NICE National Institute for Health and Care Excellence

Single Technology Appraisal

Oral azacitidine for maintenance treatment of acute myeloid leukaemia after induction therapy [ID3892]

Committee Papers

NATIONAL INSTITUTE FOR HEALTH AND CARE EXCELLENCE

SINGLE TECHNOLOGY APPRAISAL

Oral azacitidine for maintenance treatment of acute myeloid leukaemia after induction therapy [ID3892]

Contents:

The following documents are made available to consultees and commentators:

The **final scope** and **final stakeholder list** are available on the NICE website.

Pre-technical engagement documents

1. **Company submission from Celgene**

2. **Clarification questions and company responses**

3. **Patient group, professional group and NHS organisation submissions** from:
   a. Leukaemia Care
   b. Royal College of Pathologists-British Society for Haematology

4. **Evidence Review Group report** prepared by Kleijnen Systematic Reviews
   a. ERG report
   b. ERG report addendum (updated PAS discount)
   c. ERG report addendum EU-consolidation subgroup (ERG revised base case)

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

Post-technical engagement documents

6. **Technical engagement response from company**
   a. Technical engagement response form (updated PAS discount)
   b. Company model results - deterministic and probabilistic analyses (updated PAS discount)

7. **Technical engagement responses and statements from experts:**
   a. Professor Charles Craddock – clinical expert, nominated by Celgene
   b. Dr Manoj Raghavan – clinical expert, nominated by the Royal College of Pathologists
   c. Martin Burr – patient expert, nominated by Leukaemia Care

8. **Technical engagement responses from consultees and commentators:**
   a. Novartis

9. **Evidence Review Group critique of company response to technical**

engagement prepared by Kleijnen Systematic Reviews

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

NATIONAL INSTITUTE FOR HEALTH AND CARE EXCELLENCE

Single technology appraisal

Oral azacitidine for maintenance treatment of acute myeloid leukaemia after induction therapy [ID3892]

Document B

Company evidence submission

December 2021

File name	Version	Contains confidential information	Date
ID3892_Oral Azacitidine_Document B_ AIC & CIC redacted_07Jan22_Final	1.0	Yes	07 January 2022

Contents

List of tables

List of figures

Abbreviations

AE	adverse event
AESI	adverse event of special interest
AIC	Akaike information criteria
alloHSCT	allogeneic HSCT
AML	acute myeloid leukaemia
ANC	absolute neutrophil count
ANCOVA	analysis of covariance
AZA	azacitidine
BIC	Bayesian information criteria
BID	twice daily
BL	baseline
BM	bone marrow
BMS	Bristol Myers Squibb
BNF	British National Formulary
BSA	body surface area
BSC	best supportive care
BSCH	British Committee for Standards in Haematology
CC-486	azacitidine
CHMP	Committee for Medicinal Products for Human Use
CI	confidence interval
CMH	Cochran-Mantel-Haenszel
CMML	chronic myelomonocytic leukaemia
CR	complete remission
CRi	complete remission with incomplete blood count recovery
DFS	disease-free survival
DNA	deoxyribonucleic acid
DNMT	DNA Methyltransferase
DOR	duration of response
DS	Down syndrome
DSA	deterministic sensitivity analysis
DSC	Decision Support Unit
ECG	electrocardiogram

ECOG PS	Eastern Cooperative Oncology Group Performance Status
EFS	event-free survival
ELN	European LeukemiaNet
EMA	European Medicines Agency
EORTC QLQ-C30	European Organisation for Research and Treatment of Cancer
EP	extension phase
EQ-5D-3L	EuroQol 5-Dimensions 3-Level
ERG	Evidence Review Group
ESA	Erythropoiesis-stimulating agent
ESMO	European Society for Medical Oncology
FACIT	Functional Assessment of Chronic Illness Therapy
FAS	full analysis set
FLT3	fms-like tyrosine kinase 3
G-CSF	granulocyte colony-stimulating factor
GI	gastrointestinal
HCRU	healthcare resource use
HMA	hypomethylating agent
HMRN	Haematological Malignancy Research Network
HR	hazard ratio
HRQoL	health-related quality of life
HRU	healthcare resource use
HSCT	haematopoietic stem cell transplantation
HTA	health technology assessment
IC	intensive chemotherapy
ICER	incremental cost-effectiveness ratios
INV	investigator
IPD	individual patient-level data
ITD	internal tandem duplication
ITC	indirect treatment comparison
ITT	intention-to-treat
IV	intravenous
IVRS	Interactive Voice Response System

IWG	International Working Group
KM	Kaplan-Meier
LAIP	leukaemia-associated immunophenotype
LMM	linear mixed effects model
LY	life year
MAA	marketing authorisation application
Max	maximum
MAIC	matching adjusted indirect comparisons
MDS	myelodysplastic syndromes
MFC	multiparameter flow cytometry
MHRA	Medicines and Healthcare Products Regulatory Agency
MID	minimally important difference
Min	minimum
mITT	modified intent-to-treat
MMRM	mixed model repeated measures
MRC	myelodysplasia-related changes
MRD	measurable residual disease
MRR	major response rate
n	number of patients in the category
N	number of patients evaluable
N/A	not applicable
NCCN	National Cancer Comprehensive Network
NE	not evaluable
NGC	next-generation sequencing
NHS	National Health Service
NICE	National Institute of Health and Care Excellence
No.	number
NOS	not otherwise specified
NR	not reported
NRM	non-relapse mortality
NYHA	New York Heart Association
OR	overall response

ORR	objective response rate
OS	overall survival
PBO	placebo
PD	progressive disease
PFS	progression-free survival
PH	proportional hazard
PK	pharmacokinetic
PP	per-protocol
PPS	post-progression survival
PR	partial response
PRO	patient-reported outcome
PSM	partitioned survival model
PSSUR	Personal Social Services Research Unit
Pt	patient
PT	preferred term
QALY	quality-adjusted life year
QD	once daily
QoL	quality of life
RBC	red blood cell
RCT	randomised controlled trial
RDI	Relative dose intensity
RFS	relapse-free survival
RNA	ribonucleic acid
SAE	serious adverse event
SC	supportive care
Scr	screening
SCT	stem cell transplant
SD	standard deviation
SE	standard error
SLR	systematic literature review
SmPC	summary of product characteristics
SoC	standard of care

SOC	system organ class
SPM	second primary malignancy
STC	simulated treatment comparison
t-MN	therapy-related myeloid neoplasms
TA	technology appraisal
TEAE	treatment-emergent AEs
TKD	tyrosine kinase domain
TRAE	treatment-related adverse event
TSA	thrombopoiesis-stimulating agent
TSD	technical support document
UK	United Kingdom
ULN	upper limit of normal
US	United States
vs	versus
WHO	World Health Organisation
WTP	willingness-to-pay

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

B.1.1 Decision problem

The submission covers the technology's full marketing authorisation for this indication as maintenance treatment in adult patients with acute myeloid leukaemia (AML) who achieved complete remission (CR) or complete remission with incomplete blood count recovery (CRi) following induction therapy with or without consolidation treatment and who are not candidates for, including those who choose not to proceed to, haematopoietic stem cell transplantation (HSCT). The decision problem addressed within the submission is consistent with the National Institute for Health and Care Excellence (NICE) final scope with respect to the population, intervention, outcomes, and comparators with the exception of low-dose cytarabine or subcutaneous azacitidine. A summary of the decision problem and rationale for the exclusion of low-dose cytarabine and subcutaneous azacitidine as comparators is provided in Table B.1.1.

Table B.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 AML who have complete disease remission, or complete remission with incomplete blood count recovery, following induction therapy with or without consolidation treatment who are not eligible for, including those who choose not to proceed to, haematopoietic stem cell transplantation	As per final scope	N/A
Intervention	Oral azacitidine as maintenance treatment	As per final scope	N/A
Comparator(s)	• Midostaurin • Established clinical management without oral azacitidine (which may include a "watch and wait" strategy with best supportive care, low dose cytarabine or subcutaneous azacitidine)	• Midostaurin • Established clinical management without oral azacitidine (which may include a "watch and wait" strategy with best supportive care)	Low-dose cytarabine and subcutaneous azacitidine are not used in clinical practice as maintenance treatments for AML in the population eligible for maintenance treatment with oral azacitidine (as confirmed by two UK AML treating clinicians) and are therefore not considered as comparators to oral azacitidine (further detail provided in Section B.1.3.7)
Outcomes	• OS • RFS • Adverse effects of treatment • HRQoL	As per final scope	N/A
Economic analysis	The reference case stipulates that the cost effectiveness of treatments should be expressed in terms of incremental cost per quality-adjusted life year. The reference case stipulates that the time horizon for estimating clinical and cost effectiveness should be sufficiently long to reflect any differences in costs or outcomes between the technologies being compared.	As per final scope	N/A

	Final scope issued by NICE	Decision problem addressed in the company submission	Rationale if different from the final NICE scope
	Costs will be considered from an NHS and Personal Social Services perspective. The availability of any commercial arrangements for the intervention, comparator and subsequent treatment technologies will be taken into account. The availability of any managed access arrangement for the intervention will also be taken into account.		
Other considerations	The availability and cost of biosimilar and generic products should be taken into account. Guidance will only be issued in accordance with the marketing authorisation. Where the wording of the therapeutic indication does not include specific treatment combinations, guidance will be issued only in the context of the evidence that has underpinned the marketing authorisation granted by the regulator.	As per final scope	N/A

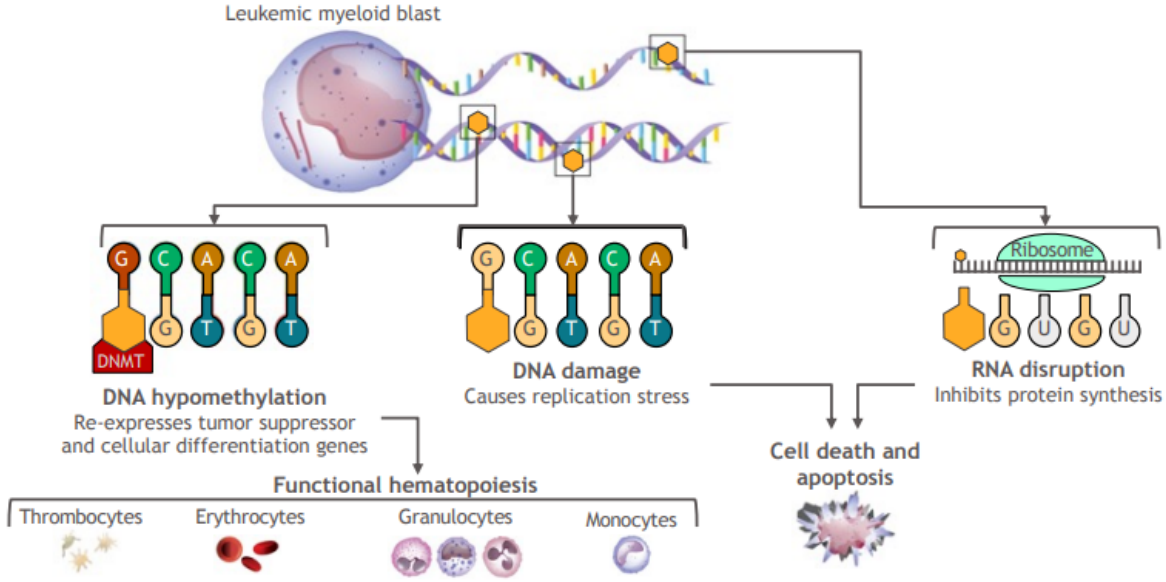
Abbreviations: AML = acute myeloid leukaemia; HRQoL = health-related quality of life; N/A = not available; NHS = National Health Service; NICE = National Institute for Health and Care Excellence; OS = overall survival; RFS = relapse free survival

B.1.2 Description of the technology being appraised

Oral azacitidine, a pyrimidine nucleoside analogue, is a hypomethylating agent (HMA) (Figure B.1.1). It is incorporated into deoxyribonucleic acid (DNA) and ribonucleic acid (RNA) following cellular uptake and enzymatic biotransformation to nucleotide triphosphates resulting in[1]:

- DNA Methyltransferase (DNMT) inhibition; resulting in DNA hypomethylation
- DNA damage; causing replication stress
- RNA disruption; inhibiting protein synthesis

Figure B.1.1. Oral azacitidine mechanism of action



Abbreviations: DNA = deoxyribonucleic acid; DNMT = DNA Methyltransferase; RNA = ribonucleic acid
Source: Wei et al., 2019[2]

Azacitidine exerts its antineoplastic effects via:[3]

- Cytotoxic effects on abnormal haematopoietic cells in the BM through inhibition of DNA, RNA, and protein synthesis; incorporation into RNA and DNA; induction of apoptosis; and activation of DNA damage pathways
- Hypomethylation of DNA from irreversible inhibition of DNMT

Incorporation of azacitidine into the RNA of AML cancer cells, including leukaemic cells, inhibited RNA methyltransferase, reduced RNA methylation, decreased RNA stability, and decreased protein synthesis.[1]

The pharmacokinetic (PK) and pharmacodynamic profiles of oral azacitidine are distinct from those of subcutaneous azacitidine. Oral azacitidine can be administered in extended dosing schedules (14 or 21 days per treatment cycle of 28 days) to sustain therapeutic activity.[4, 5] Considering this, oral azacitidine is likely to be better-suited to long-term use in the AML maintenance setting than subcutaneous azacitidine. A multicentre, open-label study assessing PK/pharmacodynamic profiles of oral azacitidine and subcutaneous azacitidine in

adults with lower-risk myelodysplastic syndrome resulted in a lower cumulative daily exposure over a prolonged period of administration of oral azacitidine vs. subcutaneous azacitidine.[5, 6]

A summary of oral azacitidine is provided in Table B.1.2 and the summary of product characteristics (SmPC) is included in Appendix C.

Table B.1.2. Technology being appraised

UK approved name and brand name	Oral azacitidine (ONUREG®)
Mechanism of action	See Section B.1.2 for oral azacitidine mechanism of action
Marketing authorisation/CE mark status	Received marketing authorisation from the MHRA on 01 July 2021 for the indication below
Indications and any restriction(s) as described in the SmPC	Oral azacitidine is indicated as maintenance therapy in adult patients with AML who achieved CR or CRi following induction therapy with or without consolidation treatment and who are not candidates for, including those who choose not to proceed to, HSCT
Contraindications	• Hypersensitivity to the active substance or to any of the excipients listed in Section 6.1 of the SmPC • Breast feeding; see Section 4.6 of the SmPC
Method of administration and dosage	**Dose schedule standard dose:** 300 mg azacitidine orally QD. Each repeated cycle consists of a treatment period of 14 days followed by a treatment free period of 14 days (28-day treatment cycle). Patients are to be treated with an anti-emetic 30 minutes prior to each dose of oral azacitidine for the first 2 treatment cycles and may be omitted after 2 cycles if there is no nausea and vomiting (see SmPC for further details). Treatment should be discontinued if more than 15% blasts are observed in peripheral blood or bone marrow, or if unacceptable toxicity. **Dose schedule modification for AML disease relapse, with 5% to 15% blasts in peripheral blood or bone marrow**, in conjunction with a clinical assessment, an extension of the dosing schedule from 14 to 21 days of repeated 28-day cycles should be considered. Dosing should not exceed 21 days during any 28-day period. **Oral azacitidine should be discontinued if more than 15% blasts are observed in either the peripheral blood or bone marrow or at the physician's discretion.** **Dose modifications for haematologic and non-haematologic adverse reactions:** Dose interruption and/or dose reduction (to 200 mg) for haematologic and non-haematologic adverse reactions are recommended based on clinical and laboratory findings (see SmPC Section 4.2 for further detail).
Additional tests or investigations	The only additional tests over and above SoC are complete blood counts which should be performed prior to initiation of oral azacitidine and every other week for the first 2 cycles (56 days), every other week for the next 2 cycles after dose adjustment, and monthly thereafter, prior to the start of subsequent cycles of treatment.

