included in this search) and the expert opinion elicitation method. Hence, the ERG was not able to assess the quality of the selection process and thus it is unclear whether the most appropriate sources have been used to inform the utility values.

(b) The selected utility values do not adhere to the NICE reference case because utility values have not been directly elicited in patients affected by APL through the EQ-5D. The CS refers to Woods et al. 2012[49] as the primary source for the utility values of several health states (Table 5.9). Woods et al. 2012 is a cost effectiveness analysis of bendamustine versus chlorambucil for the first line treatment of CLL in England and Wales.[49] Utility values in this study were obtained by mapping European Organisation for Research and Treatment of Cancer C30 quality of life data elicited in CLL patients (collected by Knauf et al.) to EQ-5D utility values. Woods et al. 2012 [49] does not provide a precise reference to the utility elicitation study of Knauf et al.. Besides Knauf et al., Woods et al. 2012[49] also used utility increments and decrements from Beusterien et al. 2010,[50] which is a vignette study in 89 members of the general population. In this study, the standard gamble technique was used to elicit utility values associated with different health states observed during CLL treatment. It was unclear to the ERG why CLL patients were the most appropriate proxy to obtain utility values for the current decision problem. Finally, Breitscheidel, 2008[52], Morton et al. 2009[53], and Cooperberg et al. 2013[54] are cost effectiveness analyses in chronic myeloid leukaemia, kidney transplant recipients, and prostate cancer patients, respectively.

(c) Although the ERG agrees that the utility values identified in the literature are not completely reflecting the population considered in the current decision problem, it is highly questionable whether the adjustments applied by the company improves the validity of the utility values. Firstly, the necessity of an age-adjustment is questionable since the impact of the disease on quality of life would outweigh the impact of age-related utility decrements. Secondly, the ERG thinks that the adjustment for the utility representing perfect health should not be applied because no evidence (nor justification) was provided to support the methodology used by the company. The company states that this adjustment ensures that the health state utility values would not be higher than the general population utility values. However, this is only true for the start of the model, i.e. not over time, which is a severe limitation given the life expectancy of the modelled patients. Thirdly, these adjustments were not applied consistently on all

health states (and the rational for doing so is missing); hence the ERG decided not to use the company's adjustment in its base-case analysis (health state utilities in column 'Original utility value' from Table 5.9 are used in the ERG base-case). In order to prevent that health state utility values exceed the general population utility values (over time), the ERG decided to cap the health state utility values in the model using the general population utility values (see scenario requested in clarification question B11).[60] Additionally, the ERG decided to use the same utility value for the first and second line treatment induction and consolidation health states (i.e. 0.70). This was adopted to ensure consistency in utility values between these health states because the company did not provide evidence to support the need for differential utility values.

5.2.9 Resources and costs

The cost categories included in the model were treatment acquisition costs, medical costs (treatment administration, supportive care, monitoring and follow-up, HSCT, palliative care), and costs of managing adverse events.

Treatment acquisition costs

The company states that publications of relevant trials were used to extract data on dosage and number of doses of intervention and comparator, and that these were validated by clinical experts. Drug costs were based on the British National Formulary (BNF) [61]. If available, container sizes minimising the costs and wastage were used to reflect real-life practice. Input values and their sources for each drug and treatment phase are presented in Tables 5.11 and 5.12.

Table 5.11: Unit treatment acquisition costs associated with the technologies studied in the economic model

		INDUCTION PHASE			CONSOLIDATION PHASE		
Model parameter	**Strategy**	**Drug**	**Value**	**Reference**	**Drug**	**Value**	**Reference**
Number of doses	AATO First line	ATRA	32	Lo-Coco et al., 2013	ATRA	15	Lo-Coco et al., 2013[12]
		ATO	32		ATO	20	
	AATO Second line	ATRA	25	Douer et al. 2005	ATRA	15	Lo-Coco et al., 2013[12]
		ATO	25		ATO	25	SPC
	AIDA	ATRA	35	Lo-Coco et al., 2013	ATRA	15	Lo-Coco et al., 2013[12]
					IDA(cycle 1)	4	
					Mitoxantrone (cycle 2)	5	
		IDA	4		IDA(cycle 3)	1	
Indicated dose per day	AATO	ATRA	45 mg/m^2	Lo-Coco et al., 2013	ATRA	45 mg/m^2	Lo-Coco et al., 2013[12]
		ATO	0.15 mg/kg		ATO	0.15 mg/kg	
	AIDA	ATRA	45 mg/m^2		ATRA	45 mg/m^2	
					IDA(cycle 1)	5 mg/m^2	

		INDUCTION PHASE			CONSOLIDATION PHASE		
Model parameter	**Strategy**	**Drug**	**Value**	**Reference**	**Drug**	**Value**	**Reference**
		IDA	12 mg/m^2		Mitoxantrone (cycle 2)	10 mg/m^2	
					IDA(cycle 3)	12 mg/m^2	
Container size	AATO	ATRA	10 mg	BNF	-	-	-
		ATO	10 mg				
	AIDA	ATRA	10 mg		Mitoxantrone	20 mg	BNF[61]
		IDA	10 mg				
Cost per container	AATO	ATRA	£1.61	BNF	-	-	-
		ATO	£292.00				
	AIDA	ATRA	£1.61		Mitoxantrone	£100.00	BNF[61]
		IDA	£174.72				

Table 5.12: Costs of technology per treatment phase

Phase	**AATO**			**AIDA**		
	ATRA	**ATO**	**Total AATO**	**ATRA**	**Chemo (IDA+MTZ)**	**Total AIDA**
First line: induction	£463.68	£16,078.58	£16,542.26	£507.15	£2,096.64	£2,603.79
First line: Consolidation	£1,521.45	£40,196.44	£41,717.89	£652.05	£1,723.04	£2,375.09
First line: Total	£1,985.13	£56,275.02	£58,260.15	£1,159.20	£3,819.68	£4,978.88
Second line: Induction	£362.25	£12,561.39	£12,923.64	£507.15	£2,096.64	£2,603.79
Second line: Consolidation	£1,521.45	£12,561.39	£14,082.84	£652.05	£1,723.04	£2,375.09
Second line: Total	£1,883.70	£25,122.77	£27,006.47	£1,159.20	£3,819.68	£4,978.88

Medical costs

Medical costs (treatment administration, supportive care, monitoring and follow-up, HSCT, palliative care), and resource use are presented in Table 5.13 below. Medical costs were mainly based on NHS reference prices, BNF and PSSRU.

Table 5.13: list of resource use per health state

Health state	**Items**	**AATO**	**Value**	**AIDA**	**References**
Induction phase	Number of bed days per patient	First line: 32 Second line: 25	£396.47	35	AATO: First line: Lo-Coco et al., 2013[12] Second line: Douer et al., 2005[39]

Health state	Items	AATO	Value	AIDA	References
					AIDA: Lo-Coco et al., 2013[12]
	Number of supportive care transfusions	15	0	22	Burnett et al., 2015[13]
	Number of annual PCR tests	5	£280	4	Expert opinion
Consolidation phase	Number of bed days per patient	0	£396.47	4	AATO: Expert opinion AIDA: assumption based on treatment schedule
	Number of ambulatory days per patient	First line: 10 Second line: 12.5	£162.00	0	AATO: Expert opinion AIDA: Inpatient treatment assumed
	Number of days of antibiotics	1	£1.65	2	Burnett et al., 2015[13]
	Number of annual PCR tests	5	£280	4	Expert opinion
Molecular remission (first, second, allo- and auto-HSCT)	Duration of follow-up	3	£210	3	First remission: Platzbecker et al., 2015[14] Others: Expert opinion
	Number of annual appointments	4	£52.50	4	First remission: Platzbecker et al., 2015[14] Others: Expert opinion
	Number of annual PCR tests	4 (0 at first remission)	£280	4	First remission: AATO: Expert opinion AIDA: Platzbecker et al., 2015[14] Others: Expert opinion
Allogeneic HSCT	Hospitalisation duration (weeks)	4	£27,907.53	4	Expert opinion
Autologous HSCT	Hospitalisation duration (weeks)	3	£7,122.97	3	Expert opinion

Costs of managing adverse events

Cost per occurrence for each type of adverse event was searched in the National Schedule of Reference Costs, 2014-2015.[62] If unavailable, recent publications reporting English or UK costs were used. In case only foreign costs could be used, these were converted to sterling using the annual exchange rates of the year the cost related to and uplifted to 2015 using inflation rates of the Office for National Statistics (Table 5.14).

Table 5.14: List of adverse reactions and summary of costs in the economic model

Adverse reactions	Value	Reference
Thrombocytopenia (grade 3-4, >15 days)	£1,746.00	NHS Reference Costs 2014–15[62]
Neutropenia (grade 3-4, >15 days)	£2,845.43	Morgan et al., 2007[63]
Infection	£253.97	Soini et al., 2016[64]
Leukocytosis	£349.44	Expert opinion
Hepatic toxicity	£5.56	Akhtar and Chung, 2014[65]
Neurotoxicity	£675.88	Calhoun et al., 2001[66]
Differentiation syndrome	£1,225.23	Milligan et al., 2006[67]; BNF[61]; National Schedule of Reference Costs[62]
Cardiac events	£1,104.02	National Schedule of Reference Costs[62]
QTc prolongation	£34.50	Expert opinion; NICE clinical guideline, No. 108[68]; NICE clinical guidelines, No. 174[69]
tMDS/AML	£6,207.00	National Schedule of Reference Costs[62]
Acute GvHD	£34,493.05	Saito et al., 2008[70]
Chronic GvHD	£8,785.25	Jones et al., 2016[71]

ERG comment: The ERG comments are in relation to (a) justification of sources for cost and resource items and (b) absence of costs related to haematological response monitoring.

(a) The ERG asked the company to provide more specific justification for each resource use and cost item. The company responded that they aimed to use NHS reference costs and the PSSRU wherever possible, supplementing this with data from studies identified through a targeted search where necessary. However, the company could not provide further justification and details about the included targeted sources, and the ERG was therefore unable to assess whether these sources were best available evidence to inform resource use and costs estimates.

(b) Costs related to monitoring of haematological response were not included in the model. The company stated that this was not mentioned by experts, most likely because the benefits of treating a molecular relapse before progression into a haematological relapse are widely recognised, so the monitoring for relapse focuses more on molecular (PCR) testing. Furthermore, the company considered the cost of haematological monitoring negligible compared to the costs of PCR testing. The ERG agrees that in first line treatment monitoring costs would be equivalent in both strategies. However, given the fact that AIDA patients relapse more frequently in second line, less monitoring would be needed in this strategy, and therefore monitoring costs would be higher in the AATO strategy.

5.2.10 Cost effectiveness results

In the deterministic base-case analysis, total QALYs and LYs gained were larger in the AATO strategy compared to the AIDA strategy. This was mostly explained by the 10.72% increase of patients in first molecular remission and the absence of tMDS/AML in the AATO arm. Furthermore, the number of APL-related deaths was reduced by around 31.85% with AATO compared to AIDA. Total costs were lower for AATO, thus the combination of AATO was dominant and no base-case ICER was calculated. Most important cost driver for AATO was treatment acquisition costs, but costs in all other categories were higher for AIDA. Transplantation costs were the highest costs for AIDA. Base-case health outcomes, discounted costs and the incremental cost-effectiveness ratio are shown in Tables 5.15-5.17.