List price and average cost of a course of treatment	**Pack (n x dose)**	**14 x 200 mg**	**14 x 300 mg**	**Total annual treatment cost/patient**
	List price	[redacted]	[redacted]	[redacted]
	PAS price	[redacted]	[redacted]	[redacted]
Patient access scheme (if applicable)	This submission includes the confidential simple patient access scheme for oral azacitidine, representing a discount to the list price of [redacted]			

Abbreviations: AML = acute myeloid leukaemia; CR = complete remission; CRi = complete remission with incomplete blood count recovery; HSCT = haematopoietic stem cell transplantation; MHRA = The Medicines and Healthcare products Regulatory Agency; NHS = National Health Service; QD, once daily; SmPC = Summary of Product Characteristics; SoC = standard of care.
Source: SmPC (Appendix C)

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

Summary

Disease overview

- AML is a rare and aggressive blood cancer associated with a very poor prognosis
- There are 3,100 new diagnoses of AML in the UK per year; incidence increases with age, the median age at diagnosis is 72.4 years[7]
- Allogeneic HSCT provides the best chance of cure for patients with AML, however <50% of patients aged <60 years and <20% of patients aged ≥60 years are estimated to undergo HSCT
- The standard of care is induction ± consolidation chemotherapy with the goal of achieving CR/CRi for patients who are fit for intensive therapy and are not candidates for HSCT
- Although most patients who are fit for induction chemotherapy are able to achieve CR/CRi, the majority of these patients will eventually relapse, many within the first year

Current treatment pathway and position of technology

- Maintenance treatment is a post-remission treatment approach that aims to delay relapse and prolong survival
- Maintenance treatment, as defined in this submission, is not currently standard of care in the UK. European and UK guidelines (that pre-date the results of the oral azacitidine QUAZAR Phase 3 study) do not recommend maintenance treatment given the lack of convincing evidence of a proven OS benefit with existing therapies
- Whilst NICE has recommended the use of midostaurin as maintenance treatment it is only recommended for use in a small subgroup of patients (*FLT3*-mutation-positive AML)
- Therefore, observation ("watch + wait") + BSC is the only available option for the majority of patients who have achieved CR/CRi post induction therapy ± consolidation therapy and are not candidates for HSCT
- In this patient population, a substantial unmet need remains for an AML maintenance treatment
- Oral azacitidine is the first and only oral HMA specifically indicated for use as maintenance treatment that provides a significant OS and RFS benefit, with no restriction on the patient population regarding genetic mutations

- As a clinically effective and convenient treatment option with a manageable safety profile, oral azacitidine addresses the high unmet need for patients with AML who have achieved CR/CRi post induction ± consolidation chemotherapy and are not candidates for HSCT and represents a new therapeutic standard of care for this patient population

B.1.3.1 Disease overview

AML is a rare (<1 case/2,000 based on the UK Rare Diseases Framework definition[8]) with a prevalence of 0.9 per 10,000.[7] It is an aggressive haematological cancer,[8-10] originating in the myeloid line of haematopoietic precursor cells, commonly as the result of a genetic aberration.[8-10] Regardless of the underlying cause, the pathophysiology of AML involves dysfunctional differentiation of myeloblasts and suppression of normal bone marrow haematopoiesis, leading to excessive proliferation of immature myeloid cells (blast cells or 'blasts') and accumulation of leukaemic cells in the bone marrow.[10-12]

The World Health Organisation (WHO) definition of AML is the presence of ≥20% leukaemic blasts in the bone marrow or peripheral blood and is diagnosed accordingly in combination with immunohistochemistry, cytogenetics, and molecular analyses, as well as prior medical history and clinical information.[13]

There are around 3,100 new diagnoses of AML in the United Kingdom (UK) per year (2016-2018).[14] The majority of cases are in older people, with >40% of new cases in people aged 75 and over.[15] Data are from the UK's population-based Haematological Malignancy Research Network (HMRN) reported a median age at diagnosis of 72.4 years.[7]

The disease is rapidly progressing and the prognosis in general is poor.[10-12] The duration of overall survival (OS) is highly dependent on a patient's fitness for intensive therapy, which is determined based on patient-specific factors including age, cytogenetic and molecular abnormalities.[9, 16-19] In patients aged ≥65 years in England the 5-year survival rate is 4.5%.[20]

Patients who respond to induction +/- consolidation treatment and achieve a CR may be considered eligible to receive allogeneic HSCT (alloHSCT) which provides the best chance of 13% cure for patients with AML.[21, 22] However, less than 50% of patients aged under 60 years and less than 20% of patients over 60 years are estimated to undergo alloHSCT,[23-25] as many are deemed ineligible. AlloHSCT is primarily an option for patients who are younger, sufficiently fit, and have high risk of disease relapse.[26] Patients are often ineligible because of older age and comorbidities.[26-29]

For patients who are fit for intensive therapy, the standard of care is induction chemotherapy with the goal of achieving CR/CRi. Induction chemotherapy is often followed by subsequent cycles of consolidation chemotherapy in patients who achieve CR/CRi.[18]

Maintenance treatment in AML is a post-remission treatment approach which aims to delay relapse and prolong survival.[30] However, maintenance treatment is not currently standard of care in the UK. Whilst the NICE has recommended the use of midostaurin in 2018 (Technology appraisal [TA] 523) as a maintenance treatment, it is only recommended in a small subgroup of *FLT3*-mutation-positive patients[a] who are in remission after previously being treated with

[a] Approximately 25% of AML patients are FLT3-mutation-positive and approximately 30-40% of will achieve first remission. A majority of these patients go on to receive HSCT leaving approximately 10% of these patients who are likely to have midostaurin maintenance in the UK (as confirmed by UK clinicians).

midostaurin in combination with chemotherapy agents during induction and consolidation chemotherapy.[31] Therefore, observation (also referred to as “watch and wait”) + best supportive care (BSC) is the only available option for the majority of patients who have achieved first remission post induction ± consolidation therapy and are not candidates for HSCT.

As most patients that achieve a CR/CRi after induction chemotherapy will experience disease relapse, a substantial unmet need remains for a well-tolerated and easily administered AML maintenance treatment, that significantly prolongs survival among patients with AML who are in remission after intensive chemotherapy (IC) without compromising health-related quality of life (HRQoL). Such a treatment could become the new standard of care for maintenance treatment in AML.

B.1.3.1.1 Aetiology

AML may arise either in the absence of prior therapy or disease (primary or de novo AML) or secondary to an antecedent haematological disorder - as the result of exposure to prior chemotherapy or after radiation therapy (secondary AML),[9, 10] AML commonly results from chromosomal abnormalities or single gene mutations: approximately 97% of patients have at least one genetic mutation and approximately 48% have at least two.[32] These mutations result in activation of proproliferative pathways (e.g.- *FLT*3), dysfunctional haematopoietic differentiation (e.g. *NPM1*), or altered epigenetic regulation (e.g. the DNA methylation related- genes *DNMT3A, TET2, IDH2, and IDH1*).[33] Cytogenetics and mutational profiles are important prognostic factors that can inform the likelihood of achieving remission, relapse rates and OS and are used to determine treatment strategies.[18]

B.1.3.1.2 Disease presentation

Most of the clinical manifestations of the disease result from the infiltration and accumulation of malignant, undifferentiated myeloid cells in the bone marrow, peripheral blood, and other tissues, contributing to impaired blood cell production and bone marrow failure.[9, 10, 33]

The signs and symptoms associated with AML are often non-specific and secondary to the development of other conditions. Flu-like symptoms may precede the diagnosis by four to six weeks. Patients may present with anaemia (low red blood cell count), neutropenia (low white cell count), and/or thrombocytopenia (low platelet count) because of impaired haematopoiesis.[9] Anaemia can lead to weakness, fatigue, feeling cold, headaches, pallor, dizziness, and dyspnoea on exertion, resulting in patients being unable to perform more than the basic activities of daily living.[34, 35] Neutropenia can result in frequent and persistent infections that are accompanied by symptoms such as fever and appetite loss. Thrombocytopenia, depending on the severity, can present with petechiae, bruising or bleeding In some cases, leukaemic cells can spread to other organs. Symptoms associated with leukaemic cell infiltration in the brain and spinal cord include headaches, weakness, seizures, vomiting, issues with balance, and blurred vision.[34, 35] Abdominal swelling is observed when leukaemic cells accumulate in the liver and spleen, and tumour-like collections of leukaemic cells can accumulate in extramedullary sites, resulting in leukaemia cutis, granulitic sarcomas, or chloromas.[35]

B.1.3.2 Diagnosis and classification

The availability of an adequate bone marrow sample (aspirate and trephine) at the time of diagnosis is important, as cytogenetic analysis and evaluation of molecular abnormalities are necessary for disease risk stratification, which is used to inform risk assessment, prognosis, and treatment selection.[13, 18]

AML is classified into six categories: AML with recurrent genetic abnormalities; AML with myelodysplasia-related changes (MRC); therapy-related myeloid neoplasms (t-MN); AML, not otherwise specified (NOS); myeloid sarcoma; and myeloid proliferations related to Down syndrome (DS).[13] These categories are used to help define risk categories (see Section B.1.3.3) and to select appropriate treatment strategies.[9]

B.1.3.3 Prognostic factors

The duration of OS in AML (see Section B.1.3.4.3) is highly dependent on several patient-specific factors, including age, cytogenetic risk status, molecular profile.[9, 16-18] These factors are also used to guide treatment decisions, as they impact both a patient's likelihood of achieving CR with intensive induction chemotherapy and their risk of relapse after achieving remission.[16, 18, 23, 30, 36, 37] Molecular risk factors with well characterised prognostic value in AML include cytogenetic abnormalities and mutations in *FLT3-ITD, WT1, Ckit-, DNMT3A, TP53, and IDH*.[18, 38]

The prognostic risk classification for AML has changed over time. The latest risk classifications systems such as the European LeukemiaNet (ELN) 2017 guidelines provide more detailed cytogenetic and molecular abnormalities for risk classification[18, 39] (Table B.1.3), than the earlier NCCN 2012 guidelines[40] (Table B.1.4).

Table B.1.3. Prognostic risk stratification of AML by genetics, ELN 2017

Risk Category	Cytogenetic and molecular features
Favourable	t(8;21)(q22;q22.1); *RUNX1-RUNXITI* inv(16)(p13.1q22) or t(16;16)(p13.1;q22); *CBFB-MYH11* Mutated *NPM1* without *FLT3*-ITD or with *FLT3*-ITD^{low} Biallelic *CEBPA* mutation
Intermediate	Mutated *NPM1* and *FLT3*-ITD^{high} Wild-type *NPM1* without *FLT3*-ITD or with *FLT3*-ITD^{low} (without adverse genetic lesions) t(9;11)(p21.3;q23.3); *MLLT3-KMT2A* Other cytogenetics, including normal karyotype, not classified as favourable or adverse
Poor	t(6;9)(p23;q34.1); *DEK-NUP214* t(v;11q23.3); *KMT2A* rearranged t(9;22)(q34.1;q11.2); *BCR-ABL1* inv(3)(q21.3q26.2) or t(3;3)(q21.3;q26.2); *GATA2, MECOM(EVII)* -5, del(5q); -7; -17/abn(17p) Complex karyotype, monosomal karyotype Wild-type *NPM1* and *FLT3-ITD^{high}* Mutated *RUNX1* Mutated *ASXL1* Mutated *TP53*

Abbreviations: AML = acute myeloid leukaemia; ELN = European LeukemiaNet.

Source: Dohner et al., 2017[18]

Table B.1.4. Prognostic risk stratification of AML by genetics, NCCN 2012

Risk Category	Cytogenetic and molecular abnormalities
Better	inv(16); t(16;16) t(8;21) t(15;17) Normal cytogenetics: with *NPM1* or isolated *CEBPA* mutation in the absence of *FLT3-ITD*
Intermediate	Normal cytogenetics +8 t(9;11) Other nondefined t(8:21); inv(;16): t(16:16): with *c-KIT* mutation
Poor	Complex (≥3 clonal chromosomal abnormalities) -5; 5q; -7; 7q 11q23 - non t(9;11) inv(3); t(3;3) t(6;9) t(9:22) Normal cytogenetics: with *FLT3-ITD* mutation

Abbreviations: AML = acute myeloid leukaemia; FLT3 = fms-like tyrosine kinase 3; NCCN = National Cancer Comprehensive Network.
Source: O'Donnell et al., 2012[40]

The 2017 ELN risk classification is a widely used risk stratification system at diagnosis and provides prognostic information in AML patients undergoing chemotherapy as well as alloHSCT. The ELN classifies genetic abnormalities in AML patients as being associated with favourable, intermediate, or poor prognostic risk, as presented in Table B.1.3 The risk stratification during disease course is adjusted based on evaluation of measurable residual disease (MRD), a post-treatment factor that help guide subsequent treatment decisions.[18]

MRD (previously referred to as minimal residual disease), is defined as post-chemotherapy persistence of leukaemic cells at levels below morphologic detection.[41] Clinical investigations of MRD have clearly shown that many patients with AML who achieve CR after induction chemotherapy have detectable residual disease, and that this is a strong independent prognostic marker of increased relapse risk and shorter survival.[41-43] MRD status can be assessed at early timepoints for e.g. after induction and consolidation therapy and sequentially beyond consolidation therapy to assess remission status and detect impending relapse.[18] MRD detection techniques include reverse transcription-quantitative polymerase chain reaction RT-qPCR, next-generation sequencing (NGS) and multiparameter flow cytometry (MFC). The suitability of technique depends on factors, such as the MRD marker to be measured.[18, 44]

B.1.3.4 Epidemiology and mortality

B.1.3.4.1 Risk factors

Several factors are associated with an increased risk of developing AML. These include older age, male sex, history of smoking, and long-term exposure to chemicals such as benzene or

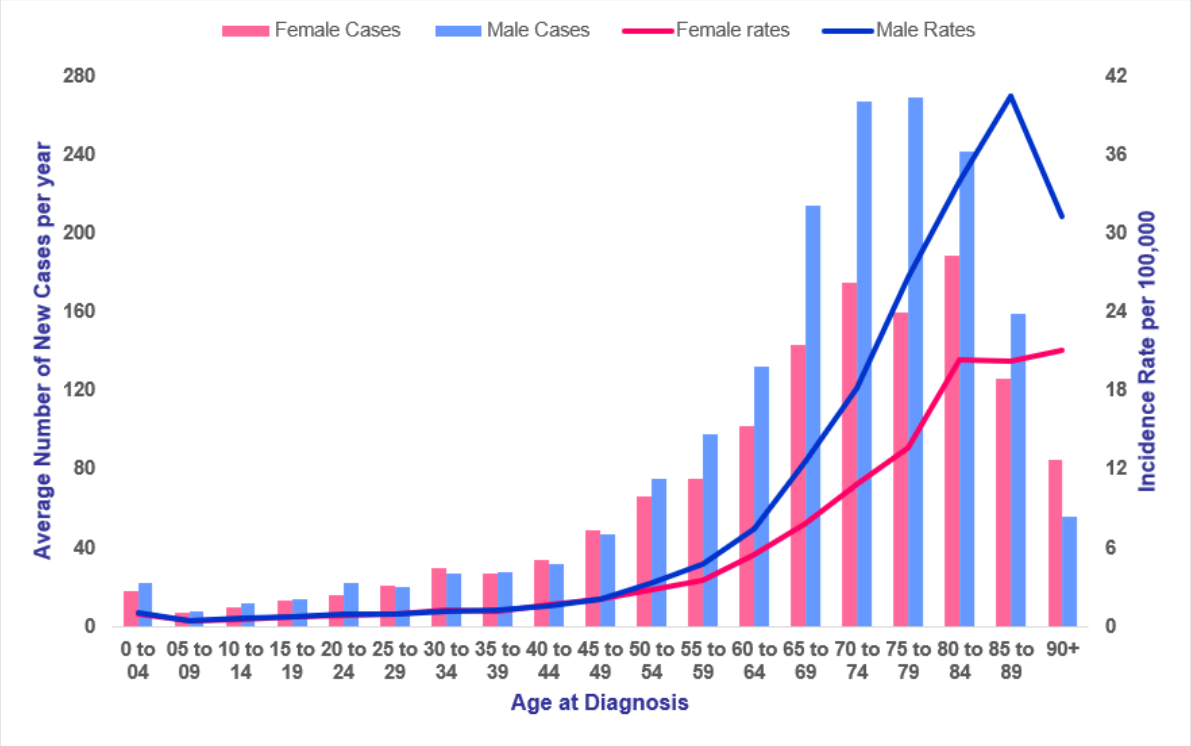
formaldehyde.[15] Development of the disease can also occur secondary to treatment of prior cancers with alkylating agents, platinum-based agents, topoisomerase II inhibitors, and radiation.[9] In addition, AML may also develop in patients with prior haematologic disorders.

B.1.3.4.2 Incidence and prevalence

Data from the UK HMRN registry reports an AML prevalence of 0.9 per 10,000.[7] Given its low prevalence AML is considered a rare disease based on the UK Rare Disease Framework definition (i.e. <1 per 2,000 people).[8] Incidence rates in the UK for females and males have remained stable (2016-2018).[14] Cancer Research UK reported there are around 3,100 new cases of AML in the UK per year (2016-2018).[14]

Incidence rates are higher among men (56%) than women (44%) and increase with age, with a drastically higher risk among men and women aged >60 years than among younger individuals (see Figure B.1.2).[14]

Figure B.1.2. Incidence of AML by age in the UK, 2016-2018



Abbreviations: AML = acute myeloid leukaemia; UK = United Kingdom.
Source: Cancer Research UK[14]

B.1.3.4.3 Mortality

AML has the lowest survival rates across all types of leukaemia (see Table B.1.5),[20] reflecting the poor prognosis of patients with AML. Overall, the median OS of patients with AML ranges from approximately 6 to 12 months.[11, 45-48] In patients aged ≥65 years, diagnosed with AML during the period 2008-2010 in England, the 5-year relative survival rate was 4.5%.[20]

Table B.1.5. Five-year relative survival rates in patients in England aged ≥65 years diagnosed with leukaemia in 2008-2010

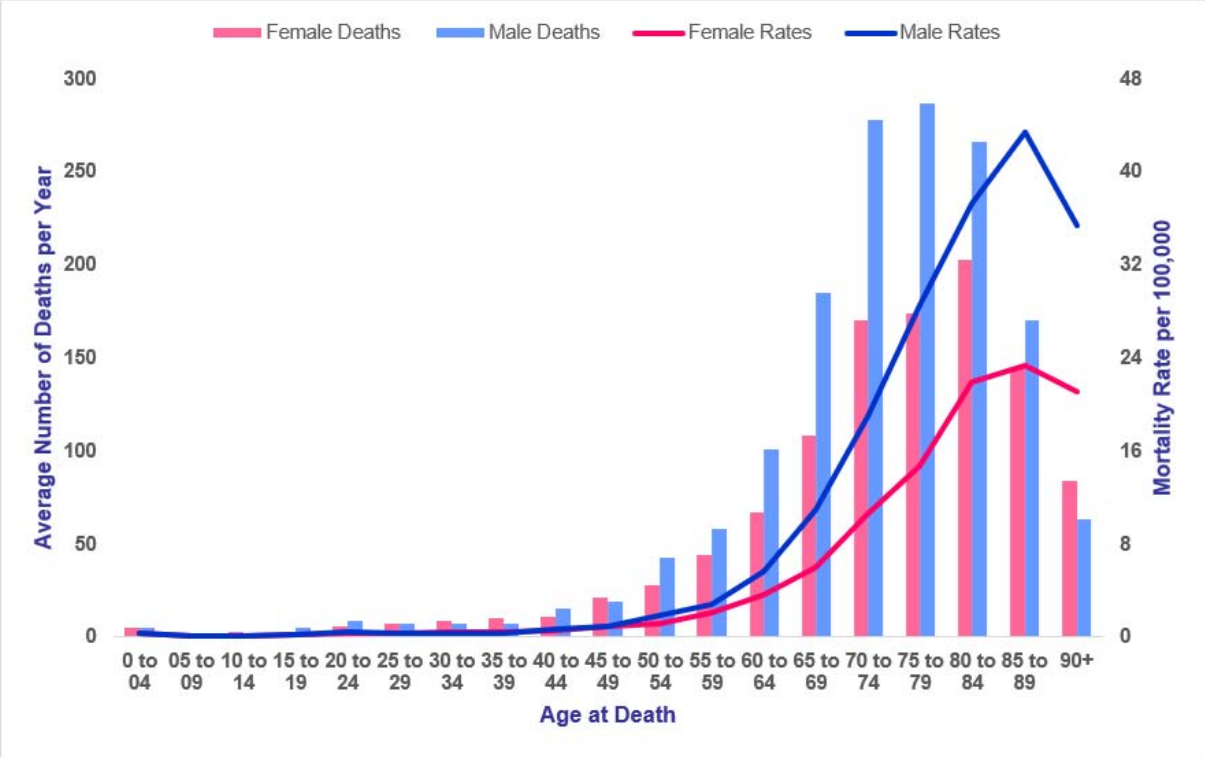
Leukaemia	5-year relative survival
Acute myeloid leukaemia	4.5%
Acute lymphoblastic leukaemia	18.8%
Chronic myeloid leukaemia	38.9%
Chronic lymphocytic leukaemia	60.2%

Source: National Cancer Intelligence Network, 2014[20]

In the UK, over the last decade (between 2006-2008 and 2016-2018), AML mortality rates for females and males combined have remained stable. Age-standardised mortality rates of 5.3 per 100,000 men and 3.1 per 100,000 women in the UK were reported in 2018.[14] There are around 2,600 AML deaths in the UK every year, approximately 7 every day (2016-2018). AML mortality is strongly related to age, with the highest mortality rates being in older people (Figure

B.1.3). In the UK in 2016-2018, on average each year more than half of deaths (53%) were in people aged 75 and over.[14] This largely reflects higher incidence and lower survival for AML in older people. The aggressive nature of AML is further substantiated by the five year OS data in this patient population (12.7%; 95%CI: 11.3-14.7) from the UK HMRN registry. However, within the AML group there is considerable variation by subtype; therapy-related AML and AML with myelodysplasia related changes being almost universally and rapidly fatal, whereas patients diagnosed with acute promyelocytic leukaemia (APL) or AML with core-binding factor mutations were more likely than not to survive for 5 years or more (range: 2.8%–58.6%).[7]

Figure B.1.3. AML average number of deaths per year and age-specific mortality rates per 100,000 population, UK, 2016-2018



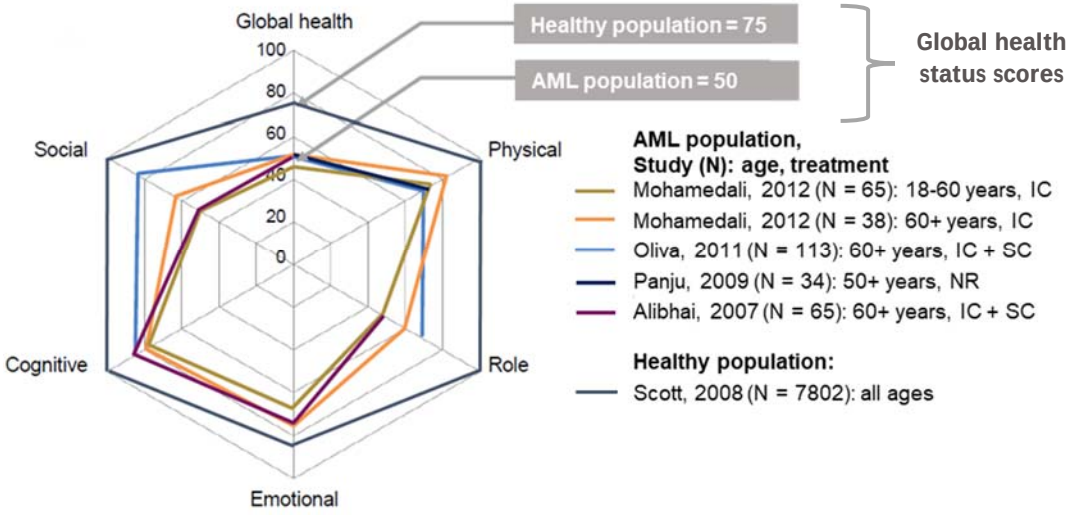
Abbreviations: AML = acute myeloid leukaemia.
Source: Cancer Research UK[49]

B.1.3.5 Disease burden

B.1.3.5.1 Patient and caregiver burden

The HRQoL of patients with AML is substantially impaired by the debilitating symptoms of the disease, the fear/anxiety of being diagnosed with a potentially fatal condition, and the inconvenience, discomfort, and side effects associated with certain AML therapies.[50-52] Patients with AML consistently report lower scores compared to general healthy population for the functional domain and global health status of the European Organization for Research and Treatment of Cancer quality of life questionnaire, 30 questions (EORTC QLQ-C30), (see Figure B.1.4).

Figure B.1.4. EORTC QLQ-C30 functional scores for patients with AML and a healthy population



Abbreviations: AML = acute myeloid leukaemia; EORTC QLQ-C30 = European Organization for Research and Treatment of Cancer quality of life questionnaire, 30 questions; IC = intensive chemotherapy; SC = supportive care.
Source: Scott et al., 2008[53]; Alibhai et al., 2007[54]; Panju et al., 2009[55]; Oliva et al., 2011[56]; Mohamedali et al., 2012[57]

Although the clinical burden of AML and many aspects of HRQoL are significantly improved among patients who achieve remission with IC, it is important to note that these patients may still feel considerable fatigue, depression, stress, and anxiety related to fear of leukaemia recurrence and death.[51, 58, 59] Therefore, the introduction of a maintenance treatment for patients in CR/CRi that significantly delays relapse and improves OS specifically among patients in remission could also reduce the stress and anxiety related to the possibility of disease recurrence and death.

Among patients with relapsed AML, HRQoL is substantially impaired by the debilitating symptoms of advanced disease, ongoing administration of therapies, and the side effects associated with treatment. These patients consistently exhibit the lowest utility values among all AML populations.[50, 60-62] In a study that evaluated the impact of disease status on patient reported outcomes (PROs) after induction chemotherapy, FACT-Leukaemia scores were significantly worse among patients with relapsed AML than among patients with de novo AML at up to 12 months after treatment initiation (mean score: 155.2 vs. 113.4; p=0.0005).[63] The high clinical burden associated with AML relapse and the substantially impaired HRQoL among patients with relapsed AML further highlights the need for a maintenance treatment that prolongs the period of remission after IC.

B.1.3.5.2 Economic burden

Management of AML is resource intensive and leads to a high economic burden. Significant direct costs are incurred related to hospitalisation, medication, treatment administration, treatment of disease-related complications and treatment-related adverse events (AEs), monitoring, and transfusions.[64, 65] In particular, inpatient care (which also accounts for patients undergoing HSCT) is an important driver of healthcare costs and can represent up to 70% of annual AML-related costs.[66] Patients frequently visit their haematologist and cancer clinic, often more than once per month, and may be hospitalised multiple times for treatment

administration, AEs, and management of symptoms and complications of the disease.[64, 66] Resource use is especially high among patients with relapsed AML, who require 87% more outpatient visits and experience 45% more hospitalisations than non-relapsed patients.[67] These findings highlight the considerable additional resource use burden associated with relapse after achieving remission. In contrast, healthcare resource use (HCRU) is much lower among patients who are in remission.[68] Therefore the introduction of a maintenance treatment that keeps patients in remission for longer may lead to a reduction in HCRU.

B.1.3.6 Treatment pathway and anticipated position of oral azacitidine

The principal guideline for the management of AML in the UK is provided by the ELN, published in 2017.[18] Other guidelines include the UK guideline provided by the British Committee for Standards in Haematology (BSCH) published in 2006[69] and the European Society for Medical Oncology (ESMO) guideline published in 2013.[70] Available European and UK guidelines concur in their suggested treatment pathway for patients with newly diagnosed AML.

The first-line treatment approach for patients with newly diagnosed AML is highly dependent on a patient's fitness for intensive therapy, which is determined based on factors such as age, performance status, cytogenetic risk status, molecular risk factors, and comorbidities.[18, 69, 70] For patients who are fit for intensive therapy, the standard of care is intensive induction chemotherapy (see Section B.1.3.6.1 for further detail) with the goal of achieving CR.[18] For patients who are not fit for intensive induction chemotherapy, treatment with non-curative, low intensity chemotherapy can be considered.[69]

For patients who achieve first remission, current guidelines recommend the continuation of chemotherapy as consolidation (see Section B.1.3.6.1 for further detail on consolidation therapy) or for candidates that are suitable for transplant as a bridge to HSCT (see Section B.1.3.6.2 for further detail on HSCT).[18] The duration of remission in AML patients achieved as a result of induction chemotherapy is an important determinant of OS.[16, 18, 36, 71] Although most patients who are fit for IC are able to achieve CR/CRi, the majority of these patients will eventually relapse,[17, 30, 36] many within the first year after achieving remission.[42, 72] There is no established SoC treatment for patients with relapsed AML, and relapse is associated with significantly reduced OS (see Section B.1.3.1 for further detail) and impaired HRQoL (see Section B.1.3.5.1 for further detail). Patients with AML who achieve remission after IC ± consolidation have historically been left with few treatment options; this is especially the case among those who are not eligible for, or choose not to undergo, HSCT. In these patients, maintenance treatment with oral azacitidine can prove beneficial in extending the period of remission which in turn is essential to prolonging life expectancy and maintaining HRQoL in this population (see Section B.1.3.7 and B.1.3.8.1 for further detail).

B.1.3.6.1 Induction and consolidation therapy

For patients who are fit for intensive therapy, the ELN and ESMO guidelines recommend induction chemotherapy as the first-line standard of care i.e typically with 3 days of anthracycline such as daunorubicin and 7 days of cytarabine (commonly referred to as "3 + 7" regimens, with the goal of achieving CR/CRi.[18, 70] The current BSCH guideline states that a "3 + 10" induction chemotherapy regimen may be used as an alternative to a "3 + 7" regimen, while noting that there is no evidence for superiority over "3 + 7".[69]

Induction chemotherapy is often followed by subsequent cycles of consolidation chemotherapy in patients who achieve CR/CRi.[18] Consolidation therapy has been an integral aspect of post-remission therapy in AML for several decades, with the goal of sustaining remission and limiting toxicity.[18]

Induction and consolidation regimens for AML may also include additional agents in combination with chemotherapy. For example, NICE in June 2018 recommended the use of midostaurin for patients with *FLT3*-positive-mutations (TA523), and in November 2018 gemtuzumab ozogamicin for patients with *CD33*-positive-mutations AML (TA545) as add-ons to chemotherapy during both induction and consolidation to target specific mutations seen in AML.[31, 73]

B.1.3.6.2 HSCT

Patients who respond to treatment and achieve a CR may be considered eligible to receive HSCT, alloHSCT provides the best chance of cure for patients with AML.[21, 22]

HSCT is primarily an option for patients who are younger, sufficiently fit, and have high risk of disease relapse.[26] Patients are often ineligible because of older age and comorbidities, which place them at high risk of complications and non-relapse mortality (NRM).[26-29] The pivotal question for clinicians when deciding whether patients receive HSCT is whether the reduction in relapse risk delivered by a transplant outweighs the attendant NRM.[26] The NRM can be reliably predicted to be ~15% using a matched sibling or unrelated donor, patients with a relapse risk >50% are therefore likely to benefit from the attendant halving of relapse risk delivered by alloHSCT, using a well-matched donor.[26] Conversely, patients with a relapse risk ≤40% are unlikely to derive any benefit from alloHSCT given the risk of transplant.[26]

Less than 50% of patients aged <60 years and <20% of patients aged ≥60 years are estimated to undergo alloHSCT,[23-25] as many are deemed ineligible. Less than half of all alloHSCT-eligible patients have a human leukocyte antigen (HLA) -matched sibling available as a donor; availability is even lower for older patients.[74]

Furthermore, a small proportion of patients who are eligible for HSCT choose not to undergo the procedure.[75-78]

B.1.3.7 Maintenance treatment in AML

Maintenance treatment in AML is a post-remission treatment approach which aims to delay relapse and prolong survival.[30] European and UK guidelines (that pre-date the results of the oral azacitidine QUAZAR Phase 3 study) do not recommend maintenance treatment for this patient population given the lack of convincing evidence of a proven OS benefit with existing therapies.[18, 79] Whilst NICE has recommended the use of midostaurin in 2018 (TA523) as a maintenance treatment it is only recommended in a small subgroup of *FLT3*-mutation-positive patients[b] who are in remission after previously being treated with midostaurin in combination with chemotherapy agents during induction and consolidation chemotherapy.[31] However, clinical experts in the UK have stipulated that midostaurin is rarely used as a maintenance

[b] Approximately 25% of AML patients are FLT3-mutation-positive and approximately [redacted] of will achieve first remission. A majority of these patients go on to receive HSCT leaving approximately [redacted] of these patients who are likely to have midostaurin maintenance in the UK (as confirmed by UK clinicians)

treatment as most patients with a *FLT3* mutation undergo HSCT. Therefore, observation (also referred to as "watch and wait") + BSC is the only available option for the majority of patients who have achieved first remission post induction ± consolidation therapy and are not candidates for HSCT.

For this appraisal, NICE has included established clinical management without oral azacitidine which may include a "watch and wait" strategy + BSC, low dose cytarabine or subcutaneous azacitidine and midostaurin as relevant comparators.

With regards to the use of low dose cytarabine and subcutaneous azacitidine, it is important to note that neither of these treatment options are recommended by NICE for the patient population eligible for maintenance treatment with oral azacitidine, nor is their use mentioned or endorsed as maintenance treatments in either the ELN (2017) or BSCH (2006) guidelines.[18, 69]

The established BMS opinion is that neither of these treatments are legitimate comparators for this appraisal. Both treatments have historically been investigated as AML maintenance options in randomised clinical trials, but neither injectable azacitidine nor low dose cytarabine have demonstrated an overall survival benefit versus comparators in the maintenance setting:

- Injectable azacitidine vs observation/no maintenance (HOVON 97 Trial); this RCT demonstrated a significant improvement in disease-free survival (DFS) after maintenance with injectable azacitidine versus observation/no maintenance (64% vs 42% at 1 year; p=0.04). This study did not show a significant OS benefit (84% vs 70% at 1 year, p=0.69).[4]
- Injectable azacitidine vs BSC (QOLESS AZA-AMLE Trial); this small study (27 patients randomised per treatment arm) did not identify statistically significant differences in DFS or OS between injectable azacitidine and BSC.[80]
- Low-dose cytarabine maintenance therapy vs observation (E5483); this trial reported statistically significant improvements for median DFS (7.4 months vs 3.3 months, p=0.084), but not for median OS (10.8 months vs 7.0 months, p=0.492).[81]

To authenticate our position, we sought expert clinical advice from two UK AML clinicians, who unequivocally confirmed that these treatments are not used in UK clinical practice for AML maintenance. The clinical experts could only provide very limited examples where these treatments could be used in situations resembling maintenance treatment, such as those patients whose disease was in partial remission, or patients who showed signs of early relapse. We believe that these situations might be miscategorised as maintenance treatment.