Table 5.15: Discounted health outcomes in the model

	AATO	AIDA	AATO vs. AIDA
Number of QALYs	16.34	13.72	2.62
Number of LYs	19.56	16.56	3.00
First remission*	99.83%	89.11%	10.72%
First long remission (> 2 years)*	92.84%	76.11%	16.73%
MDS*	0.00%	1.39%	-1.39%
Death* (not discounted)	57.04%	74.13%	-17.09%
APL related death* (not discounted)	7.54%	39.38%	-31.85%
Background death* (not discounted)	49.51%	34.75%	14.76%

AATO= arsenic trioxide + all-trans retinoic acid; LY=life years; MDS=myelodysplastic syndrome; QALY=quality-adjusted life years
*proportion of patients ever, transiting to this health state

Table 5.16: Discounted disaggregated and total costs – base-case scenario

Cost category	AATO	ATRA+AIDA	AATO vs. AIDA
Treatments	£60,336	£21,604	£38,731
Administration	£25,402	£31,660	-£6,259
Supportive care and antibiotics	£3,575	£6,487	-£2,912
Follow-up and monitoring	£2,991	£10,389	-£7,398
Adverse Events	£4,142	£12,378	-£8,236
MDS	£0	£226	-£226
HSCT	£7,645	£48,326	-£40,681
Palliative care	£906	£5,196	-£4,290
Total	£104,996	£136,267	-£31,270

AATO= arsenic trioxide + all-trans retinoic acid; HSCT=haematopoietic stem cell transplantation; MDS=myelodysplastic syndrome

Table 5.17: Base-case incremental results (discounted)

Treatment	Total costs (£)	Total LYG	Total QALYs	Incremental costs (£)	Incremental LYG	Incremental QALYs	ICER versus baseline (£/QALY)	ICER incremental (£/QALY)
AATO	104,996	19.56	16.34	-£31,270	3.00	2.62	Dominant	Dominant
AIDA	136,267	16.56	13.72				-	-

ERG comment: The undiscounted LYs and QALYs provided by the company were 33.22 and 27.91 for AATO and 26.84 and 22.38 for AIDA, respectively. The ERG perceives the life expectancy in the model to be relatively long. This might be related to a lack of disease-specific mortality in the first and second line health states, as well as assumptions concerning (extrapolation of) treatment benefits. More details can be found in section 5.2.12 about model validation and face validity.

5.2.11 Sensitivity analyses

The company performed and presented a probabilistic sensitivity analysis (PSA) and deterministic sensitivity analyses (DSA) in order to quantify the uncertainty surrounding the base case results.

The PSA with 1,500 Monte Carlo simulations showed similar incremental costs and QALYs compared with the deterministic results, the AATO strategy still being dominant (Table 5.18). The cost effectiveness acceptability curve showed that the probability of AATO being cost effective at a willingness-to-pay (WTP) of £0 per QALY was 81%. This probability increased to 94% at a WTP of £30,000 per QALY (Figure 5.2).

The company conducted a one-way DSA to study the impact of varying individual parameter values on incremental costs, incremental QALYs and ICER of AATO compared with AIDA. Parameters that most affected incremental costs were the probability of relapse at 48 months after first remission in the AIDA arm, discount rate for costs, time horizon, complete haematological remission rate following AIDA in first line, and the probability of relapse at 24 months after first remission in the AIDA strategy. Incremental effectiveness was mostly affected by changes in the discount rate for health outcomes, time horizon, probability of relapse observed at 48 months for AIDA, first line haematological remission rate associated with AIDA treatment and the utility value in the first molecular remission (2> years) health state (Table 5.19). The ICER could only be computed in four sensitivity analyses, in all other analyses AATO dominated AIDA. The ICER based on a time-horizon of five years was 148,179 £/QALY. This can mostly be explained by the high treatment acquisition costs in the first year and the inability to capture the full HRQoL benefits within a time horizon this short.

The following scenario analyses were performed by the company (Table 5.20):

- Scenario 1: AIDA used in second line following both first line treatments
- Scenario 2: Utilities from Tallman et al.[31]
- Scenario 3: Societal perspective
- Scenario 4: AML17 protocol: a scenario using the schedule, dosage, efficacy and safety inputs based on the AML17 clinical study.
- Scenario 5: “Worst-case” scenario: a scenario accumulating unfavourable inputs for the AATO strategy.
- Scenario 6: Probability of undergoing HSCT reflecting clinical practice, with a lower proportion of patients undergoing autologous HSCT and allogeneic HSCT reserved for patients who did not achieve molecular remission after second line induction.

The ICER was dominant across all scenarios. Scenario 4 had the largest impact on both incremental costs (£66,384) and QALYs (3.39).

Table 5.18: Probabilistic sensitivity analysis results

	Incremental costs (£)	Incremental QALYs
Mean	-31,088	2.546
Median	-28,654	2.435
Min	-169,499	-8.570
Q 0.025	-110,732	-1.746
Q 0.975	32,992	6.771
Max	109,569	14.167

Figure 5.2: Cost effectiveness acceptability curve

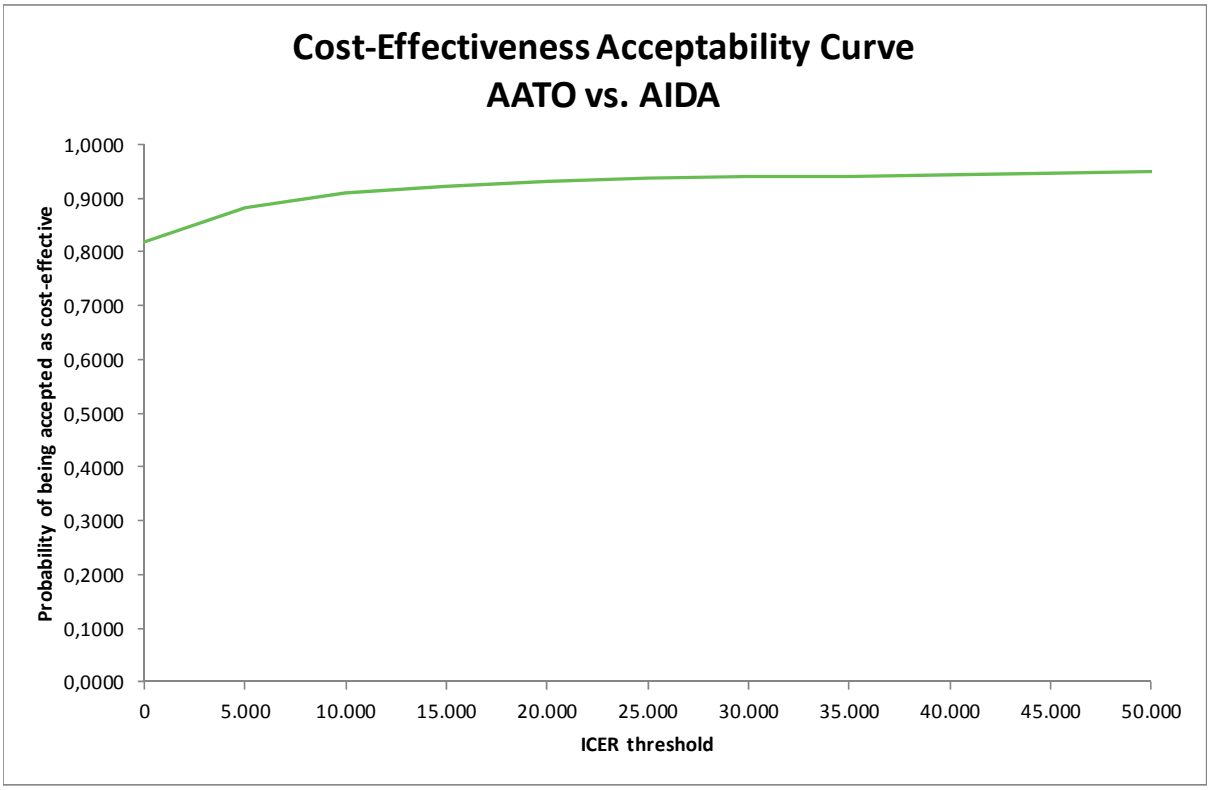


Table 5.19: Results of the DSA – incremental costs and QALYs

Parameters	**Incremental Costs (£)**		
	Low case	**High case**	**Distance***
Relapse following remission (48 months) - AIDA	25,701	-66,546	92,246
Discount rate for costs	-79,401	-19,602	59,799
Time Horizon	22,128	-25,208	59,461
CHR rate - First line - AIDA	-85,568	-27,157	58,411
Relapse following remission (24 months) - AIDAAIDA	-51,405	-510	50,896
Parameters	**Incremental QALYs**		
	Low case	**High case**	**Distance***
Discount rate for health outcomes	5.52	1.98	3.54
Time Horizon	0.15	2.07	3.02
Relapse following remission (48 months) - AIDA	1.00	3.78	2.78
CHR rate - First line - AIDA	4.84	2.45	2.39

Utilities - First line - Molecular remission (>2 years)	1.27	3.58	2.31
* Distance is ABS(Low case – Base case)+ABS(High case – Base case)			

Table 5.20: Results of scenario analyses – incremental costs, effectiveness and ratio

	AATO vs. AIDA					
	Incremental costs		**Incremental effectiveness**		**Incremental ratio**	
Scenario's	**Not discounted**	**Discounted**	**Not discounted**	**Discounted**	**Not discounted**	**Discounted**
Base-case	-	-£31,270	-	2.62	-	Dominant
Scenario 1	-£65,974	-£21,593	5.70	2.72	Dominant	Dominant
Scenario 2	-	-	6.03	2.93	-	-
Scenario 3	-	-£32,833	-	-	-	-
Scenario 4	-£125,336	-£66,384	7.10	3.39	Dominant	Dominant
Scenario 5	-£53,471	-£9,986	3.52	1.58	Dominant	Dominant
Scenario 6	-£76,110	-£28,664	5.20	2.43	Dominant	Dominant

ERG comment: The ERG concerns are related to a) the inclusion of patient characteristics and b) the approach to incorporate resource use in the PSA.

a) The ERG had minor concerns regarding the inclusion of patient characteristics (percentage male, age of patients, average height) in the PSA. Given that these parameters reflect first order uncertainty which should not be incorporated in the PSA. This is corrected in the ERG base-case.

b) The company's approach to incorporate resource use in the PSA using a normal distribution, generating new random numbers in case the resource use was negative (see response to clarification question B19b for more details), is flawed as removing these negative numbers (i.e. lower part of the distribution) will artificially increase the average of the distribution. However, given that using a Gamma distribution for resource use does not substantially influence the outcomes (see response to clarification question B19b), the ERG did not alter this in its base-case.