This does not align with the definition of maintenance treatment considered in this appraisal, therefore, low dose cytarabine and subcutaneous azacitidine have been disregarded as relevant comparators to oral azacitidine.

B.1.3.8 Unmet need

The main goal of maintenance treatment in AML patients is to delay relapse and prolong survival.[30] Maintenance treatment is not currently standard of care in the UK. Whilst NICE recommends midostaurin for maintenance treatment, its use is restricted to a small subgroup of *FLT3*-mutation-positive patients who are in remission after previously being treated with midostaurin in combination with chemotherapy agents during induction and consolidation

chemotherapy.[31] Although midostaurin has shown improvements in OS, the pivotal midostaurin study (RATIFY) was not prospectively designed to determine the independent effect of midostaurin as maintenance treatment (i.e. initiated only after achievement of CR following induction/consolidation therapy).[82] Additionally in a subset analysis of data from the RATIFY study, that specifically evaluated the maintenance phase, there were no significant differences between the midostaurin and placebo groups in terms DFS or OS.[83] The study authors concluded that the results of the post hoc analysis did not allow for conclusions on the value of midostaurin as maintenance treatment. For the majority of patients with AML, (i.e. non *FLT3*-mutation-positive), "watch and wait" + BSC is the only available option for those who have achieved first remission post induction ± consolidation therapy and are not candidates for HSCT.

As most patients with AML experience disease relapse after induction chemotherapy (Section B.1.3.4.3), effective maintenance treatment for patients who attain remission (CR/CRi) may play a role in preventing disease relapse and prolonging OS. A substantial unmet need remains for a well-tolerated and easily administered AML maintenance treatment, that significantly prolongs survival among patients with AML who are in remission after IC without compromising HRQoL. Such a therapy could become the new standard of care for maintenance treatment in AML.

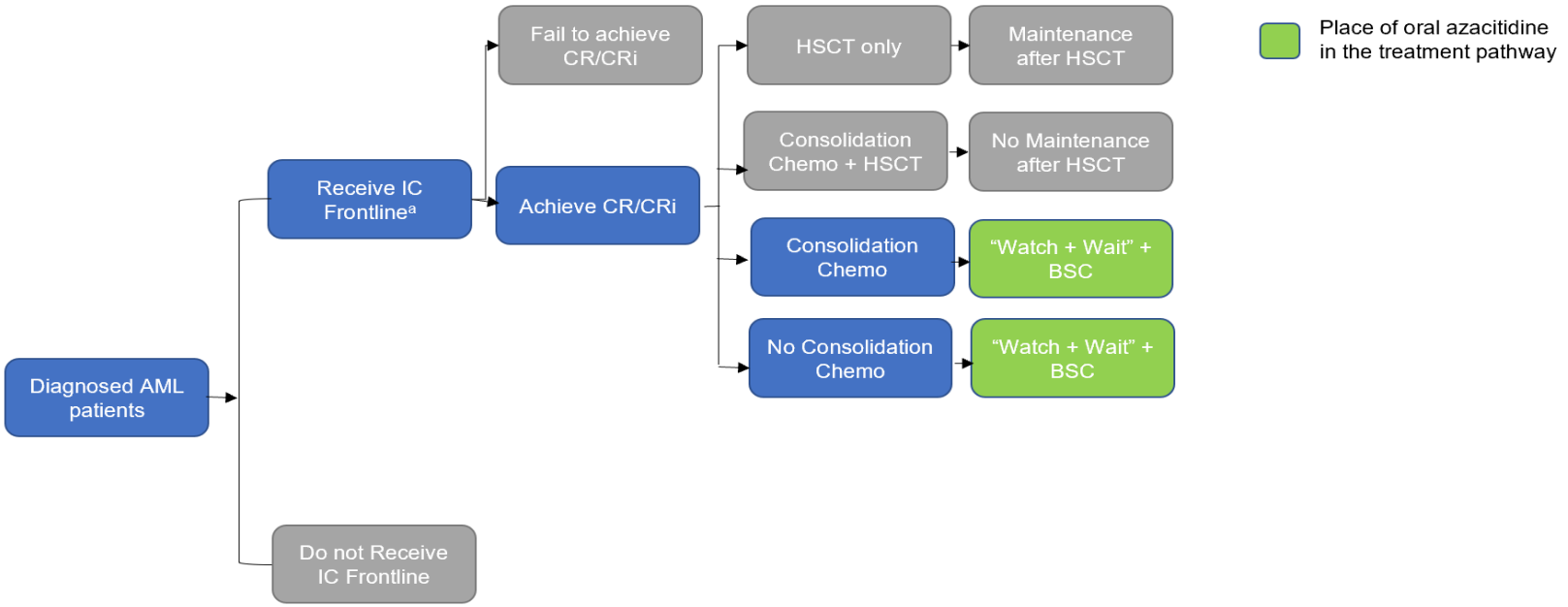
Oral azacitidine is the first and only oral HMA specifically indicated for use as maintenance treatment in adult patients with AML who have achieved CR or CRi following induction therapy with or without consolidation treatment and who are not candidates for, including those who choose not to proceed to, HSCT.[1] Notably, there are no restrictions on the patient population eligible for oral azacitidine treatment in terms of genetic mutations. Oral azacitidine is the first and only maintenance treatment to provide a significant OS and RFS benefit in the AML maintenance setting.[84] In addition, its mode of administration offers the convenience of an oral pill as demonstrated from a discrete choice experiment conducted among patients from Germany and Italy (patients preferred therapies with oral administration once daily for 14 or 21 days per month followed by IV administration 5 days per month, and subcutaneous injection 7 days per month, compared with IV administration in a hospital or clinic 7 days per month).[85] In the context of the COVID-19 pandemic, these benefits of an oral treatment could also translate into a reduction in NHS burden in the long-term beyond the initial months of monitoring, i.e. reducing regular visits to the hospital for treatment. Therefore, as a clinically effective treatment option with a manageable safety profile and its convenient mode of administration, oral azacitidine addresses the high unmet need for patients with AML who have achieved CR/CRi after intensive induction chemotherapy ± consolidation chemotherapy and are not candidates for HSCT.

Maintenance treatment with oral azacitidine represents a new potential therapeutic standard for adult patients with AML in first remission.

B.1.3.8.1 Position of oral azacitidine in the AML treatment pathway

The current treatment pathway reflecting currently available UK and European guidelines.[18, 31, 69, 70] for newly diagnosed AML patients (non *FLT3*-mutation-positive and *FLT3*-mutation-positive) who are fit for intensive induction chemotherapy, and the place of oral azacitidine within this pathway (green boxes) is shown in Figure B.1.5.

Figure B.1.5. Current treatment pathway and place of oral azacitidine



Abbreviations: BSC = best supportive care; IC = intensive chemotherapy; CR = complete remission; CRi = complete remission with incomplete blood count recovery; FLT3 = fms-like tyrosine kinase 3; HSCT = haematopoietic stem cell transplantation.
[a] NICE TA523 recommends the use of midostaurin for patients with a FLT3 mutation, with standard daunorubicin and cytarabine as induction therapy and with high-dose cytarabine as consolidation therapy, these patients would receive midostaurin alone as a maintenance treatment after a complete response and would not undergo a "watch and wait" + BSC strategy.
Source: Current treatment UK and European guidelines[18, 31, 69, 70]

B.1.4 Equality considerations

There are no known equality issues relating to the use of oral azacitidine as maintenance treatment in patients with AML.

B.2 Clinical effectiveness

Summary

- The QUAZAR AML-001 study met its primary outcome based on the primary database lock (15 July 2019) and this effect was maintained in the long term (8 September 2020 data cut)
    - At a median follow-up of 41.2 months (data cut-off date, 15 July 2019) maintenance treatment with oral azacitidine significantly extended OS compared with placebo (p<0.001) (Table B.2.1)
    - In the most recent data cut for OS (data cut-off date, 8 September 2020) at a median follow-up of 51.7 months, oral azacitidine continued to be associated with a significantly longer OS compared with placebo (p<0.001) (Table B.2.1)
- Based on these OS results oral azacitidine meets the EOL criteria i.e. offers an extension to life of >3 months (median OS 24.7 months with oral azacitidine compared with placebo 14.8 months, overall difference 9.9 months) (Table B.2.26)

Table B.2.1. OS results

	15 July 2019 data cut			8 September 2020 data cut		
Parameter	**Oral azacitidine (N=238)**	**Placebo (N=234)**	**Difference (95% CI)**	**Oral azacitidine (N=238)**	**Placebo (N=234)**	**Difference (95% CI)**
Patients with event (death), n (%)	158 (66.4)	171 (73.1)	-	[redacted]	[redacted]	-
Patients censored, n (%)	80 (33.6)	63 (26.9)	-	[redacted]	[redacted]	-
Median OS, months (95% CI)	24.7 (18.7-30.5)	14.8 (11.7-17.6)	9.9 (4.6-15.3)	24.7 (18.7-30.5)	14.8 (11.7-17.6)	9.9 (4.5-15.4)
HR (95% CI)	0.69 (0.55-0.86)[a]		-	0.69 (0.56, 0.86)		-
p-value	0.0009		-	0.0008		-
1-year survival estimate (95% CI)	0.728 [redacted]	0.558 [redacted]	0.170 (0.084-0.256)	0.728 [redacted]	0.558 [redacted]	0.170 (0.084-0.256)
2-year survival estimate (95% CI)	0.506 [redacted]	0.371 [redacted]	0.135 (0.045-0.225)	0.506 [redacted]	0.371 [redacted]	0.135 (0.045-0.225)

Abbreviations: CI = confidence interval; HR = hazard ratio; ITT = intention-to-treat; n = number of patients in the category; N = number of patients evaluable; OS = overall survival

[a]Hazard ratios were not provided in the primary publication, as the proportional hazards assumption appeared to be violated, as indicated by the significant treatment-by-time interaction

- At a median follow-up of 41.2 months (data cut-off date, 15 July 2019):
  - Treatment with oral azacitidine significantly extended RFS (key secondary outcome) compared with placebo, with a clinically meaningful difference in median RFS of 5.3 months (median RFS: 10.2 months vs. 4.8 months; HR 0.65 [95% CI: 0.52-0.81], p<0.001)
  - Subgroup analyses showed that OS and RFS benefits provided by oral azacitidine were consistent across demographic and disease-related subgroups
  - Lower relapse rates were observed in the oral azacitidine group than in the placebo group at six months, one year and two years (0.31 vs. 0.54, 0.53 vs. 0.72, and 0.69 vs. 0.82, respectively)
  - Median time to discontinuation from study treatment was longer with oral azacitidine compared with placebo (11.4 vs. 6.1 months, respectively)
  - Lower treatment discontinuation rates were observed in the oral azacitidine group than in the placebo group at six months, one year and two years ([redacted] respectively)
  - Maintenance treatment with oral azacitidine significantly improved survival without compromising HRQoL, at a level similar to placebo and the general population
  - The percentage of patients hospitalised, the number of hospitalisation events (per person per year), and the number of hospitalisation days was lower with oral azacitidine than placebo (45.8% vs. 50.6%, 0.48 vs. 0.64, and 2872 vs. 3139 days, respectively).[86]
  - Patients treated with oral azacitidine achieved or maintained post-baseline MRD-negative status than patients who received placebo and a survival benefit was observed independently of MRD status at baseline

B.2.1 Identification and selection of relevant studies

A systematic Literature Review (SLR) was conducted to identify and summarise the available randomised controlled trial (RCT) evidence for maintenance treatment options for adult patients (≥18 years) with AML who have achieved CR/CRi and are not candidates for stem cell transplant (SCT). The literature search was originally performed on 18 January 2020 and updated twice, on 19 February 2021 and on 11 June 2021. Full details of the methodology and the results of the SLR are detailed in Appendix D.

B.2.2 List of relevant clinical effectiveness evidence

The SLR included 24 unique studies, of these one unique RCT provided evidence for the efficacy and safety of oral azacitidine, the pivotal Phase 3 QUAZAR AML-001 study.[84] An overview of this pivotal study is provided in Table B.2.2.

Table B.2.2. Clinical effectiveness evidence

Study	QUAZAR AML-001 (NCT01757535)[84]				
Study design	International, multicentre, Phase 3, randomised, two-arm, double-blind, placebo-controlled, parallel group				
Population	Patients with AML in CR/CRi after IC with or without consolidation chemotherapy, and were not candidates for HSCT				
Intervention(s)	Oral azacitidine 300 mg QD + BSC[a]				
Comparator(s)	Matching placebo + BSC				
Indicate if trial supports application for marketing authorisation	Yes	☒	Indicate if trial used in the economic model	Yes	☒
	No	☐		No	☐
Rationale for use/non-use in the model	This study was used in the model as this is the Pivotal Phase 3 study supporting the licensed indication and was used for marketing authorisation submissions. The study provides the primary source of evidence for the clinical efficacy and safety of oral azacitidine relevant to the decision problem.				
Reported outcomes specified in the decision problem[b]	• OS • RFS • Adverse effects of treatment • HRQoL				
All other reported outcomes[b]	• Time to relapse from CR/CRi • Time to discontinuation from treatment • Healthcare resource utilisation				

Abbreviations: AML = acute myeloid leukaemia; BSC = best supportive care; CR = complete remission, CRi = complete remission with incomplete blood count recovery, ESA = erythropoiesis-stimulating agent; G-CSF = granulocyte colony-stimulating factor; HRQoL = health-related quality of life; HSCT = haematopoietic stem cell transplant; IC = intensive chemotherapy; OS = overall survival; QD = once daily; RBC = red blood cell; RFS = relapse-free survival

[a] Throughout the treatment period of the QUAZAR AML-001 study, patients in both the placebo and oral azacitidine treatment groups were permitted to receive BSC, which may have included RBC and platelet transfusions; use of an ESA; antibiotic, antiviral, and/or antifungal therapy; nutritional support; and/or G-CSF for patients experiencing neutropenic infections.

[b] All outcomes are used in the model.

Source: Wei et al., 2020[84]

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

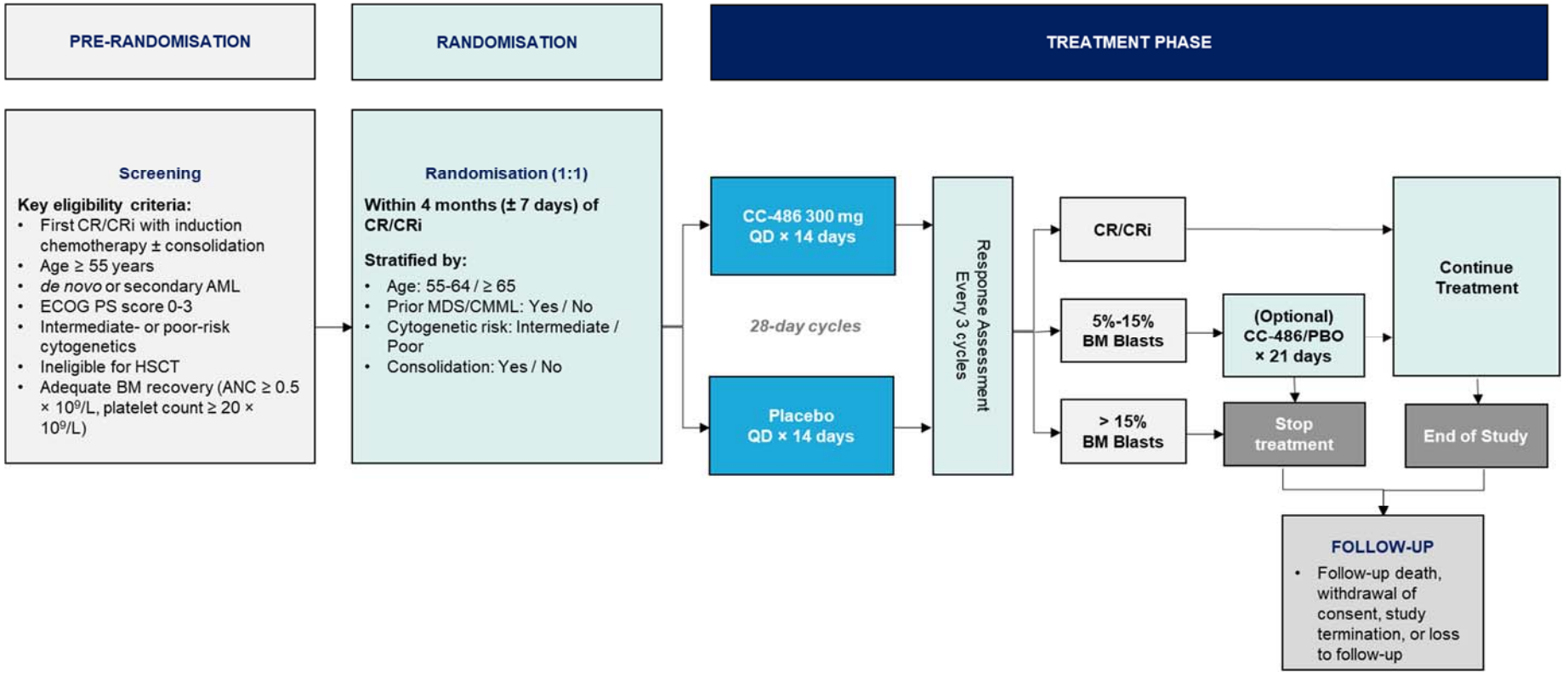
QUAZAR AML-001 is the pivotal study supporting this submission, a summary of the study design and methodology is provided in this section and results are presented in Section B.2.6.

B.2.3.1 Study design

QUAZAR AML-001 is an ongoing, international, multicentre, Phase 3, randomised, two-arm, double-blind, placebo-controlled, parallel group study of oral azacitidine + BSC vs. placebo + BSC as maintenance treatment in patients with AML who have achieved CR or CRi after induction with IC with or without consolidation chemotherapy, and were not candidates for HSCT.[84, 87]

An overview of the study design is presented in Figure B.2.1. The study consisted of 4 phases; Pre-randomisation Phase (Screening Phase), Treatment Phase, Follow-up Phase and Extension Phase (EP).

Figure B.2.1. Study design, QUAZAR AML-001 study



Abbreviations: AML = acute myeloid leukaemia; ANC = absolute neutrophil count; BM = bone marrow; CC-486 = azacitidine; CMML = chronic myelomonocytic leukaemia; CR = complete remission; CRi = complete remission with incomplete blood count recovery; ECOG PS = Eastern Cooperative Oncology Group performance status; HSCT = haematopoietic stem cell transplant; IWG = International Working Group; MDS = myelodysplastic syndromes; QD = once daily.
Source: Supplementary appendix to Wei et al., 2020[88]

B.2.3.1.1 Pre-randomisation Phase (Screening Phase)

All screening procedures were performed within the 28 days before randomisation, such as confirmation of diagnosis, verification of CR/CRi status (by central pathology and cytogenetics review), documentation of induction and consolidation therapies, and determination of transplant eligibility.[87] See section **Error! Reference source not found.** for key eligibility criteria and a full list of inclusion and exclusion criteria is provided in Appendix L.

B.2.3.1.2 Randomisation and Treatment Phase

After confirmation of eligibility at screening, patients were randomised 1:1 to receive 300 mg oral azacitidine QD or matching placebo for the first 14 days of each 28-day cycle. Treatment was assigned by a central randomisation procedure using an Interactive Voice Response System (IVRS). Randomisation must have occurred within 4 months (± 7 days) of achieving the first CR/CRi[c] for two reasons:

1. To allow time for eligible patients to receive (and recover from) up to four cycles of consolidation chemotherapy.
2. To avoid enriching the study population with patients with better prognoses.

Randomisation was stratified by the following key prognostic factors:

- Age at the time of induction therapy (55 to 64 years or ≥65 years)
- Prior history of MDS or CMML (yes or no)
- Cytogenetic risk status at the time of induction therapy (intermediate or poor risk)
- Receipt of consolidation therapy (yes or no)

After randomisation, no crossover between the treatment arms was permitted. Patients, investigators, site staff and clinical and medical personnel were unaware of treatment assignments until study closure and database lock.

Bone marrow aspirate (or biopsy if adequate aspirate was not attainable) samples during the double-blind treatment phase were collected on Day 1 (± 7 days) of Cycles 3, 6, 9, 12, 15, 18, 21, 24, 30, 36, and the Treatment Discontinuation Visit. After Cycle 36, bone marrow aspiration collection and evaluation were performed if clinically indicated at the discretion of the investigator.

Additional bone marrow samples were collected as clinically indicated. A bone marrow biopsy was performed if adequate aspirate was not attainable. When a bone marrow sample was collected, a peripheral blood smear was also prepared.

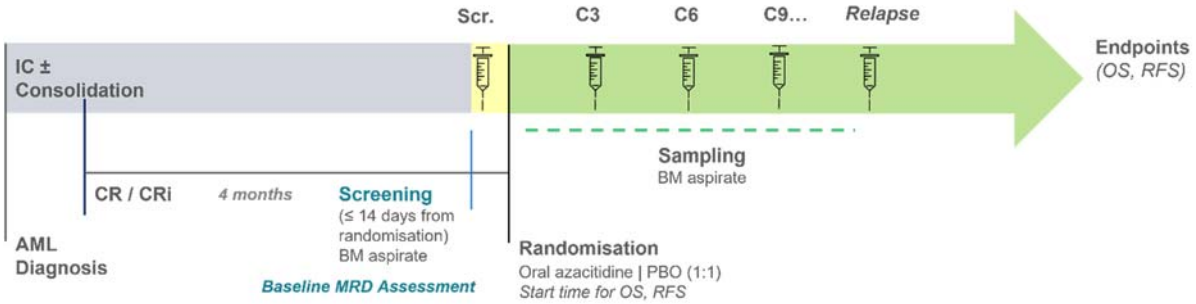
CR/CRi status and disease relapse were defined according to the International Working Group (IWG) 2003 response criteria for AML.[84] A central review of all bone marrow aspirates, bone marrow biopsies, and peripheral blood smears was conducted by an independent pathologist

[c] As evidenced by the following: 5% blasts in bone marrow, absence of blasts with Auer rods, absence of extramedullary disease, independent of blood transfusions, platelet count ≥100 x 10^9/L (for CR) and ANC <1.0 x 10^9/L or platelet count <100 x 10^9/L for (CRi).

who was blinded to treatment to confirm CR/CRi status at screening and during treatment. Status assessments for maintenance of CR/CRi occurred every three cycles up to Cycle 24, every six cycles from Cycles 24 to 36, at the investigators discretion thereafter, and at the treatment discontinuation visit (regardless of the number of cycles completed).[84]

The presence of MRD was assessed by flow cytometry, with the use of a leukaemia-associated immunophenotype (LAIP)–based "different-from-normal" method with a 0.1% threshold for MRD.[84] Bone marrow aspirates were collected at screening (i.e. after CR/CRi and any consolidation), at cycles 3, 6, 9, 12, 15, 18, 21, 24, 30 and 36 (and as clinically indicated), until time of relapse (Figure B.2.2).

Figure B.2.2. MRD assessments, QUAZAR AML-001 study



Abbreviations: AML = acute myeloid leukaemia; BM = bone marrow; CR = complete remission; CRi = complete remission with incomplete blood count recovery; IC = intensive chemotherapy; MRD = measurable residual disease; OS = overall survival; PBO = placebo; RFS = relapse-free survival; Scr. = screening
Source: Adapted from Roboz et al., 2020[89]

Study assessments were conducted during the Treatment Phase, included monitoring for AEs, maintenance of CR/CRi or relapse, completion of PROs for HRQoL, utilisation of healthcare resources, and evaluation of physical/clinical status.[87]

Safety assessments consisted of evaluation of AEs and SAEs (and concomitant medication/therapies used to treat them), second primary malignancy (SPM), haematology/serum chemistry parameters, body weight measurements, physical examination findings, and pregnancy testing (for females of childbearing potential). Urinalysis and electrocardiogram (ECG) were repeated whenever clinically indicated during the double-blind Treatment Phase. AML relapse was not considered an AE for the purposes of the safety analysis.[87]

The FACIT-Fatigue Scale and the EQ-5D-3L questionnaires were ideally completed prior to dosing and prior to interaction with study personnel on Day 1 of every cycle, beginning on Day 1 of Cycle 1 and the Treatment Discontinuation Visit.[87]

Information on HCRU (medications, hospitalisations, clinic visits, medical/diagnostic procedures, and treatments received) was collected after a patient signed informed consent through 28 days after the last dose of study treatment or until the date of the last study visit.[87]

Patients with subsequent evidence of AML relapse (≥5% and ≤15% blasts in the peripheral blood or bone marrow) had the option to continue treatment with an extended dose schedule to 300 mg QD for 21 days, provided it was in the best interest of the patient to do so as judged by the INV. Similarly, if patients experienced toxicity considered possibly related to treatment,

dosing with investigational product could be interrupted or delayed, reduced to 200 mg QD for 14 days or 200 mg QD for 7 days in a stepwise fashion.[84, 87]

Patients were discontinued from treatment following AML relapse when they had >15% blasts in the bone marrow or peripheral blood, which was attributable to relapse following CR/CRi, and not attributable to any other cause (e.g. bone marrow regeneration after consolidation therapy).[87]

Throughout the Treatment Phase, patients in both the placebo and oral azacitidine treatment groups were permitted to receive BSC, which may have included RBC and platelet transfusions; use of an ESA; antibiotic, antiviral, and/or antifungal therapy; nutritional support; and/or G-CSF; for patients experiencing neutropenic infections.[87] The inclusion of BSC in the study design minimised the risk of providing patients with inadequate care and was consistent with current practice for this cohort of patients when the study was conducted.

B.2.3.1.3 Follow-up Phase

During the Follow-up Phase, all patients who discontinued study treatment underwent discontinuation visit procedures at the time they left the study. Patients had a follow-up visit for collection of AEs up to 28 days after the last dose of study treatment or up to the treatment discontinuation visit, whichever was longer. Patients were subsequently followed for survival every month for the first year and then every three months until death, withdrawal of consent for further follow-up, study end, or loss to follow-up.[87]

B.2.3.1.4 Extension Phase

The EP allowed patients receiving oral azacitidine who were demonstrating clinical benefit as assessed by the investigator to continue to receive oral azacitidine after unblinding by the Sponsor until they met the criteria for study discontinuation or until oral azacitidine became commercially available. Patients who discontinued treatment but remained in the study were (or are being) followed for survival subject to additional consent.[87]

B.2.3.2 Study methodology

B.2.3.2.1 Summary of study methodology

A summary of the QUAZAR AML-001 study methodology is provided in Table B.2.3.

Table B.2.3. Summary of trial methodology, QUAZAR AML-001 study

Study design	International, multicentre, Phase 3, randomised, two-arm, double-blind, placebo-controlled, parallel group
Study objective	**Primary objective:** • Evaluate whether maintenance treatment with oral azacitidine improved OS compared with placebo **Secondary objectives:** • Determine the effect of oral azacitidine on RFS, safety and tolerability, HRQoL and HCRU (hospitalisations, medications, clinic visits, medical/diagnostic procedures, and treatment for AEs)
Locations	Conducted at 148 sites in 23 countries (Australia, Austria, Belgium, Brazil, Canada, Czech Republic, Finland, France, Germany, Ireland, Israel, Italy,

	Korea, Lithuania, Mexico, Poland, Portugal, Russia, Spain, Taiwan, Turkey, United Kingdom [N=8], and United States)
Study status	Ongoing • First patient, first visit: 10 May, 2013 • First data cut-off date: 15 July, 2019 (All outcomes) • Second data cut-off date: 8 September, 2020 (OS only)
Study treatments	Treatment was assigned by a central randomisation procedure using an IVRS. Patients were randomly assigned in a 1:1 ratio to receive: • Oral azacitidine tablets 300 mg QD (N=238) or • Matching placebo (N=234)
Blinding	Patients, investigators, site staff and clinical and medical personnel were unaware of treatment assignments until study closure and database lock for the primary analysis (data cut 15 July 2019). The EP until the most recent data cut (8 September 2020) was unblinded.
Concomitant medication	**Permitted:** • BSC (including, but not limited to RBC and platelet transfusions, use of an ESA, antibiotic, antiviral, and antifungal therapy, nutritional support, and G-CSFs for patients experiencing neutropenic infections, pre-treatment or post-treatment with a serotonin (5-HT3) receptor antagonist (or other anti-emetic medication) **Disallowed:** • Cytotoxic chemotherapeutic agents or experimental agents, romiplostim and other TSAs (e.g. interleukin-11), hydroxyurea, lenalidomide, pomalidomide, thalidomide, arsenic trioxide, interferon and retinoids
Primary outcome	• OS
Secondary outcomes	**Key secondary outcome:** • RFS **Other secondary efficacy outcomes:** • Time to relapse from CR/CRi • Time to discontinuation from treatment • HRQoL assessment (FACIT-Fatigue Scale, EQ-5D-3L) • HCRU assessment
Exploratory outcomes	• Flow cytometric analysis of hematopoietic cell immunophenotypes (MRD analysis) • Analysis of genetic alterations, including gene sequencing for recurrent gene aberrations in AML • HRQoL exploratory assessment • Biomarker outcomes
Pre-planned subgroups	Additional pre-planned exploratory subgroup analyses were performed where an adequate number of patients were available to allow meaningful interpretation. Analyses were performed within the following subgroups for the OS and RFS outcomes:

	• Age (<65, ≥65, ≥75 years) • Gender • Race • CR/CRi status at randomisation and at first achieving response • CR/CRi status at randomisation and use of consolidation • Prior history of MDS or CMML • Cytogenic risk category at induction therapy • MRD status at screening (prior to randomisation) • CR/CRi status at randomisation and MRD status at screening • Consolidation therapy following induction • Geographic region • ECOG PS • WHO AML classification • Type of first line subsequent therapy

Abbreviations: AEs = adverse event; AML = acute myeloid leukaemia; BSC = best supportive care; CMML = chronic myelomonocytic leukaemia; CR = complete remission; CRi = complete remission with incomplete blood count recovery; ECOG PS = Eastern Cooperative Oncology Group performance status; EP = extension phase; EQ-5D-3L = European Quality of Life–5 Dimensions; ESA = erythropoiesis stimulating agent; FACIT = Functional Assessment of Chronic Illness Therapy; G-CSF = granulocyte colony-stimulating factor; HCRU = health care resource utilisation; HMA = hypomethylating agent; HRQoL = health-related quality of life; INV, investigator; IVRS = interactive voice response system; MRD = measurable residual disease; OS = overall survival; QD = once daily; RBC = red blood cell; RFS = relapse-free survival; TSA = thrombopoiesis-stimulating agent; WHO = World Health Organisation
Source: Wei et al., 2020[84]; QUAZAR AML-001 CSR (Data on File)[87]

B.2.3.2.2 Key eligibility criteria

Key inclusion and exclusion criteria are provided in Table B.2.4. A full list of inclusion and exclusion criteria is provided in Appendix L.

Table B.2.4. Key inclusion and exclusion criteria, QUAZAR AML-001 study

<table>
<tr><td>Inclusion criteria</td><td>
• Men and women aged ≥55 years
• Newly diagnosed, histologically confirmed de novo acute AML or AML secondary to MDS or CMML
• Had undergone induction with IC with or without consolidation therapy
• Achieved first CR or CRi status within 4 months (± 7 days) prior to randomisation, as evidenced by the criteria in the Table below.
<table>
<tr><th>CR</th><th>CRi</th></tr>
<tr><td>• < 5% blasts in bone marrow
• absence of blasts with Auer rods
• absence of extramedullary disease
• independent of blood transfusions
• peripheral neutrophil count >1.0 x 10^9/L
• platelet count ≥ 100 x 10^9</td><td>• < 5% blasts in bone marrow
• absence of blasts with Auer rods
• absence of extramedullary disease
• independent of blood transfusions
• peripheral neutrophil count <1.0 x 10^9/L platelet count < 100 x 10^9</td></tr>
</table>
• ECOG PS of 0–3
• Adequate bone marrow function based on ANC ≥0.5 x 10^9/L and platelet count ≥20 x 10^9/L
</td></tr>
</table>

Exclusion criteria	• Candidate for HSCT at screening • AML associated with inv(16), t(8;21), t(16;16), t(15;17), or t(9;22) karyotypes or molecular evidence of such translocations • Prior bone marrow or SCT • Achieved CR/CRi following therapy with hypomethylating agents

Abbreviations: AML = acute myeloid leukaemia; ANC = absolute neutrophil count; CMML = chronic myelomonocytic leukaemia; CR = complete remission; CRi = complete remission with incomplete blood count recovery; ECOG PS = Eastern Cooperative Oncology Group performance status; IC = intensive chemotherapy; MDS = myelodysplastic syndromes; SCT = stem cell transplantation
Source: QUAZAR AML-001 CSR (Data on File)[87]; Supplementary appendix to Wei et al., 2020[88]

B.2.3.2.3 Outcome definitions

A summary of study outcome definitions is provided in Table B.2.5.