5.2.12 Model validation and face validity check

Internal validation

Internal validation was performed to identify programming errors, data entry issues and logical inconsistencies in the model. For this purpose, a variety of extensive tests were performed considering the following aspects of the model: efficacy and safety of compared strategies, treatment schedules, treatment costs, resource use and mortality in the modelled population. See CS Table 3.27 for a summary of the internal validation. In addition, data, calculations and formulae were verified by a person not involved in the initial project.

This part of the validation included quality control conducted following the methodology proposed by the York Health Economics Consortium (YHEC). A summary of evidence on the internal validity of the model is reported in CS Table 3.27.

External validation

An external validation was conducted, comparing the outcomes from the model to those observed in clinical trials at different time points (24 and 50 months):

- DFS in terms of proportion of patients in first remission health states (molecular remission and +2y remission).
- DFS in terms of proportion of patients in all remission health states (molecular remission, +2y remission, second line molecular remission and HSCT remission)
- OS estimated as the proportion of patients alive at a given time point in the model

None of the "absolute differences" between the trial and economic model exceeded 10 percentage points (CS Table 3.28). In general, the model overestimates outcomes (i.e. produces higher DFS and OS than observed at 24 and 50 months in the trial), with the exceptions of the AIDA first line treatment arm, in which DFS, in terms of proportion of patients in first line remission is underestimated (absolute difference ranged from -7.39% to -4.19%) and for the AATO, in which 50 months OS is underestimated (absolute difference: -0.11%). Similarly, the relative difference between AATO and AIDA is generally underestimated in the model (compared with the trial) except for DFS in terms of proportion of patients in first line remission. The company states that this might be due to the assumption in the model that patients experiencing cardiac events or failing to reach molecular remission after first line consolidation switch to second line.

ERG comment: The ERG's concerns relate to a) the lack of detailed descriptions and justification for calculations, assumptions and selected sources; b) the long life expectancy in the model; c) overestimation of proportion of patients in first line remission (illustrated in external validation); d) the lack of cross-validation and; e) the inability to perform probabilistic analyses (without errors) using the model file received in response to the clarification questions (named "ID446 arsenic trioxide TEVA CEM_v4.2_rem26 v0.1 170118 SC [noACIC].xlsm").

a) As mentioned in the preceding sections, the CS lacked transparency and appropriate justifications. This included the lack of detailed descriptions and justification for calculations, assumptions and selected sources. This also includes the lack of elaborate consideration in the CS regarding assumptions related to the extrapolation (e.g. extrapolation of treatment benefit). Currently, treatment benefits, in terms of different transition probabilities for AATO and AIDA, are maintained for the entire time horizon. Alternative assumptions regarding the extrapolation could be influential as illustrated by the analyses performed by the company in response to clarification question B7. In this scenario, assuming no relapse after 24 months in the first line remission health state, AATO did not remain dominant as it became more expensive than AIDA with an ICER of £7,610 per QALY gained. Moreover, the external validation efforts, reported in CS Table 3.28, do not consider long-term outcomes beyond 50 months. Neither did the validation section in the CS (section B3.10) include specific comments regarding the face validity of the long-term extrapolation. Hence, the long-term validity of the outcomes should be regarded as a major and potentially influential uncertainty. Acknowledging this uncertainty, the ERG added a scenario analyses assuming an equal relapse probability two year after first line remission for AATO and AIDA.

b) Related to the long-term extrapolation, the ERG perceives the life expectancy estimated in the model to be relatively long. This is likely linked to the lack of disease-related mortality in the model during the first line and second line health states (only general population mortality is considered) as well as assumptions concerning (extrapolation of) treatment benefits. The undiscounted LYs and QALYs for AATO, estimated in the model, are 33.22 and 27.91 respectively. When extending the model time horizon (to 56 years, to represent a life time horizon, which is consistent with the NICE reference case),

this would increase to 35.83 and 30.12 respectively. In the general UK population, the LY and QALYs estimated for patients aged 45 (with 48.7% being male) are 37.62 and 29.62 respectively. Hence, the outcomes estimated by the model are ~2 LYs below and ~0.5 QALYs above those for the general population.[72, 73] The latter (i.e. higher QALYs than for the general population) is likely the result of the use of utility values that exceed those for the general population over time (this is corrected in the ERG base-case). The ERG is uncertain whether these LY and QALYs, as calculated in the model, have face validity.

c) The external validation showed that the model overestimates the relative difference between AATO and AIDA (compared with the trial) for DFS in terms of proportion of patients in first line remission. This is likely related to the overestimation of the proportion of patients having cardiac events (and thus requiring a treatment switch) after AIDA treatment. This overestimation of cardiac events (illustrated in Table 2.2 of CS Appendix J) and the ERG's approach to correct this is discussed in section 5.2.6 of the ERG report.

d) As stated in section 5.2.2 of the ERG report, the company adopted a model structure that diverges from those used in other economic evaluations in this condition. Additionally, the other CEAs identified in the company's SLR resulted in positive incremental costs, while in the CS base-case AATO was cost saving. Unfortunately, the company could not perform a cross-validation to explore the exact sources for the differences in the outcomes. In response to clarification question B18 the company stated: "Since we have developed a more comprehensive approach [than previously published models], it is not possible to perform an exact cross-validation of the assumptions, inputs and outputs." The ERG believes this is reasonable.

e) The ERG was unable to perform probabilistic analyses (without errors) using the model file received in response to the clarification questions (named "ID446 arsenic trioxide TEVA CEM_v4.2_rem26 v0.1 170118 SC [noACIC].xlsm"). The ERG made its adjustments using this model file given this version of the model incorporated structural adjustments that the ERG preferred to use in its base-case. Unfortunately, given the complex implementation of the PSA in the company's model, the ERG was not able to correct the cause of this error. However, the deterministic and probabilistic results, produced by the model initially submitted by the company, are relatively similar and thus the ERG will rely on deterministic results.

5.3 Exploratory and sensitivity analyses undertaken by the ERG

Based on all considerations from Section 5.2, the ERG defined a new base-case. This base-case included multiple adjustments to the original base-case presented in the previous sections. These adjustments made by the ERG form the ERG base-case and were subdivided into three categories (derived from Kaltenthaler 2016[74]):

- Fixing errors (correcting the model where the company's submitted model was unequivocally wrong)
- Fixing violations (correcting the model where the ERG considered that the NICE reference case, scope or best practice had not been adhered to)
- Matters of judgement (amending the model where the ERG considers that reasonable alternative assumptions are preferred)

Additionally, exploratory scenario analyses were performed by the ERG to examine the potential impact of alternative assumptions on the cost effectiveness estimates (section 5.3.2). Moreover, a subgroup analysis was performed to reflect the second line population, i.e. refractory/relapsed APL (section 5.3.3).

Fixing errors

1. Number of tunnel states for the "first line molecular remission" health state (section 5.2.2).
   The ERG increased the number of tunnel states to 26 model cycles to reflect two years.
2. Overestimation of proportion of patients switching to second line induction due to experiencing a cardiac event (section 5.2.6).
   The ERG corrected the probability related to cardiac events.
3. Calculation error related to proportion of patients transiting to first line consolidation early for AIDA (section 5.2.6).
   The ERG corrected this calculation error.
4. Assumptions and calculation concerning the relapse probabilities (section 5.2.6).
   The ERG corrected errors related to the assumptions and calculation (i.e. incorrectly using unconditional probabilities as conditional probabilities as well as the lack of time correction for the 48 month relapse probability).
5. Reference error related to the transition probability from second line induction to consolidation for AIDA (section 5.2.6).
   The ERG corrected this reference error.
6. Calculation error related to transition probability from second line AIDA consolidation to "tMDS/AML" (section 5.2.6).
   The ERG corrected this calculation error.

Fixing violations

7. Time horizon not reflecting lifetime (section 5.2.5).
   The ERG extended the time horizon to 56 years to reflect a lifetime time horizon.
8. Utility adjustments (section 5.2.8).
   The ERG removed the utility adjustments made by the company and assumed a different utility for the second line consolidation phases (consistent utility as used for the other induction and consolidation phases).
9. Utility values higher than the general population utility values over time (section 5.2.8).
   The ERG capped the utility values to ensure that these would not exceed the general population utility values over time.
10. Inappropriate parameters in PSA: patient characteristics were included in the PSA (section 5.2.11).
    The ERG removed patient characteristics from the PSA.

Matters of judgment

11. Calculations and assumptions regarding the remission probability (sections 5.2.2 and 5.2.6).
    The ERG informed the remission probability based on APL0406 data and used the molecular remission rate to inform the probability of transiting to remission for patients that are evaluable with PCR (removing the additional multiplication with the haematological response rate).

Table 6.1 shows how individual adjustments impact the results plus the combined effect of all abovementioned adjustments simultaneously, resulting in the (deterministic) ERG base-case. The 'fixing error' adjustments were combined and the other ERG analyses were performed also incorporating these 'fixing error' adjustments given the ERG considered that the 'fixing error' adjustments corrected unequivocally wrong issues.

5.3.1 Deterministic ERG base-case

In the ERG base-case, incorporating all abovementioned adjustments, AATO resulted in costs savings of £23,502 and yielded 2.254 more QALYs than AIDA and hence remained dominant (see Table 6.1).

As highlighted in section 5.2.12, the ERG was unable to perform probabilistic analyses. However, the deterministic and probabilistic results, produced by the model initially submitted by the company, are relatively similar. Hence, AATO is likely to remain dominant if the ERG would be able to produce probabilistic results for its base-case.

5.3.2 Deterministic scenario analyses performed conditional on the ERG base-case

Deterministic scenario analyses were performed, conditional on the ERG base-case, to examine the potential impact of the following alternative assumptions on the cost effectiveness estimates:

12. Adding disease-related mortality, in addition to general population mortality, during the induction phases (both first and second line) using the 60 day mortality from the AML17 trial (sections 5.2.2 and 5.2.12)
13. Assuming an equal relapse probability two years after first line remission for AATO and AIDA (sections 5.2.6 and 5.2.12)
14. Replacing transitions to alloHSCT for transitions to autoHSCT (section 5.2.2)
15. Removing the transitions to the HSCT states from second line remission (section 5.2.6)
16. Assuming 'uncorrected' transitions to HSCT states from second line remission (section 5.2.6)
17. Incorporating 2% cardiac events for AATO during the induction phase, reflecting treatment switching due to potential arrhythmia (sections 5.2.2 and 5.2.6)

AATO remained dominant in these deterministic scenario analyses except for the exploratory scenario wherein the relapse probability was assumed to be equal for AATO and AIDA two years after first line remission. This scenario acknowledges uncertainty in the extrapolation of treatment benefits and hence indicates that this might be influential given this scenario resulted in an ICER of £19,734 per QALY gained (see Table 6.2 for more detailed results of the scenario analyses performed by the ERG).

Additionally, the ERG performed a worst-case scenario, implementing all scenarios analyses listed above simultaneously (except for analysis 16). This deterministic worst-case scenario resulted in an ICER of £21,622 per QALY gained.