Table B.2.5. Summary of outcome definitions in the QUAZAR AML-001 study

Category	Outcome	Definition or analysis
Primary efficacy outcome	OS[a]	The number of days from the date of randomisation until the date of death from any cause, calculated as (date of death – date of randomisation + 1). Patients surviving at the end of the follow-up period or who were lost to follow-up were censored at the date last known to be alive. For patients who withdrew consent, the last date known alive was considered the date of consent withdrawal from the study.
Key secondary efficacy outcome	RFS[a]	The time from the date of randomisation to the date of documented relapse or death, whichever occurred first. Patients who were still alive without documented relapse, or who were lost to follow-up or withdrew consent without documented relapse, were censored at the date of their last response assessment. Documented relapse was defined as the earliest date of any of the following (according to IWG for AML criteria): • ≥5% BM blasts from the central pathology report; • The appearance of >0% blasts in the peripheral blood with a later BM confirmation (BM blasts ≥5%) within 100 days; or • At least two peripheral blasts ≥5% within 30 days
Additional secondary efficacy outcomes	Time to relapse[a]	The time from the date of randomisation to the date of documented relapse. Estimates of relapse rates at different times from randomization were based on the cumulative incidence function from a competing risk analysis with death as a competing risk of relapse from CR/CRi; this differs from the censoring approach used for RFS.

	Time to discontinuation from treatment[a]	The time from the date of randomisation to the date of discontinuation from investigational product.
	HRQoL: FACIT Fatigue scale[b]	Analysed as both change from baseline and the proportion of patients with clinically meaningful improvement based on a prespecified MID. A ≥3 point change from baseline was used to define clinically meaningful improvement.[c]
	HRQoL: EQ-5D-3L health utility index[b]	Analysed as both change from baseline and the proportion of patients with clinically meaningful improvement based on a prespecified MID. A 0.08- and 0.10-point or greater change from baseline was used to define clinically meaningful improvement.[d]
	HCRU: Hospitalisations	Analysed as the total number of hospitalisations, total number of days hospitalised, rate of hospitalisations, days of hospitalisation per person-year of exposure, and associated relative risk of hospitalisation (with 95% CI).

Abbreviations: AML = acute myeloid leukaemia; BM = bone marrow; CI = confidence interval; EQ-5D-3L = European Quality of Life - 5 dimensions, 3 levels; FACIT = Functional Assessment of Chronic Illness Therapy; HCRU = healthcare resource utilisation; HRQoL = health-related quality of life; ITT = intention-to-treat; IWG = International Working Group; MID = minimally important difference; OS = overall survival; RFS = relapse-free survival

[a] Analysed using the ITT population.

[b] Analysed using the HRQoL-evaluable population, defined as all randomised patients who had a valid (i.e. not missing) assessment at baseline (i.e. Cycle 1 Day 1) and at least one valid post-baseline assessment.

[c] Prespecified MID from Cella et al., 2002[90]

[d] Prespecified MID from Kvam et al., 2001[91]

Source: QUAZAR AML-001 CSR (Data on File)[87]

B.2.3.2.4 Patient disposition

A total of 472 patients were randomly assigned to receive oral azacitidine (238 patients) or placebo (234 patients). Three patients (2 in the oral azacitidine group and 1 in the placebo group) did not receive oral azacitidine or placebo and were excluded from analyses.[84] A CONSORT diagram of patient disposition for the primary database lock (15 July 2019) is provided in Appendix D.

B.2.3.2.5 Demographics and baseline characteristics

A summary of baseline demographics is provided in Table B.2.6. Overall, the baseline demographic and disease characteristics between the two treatment groups were comparable.

The median age for all patients was 68.0 years, with 60.6% of patients in the age range ≥ 65 to <75 years; 11.0% of subjects were ≥75 years.[92] The proportion of male patients was 50.0% in the oral azacitidine group, and 54.0% in the placebo group. Most (>84%) patients in each treatment group were White. The median weight (73.0 kg), height (166.0 cm), and body mass index (BMI), (25.8 kg/m^2) were also similar between treatment groups. The majority of patients (67%) were from Europe.[87, 92]

Table B.2.6. Baseline demographics, QUAZAR AML-001 study (ITT population)

Parameter	Oral azacitidine (N=238)	Placebo (N=234)	Total (N=472)
Age (years)			
Median (range)	68 (55–86)	68 (55–82)	68 (55–86)
Age category, n (%)			
≥55 to <65 years	66 (28)	68 (29)	134 (28)
≥65 to <75 years	144 (61)	142 (61)	286 (61)
≥75 years	27 (11)	24 (10)	51 (11)
≥85 years	1 (0)	0	1 (0)
Sex, n (%)			
Male	118 (50)	127 (54)	245 (52)
Female	120 (50)	107 (46)	227 (48)
Race, n (%)			
White	216 (90.8)	197 (84.2)	413 (87.5)
Black or African-American	2 (0.8)	6 (2.6)	8 (1.7)
Asian	6 (2.5)	20 (8.5)	26 (5.5)
Other	12 (5.0)	11 (4.7)	23 (4.9)
Missing	2 (0.8)	0	2 (0.4)
Ethnicity, n (%)			
Hispanic/Latino	20 (8.4)	14 (6.0)	34 (7.2)
Non-Hispanic/Latino	196 (82.4)	202 (86.3)	398 (84.3)
Unknown	22 (9.2)	18 (7.7)	40 (8.5)
Geographical region[a], n (%)			
North America	37 (16)	42 (18)	79 (17)
Europe (UK = [redacted] patients)	167 (70)	147 (63)	314 (67)
Asia	6 (3)	17 (7)	23 (5)
Australia	26 (11)	23 (10)	49 (10)
South America	2 (1)	5 (2)	7 (1)

Abbreviations: ITT = intention-to-treat; max = maximum; min = minimum; n = number of patients in the category; N = number of patients evaluable; SD = standard deviation.
[a] North America includes Canada, Mexico, and the United States; Asia includes South Korea and Taiwan; Australia includes Australia; Europe includes Austria, Belgium, Czech Republic, Finland, France, Germany, Ireland, Israel, Italy, Lithuania, Poland, Portugal, the Russian Federation, Spain, and Turkey; South America includes Brazil.
Source: QUAZAR AML-001 CSR (Data on File)[87]; Supplementary appendix to Wei et al., 2020[88]; ClinicalTrials.gov[93]; FDA, 2020[92]

A summary of baseline disease characteristics is provided in Table B.2.7. Baseline disease characteristics were generally similar between treatment groups. For the ITT population, 91% of patients had de novo AML. The median time since original AML diagnosis to randomisation was 4.2 months (range: 1.4 to 10.9), with 8.3% having a prior history of myelodysplastic syndromes (MDS) / chronic myelomonocytic leukaemia (CMML). The majority of patients had an ECOG performance status score of 0 (48.1%) or 1 (43.9%). Cytogenetic risk category at diagnosis was intermediate for 86.0% of patients. Approximately half (51.7%) of patients were MRD negative at randomisation and 46.4% were MRD positive. Following induction therapy, 80.1% of patients received consolidation therapy, with most receiving 1 cycle (44.9%) or 2 cycles (31.1%). Patients were ineligible for transplant primarily due to age (64.8%), comorbidities (21.6%), and no available donor (15.3%). All patients in the ITT population achieved CR or CRi after induction therapy, 96.2% were in CR or CRi at randomisation, and the median time from first achieving CR or CRi to randomisation was 85.0 days.[87]

At least 1 subsequent AML therapy[d] was reported for 57.6% and 72.6% of subjects in the oral azacitidine and placebo groups, respectively. The most frequently reported (≥ 10% in the either group) subsequent AML therapies were in the Anatomical Therapeutic Chemical (ATC) classes of antineoplastic and immune modulating agents, specifically, cytarabine, fludarabine, azacitidine, hydroxycarbamide and idarubicin (Table B.2.7).[87, 92]

Table B.2.7. Baseline disease characteristics, QUAZAR AML-001 study (ITT population)

Parameter	Oral azacitidine (N=238)	Placebo (N=234)	Total (N=472)
Initial AML classification, n (%)			
AML with recurrent genetic abnormalities	39 (16)	46 (20)	85 (18)
AML with myelodysplasia - related changes	49 (21)	42 (18)	91 (19)
Therapy-related myeloid neoplasms	2 (1)	0	2 (0.4)
AML not otherwise specified	148 (62)	145 (62)	293 (62)
Missing	0	1 (0.4)	1 (0.2)
Type of AML, n (%)			
Primary (de novo)	213 (89)	216 (92)	429 (91)

[d]All subsequent therapies for AML as documented on the Case Report Form were included in the analysis. All AML therapies were considered disease modifying, except for hydroxycarbamide.

Secondary	25 (11)	18 (8)	43 (9)
Time since original AML diagnosis (months) to randomisation			
Median (range)	4.2 (1.5–9.2)	4.2 (1.4–10.9)	4.2 (1.4–10.9)
Time from CR/CRi to randomization, days			
Median (range)	84 (7–154[b])	86 (7–263[b])	N/A
Prior history of MDS/CMML, n (%)			
Primary	20 (8)	17 (70)	37 (8)
Secondary	0	0	0
Missing	2 (1)	0	2 (0.4)
ECOG performance status, n (%)			
Grade 0	116 (49)	111 (47)	227 (48)
Grade 1	101 (42)	106 (45)	207 (44)
Grade 2–3	21 (9)	17 (7)	38 (8)
Cytogenetic risk category defined by NCCN at diagnosis, n (%)			
Intermediate	203 (85)	203 (87)	406 (86)
Poor	35 (15)	31 (13)	66 (14)
MRD status at randomisation[a], n (%)			
Negative	133 (56)	111 (47)	244 (52)
Positive	103 (43)	116 (50)	219 (46)
Missing	2 (1)	7 (3)	9 (2)
Reason ineligible for transplant[b], n (%)			
Age	154 (65)	152 (65.)	306 (65)
Comorbidities	52 (22)	50 (21)	102 (22)
Performance Status	14 (6)	9 (4)	23 (5)
Not acceptable or available donor	37 ()	35 (15.0)	72 (15)
Patient decision	19 (8)	32 (14)	51 (11)
Unfavourable cytogenetics	6 (3)	10 (4)	16 (3)
Other	28 (12)	21 (9)	49 (10)
Received subsequent HSCT	15 (6.3)	32 (13.7)	47 (10.0)

Received consolidation therapy following induction therapy			
Yes	186 (78)	192 (82)	378 (80)
1 Cycle	110 (46)	102 (44)	212 (45)
2 Cycles	70 (29)	77 (33)	147 (31)
3 Cycles	6)	13 (6)	19 (4)
4 Cycles	0	0	0
No	52 (22)	42 (18)	94 (20)
Response achieved after induction therapy (with or without consolidation therapy), n (%)			
CR	187 (79)	197 (84)	384 (81)
CRi	51 (21)	37 (16)	88 (19)
CR/CRi status at randomisation[c], n (%)			
CR	183 (77)	177 (76)	360 (76)
CRi	50 (21)	44 (19)	94 (20)
Not in CR/CRi	5 (2)	11 (5)	16 (3)
Missing	0	2 (1.0)	2 (0.4)
Time from start of induction therapy to randomisation, months			
Median (range)	4.0 (1.4–8.8)	4.0 (1.3–15.1)	4.0 (1.3–15.1)
Time from induction therapy to first achieving CR/CRi, days			
Median (range)	36.0 (13.0–242.0)	35.0 (14.0–455.0)	35.0 (13.0–455.0)
Time since first achieving CR/CRi to randomisation, days			
Median (range)	84.5 (7.0–154.0)	86.0 (7.0–263.0)	85.0 (7.0–263.0)
Bone marrow blasts, %			
Median (range)	2.0 (0.0– 5.0)	2.0 (0.0–6.5)	2.0 (0.0–6.5)
Peripheral blood blasts, %			
Median (range)	0.0 (0.0–2.0)	0.0 (0.0–2.0)	0.0 (0.0–2.0)
ATC Dictionary Level Preferred name[d], n (%)			
Subjects with at least 1 subsequent AML therapy	137 (57.6)	170 (72.6)	307 (65.0)

Intensive chemotherapy	69 (29)	88 (38)	157 (33)
Low intensity therapy	94 (40)	110 (47)	204 (43)
Other	[redacted]	[redacted]	[redacted]
Missing	[redacted]	[redacted]	[redacted]
Subsequent AML therapies reported for ≥ 10% of subjects in either treatment group, n (%)			
Antineoplastic and immunomodulating agents	[redacted]	[redacted]	[redacted]
Cytarabine	83 (34.9)	92 (39.3)	175 (37.1)
Fludarabine	32 (13.4)	48 (20.5)	80 (16.9)
Azacitidine	31 (13.0)	47 (20.1)	78 (16.5)
Hydroxycarbamide	28 (11.8)	34 (14.5)	62 (13.1)
Idarubicin	20 (8.4)	33 (14.1)	53 (11.2)

Abbreviations: AML = acute myeloid leukaemia; ANC = absolute neutrophil count; ATC= Anatomical Therapeutic Chemical; BM = bone marrow; CMML = chronic myelomonocytic leukaemia; CR = complete remission; CRi = complete remission with incomplete blood count recovery; ECOG = Eastern Cooperative Oncology Group; HSCT = haematopoietic stem cell transplant; IWG = International Working Group; MDS = myelodysplastic syndrome; MRD = measurable residual disease; max = maximum; min = minimum; n = number of patients in the category; N = number of patients evaluable; NCCN = National Comprehensive Cancer Network; SD = standard deviation; WHO = World Health Organisation.
[a] During the screening period.
[b] A patient may have had more than one reason.
[c] CR/CRi at randomisation was programmatically derived based on IWG for AML response criteria using BM data collected during screening, and ANC and platelets closest to randomisation date. For a patient with BM blasts <5%, and both ANC <1.0 x 10^9/L and platelet count <100 x 10^9/L, the patient was considered not in CR/CRi
[d] Coded using WHODrug Dictionary version March 2019. A subject with multiple occurrences of a drug class or drug preferred name is counted only once in the specific ATC classification or preferred name, respectively
Note: time interval in days was calculated as the difference between the randomisation date and the date of interest (e.g. date of original AML diagnosis) plus one day. Time interval presented in months is transformed from days to months by using the conversion formula: months = days/30.4375.
Source: QUAZAR AML-001 CSR (Data on File)[87]; Supplementary appendix to Wei et al., 2020[88]; Ravandi et al., 2021[94]

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

B.2.4.1 Analysis populations in the QUAZAR AML-001 study

Analysis sets in the QUAZAR AML-001 study included the intention-to-treat (ITT) population, the safety population and the modified intent-to-treat (mITT) population and are outlined in Table B.2.8.

Table B.2.8. Analysis populations in the QUAZAR AML-001 study

Analysis Population	Oral azacitidine (N=238) n (%)	Placebo (N=234) n (%)	Total (N=472) n (%)	Definition
ITT population	238 (100.0)	234 (100.0)	472 (100.0)	The ITT population included all randomised patients, regardless of whether they received study treatment; this population was used for analyses of the primary and secondary efficacy endpoints (other than HRQoL endpoints). Patients were analysed based on randomised treatment group as assigned by the IVRS.
Safety population	236 (99.2)	233 (99.6)	469 (99.4)	The safety population included all randomised patients who received at least one dose of study treatment; this population was used for drug exposure and all safety analyses unless otherwise specified. Patients were analysed based on the initial treatment received.
mITT population	223 (93.7)	217 (92.7)	440 (93.2)	The mITT population included all patients who met the inclusion/exclusion criteria, experienced no protocol violations during the study, and received a minimum of one cycle of treatment. This population was used for sensitivity analyses of primary and secondary efficacy endpoints. Patients were analysed based on randomised treatment group.

Abbreviations: HRQoL = health-related quality of life; ITT = intention-to-treat; IVRS = interactive voice response system; mITT = modified intent-to-treat.
Source: Wei et al., 2020[84]; EMA/308711/2021[95]

A summary of statistical analyses is provided in **Error! Not a valid bookmark self-reference.**.

Table B.2.9. Summary of statistical analyses, QUAZAR AML-001 study

Hypothesis objective	• The null hypothesis for testing the primary efficacy outcome is that the OS distributions for oral azacitidine and placebo are equivalent
Statistical analysis	**Primary outcome (OS)[a]** • KM methods were used to estimate the survival distribution functions for each treatment group • Survival distributions were compared using a stratified log-rank test, stratifying by age at time of induction therapy, prior history of MDS, cytogenic risk category, received consolidation therapy following induction therapy • A stratified Cox proportional hazards model was used to estimate the hazard rate ratio with interaction terms of treatment and time and with a p-value of 0.006 • CIs for survival estimates at 6 months, 1-year and 2-years were calculated with Greenwood's variance formula • A sequential gate-keeping approach was used to control the overall type 1 error in order to perform hypothesis testing on multiple outcomes, OS was tested first at the two-side 0.05 significance level • Other than the pre-specified sequential testing of OS and RFS, no additional alpha adjustments for multiplicity were made **Key secondary outcome (RFS)[a]** • RFS was analysed using the same methods as those for OS • In order to preserve the overall alpha level at 0.05 across the OS and RFS outcomes, formal statistical inference for the RFS analyses can only be made if superiority of oral azacitidine is demonstrated for OS, at the two-sided 0.05 significance level **Additional time-to-event secondary efficacy outcomes (time to relapse and time to discontinuation from treatment)[a]** • Time-to-event secondary efficacy outcomes were analysed similarly to the primary outcome without stratification • KM methods were used to estimate time-to-event curves, unless otherwise specified **HRQoL secondary efficacy outcomes (FACIT-Fatigue and EQ-5D-3L)[b]** • The secondary HRQoL change from baseline outcomes were analysed using ANCOVA and longitudinally using a MMRM method • MID, which is the threshold of a clinically meaningful difference for a given scale, was used to determine whether or not the between-group difference was considered clinically meaningful in the MMRM analysis

	• If the lower bound of the two-sided 95% CI of the between-group difference in the overall Least Squares Mean changes from baseline was greater than the MID, then the difference (or worsening in this case) was not considered clinically meaningful • A change from baseline of ≥3 points was used to define clinically meaningful improvement and worsening at the individual level for the FACIT-F Scale • For the EQ-5D-3L health utility index, a 0.08- and 0.10-point or greater change from baseline was used to define clinically meaningful improvement and worsening, respectively • The CMH test, stratified by levels of randomisation factors, was used to compare proportions of patients with clinically meaningful improvement/worsening between treatment groups • The p-values for all HRQoL analyses were considered descriptive
Sample size, power calculation	• The equality of the OS curves were compared between the oral azacitidine and placebo treatment groups using a stratified log-rank test • Assuming a median OS of 16 months in the placebo group, a median OS of 22.9 months in the oral azacitidine group (43% improvement), and a study duration of 60 months with a drop-out rate of 5% from both treatment groups, over the duration of the study, the design requires 330 deaths and approximately 460 patients (230 per treatment group) to be randomised in order achieve at least 90% power to detect a constant HR of 0.70 and demonstrate a statistically significant difference in OS • It was assumed that the OS distribution was exponential with a constant failure (hazard) rate and that accrual was non-uniform during an accrual period of 36 months with 25% of the patients accrued during each of the first 2 years of enrolment (50% accrued at 24 months) and the remaining 50% accrued during the last year of enrolment • Sample size calculations were based on a one-sided alpha of 0.025 with one interim analysis for futility after 30% of the events have occurred.

Data management, patient withdrawals	**Missing data** • Missing individual data were generally treated as missing and no values were imputed. **Discontinuations** • Patients who discontinued study treatment for any reason were to undergo End-of-Treatment procedures • Additionally, all discontinued patients were followed for 28 days following the last dose of study treatment or until the date of the last study visit (whichever was longer) for AEs • After the follow-up visit, patients were followed for survival by telephone, every month for the first year and then every 3 months until death, withdrawal of consent for further follow-up, study end, or until the patient was lost to follow-up • Discontinued patients were not replaced

Abbreviations: AE = adverse event; ANCOVA = analysis of covariance; CI = confidence interval; CMH = Cochran-Mantel-Haenszel; EQ-5D-3L = European Quality of Life - 5 dimensions, 3 levels; FACIT = Functional Assessment of Chronic Illness Therapy; HCRU = healthcare resource utilisation; HR = hazard ratio; HRQoL = health-related quality of life; ITT = intention-to-treat; KM = Kaplan-Meier; MDS = myelodysplastic syndromes; MID = minimally important difference; MMRM = mixed model repeated measures; OS = overall survival; RFS = relapse-free survival.

[a] The primary efficacy analysis was performed using the ITT population.

[b] The HRQoL outcomes were evaluated for the HRQoL-evaluable population.

Source: Wei et al., 2020[84]; QUAZAR AML-001 CSR (Data on File)[87]; Supplementary appendix to Wei et al., 2020[88]; QUAZAR AML-001 SAP (Data on File)[96]

B.2.5 Quality assessment of the relevant clinical effectiveness evidence

The quality assessment of the QUAZAR AML-001 study is provided in Appendix D.

B.2.6 Clinical effectiveness results of the relevant trials

The primary analysis for OS and analyses for all other outcomes were conducted based on the primary data base lock (15 July 2019 cut-off), when 329 events occurred to allow for a fully powered OS analysis.[87] The primary analysis was used for EMA/MHRA regulatory submissions and subsequent approval.

Since then a further data cut has become available for OS (8 September 2020), providing longer follow-up and more mature survival data. Therefore, in this section all of the data presented are from the primary analysis (15 July 2019), except for the OS data which are presented for the primary analysis (July 2019) and the most recent data cut (September 2020).

A summary of results for primary and secondary efficacy outcomes is provided in Table B.2.10.

Table B.2.10. Summary of results (primary and secondary efficacy outcome), QUAZAR AML-001 study

Parameter	ITT population		
	Oral azacitidine (N=238)	**Placebo (N=234)**	**Difference (95% CI)**
Primary outcome (OS)			
15 July 2019			
Median OS, months (95% CI)	24.7 (18.7-30.5)	14.8 (11.7-17.6)	9.9 (4.6-15.3)
HR (95% CI)	0.69 (0.55-0.86)[a]		-
p-value	0.0009		-
8 September 2020			
Median OS, months (95% CI)	24.7 (18.7-30.5)	14.8 (11.7-17.6)	9.9 (4.5-15.4)
HR (95% CI)	0.69 (0.56-0.86)		-
p-value	0.0008		-
Key secondary outcome (RFS) (15 July 2019)			
Median RFS, months (95% CI)	10.2 (7.9-12.9)	4.8 (4.6-6.4)	5.3 (3.1-7.5)
HR (95% CI)	0.65 (0.52-0.81)		-
p-value	0.0001		-
Secondary efficacy outcomes (15 July 2019)			
Median time to relapse, months (95% CI)	10.2 (8.3-13.4)	4.9 (4.6-6.4)	-
Median time to treatment discontinuation, months (95% CI)	11.4 (9.8-13.6)	6.1 (5.1-7.4)	5.4 [redacted]
Parameter	**HRQoL-evaluable population**		
	Oral azacitidine (N=225)	**Placebo (N=219)**	**Difference (95% CI)**
Secondary efficacy outcomes (15 July 2019)			
FACIT-Fatigue scale, mean (SD)	[redacted]	[redacted]	-
EQ-5D-3L health utility index, mean (SD)	[redacted]	[redacted]	-
Parameter	**Safety population**		
	Oral azacitidine (N=236)	**Placebo (N=233)**	**Difference (95% CI)**

Secondary efficacy outcomes (15 July 2019)			
Number of patients hospitalised, n (%)	108 (45.8)	118 (50.6)	-
Number of hospital events	173	151	-
Rate/person-year (2-sided 95% CI)	0.48 [redacted]	0.64 [redacted]	-
Relative risk (2-sided 95% CI)	0.740 (0.595-0.920)		-
Two-sided p-value	0.0068		-
Number of days hospitalised	2872	3139	-
Rate/person-year (2-sided 95% CI)	7.89 [redacted]	13.36 [redacted]	-
Relative risk (2-sided 95% CI)	0.591 (0.562-0.621)		-
Two-sided p-value	<0.0001		-

Abbreviations: CI = confidence interval; EQ-5D-3L = European Quality of Life - 5 dimensions, 3 levels; FACIT = Functional Assessment of Chronic Illness Therapy; HR = hazard ratio; ITT = intention-to-treat; n = number of patients in the category; N = number of patients evaluable; OS = overall survival; RFS = relapse-free survival; SD = standard deviation.
[a]Hazard ratios were not provided in the primary publication, as the proportional hazards assumption appeared to be violated, as indicated by the significant treatment-by-time interaction

Source: Wei et al., 2019[2]; Wei et al., 2020[84]; Oliva et al, 2020[86]; QUAZAR AML-001 SAP (Data on File)[96]; Wei et al, 2021[97]

B.2.6.1 Primary outcome: OS

The QUAZAR AML-001 study met its primary outcome based on the primary database lock (15 July 2019) and this effect was maintained in the long term (8 September 2020 data cut) (Table B.2.11). Whilst median OS was unchanged in the latest data cut, the tails of the oral azacitidine and placebo OS curves showed greater separation than in the primary analysis, indicating a sustained, long-term OS benefit with oral azacitidine.[97]

At a median follow-up of 41.2 months (primary database lock), oral azacitidine was associated with a significantly longer OS compared with placebo, with a clinically meaningful difference in median OS of 9.9 months (median OS: 24.7 months vs. 14.8 months; HR 0.69 [95% CI: 0.55-0.86], p<0.001) (see Figure B.2.3 and Table B.2.11).[84] A lower death rate was observed in the oral azacitidine group than in the placebo group as early as 90 days after randomisation ([redacted], respectively).[87] Survival rates were higher in the oral azacitidine group than in the placebo group at one year after randomisation (72.8% vs. 55.8%; difference 17.0 percentage points [95% CI: 8.4-25.6])(Table B.2.11).[84]

The OS findings were consistent across key demographic and disease-related subgroups (see Section **Error! Reference source not found.** for the forest plot).

In the most recent data cut for OS at a median follow-up of 51.7 months, oral azacitidine continued to be associated with a significantly longer OS compared with placebo, maintaining the median OS observed at the primary analysis (Figure B.2.4 and Table B.2.11).[92, 92, 97, 98]

Survival rates continued to be higher in the oral azacitidine group at one year (72.8% vs. 55.8%; difference 17 percentage points) after randomisation (Table B.2.11).[97, 98]

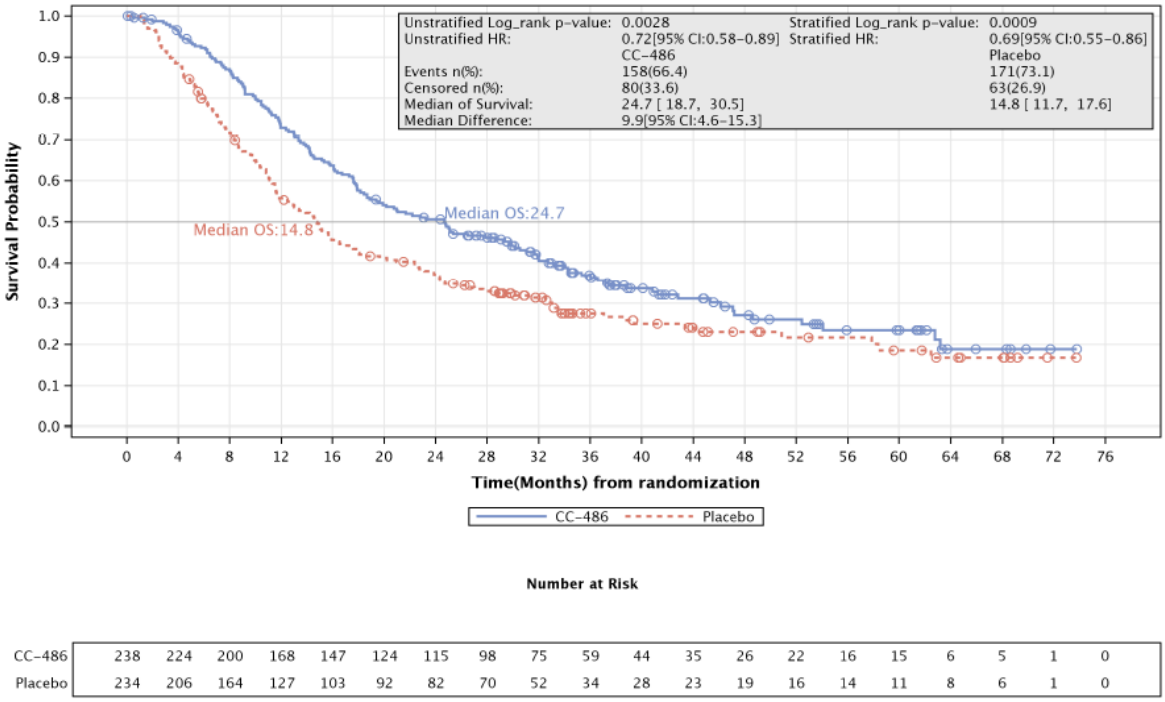
Table B.2.11. Summary of OS, QUAZAR AML-001 study (ITT population)

Parameter	Oral azacitidine (N=238)	Placebo (N=234)	Difference (95% CI)
15 July 2019			
Patients with event (death), n (%)	158 (66.4)	171 (73.1)	-
Patients censored, n (%)	80 (33.6)	63 (26.9)	-
Median OS, months (95% CI)[a]	24.7 (18.7-30.5)	14.8 (11.7-17.6)	9.9 (4.6-15.3)
HR (95% CI)[b]	0.69 (0.55, 0.86)[e]		-
p-value[c]	0.0009		-
1-year survival estimate (95% CI)[d]	0.728 [redacted]	0.558 [redacted]	0.170 (0.084-0.256)
2-year survival estimate (95% CI)[d]	0.506 [redacted]	0.371 [redacted]	0.135 (0.045-0.225)
8 September 2020			
Patients with event (death), n (%)	[redacted]	[redacted]	[redacted]
Patients censored, n (%)	[redacted]	[redacted]	[redacted]
Median OS, months (95% CI)[a]	24.7 (18.7-30.5)	14.8 (11.7-17.6)	9.9 (4.5-15.4)
HR (95% CI)[b]	0.69 (0.56-0.86)	-	
p-value[c]	0.0008	-	
1-year survival estimate (95% CI)[d]	0.728 [redacted]	0.558 [redacted]	0.170 (0.084-0.256)
2-year survival estimate (95% CI)[d]	0.506 [redacted]	0.371 [redacted]	0.135 (0.045-0.225)
3-year survival estimate (95% CI)[d]	[redacted]	[redacted]	[redacted]

Abbreviations: CI = confidence interval; HR = hazard ratio; ITT = intention-to-treat; KM = Kaplan–Meier; n = number of patients in the category; N = number of patients evaluable; OS = overall survival.
[a] Median estimate of OS was derived using the KM method. Difference was calculated as oral minus placebo. The CI for the difference was derived using Kosorok's method.
[b] The HR is from a Cox proportional hazards model stratified by age, cytogenetic risk category, and receipt of consolidation therapy or not.
[c] The p-value is two-sided from a log-rank test stratified by age, cytogenetic risk category, and receipt of consolidation therapy or not.
[d] KM methods were used to estimate the one-year, two-year and three-year survival probabilities. The CIs for the difference in the one-year and two-year survival probabilities were derived using Greenwood's variance estimate.
[e]Hazard ratios were not provided in the primary publication, as the proportional hazards assumption appeared to be violated, as indicated by the significant treatment-by-time interaction
Note: percentages are based on the number of patients in each treatment group, unless otherwise specified.

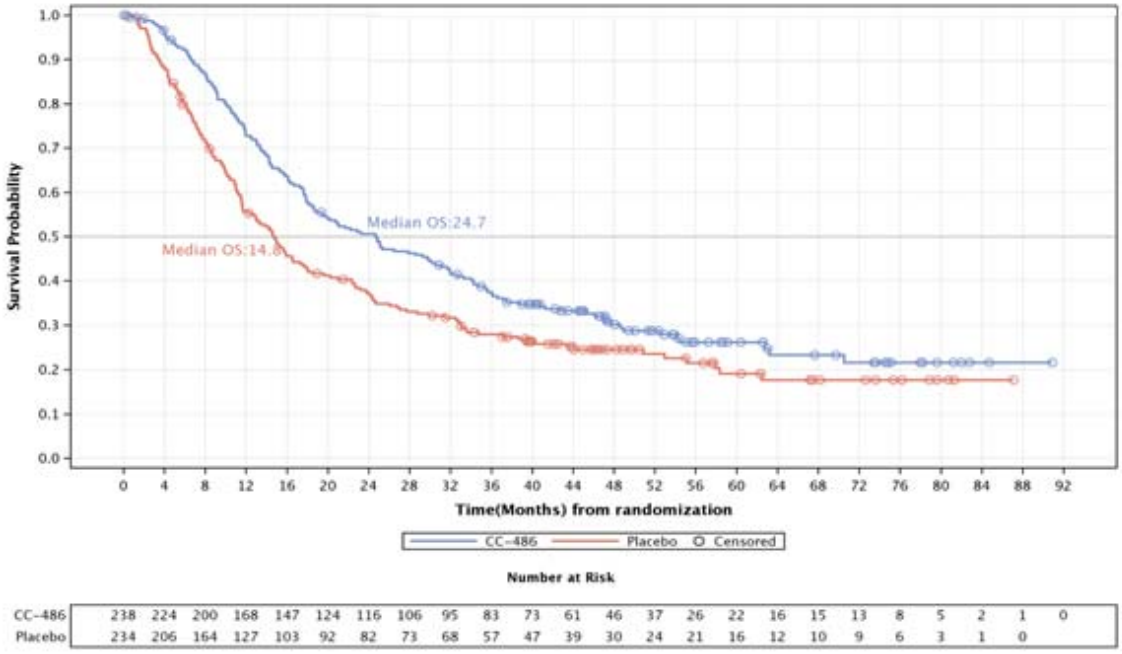
Source: Wei et al, 2020[84]; QUAZAR AML-001 CSR (Data on File)[87]; Wei et al, 2021[97]; BMS, 2021 (Data on File)[98]; FDA, 2020[92]; EMA/308711/2021[95]

Figure B.2.3. KM analysis of OS (data cut-off date, 15 July 2019), QUAZAR AML-001 study (ITT population)



Abbreviations: CC-486 = oral azacitidine; CI = confidence interval; HR = hazard ratio; ITT = intention-to-treat; KM = Kaplan-Meier; OS = overall survival.
Source: EMA/308711/2021[95]

Figure B.2.4. KM analysis of OS (data cut-off date, 8 September 2020), QUAZAR AML-001 study (ITT population)



Abbreviations: CC-486 = oral azacitidine; CI = confidence interval; HR = hazard ratio; ITT = intention-to-treat; KM = Kaplan-Meier; OS = overall survival.
Source: Wei et al, 2021[97]

The 15 July 2019 data cut was taken at the end of the Follow-up Phase, and was the dataset submitted to EMA/MHRA. The 8 September 2020 data cut was taken during the EP, after study unblinding. Patients randomised into the placebo group were discontinued from treatment and did not subsequently receive oral azacitidine in the EP. OS data were still routinely collected through the EP subject to additional consent from patients and so the September 2020 data were therefore considered robust and the most mature data to use in the cost-effectiveness model (section B.3.3.1.1). This was considered appropriate given several findings. Firstly, the September 2020 data are consistent with the July 2019 data, with unchanged median OS and HR. Secondly, the September 2020 data provide additional reliability for the tail end of the OS KM curves (Figure B.2.4). At the July 2019 cut-off date, the number of patients at risk at 48 months were 26 and 19 for the oral azacitidine and placebo arms, respectively; by month 64, there were just 6 patients at risk in the oral azacitidine arm, and 8 for the placebo arm (Figure B.2.3). With fewer patients remaining at risk after 48 months, survival estimates beyond this point become less reliable and additional follow-up may influence the tail end of the curves. The September 2020 data provides an additional ~14 months of follow-up and greater reliability to the shape of the tails (Figure B.2.4).