5.3.3 Deterministic subgroup analyses performed conditional on the ERG base-case

The ERG performed a deterministic subgroup analysis, conditional on the ERG base-case, reflecting the second line population, i.e. refractory/relapsed APL (section 5.2.3). This was implemented by removing the first line health states. This analysis indicated that for this subgroup AATO would cost £18,207 more and gain 0.584 more QALYs compared with AIDA, resulting in an ICER of £31,184 per QALY gained.

5.4 Conclusions of the cost effectiveness section

The cost effectiveness searches in the company submission were all documented and reproducible. However, there were a number of inconsistencies and inaccuracies, and some redundancy. The MEDLINE and Embase search strategies used an inappropriate 'animals' limit, and it is possible that relevant evidence may have been missed as a consequence.

Although the SLR identified CEAs in the literature, the company decided to develop a de novo model. The model structure proposed by the company however diverges from those identified in the SLR. The company justified the more complex model structure by stating that the existing economic evaluations did "not adequately reflect the trajectory of APL patients". According to the company, the aim in this economic evaluation was to "offer more granularity with treatment phases, molecular remission and HSCT" to better reflect the clinical trajectory of APL patients. The ERG considers that the model

structure is appropriate to reflect this condition and treatment pathway. The economic model described in the CS is considered by the ERG to partly meet the NICE reference case. Deviations from the NICE reference case included that the population and comparators considered in the scope were not fully considered. Moreover, the HRQoL used as well as the time horizon adopted by the company deviated from the NICE reference case. The transition probabilities from the first line phase of the model were informed by the APL0406 trial. The evidence to inform transitions from second line health states was weak and it was frequently not transparently reported how the transition probabilities were obtained. Similarly, most of the evidence sources to inform transition probabilities from the remaining HSCT health states are not described in the CS (neither are the transition probabilities). This includes assumptions regarding the extrapolation of treatment effectiveness which is not extensively discussed in the CS. The lack of detailed description and justification is worrying, given treatment effectiveness (including implicit assumptions made and selection of evidence sources to obtain transition probabilities) is often an influential part of cost effectiveness models.

In the company base-case (probabilistic) AATO is less expensive (£31,088 saved) and more effective (2.546 QALYs gained) than AIDA and thus the dominating strategy for newly diagnosed low-to-intermediate risk APL (i.e. the first line population). AATO remained dominant in most of the sensitivity and scenario analyses conducted by the company. The ERG has incorporated various adjustments to the company base-case this resulted in the (deterministic) ERG base-case wherein AATO remained dominant. Moreover, the ERG produced a worst-case scenario (combination of some of the scenario analyses explored by the ERG), to acknowledge the uncertainties discussed in section 5.2 of this report. This resulted in an ICER of £21,622 per QALY gained (deterministic). The ERG was unable to perform probabilistic analysis for its base-case. However, the ERG does not consider this to be a major issue as AATO is likely to remain dominant if the ERG would be able to produce probabilistic results for its base-case.

In conclusion, despite the ERG's criticism of the economic model and several highlighted uncertainties, it is reassuring that AATO for the first line population remained dominant in the ERG base-case, and that the worst-case scenario produced by the ERG resulted in an ICER of £21,622. However, as indicated by the subgroup analysis performed by the ERG, the cost effectiveness of AATO for the second line might be substantially different (estimated ICER of £31,184 per QALY gained).

6. IMPACT ON THE ICER OF ADDITIONAL CLINICAL AND ECONOMIC ANALYSES UNDERTAKEN BY THE ERG

In Section 5.3 the ERG base-case was presented, which was based on various changes compared to the company base-case. Table 6.1 shows how individual changes impact the results plus the combined effect of all changes simultaneously. The exploratory scenario analyses are presented in Table 6.2. These are all conditional on the ERG base-case. The analyses numbers in Tables 6.1 and 6.2 correspond to the analyses numbers reported in Section 5.3. Finally, Table 6.3 provides the results of the subgroup analysis (described in Section 5.3.3). The submitted model file contains technical details on the analyses performed by the ERG (e.g. the "ERG" sheet provides an overview of the cells that were altered for each adjustment).

Table 6.1: Deterministic ERG base-case

CS base-case					
Technologies	**Total costs**	**Total QALYs**	**Incremental costs**	**Incremental QALYs**	**ICER (£/QALY)**
AATO	£104,996	16.336			
AIDA	£136,267	13.717	-£31,270	2.618	Dominance
Fixing errors (1-6)					
AATO	£105,847	16.287			
AIDA	£131,760	13.859	-£25,914	2.428	Dominance
Extend time horizon (1-6, 7)[a]					
AATO	£106,722	16.777			
AIDA	£134,262	14.149	-£27,540	2.629	Dominance
Alternative utility values (1-6, 8)[a]					
AATO	£105,847	16.527			
AIDA	£131,760	14.116	-£25,914	2.411	Dominance
Capping utility values (1-6, 9)[a]					
AATO	£105,847	15.598			
AIDA	£131,760	13.338	-£25,914	2.260	Dominance
Remove inappropriate parameters from PSA (1-6, 10)[a]					
AATO	£105,847	16.287			
AIDA	£131,760	13.859	-£25,914	2.428	Dominance
Alternative remission probabilities (1-6, 11)[a]					
AATO	£106,055	16.280			
AIDA	£127,908	14.015	-£21,853	2.265	Dominance
ERG base-case (1-11)					
AATO	£106,931	16.135			
AIDA	£130,432	13.881	-£23,502	2.254	Dominance
[a]Analyses performed conditional on the fixing error analysis.					

Table 6.2: Deterministic scenario analyses conditional on ERG base-case

ERG base-case (1-11)					
Technologies	Total costs	Total QALYs	Incremental costs	Incremental QALYs	ICER (£/QALY)
AATO	£106,931	16.135			
AIDA	£130,432	13.881	-£23,502	2.254	Dominance
Disease-related mortality during the induction phase (1-11, 12)					
AATO	£103,532	15.530			
AIDA	£120,599	12.848	-£17,066	2.682	Dominance
Relapse probability equal for all treatments two year after first line remission (1-11, 13)					
AATO	£106,931	16.135			
AIDA	£86,524	15.100	£20,407	1.034	£19,734
Transitions to alloHSCT replaced for transitions to autoHSCT (1-11, 14)					
AATO	£103,523	16.283			
AIDA	£113,388	14.659	-£9,865	1.624	Dominance
Transitions from second line remission to HSCT states removed (1-11, 15)					
AATO	£107,200	16.129			
AIDA	£132,049	13.849	-£24,848	2.281	Dominance
Transitions from second line remission to HSCT states ‘uncorrected (1-11, 16)					
AATO	£106,773	16.137			
AIDA	£129,496	13.895	-£22,723	2.242	Dominance
Cardiac events added for AATO to reflecting treatment switching due to (potential) arrhytmia (1-11, 17)					
AATO	£107,285	16.121			
AIDA	£130,891	13.836	-£23,606	2.285	Dominance
Worst-case scenario (1-15, 17)					
AATO	£100,561	15.662			
AIDA	£73,494	14.410	£27,067	1.252	£21,622

Table 6.3: Deterministic subgroup analysis reflecting the second line population

Second line population (conditional on ERG base-case)					
Technologies	**Total costs**	**Total QALYs**	**Incremental costs**	**Incremental QALYs**	**ICER (£/QALY)**
AATO	£209,365	9.204			
AIDA	£191,158	8.620	£18,207	0.584	£31,184

7. OVERALL CONCLUSIONS

7.1 Statement of principal findings

The company presented evidence from three RCTs: Two of these were trials in newly diagnosed APL (APL0406 and AML17) and the third was a study in patients with relapsed APL (Raffoux, et al. 2003).

Untreated APL

Both trials in newly diagnosed APL (APL0406 and AML17) compared AATO (all-trans retinoic acid (ATRA) + ATO) with AIDA (ATRA + idarubicin). The dosing and regimens for AATO in AML17 were not in accordance with the licence. As the dosing and regimens for AATO in APL0406 were in accordance with the licence the ERG focused on this trial. APL0406 included 266 patients with newly-diagnosed, low- to intermediate-risk APL aged 18 to 71 years and took place in Italy and Germany.

Results from APL0406 showed that AATO significantly improved overall survival (OS) at 50 months compared with AIDA (99.2% vs 92.6% respectively, p=0.007). The primary endpoint of this trial was event-free survival (EFS) at two years in the initial cohort of 156 patients (97% with AATO vs 86% with AIDA, p<0.001 for non-inferiority, p=0.02 for superiority). EFS was significantly better in the AATO group across all subsequent analyses to reach 97.3% at 50 months in the full cohort of 266 patients, compared with 80.0% in the AIDA group (p<0.001). At 50 months, the cumulative incidence of relapse was 1.9% in the AATO group compared with 13.9% in the AIDA group (p=0.0013). In the AATO group patients experienced fewer haematological adverse events including fever and infection episodes and grade 3 to 4 neutropenia and thrombocytopenia lasting over 15 days. AATO was also more favourable than AIDA for grade 3-4 gastrointestinal toxicity. Other adverse events were more common with AATO mainly in the induction phase of treatment. In the AATO group incidence of leukocytosis was 43%. A greater number of patients experienced QTc prolongation with AATO (11% vs 0.7%). In addition, a greater number of patients experienced grade 3 to 4 hepatic toxicity, (40% vs. 3%). There were no instances of neurotoxicity with AIDA but 19 instances were noted with AATO.

Relapsed or refractory APL

The only trial presented for relapsed/refractory patients was by Raffoux et al. (2003). This small trial compared AATO with ATO, which is not a relevant comparison according to the scope. OS was similar between the AATO and ATO study arms. Across both groups, the estimated two-year OS was 59% (95% CI: 35%–77%). EFS was not reported in this study.

Economic evaluation

In the company base-case (probabilistic) AATO is less expensive (£31,088 saved) and more effective (2.546 QALYs gained) than AIDA and thus the dominating strategy for newly diagnosed low-to-intermediate risk APL (i.e. the first line population). AATO remained dominant in most of the sensitivity and scenario analyses conducted by the company. The ERG has incorporated various adjustments to the company base-case this resulted in the (deterministic) ERG base-case wherein AATO remained dominant. Moreover, the ERG produced a worst-case scenario (combination of some of the scenario analyses explored by the ERG), to acknowledge the uncertainties discussed in section 5.2 of this report. This resulted in an ICER of £21,622 per QALY gained (deterministic). The ERG was unable to perform probabilistic analysis for its base-case. However, the ERG does not consider this to be a major issue as AATO is likely to remain dominant if the ERG would be able to produce probabilistic results for its base-case.

In conclusion, despite the ERG's criticism of the economic model and several highlighted uncertainties, it is reassuring that AATO for the first line population remained dominant in the ERG base-case, and

that the worst-case scenario produced by the ERG resulted in an ICER of £21,622. However, as indicated by the subgroup analysis performed by the ERG, the cost effectiveness of AATO for the second line might be substantially different (estimated ICER of £31,184 per QALY gained).

7.2 Strengths and limitations of the assessment

Overall, the company submission searches were well presented and reproducible. Searches were carried out on a range of databases and supplementary resources. However, the ERG was concerned about the overall quality of the searches conducted, as there were numerous inconsistencies, inaccuracies and redundancy throughout. It is, thus, possible that relevant evidence may have been missed. However, the main weakness of the submission is that only one trial is directly relevant to the appraisal (APL0406) which provides data on an untreated population only. The trial does not have any UK patients. The company presented one trial in relapsed/refractory patients. However, the trial did not present a relevant comparison according to the NICE scope. The committee will need to consider whether it is necessary to explore further the evidence for relapsed/refractory patients or whether it is sufficiently well-established in routine clinical practice.

Strengths related to the economic evaluation include the granularity the model structure provides in comparison with other CEAs identified in the SLR. However, related to this, the (lack of) data to inform post first line transition probabilities can be regarded as a limitation. Additionally, the lack of (EQ5D) utility values for the APL population is a concern. Nevertheless, it is reassuring that AATO for the first line population remained dominant in the ERG base-case, and that the worst-case scenario produced by the ERG resulted in an (deterministic) ICER of £21,622 per QALY gained.

7.3 Suggested research priorities

Although decision uncertainty in the economic evaluation is relatively low, suggested research priorities regarding the cost effectiveness might be focused on obtaining health state utility values for the APL population as well as transition probabilities from and to the HSCT health states reflective of UK clinical practice.

8. REFERENCES

[1] Teva Pharma B.V. *Arsenic trioxide for treating acute promyelocytic leukaemia - TA10216. Document B. Company submission from Teva. Single technology appraisal (STA)*: Teva Pharma B.V., 2017

[2] National Institute for Health and Care Excellence. *Arsenic trioxide for treating acute promyelocytic leukaemia. Final scope [Internet]*. London: National Institute for Health and Care Excellence, 2017 [accessed 8.12.17] Available from: https://www.nice.org.uk/guidance/gid-ta10216/documents/final-scope

[3] Sant M, Allemani C, Tereanu C, De Angelis R, Capocaccia R, Visser O, et al. Incidence of hematologic malignancies in Europe by morphologic subtype: results of the HAEMACARE project. *Blood* 2010;116(19):3724-34.

[4] Visser O, Trama A, Maynadié M, Stiller C, Marcos-Gragera R, De Angelis R, et al. Incidence, survival and prevalence of myeloid malignancies in Europe. *Eur J Cancer* 2012;48(17):3257-66.

[5] Dores GM, Devesa SS, Curtis RE, Linet MS, Morton LM. Acute leukemia incidence and patient survival among children and adults in the United States, 2001-2007. *Blood* 2012;119(1):34-43.

[6] Rashidi A, Riley M, Goldin TA, Sayedian F, Bayerl MG, Aguilera NS, et al. Delay in the administration of all-trans retinoic acid and its effects on early mortality in acute promyelocytic leukemia: final results of a multicentric study in the United States. *Leuk Res* 2014;38(9):1036-40.

[7] Sanz MA, Lo Coco F, Martín G, Avvisati G, Rayón C, Barbui T, et al. Definition of relapse risk and role of nonanthracycline drugs for consolidation in patients with acute promyelocytic leukemia: a joint study of the PETHEMA and GIMEMA cooperative groups. *Blood* 2000;96(4):1247-53.

[8] Fey MF, Buske C, ESMO Guidelines Working Group. Acute myeloblastic leukaemias in adult patients: ESMO Clinical Practice Guidelines for diagnosis, treatment and follow-up. *Ann Oncol* 2013;24 Suppl 6:vi138-43.

[9] Sanz MA, Grimwade D, Tallman MS, Lowenberg B, Fenaux P, Estey EH, et al. Management of acute promyelocytic leukemia: recommendations from an expert panel on behalf of the European LeukemiaNet. *Blood* 2009;113(9):1875-91.

[10] Lengfelder E, Lambert J-F, Nachbaur D, Niederwieser D, Platzbecker U, Schlenk RF, et al. *Onkopedia. Akute promyelozyten leukämie (APL) [Internet]*. Berlin: DGHO Deutsche Gesellschaft für Hämatologie und Medizinische Onkologie e.V., 2014 [accessed 15.12.17] Available from: https://www.onkopedia.com/de/onkopedia/guidelines/akute-promyelozyten-leukaemie-apl/@@view/html/index.html

[11] Kompetenz bei Akuten und Chronischen Leukämien. *German Intergroup recommendations on the diagnostic and therapeutic management of acute promyelocytic leukemia (APL). Version: Final 3.0 [Internet]*. Mannheim: Kompetenz bei Akuten und Chronischen Leukämien, 2015 [accessed 15.12.17] Available from: https://www.kompetenznetz-leukaemie.de/sites/kompetenznetz-leukaemie/content/e53457/e55758/e57007/NapoleonRegister_TreatmentRecommendation_Final_V3.0_20150630.pdf

[12] Lo-Coco F, Avvisati G, Vignetti M, Thiede C, Orlando SM, Iacobelli S, et al. Retinoic acid and arsenic trioxide for acute promyelocytic leukemia. *N Engl J Med* 2013;369(2):111-21.

[13] Burnett AK, Russell NH, Hills RK, Bowen D, Kell J, Knapper S, et al. Arsenic trioxide and all-trans retinoic acid treatment for acute promyelocytic leukaemia in all risk groups (AML17): results of a randomised, controlled, phase 3 trial. *Lancet Oncol* 2015;16(13):1295-305.

[14] Platzbecker U, Avvisati G, Cicconi L, Thiede C, Paoloni F, Vignetti M, et al. Improved outcomes with retinoic acid and arsenic trioxide compared with retinoic acid and chemotherapy in non-high-risk

acute promyelocytic leukemia: final results of the randomized Italian-German APL0406 trial. *J Clin Oncol* 2017;35(6):605-12.

[15] Russell NH, Burnett AK, Hills RK, Betteridge S, Dennis M, Dillon R, et al. Long term follow up from the NCRI AML17 trial of attenuated arsenic trioxide and ATRA therapy for newly diagnosed and relapsed acute promyelocytic leukaemia. *Blood* 2016;128(22):897.

[16] Lengfelder E, European LeukemiaNet. *European recommendation for salvage therapy of relapsed acute promyelocytic leukemia (APL) including arsenic trioxide (ATO) [Internet]*: European LeukemiaNet, 2007 [accessed 15.12.17] Available from: https://www.leukemia-net.org/content/leukemias/aml/apl/apl_recommendations/e6479/infoboxContent6480/APLrelapserecommendation121207.pdf.

[17] Lengfelder E, Lo-Coco F, Ades L, Montesinos P, Grimwade D, Kishore B, et al. Arsenic trioxide-based therapy of relapsed acute promyelocytic leukemia: registry results from the European LeukemiaNet. *Leukemia* 2015;29(5):1084-91.

[18] Raffoux E, Rousselot P, Poupon J, Daniel MT, Cassinat B, Delarue R, et al. Combined treatment with arsenic trioxide and all-trans-retinoic acid in patients with relapsed acute promyelocytic leukemia. *J Clin Oncol* 2003;21(12):2326-34.

[19] Teva Pharma B.V. *Arsenic trioxide for treating acute promyelocytic leukaemia [ID446]. Response to request for clarification from the ERG*: Teva Pharma B.V., 2018

[20] European Medicines Agency. *Assessment report - variation. Trisenox. International non-proprietary name: arsenic trioxide. Procedure no. EMEA/H/C/000388/II/0058 [Internet]*. London: European Medicines Agency, 2016 [accessed 8.1.18] Available from: http://www.ema.europa.eu/docs/en_GB/document_library/EPAR_-_Assessment_Report_-_Variation/human/000388/WC500220071.pdf

[21] U.S. National Library of Medicine. Genetics Home Reference. Acute promyelocytic leukemia [Internet]. U.S. National Library of Medicine, [accessed 15.12.17]. Available from: https://ghr.nlm.nih.gov/condition/acute-promyelocytic-leukemia

[22] Higgins JPT, Green S, eds. *Cochrane handbook for systematic reviews of interventions [Internet]*. Version 5.1.0 [updated March 2011]: The Cochrane Collaboration, 2011 [accessed 8.1.18]. Available from: http://handbook.cochrane.org/

[23] Centre for Reviews and Dissemination. *Systematic Reviews: CRD's guidance for undertaking reviews in health care [Internet]*. York: University of York, 2009 [accessed 8.1.18] Available from: http://www.york.ac.uk/inst/crd/SysRev/!SSL!/WebHelp/SysRev3.htm

[24] Thomas X, Pigneux A, Raffoux E, Huguet F, Caillot D, Fenaux P. Superiority of an arsenic trioxide-based regimen over a historic control combining all-trans retinoic acid plus intensive chemotherapy in the treatment of relapsed acute promyelocytic leukemia. *Haematologica* 2006;91(7):996-7.

[25] Yanada M, Yano S, Kanamori H, Gotoh M, Emi N, Watakabe K, et al. Autologous hematopoietic cell transplantation for acute promyelocytic leukemia in second complete remission: outcomes before and after the introduction of arsenic trioxide. *Leuk Lymphoma* 2017;58(5):1061-7.

[26] Lo-Coco F, Avvisati G, Vignetti M, Breccia M, Gallo E, Rambaldi A, et al. Front-line treatment of acute promyelocytic leukemia with AIDA induction followed by risk-adapted consolidation for adults younger than 61 years: results of the AIDA-2000 trial of the GIMEMA Group. *Blood* 2010;116(17):3171-9.

[27] Ravandi F, Estey E, Jones D, Faderl S, O'Brien S, Fiorentino J, et al. Effective treatment of acute promyelocytic leukemia with all-trans-retinoic acid, arsenic trioxide, and gemtuzumab ozogamicin. *J Clin Oncol* 2009;27(4):504-10.

[28] Efficace F, Mandelli F, Avvisati G, Cottone F, Ferrara F, Di Bona E, et al. Randomized phase III trial of retinoic acid and arsenic trioxide versus retinoic acid and chemotherapy in patients with acute promyelocytic leukemia: health-related quality-of-life outcomes. *J Clin Oncol* 2014;32(30):3406-12.

[29] Burnett AK, Hills RK, Grimwade D, Jovanovic JV, Craig J, McMullin MF, et al. Inclusion of chemotherapy in addition to anthracycline in the treatment of acute promyelocytic leukaemia does not improve outcomes: results of the MRC AML15 trial. *Leukemia* 2013;27(4):843-51.

[30] US Food and Drug Administration. Search orphan drug designations and approvals. Arsenic trioxide [Internet]. Silver Spring, MD: US Food and Drug Administration, [accessed 15.12.17]. Available from: https://www.accessdata.fda.gov/scripts/opdlisting/oopd/detailedIndex.cfm?cfgridkey=109697

[31] Tallman M, Lo-Coco F, Barnes G, Kruse M, Wildner R, Martin M, et al. Cost-effectiveness analysis of treating acute promyelocytic leukemia patients with arsenic trioxide and retinoic acid in the United States. *Clin Lymphoma Myeloma Leuk* 2015;15(12):771-7.

[32] Russell N, Hills R, Betteridge S, Dennis M, Dillon R, Grimwade D, et al. *The feasibility of more attenuated dosage schedules of ATO in newly diagnosed and relapsed acute promyelocytic leukaemia. Data on file*, 2017

[33] Hosing C, Saliba RM, McLaughlin P, Andersson B, Rodriguez MA, Fayad L, et al. Long-term results favor allogeneic over autologous hematopoietic stem cell transplantation in patients with refractory or recurrent indolent non-Hodgkin's lymphoma. *Ann Oncol* 2003;14(5):737-44.

[34] Ramadan SM, Di Veroli A, Camboni A, Breccia M, Iori AP, Aversa F, et al. Allogeneic stem cell transplantation for advanced acute promyelocytic leukemia in the ATRA and ATO era. *Haematologica* 2012;97(11):1731-5.

[35] de Botton S, Fawaz A, Chevret S, Dombret H, Thomas X, Sanz M, et al. Autologous and allogeneic stem-cell transplantation as salvage treatment of acute promyelocytic leukemia initially treated with all-trans-retinoic acid: a retrospective analysis of the European Acute Promyelocytic Leukemia Group. *J Clin Oncol* 2005;23(1):120-6.

[36] Holter Chakrabarty JL, Rubinger M, Le-Rademacher J, Wang HL, Grigg A, Selby GB, et al. Autologous is superior to allogeneic hematopoietic cell transplantation for acute promyelocytic leukemia in second complete remission. *Biol Blood Marrow Transplant* 2014;20(7):1021-5.

[37] Ma X, Does M, Raza A, Mayne ST. Myelodysplastic syndromes: incidence and survival in the United States. *Cancer* 2007;109(8):1536-42.

[38] Lachaine J, Mathurin K, Barakat S, Schuh AC. Economic evaluation of arsenic trioxide for treatment of newly diagnosed acute promyelocytic leukaemia in Canada. *Hematol Oncol* 2015;33(4):229-38.

[39] Douer D, Tallman MS. Arsenic trioxide: new clinical experience with an old medication in hematologic malignancies. *J Clin Oncol* 2005;23(10):2396-410.

[40] Wolff SN, Herzig RH, Fay JW, Phillips GL, Lazarus HM, Flexner JM, et al. High-dose cytarabine and daunorubicin as consolidation therapy for acute myeloid leukemia in first remission: long-term follow-up and results. *J Clin Oncol* 1989;7(9):1260-7.

[41] Fenaux P, Le Deley MC, Castaigne S, Archimbaud E, Chomienne C, Link H, et al. Effect of all transretinoic acid in newly diagnosed acute promyelocytic leukemia. Results of a multicenter randomized trial. European APL 91 Group. *Blood* 1993;82(11):3241-9.

[42] WebMD. Pneumonia - What happens [Internet]. [accessed 29.1.18]. Available from: http://www.webmd.com/lung/tc/pneumonia-what-happens

[43] Shoenfeld Y, Gurewich Y, Gallant LA, Pinkhas J. Prednisone-induced leukocytosis. Influence of dosage, method and duration of administration on the degree of leukocytosis. *Am J Med* 1981;71(5):773-8.

[44] Zhu HH, Wu DP, Jin J, Li JY, Ma J, Wang JX, et al. Oral tetra-arsenic tetra-sulfide formula versus intravenous arsenic trioxide as first-line treatment of acute promyelocytic leukemia: a multicenter randomized controlled trial. *J Clin Oncol* 2013;31(33):4215-21.

[45] Ratnaike RN. Acute and chronic arsenic toxicity. *Postgrad Med J* 2003;79(933):391-6.

[46] Breccia M, Latagliata R, Carmosino I, Cannella L, Diverio D, Guarini A, et al. Clinical and biological features of acute promyelocytic leukemia patients developing retinoic acid syndrome during induction treatment with all-trans retinoic acid and idarubicin. *Haematologica* 2008;93(12):1918-20.

[47] Mathews VB, P., Shaji RV, George B, Chandy M, Srivastava A. Arsenic trioxide in the treatment of newly diagnosed acute promyelocytic leukemia: a single center experience. *Am J Hematol* 2002;70(4):292-9.

[48] Siu CW, Au WY, Yung C, Kumana CR, Lau CP, Kwong YL, et al. Effects of oral arsenic trioxide therapy on QT intervals in patients with acute promyelocytic leukemia: implications for long-term cardiac safety. *Blood* 2006;108(11):103-6.

[49] Woods B, Hawkins N, Dunlop W, O'Toole A, Bramham-Jones S. Bendamustine versus chlorambucil for the first-line treatment of chronic lymphocytic leukemia in England and Wales: a cost-utility analysis. *Value Health* 2012;15(5):759-70.

[50] Beusterien KM, Davies J, Leach M, Meiklejohn D, Grinspan JL, O'Toole A, et al. Population preference values for treatment outcomes in chronic lymphocytic leukaemia: a cross-sectional utility study. *Health Qual Life Outcomes* 2010;8:50.

[51] Szende A, Janssen B, Cabases J, eds. *Self-reported population health: an international perspective based on EQ-5D [Internet]*. Dordrecht: SpringerOpen, 2014 [accessed 20.12.17]. Available from: https://link.springer.com/book/10.1007/978-94-007-7596-1

[52] Breitscheidel L. Cost utility of allogeneic stem cell transplantation with matched unrelated donor versus treatment with imatinib for adult patients with newly diagnosed chronic myeloid leukaemia. *J Med Econ* 2008;11(4):571-84.

[53] Morton RL, Howard K, Webster AC, Wong G, Craig JC. The cost-effectiveness of induction immunosuppression in kidney transplantation. *Nephrol Dial Transplant* 2009;24(7):2258-69.

[54] Cooperberg MR, Ramakrishna NR, Duff SB, Hughes KE, Sadownik S, Smith JA, et al. Primary treatments for clinically localised prostate cancer: a comprehensive lifetime cost-utility analysis. *BJU Int* 2013;111(3):437-50.

[55] Attard CL, Brown S, Alloul K, Moore MJ. Cost-effectiveness of folfirinox for first-line treatment of metastatic pancreatic cancer. *Curr Oncol* 2014;21(1):e41-51.

[56] Stevenson SM, Danzig MR, Ghandour RA, Deibert CM, Decastro GJ, Benson MC, et al. Cost-effectiveness of neoadjuvant chemotherapy before radical cystectomy for muscle-invasive bladder cancer. *Urol Oncol* 2014;32(8):1172-7.

[57] Choi HW, Miele K, Dowdy D, Shah M. Cost-effectiveness of Xpert(R) MTB/RIF for diagnosing pulmonary tuberculosis in the United States. *Int J Tuberc Lung Dis* 2013;17(10):1328-35.

[58] Prica A, Chan K, Cheung M. Combined modality therapy versus chemotherapy alone as an induction regimen for primary central nervous system lymphoma: a cost-effectiveness analysis. *Neuro Oncol* 2014;16(10):1384-91.

[59] Nshimyumukiza L, Duplantie J, Gagnon M, Douville X, Fournier D, Lindsay C, et al. Dabigatran versus warfarin under standard or pharmacogenetic-guided management for the prevention of stroke and systemic thromboembolism in patients with atrial fibrillation: a cost/utility analysis using an analytic decision model. *Thrombosis Journal* 2013;11(1):14.

[60] National Institute for Health and Care Excellence. *Arsenic trioxide for treating acute promyelocytic leukaemia [ID446]. Clarification letter*. London: National Institute for Health and Care Excellence, 2018

[61] Joint Formulary Committee. *British National Formulary. 73*. London: BMJ Group and Pharmaceutical Press, 2017.

[62] Department of Health. *NHS reference costs 2014-2015 [Internet]*. London: Department of Health, 2015 [accessed 20.12.17] Available from: https://www.gov.uk/government/publications/nhs-reference-costs-2014-to-2015

[63] Morgan A, Sutton A, Wailoo A. *The risk and costs of febrile neutropenia in patients with non-small cell lung cancer treated with docetaxel. Report by the NICE Decision Support Unit [Internet]*. Sheffield: NICE Decision Support Unit, 2007 [accessed 20.12.17] Available from: http://nicedsu.org.uk/

[64] Soini E, Hautala A, Poikonen E, Becker U, Kyttälä M, Martikainen J. Cost-effectiveness of first-line chronic lymphocytic leukemia treatments when full-dose fludarabine is unsuitable. *Clin Ther* 2016;38(4):889-904.e14.

[65] Akhtar W, Chung Y. Saving the NHS one blood test at a time. *BMJ Qual Improv Rep* 2014;2(2):pii: u204012.w1749.

[66] Calhoun EA, Chang CH, Welshman EE, Fishman DA, Lurain JR, Bennett CL. Evaluating the total costs of chemotherapy-induced toxicity: results from a pilot study with ovarian cancer patients. *Oncologist* 2001;6(5):441-5.

[67] Milligan DW, Grimwade D, Cullis JO, Bond L, Swirsky D, Craddock C, et al. Guidelines on the management of acute myeloid leukaemia in adults. *Br J Haematol* 2006;135(4):450-74.

[68] National Institute for Health and Care Excellence. *Chronic heart failure: costing report*. London: National Institute for Health and Care Excellence, 2010

[69] National Institute for Health and Care Excellence. *Intravenous fluid therapy in adults in hospital. NICE Clinical guideline 174 [Internet]*. London: National Institute for Health and Care Excellence, 2013 [accessed 20.12.17] Available from: https://www.nice.org.uk/guidance/cg174

[70] Saito AM, Cutler C, Zahrieh D, Soiffer RJ, Ho VT, Alyea EP, et al. Costs of allogeneic hematopoietic cell transplantation with high-dose regimens. *Biol Blood Marrow Transplant* 2008;14(2):197-207.

[71] Jones CA, Fernandez LP, Weimersheimer P, Zakai NA, Sharf M, Mesa OA, et al. Estimating the burden of cost in chronic graft-versus-host disease: a human capital approach. *JHEOR* 2016;4(2):113-8.

[72] Office for National Statistics (ONS). *National Life Tables, England & Wales, 1980-82 to 2014-16*: Office for National Statistics

[73] Kind P, Hardman G, Macran S. *UK population norms for EQ-5D. Discussion paper 172 [Internet]*. York: Centre for Health Economics, University of York, 1999 [accessed 15.2.18] Available from:

https://www.york.ac.uk/media/che/documents/papers/discussionpapers/CHE%20Discussion%20Paper%20172.pdf

[74] Kaltenthaler E, Carroll C, Hill-McManus D, Scope A, Holmes M, Rice S, et al. The use of exploratory analyses within the National Institute for Health and Care Excellence single technology appraisal process: an evaluation and qualitative analysis. *Health Technol Assess* 2016;20(26):1-48.

Appendix 1: ERG search strategies

Detailed critique of clinical effectiveness searches:

- The database searches were clearly structured (population and study design), using a combination of subject heading indexing and free text terms, with synonyms, adjacency operators and truncation.
- The population facet would have been improved by introducing a specified number to the adjacency operator. When using *adj* without a number affixed the search terms must appear next to each other in that order; affixing a number finds the search terms in any order, within the specified number of words. For example, using ***adj3*** in the search line '(promyelocyt* adj (leukaemia or leukemia)).mp.' would have increased sensitivity by identifying records with 'promyelocytic acute leukaemia' and 'leukemia, acute promyelocytic'.
- The search terms for leukaemia could have been truncated to increase sensitivity, e.g. leukaemia$ or leukemia$
- Additional synonyms and acronyms could have been included in the search strategies, e.g. APML, APL, AML M3, ANLL M3, progranulocytic leukaemia.
- The full date span for the databases searched was not provided.
- The eligibility criteria provided in Table 2.1 of the company submission included systematic reviews and meta-analyses, but no attempt to search for these study designs was made. Search terms for systematic reviews were not included in the strategies, and systematic review specific resources such as the Cochrane Database of Systematic Reviews (CDSR) and the Database of Abstracts of Reviews of Effects (DARE) were not searched.
- The search strategy provided in Appendix D reported a simultaneous search of MEDLINE and Embase using the Ovid interface. A simultaneous multi-file search such as this should include both MeSH and EMTREE subject headings to ensure that all subject indexing terms are searched; the search strategy only included the EMTREE term 'promyelocytic leukemia/' for the initial search, then only the MeSH term 'exp Leukemia, Promyelocytic, Acute/' for the update search. In this case, the EMTREE term does not map to the equivalent MeSH term when conducting a simultaneous multi-file search, whereas the MeSH term does map to the EMTREE term. Although the EMTREE term 'promyelocytic leukemia/' is reported in Appendix D, Table 1.1, the results would indicate that the MeSH term was actually used in the search. Indeed, when a simultaneous search of MEDLINE and Embase is conducted in Ovid the following message appears: *[Ovid MEDLINE] – The subject heading 'promyelocytic leukemia' is invalid in this database.*
- It appears that the RCT filter used was based on the Cochrane Highly Sensitive Search Strategy for identifying randomized trials in MEDLINE; this was not explicitly stated in either the clinical effectiveness section of the company submission (B.2) or in Appendix D. This search filter was designed specifically for use in MEDLINE, and does not translate to work efficiently in Embase.
- The company submission attempted to identify safety data as well as clinical effectiveness data by conducting two separate literature searches. The study design terms used to search for non-RCTs were possibly too restrictive to capture all safety data. It is not clear where the search terms used to search for non-RCTs were derived from. EMTREE subject heading terms were included, but not MeSH terms. Although the strategy was not limited to RCTs, it was still limited to study designs that do not necessarily capture safety data (longitudinal studies, retrospective studies, prospective studies, follow-up studies). CRD guidance[23] recommends that if searches have been limited by a study design filter, additional searches should be undertaken to ensure that adverse events that are long-term, rare or unanticipated are not missed. Ideally,

this would entail searching without any study design terms, or would include generic and specific adverse event and safety search terms.

- There were a number of redundant lines included in the search strategies, e.g. Line 3 in Table 1.1, Lines 3 and 4 in Table 1.2 (Appendix D)
- The method used to limit the MEDLINE and Embase searches to human studies was incorrect. The strategy included the line '*Animals.sh*.' and then used the Boolean operator *NOT* to remove the records identified. The correct limit should be '*exp animals/ not humans.sh.*' when searching MEDLINE, or for a simultaneous multi-file search the automatic limit provided by Ovid should be used: *Human*. It is possible that potentially relevant studies were excluded from the final search results using this approach, as records including terms for both human and animal would have been omitted.
- The MEDLINE and Embase search strategy used a variety of different field tags (mp, tw, ti, ab) when a more consistent approach is used in current best practice.
- It is not clear which database was searched in the Cochrane Library for RCTs. CENTRAL should have been searched for RCTs, but from the results reported in the strategy it would appear that CDSR was searched instead. In response to the ERG clarification letter, the company confirmed that CENTRAL was searched (though the results reported in Table 1.3 of Appendix D would suggest otherwise).
- For the searches of conference proceedings the company submission did not provide full details of the search terms used, the precise date of the searches or the number of records retrieved. Details were provided by the company in response to the ERG clarification letter.
- The October 2017 update searches did not include the specific date ranges searched in MEDLINE and Embase.
- The date limit used for the MEDLINE and Embase update searches was unusual: a line for publication date 1968 to 2015 was combined with Boolean *NOT* to identify studies published from 2016 to 2017, when simply limiting to 2016 to 2017 would have been sufficient.
- The results reported in the search strategies did not correspond with those presented in the PRISMA flow charts: Figure 2.1 and Figure 2.2 in section B.2.

Detailed critique of cost effectiveness searches:

- The database searches were clearly structured (population and study design), using a combination of subject heading indexing and free text terms, with synonyms, adjacency operators and truncation.
- As per the clinical effectiveness search comments above (4.1.1), better use of adjacency, truncation and synonyms would have increased the sensitivity of the searches.
- The full date span for the databases searched was not provided.
- The search strategy reported a simultaneous search of MEDLINE and Embase using the Ovid interface. A simultaneous multi-file search such as this should ideally include both MeSH and EMTREE subject headings. EMTREE subject heading terms were included in the population facet, but not MeSH terms; whereas MeSH terms were included in the cost effectiveness facet, but not EMTREE. As with the clinical effectiveness searches, the EMTREE term '*exp promyelocytic leukemia/*' was reported in the initial 2016 search (Appendix G, Table 1.1), whilst the MeSH term '*exp Leukemia, Promyelocytic, Acute/*' was reported in the update search of October 2017 (Appendix G, Table 1.3).
- It is not clear where the search terms used for the cost effectiveness facet were derived from.
- The method used to limit the MEDLINE and Embase search to human studies was incorrect. The strategy included the line '*Animals.sh.*' and then used the Boolean operator *NOT* to remove

the records identified. The correct limit should be *'exp animals/ not humans.sh.'* when searching MEDLINE. The automatic Ovid limit '*Human'* would have been a better option for this simultaneous multi-file search. It is possible that potentially relevant studies were excluded from the final search results using this approach, as records using terms for both human and animal would have been omitted.

- The MEDLINE and Embase search strategy used a variety of different field tags (mp, tw, ti, ab) when a more consistent approach is used in current practice.
- The host provider used to search NHS EED was not reported. The company responded to the ERG clarification letter to confirm that the CRD interface was used to search NHS EED.
- There was no reason to conduct an update search of NHS EED in October 2017, as this database ceased in April 2015 (Issue 2 of 4). The initial search was conducted in July 2016.
- The date limit used for the MEDLINE and Embase update searches was unusual: a line for publication date 1968 to 2015 was combined with Boolean *NOT* to identify studies published from 2016 to 2017. Limiting to 2016 to 2017 would have been sufficient.
- A search of health economic databases, such as the Cost Effectiveness Analysis (CEA) Registry (www.cearegistry.org) and ScHARRHUD (http://www.scharrhud.org/), would have been a useful addition to the literature searches.

Detailed critique of measurement and valuation of health effects searches:

- The database searches were clearly structured (population and study design), using a combination of subject heading indexing and free text terms, with synonyms, adjacency operators and truncation.
- Again, as with the searches conducted for clinical effectiveness and cost effectiveness, better use of adjacency, truncation and synonyms could have been made to increase the sensitivity of the search results.
- The full date span for the databases searched was not provided.
- The search strategy reported a simultaneous search of MEDLINE and Embase using the Ovid interface. A simultaneous multi-file search such as this should include both MeSH and EMTREE subject headings. EMTREE subject heading terms were included in the population facet, but not MeSH terms; whereas MeSH terms were included in the health-related quality-of-life and utilities facet, but no EMTREE terms were included.
- It is not clear where the search terms used for the health-related quality-of-life and utilities facet were derived from.
- There were a number of redundant lines included in the search strategies, e.g. Lines 5, 25, 26, 28 and 36 in Table 1.1, Lines 1 and 3 in Table 1.2 (Appendix H).
- The method used to limit the MEDLINE and Embase search to human studies was incorrect. The strategy included the line '*Animals.sh.*' and then used the Boolean operator *NOT* to remove the records identified. The correct limit should be '*exp animals/ not humans.sh.*' for searching in MEDLINE. The automatic Ovid limit '*Human'* would have been a better option for this simultaneous multi-file search. It is possible that potentially relevant studies were excluded from the final search results using this approach, as records using terms for both human and animal would have been omitted.
- The host provider used to search NHS EED was not reported. The company responded to the ERG clarification letter to confirm that the CRD interface was used to search NHS EED. A mixture of both CRD and Ovid search syntax was reported in the strategy.
- The NHS EED strategy included a facet of search terms for health-related quality-of-life and utilities; restricting the search unnecessarily.

- There was no reason to conduct an update search of NHS EED, as this database ceased in April 2015 (Issue 2 or 4), and the initial search was conducted in July 2016.
- The date limit used for the MEDLINE and Embase update searches was unusual: a line for publication date 1968 to 2015 was combined with Boolean *NOT* to identify studies published from 2016 to 2017. Limiting to 2016 to 2017 would have been sufficient.
- A search of health economic databases, such as Cost Effectiveness Analysis (CEA) Registry (www.cearegistry.org) and ScHARRHUD (http://www.scharrhud.org/), for utilities data would have been a useful addition to the literature searches

Detailed critique of cost and healthcare resource identification searches:

- The database searches were clearly structured (population and study design), using a combination of subject heading indexing and free text terms, with synonyms, adjacency operators and truncation.
- The full date span for the databases searched was not provided.
- The search strategy reported a simultaneous search of MEDLINE and Embase using the Ovid interface. A simultaneous multi-file search such as this should include both MeSH and EMTREE subject headings; the strategy only included the EMTREE term '*promyelocytic leukemia/'* in the initial 2016 search, whilst only including the MeSH term '*exp Leukemia, Promyelocytic, Acute/'* in the October 2017 update search.
- It is not clear where the search terms used for the resource use and costs facet were derived from.
- The method used to limit the MEDLINE and Embase search to human studies was incorrect. The strategy included the line '*Animals.sh*.' and then used the Boolean operator *NOT* to remove the records identified. The correct limit should be '*exp animals/ not humans.sh.'* when searching MEDLINE, whilst the automatic Ovid limit '*Human'* would have been preferable for a simultaneous multi-file search such as this. It is possible that potentially relevant studies were excluded from the final search results using this approach, as records using terms for both human and animal would have been omitted.
- The host provider used to search NHS EED was not reported. The company responded to the ERG clarification letter to confirm that the CRD interface was used to search NHS EED.
- The update search of NHS EED was unnecessary as this database ceased in April 2015 (Issue 2 of 4).
- The date limit used for the MEDLINE and Embase update searches was unusual: a line for publication date 1968 to 2015 was combined with Boolean *NOT* to identify studies published from 2016 to 2017. Limiting to 2016 to 2017 would have been sufficient.

National Institute for Health and Care Excellence
Centre for Health Technology Evaluation

Pro-forma Response

ERG report

Arsenic trioxide for treating acute promyelocytic leukaemia [ID446]

You are asked to check the ERG report from Kleijnen Systematic Reviews Ltd to ensure there are no factual inaccuracies contained within it.

If you do identify any factual inaccuracies you must inform NICE by **5pm on Monday 5 March 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.

Issue 1 Treatment duration in the model

Description of problem	Description of proposed amendment	Justification for amendment	ERG Response
Page 57: 'As a note, the model implementation differs from the company's description in the CS, in which a maximum of three model cycles of induction and 10 cycles for consolidation phases was stated (although this might be technically possible in the economic model, first line treatment is restricted to fewer cycles).' Note this issue may also be relevant to the following fragment on page 57: 'b) because treatment failed after completion of the first line consolidation phase (40 weeks for AATO or 20 weeks for AIDA)'.	We believe there is a misunderstanding between what the programming allows for providing a flexible model and the schedules considered in the UK model submitted to NICE. In section 5.2.2. (Model Structure, page 97) the CS stated that the first-line induction state 'was compounded of three tunnels (sub-health states representing a period of 4 weeks spent in the health state), ensuring that patients could not remain in the first-line induction state for more than 12 weeks (3 cycles of 4 weeks)'. For first-line consolidation, the CS stated on page 98 that 'This health state consisted of ten tunnel states, allowing the consolidation phase to comprise up to five treatment cycles and assuring that patients remained in this phase for the right amount of time.' These fragments provide only a description of the model structure (i.e. what is feasible in the model). In turn, section 3.3.3.2 (Treatment schedule) describes the relevant inputs that were used to populate the model. As listed on pages 107 and 108 of the CS, the base case duration of induction and consolidation phases was based on the SPC and the APL0406 trial. The numbers used to populate the model were presented in	Amended to clarify the issue flagged by the ERG.	Not a factual error.

	Appendix J. Briefly, the maximum duration of induction in APL0406 was 60 days (i.e. 2 model cycles) for both the AATO and AIDA arm. Consolidation in the AATO arm included 4 cycles of treatment (4 weeks on-4 weeks off) so 7 model cycles (28 weeks), while consolidation in the AIDA arm was divided into 3 monthly cycles (i.e. 3 model cycles, or 12 weeks), as described in Lo-Coco 2013 for the APL0406 trial.		

Issue 2 Molecular evaluation

Description of problem	Description of proposed amendment	Justification for amendment	ERG Response
Page 57: 'In case of molecular remission after the "first line consolidation" phase, or, according to the company, if the patient could not be evaluated for remission (i.e. in the absence of evidence for treatment failure), the patient moves to the molecular remission health state. However, the latter is true only for the AATO+AATO strategy in the model. There is an inconsistency in that, in the AIDA+AATO and AATO+AIDA strategies in the model, patients that could not be evaluated with	We are struggling to understand the inconsistency that the ERG refers to. In line with the APL0406 trial, patients would only move to consolidation states once they have achieved haematological CR. According to an expert opinion from Prof. Lo-Coco, following consolidation patients in haematological remission but not evaluable for PCR are considered in remission and move to the first-line remission state. Thus, the first line remission health state contains patients in molecular remission and patients in haematological remission not evaluable for PCR. Patients who are not in haematological remission move to second line in the model. In general, the prognosis of patients who do not achieve	The comment from the ERG was unclear to us.	Not a factual error. For AATO+AATO the company assumed that all patient that could not be evaluated for remission (based on PCT) would move to the molecular remission health state For AIDA+AATO and AATO+AIDA, instead of assuming that all patient that could not be evaluated for remission (based on PCT) would move to the molecular remission health state, the proportion of

PCR would be evaluated based on haematological response instead of being assumed to move to the molecular remission health state, and only in the case of haematological response would they move to the molecular remission health state.' This issue appears again on page 59: 'There is an inconsistency in what happens in the model when patients could not be evaluated for molecular remission. Patients in the AATO+AATO strategy would be assumed to be in molecular remission, while patients in the AATO+AIDA and AIDA+AATO strategies that could not be evaluated with PCR would be evaluated based on haematological response, and only if this was given patients were assumed to be in molecular remission (**instead of assuming that all patients, regardless of haematological response, are in molecular remission**). This was not	haematological remission is poor – in the AML17 trial, only four patients survived beyond day 60 without achieving haematological remission (Burnett, et al. 2015). Further, haematological remission often occurs earlier during treatment than molecular remission, so implementing the suggestion from the ERG (highlighted in bold to the left) appears unreasonable, and would not be in line with the description of the treatment pathway provided by Prof. Lo-Coco.		patients moving to the molecular remission health state is based on haematological response. Hence, this is inconsistent.

justified and the ERG prefers to implement this in a consistent manner across treatment strategies. This is further explored in the treatment effectiveness section (Section 5.2.6) of this report.'			

Issue 3 Treatment effectiveness for AATO

Description of problem	Description of proposed amendment	Justification for amendment	ERG Response
Page 64: 'Treatment effectiveness of the AATO strategy was estimated by separately estimating Markov traces (as well as costs and QALYs) for AATO+AATO and AATO+AIDA. Subsequently, **a weighted average was calculated with weights of 98% and 2% for AATO+AATO and AATO+AIDA respectively. No justification was provided for these weights.**'	The explanation for the weighted average (shown in bold to the left) is as follows: Patients following the ATRA+ATO/ATRA+ATO pathway were those who were in molecular remission or in haematological remission and did not relapse before 24 months. The others followed the ATRA+ATO/AIDA pathway.	It appears we have omitted this calculation from the company submission.	Not a factual error.

Issue 4 Second-line treatment choice

Description of problem	Description of proposed amendment	Justification for amendment	ERG Response
Page 66: 'In the CS it is assumed for AATO+AATO, that all surviving patients would transit to remission. This was done despite available evidence from the APL0406 trial to inform this parameter in the model'	The ERG model does not consider the decision node for 2nd-line treatment and patients may receive AATO as 2nd-line therapy even if they relapse earlier than 2 years from achieving remission. The company model was programed to reflect the treatment pathway described by Prof. Lo-Coco, where patients treated with AATO first line could receive AATO only for delayed relapse (2+ years after remission).	We feel the solution proposed by the ERG does not adequately reflect treatment choices for patients who relapse.	Not a factual error.

Issue 5 Relapse probabilities for AATO

Description of problem	Description of proposed amendment	Justification for amendment	ERG Response
Page 66: 'By using unconditional relapse probabilities , the company likely underestimates the relapse probability for AATO'	We used conditional probabilities for post-24 months relapses. We computed the probability of relapse between 24 and 48 months, meaning that patients did not relapse before 24 months.	The comment from the ERG was unclear to us.	Not a factual error.

Issue 6 Transition to HSCT states

Description of problem	Description of proposed amendment	Justification for amendment	ERG Response
Page 69: 'The transition from "second line molecular	In order to compute a per-cycle probability, it was assumed that the probability of	Amended to provide information requested by the ERG.	Not a factual error.

remission" to the HSCT states is adjusted using the median time to relapse following second line remission. The rationale of this adjustment is unclear to the ERG.'	relapse in the remission state was observed over a certain period, defined by the median time to relapse as a proxy for the average time to relapse.		

Issue 7 AE probabilities

Description of problem	Description of proposed amendment	Justification for amendment	ERG Response
Page 74: 'The AE probabilities were not mentioned in the CS.'	They were, however, listed in Appendix J.	Amended to provide information requested by the ERG.	Not a factual error, this information was not provided in the CS (i.e. main document). Additionally Table 2.2 of Appendix J provides specific model results (i.e. not the specific input parameters of interest).

Issue 8 AEs included in the model

Description of problem	Description of proposed amendment	Justification for amendment	ERG Response
Page 74: 'Moreover, the selection of these specific AEs is unclear to the ERG; this includes that is was unclear	These are most of the AEs described in detail in the APL0406 trial. GI toxicity was not considered in the model, but this is conservative given that it was significantly	Amended to provide NICE with additional insight the rationale for AE selection.	Not a factual error.

why reversible arrhythmia was not considered in the model (as discussed in section 5.2.6).'	more frequent with AIDA than AATO. As for 'reversible arrhythmia' the model did consider QTc prolongation, albeit it was assumed not to prompt a treatment switch.		

Issue 9 Utility sources

Description of problem	Description of proposed amendment	Justification for amendment	ERG Response
Page 76: 'The primary source for the "CML after HSCT without GvHD" utility value and the rationale for adjusting it were not provided.'	This is listed in Table 3.3 page 119 of the CS. The primary source was Breitscheidel L., 2008. Details of adjustments are also provided therein.	Amended to provide information requested by the ERG.	Not a factual error.

Issue 10 AE frequency

Description of problem	Description of proposed amendment	Justification for amendment	ERG Response
Page 78: 'The proportion of patients experiencing each adverse event was not reported in the CS (Section 5.2.7).'	This was reported in Appendix J.	Amended to provide information requested by the ERG.	Not a factual error, this information was not provided in the CS (i.e. main document). Additionally Table 2.2 of Appendix J provides specific model results (i.e. not the specific input parameters of interest).

Issue 11 GvHD in the model

Description of problem	Description of proposed amendment	Justification for amendment	ERG Response
Page 78: 'The proportion of patients experiencing GvHD , the duration of GvHD , and the duration of hospitalisation in the above-mentioned health states were not reported in the CS. The CS also emphasised that patients could experience acute and chronic GvHD but did not describe how these were differentiated in the cost effectiveness model.'	The probability of GvHD was provided in Appendix J. The duration of GvHD was reported in reported in Table 3.4 page 121 of the CS. The duration of hospitalisation was provided in Appendix J. Finally, the differentiation between acute and chronic GvHD was provided in Table 3.4 page 121 of the CS.	Amended to provide information requested by the ERG.	Not a factual error, this information was not provided in the CS (i.e. main document). Additionally Table 3.4 of Appendix J is the Markov trace (i.e. not the specific input parameters of interest).

Issue 12 PSA

Description of problem	Description of proposed amendment	Justification for amendment	ERG Response
Page 91: 'The ERG was unable to perform probabilistic analyses (without errors) using the model file received in response to the clarification questions (named "ID446 arsenic trioxide TEVA CEM_v4.2_rem26 v0.1 170118 SC [noACIC].xlsm"). The ERG made its adjustments using this model file given this version of the model incorporated	Indeed, the PSA in the model included placeholders for some parameters, but information on distribution was not filled or was filled with inconsistent data (e.g. mean and standard deviation equal to 0). This generated the errors described by the ERG.	Clarified and provided additional information on PSA.s	Not a factual error.

structural adjustments that the ERG preferred to use in its base-case. Unfortunately, given the complex implementation of the PSA in the company's model, the ERG was not able to correct the cause of this error'			