The more mature OS data support the conclusion that maintenance treatment with oral azacitidine provides a significant OS benefit to patients who achieved CR/CRi following IC, i.e. standard of care in AML. Moreover, with regards to treatment duration, the extended OS data are considered to be reflective of the expected outcomes in UK clinical practice.

B.2.6.1.1 OS sensitivity analysis (primary database lock)

In the sensitivity analysis of the primary outcome in the mITT population (oral azacitidine N=223, placebo N=217) the results were highly consistent with those in the primary analysis: oral azacitidine was associated with significantly improved OS compared with placebo, with a clinically meaningful difference in median OS of 10.2 months (median OS: 24.8 months vs. 14.6 months; HR 0.66 [95% CI: 0.53-0.83], [redacted]).[87, 95]

Additional sensitivity analyses were conducted to determine whether subsequent therapy for AML may have impacted the findings for OS. Censoring for any subsequent AML therapy[e] [redacted] or disease-modifying AML therapy[f] [redacted]) led to results that were generally consistent with those of the primary analysis, although they did not reach statistical significance. However, these results should be interpreted with caution, due to censoring >80% of patients in each treatment group.[87]

A small proportion of patients underwent post-treatment HSCT (15 [6.3%] patients in the oral azacitidine group and 32 [13.7%] patients in the placebo group).[88] When censoring patients that underwent HSCT, those in the oral azacitidine group had a median overall survival of [redacted] compared with [redacted] in the placebo group [redacted],[87] which was highly consistent with primary analysis of the ITT population. This is especially important since transplant is known to impact OS. Whilst the number of post-treatment HSCT's was imbalanced across treatment groups (more patients in the placebo group underwent post-treatment HSCT than in the oral azacitidine group), this did not impact the OS results: oral azacitidine was still associated with a significant improvement in OS compared with placebo.

B.2.6.2 Key secondary outcome: RFS

RFS (data cut-off date, 15 July 2019) was significantly longer with oral azacitidine compared with placebo, with a clinically meaningful difference in median RFS of 5.3 months (median RFS: 10.2 months vs. 4.8 months; HR 0.65 [95%CI: 0.52-0.81], p<0.0001) (Figure B.2.5 and

Abbreviations: CC-486 = oral azacitidine; CI = confidence interval; HR = hazard ratio; ITT = intention-to-treat; KM = Kaplan-Meier; RFS = relapse-free survival.

Source: EMA/308711/2021[95]

[e]All subsequent therapies for AML as documented on the Case Report Form were included in the analysis.
[f]All AML therapies were considered disease modifying, except for hydroxycarbamide.

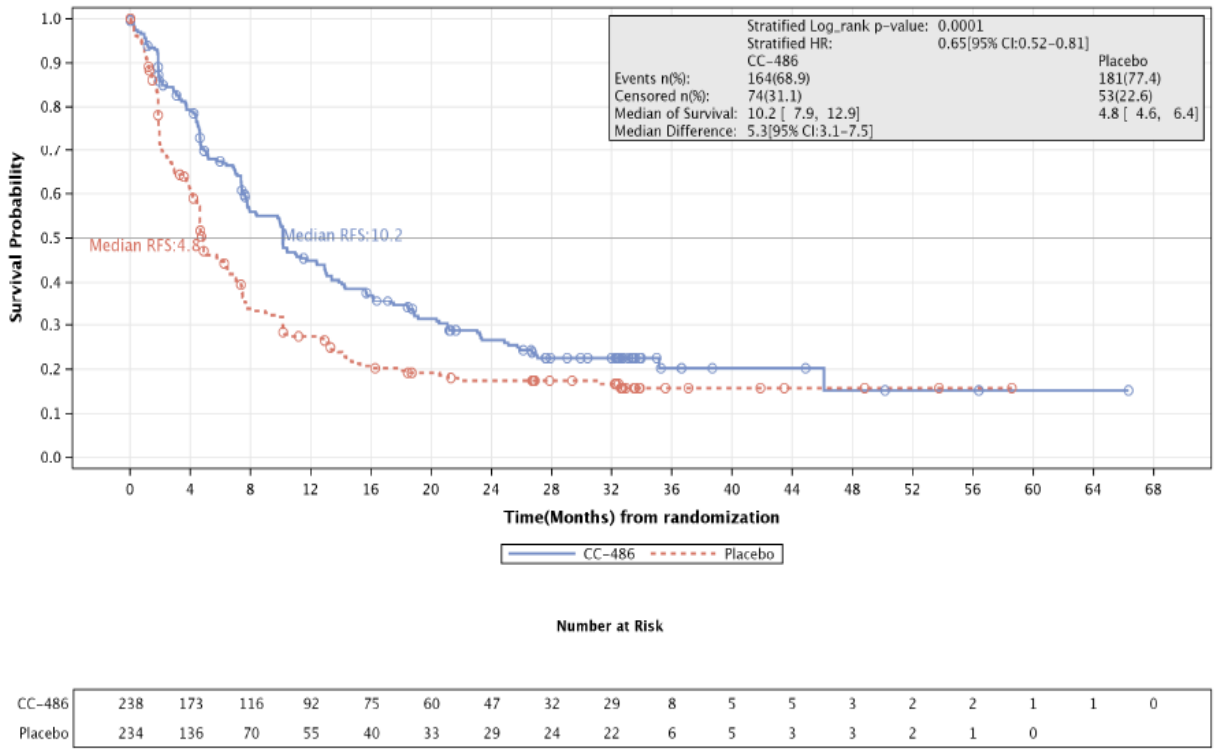
Table B.2.12).[84] Higher RFS rates were observed in the oral azacitidine group than in the placebo group at six months (67.4% vs. 45.2%), one year (44.9% vs. 27.4%), and two years (26.6% vs. 17.4%) (Abbreviations: CC-486 = oral azacitidine; CI = confidence interval; HR = hazard ratio; ITT = intention-to-treat; KM = Kaplan-Meier; RFS = relapse-free survival.

Source: EMA/308711/2021[95]

Table B.2.12).[84, 87, 95]

The RFS findings were consistent across key demographic and disease-related subgroups (See Section B.2.7.2).

Figure B.2.5. KM analysis of RFS (data cut-off date, 15 July 2019), QUAZAR AML-001 study (ITT population)



Abbreviations: CC-486 = oral azacitidine; CI = confidence interval; HR = hazard ratio; ITT = intention-to-treat; KM = Kaplan-Meier; RFS = relapse-free survival.
Source: EMA/308711/2021[95]

Table B.2.12. Summary of RFS, (data cut-off date, 15 July 2019), QUAZAR AML-001 study (ITT population)

Parameter	Oral azacitidine (N=238)	Placebo (N=234)	Difference (95% CI)
Patients with event (relapse or death), n (%)	164 (68.9)	181 (77.4)	-
Patients censored, n (%)	74 (31.1)	53 (22.6)	-
Median RFS, months (95% CI)[a]	10.2 (7.9-12.9)	4.8 (4.6-6.4)	5.3 (3.1-7.5)
HR (95% CI)[b]	0.65 (0.52-0.81)		-
p-value[c]	0.0001		-
6-month RFS estimate (95% CI)[d]	0.674 [redacted]	0.452 [redacted]	0.222 [redacted]
1-year RFS estimate (95% CI)[d]	0.449 [redacted]	0.274 [redacted]	0.175 [redacted]
2-year RFS estimate (95% CI)[d]	0.266 [redacted]	0.174 [redacted]	0.092 [redacted]

Abbreviations: CI = confidence interval; HR = hazard ratio; ITT = intention-to-treat; n = number of patients in the category; N = number of patients evaluable; RFS = relapse-free survival.
[a] Median estimate of RFS was derived using the Kaplan–Meier method. Difference was calculated as oral azacitidine minus placebo. The CI for the difference was derived using Kosorok's method.
[b] The hazard ratio is from a Cox proportional hazards model stratified by age, cytogenetic risk category, and receipt of consolidation therapy or not.
[c] The p-value is two-sided from a log-rank test stratified by age, cytogenetic risk category, and receipt of consolidation therapy or not.
[d] Kaplan–Meier methods were used to estimate the six-month, one-year, and two-year RFS probabilities. The CIs for the difference in these RFS probabilities were derived using Greenwood's variance estimate.
Note: percentages are based on the number of patients in each treatment group, unless otherwise specified.
Source: Wei et al., 2020[84]; QUAZAR AML-001 CSR (Data on File)[87]; FDA, 2020[92]; EMA/308711/2021[95]

Neither bone marrow nor peripheral blood samples were required to be collected in the EP of the study (this is a requirement for RFS assessment as outlined in section B.2.3.2.3), and therefore only OS data is available to be interpreted from the September 2020 data. There were isolated bone marrow or peripheral blood samples collected after the July 2019 database lock, [redacted]. Further details are presented in Appendix M.

B.2.6.2.1 RFS sensitivity analysis

In the sensitivity analysis of the key secondary outcome (data cut-off date, 15 July 2019) that used the mITT population (oral azacitidine N=223, placebo N=217) the results were consistent with those in the primary analysis: oral azacitidine was associated with significantly improved RFS compared with placebo, with a clinically meaningful difference in median RFS of [redacted] (median RFS: [redacted] vs. [redacted]; [redacted][87, 93]

B.2.6.3 Other secondary efficacy outcomes

Overall, secondary efficacy outcomes demonstrated important benefits for patients with AML receiving oral azacitidine compared with placebo, for example lower relapse rates and longer time on study treatment. The favourable HRQoL of patients in remission was not compromised

by oral azacitidine whilst providing a significant survival benefit for these patients. Moreover, significantly lower hospitalisation rates in the oral azacitidine arm compared with placebo suggest that the direct burden of disease is reduced in patients receiving oral azacitidine.[87]

B.2.6.3.1 Time to relapse

A programmatically derived documented relapse occurred in 154 (64.7%) patients in the oral azacitidine group and 179 (76.5%) patients in the placebo group.[87] Ten (4.2%) patients in the oral azacitidine group and two (0.9%) patients in the placebo group died without documented relapse.[87, 92]The median time to relapse was 10.2 months in the oral azacitidine group and 4.9 months in the placebo group.[87] Lower relapse rates were observed in the oral azacitidine group than in the placebo group at six months (31.3% vs. 54.4%), one year (52.8% vs. 71.7%), and two years (69.1% vs. 81.7%) (Table B.2.13).[87]

Table B.2.13. Summary of time to relapse, QUAZAR AML-001 study (ITT population)

Parameter	Oral azacitidine (N=238)	Placebo (N=234)
Patients relapsed, n (%)	154 (65)	179 (76)
Patients died without relapse, n (%)	10 (4.2)	2 (0.9)
Patients censored, n (%)	74 (31.1)	53 (22.6)
Median time to relapse, months (95% CI)[a]	10.2 (8.3-13.4)	4.9 (4.6-6.4)
6-month relapse rate estimate (95% CI)[b]	0.31 (0.25-0.37)	0.54 (0.48-0.61)
1-year relapse rate estimate (95% CI)[b]	0.53 (0.46-0.59)	0.72 (0.65-0.77)
2-year relapse rate estimate (95% CI)[b]	0.69 (0.62-0.75)	0.82 (0.76-0.86)

Abbreviations: CI = confidence interval; CR = complete remission; CRi = complete remission with incomplete blood count recovery; ITT = intention-to-treat; n = number of patients in the category; N = number of patients evaluable; RFS = relapse-free survival.
[a] Unstratified Kaplan–Meier analysis.
[b] Estimates of relapse rates are based on the cumulative incidence function from a competing risk analysis with death as a competing risk of relapse from CR/CRi.
Time to relapse is defined as the interval from the date of randomisation to the date of documented relapse.
Note: percentages are based on the number of patients in each treatment group, unless otherwise specified.
Source: Supplementary appendix to Wei et al, 2020[88]; FDA, 2020[92]; EMA/308711/2021[95]

B.2.6.3.2 Time to discontinuation from treatment

At the time of the primary analysis, the majority of patients in both the oral azacitidine group (81.1%) and the placebo group (88.9%) had discontinued from study treatment.[92] Those receiving oral azacitidine remained on study treatment for longer than patients in the placebo group; median time to discontinuation for any reason was 11.4 months in the oral azacitidine group and 6.1 months in the placebo group.[84, 87] Lower treatment discontinuation rates were observed in the oral azacitidine group than in the placebo group at six months [redacted] one year [redacted] and two years [redacted] (**Error! Not a valid bookmark self-reference.**).[87]

Table B.2.14. Summary of time to discontinuation from treatment, QUAZAR AML-001 study (ITT population)

Parameter	Oral azacitidine (N=238)	Placebo (N=234)	Difference (95% CI)
Patients with treatment discontinuation, n (%)	193 (81.1)	208 (88.9)	-
Patients censored, n (%)	45 (18.9)	26 (11.1)	-
Median time to treatment discontinuation, months (95% CI)[a]	11.4 (9.8-13.6)	6.1 (5.1-7.4)	5.4 [redacted]
6-month treatment discontinuation rate estimate (95% CI)[b]	[redacted]	[redacted]	[redacted]
1-year treatment discontinuation rate estimate (95% CI)[b]	[redacted]	[redacted]	[redacted]
2-year treatment discontinuation rate estimate (95% CI)[b]	[redacted]	[redacted]	[redacted]

Abbreviations: CI = confidence interval; ITT = intention-to-treat; KM = Kaplan–Meier; n = number of patients in the category; N = number of patients evaluable.
[a] Median estimate of time to discontinuation is from an unstratified Kaplan–Meier analysis. Differences were calculated as oral azacitidine minus placebo. The CIs for the differences were derived using Kosorok's method.
[b] KM methods were used to estimate the treatment discontinuation rate. Differences were calculated as oral azacitidine minus placebo. The CIs for the difference were derived using Greenwood's variance estimate.
Note: percentages are based on the number of patients in each treatment group, unless otherwise specified.
Source: Wei et al., 2020[84]; QUAZAR AML-001 CSR (Data on File)[87]; FDA, 2020[92]

The analysis was further refined by evaluating time to treatment discontinuation due to disease relapse using a competing risk method. At the time of the primary analysis, 143 (60.1%) patients in the oral azacitidine group and 180 (76.9%) patients in the placebo group had discontinued treatment because of relapse.[87, 92] The median time to discontinuation was [redacted] in the oral azacitidine group and [redacted] in the placebo group.[87]

B.2.6.3.3 HRQoL

The HRQoL outcome was a key secondary endpoint. Only randomised patients who had a valid QoL assessment at baseline (Cycle 1 Day 1) and at least one valid post baseline assessment were considered (HRQoL evaluable population)[g]; 225 (94.5%) patients in the oral azacitidine group and 219 (93.6%) patients in the placebo group were included in the HRQoL-evaluable population for the FACIT-Fatigue scale.[87] Similarly, 225 (94.5%) patients in the oral azacitidine group and 217 (92.7%) patients in the placebo group were included in the HRQoL-evaluable population for the EQ-5D-3L scores.[87] Baseline demographic and disease characteristics were comparable between treatment groups for the HRQoL-evaluable population.[87]

At baseline, mean scores on both the FACIT-Fatigue scale and the EQ-5D-3L health utility index were similar across the oral azacitidine and placebo groups (Table B.2.15).[99, 100] Low levels of fatigue were reported at baseline for both treatment groups based on the mean FACIT-Fatigue scores. Subjects in both treatment groups had a good health state at baseline

[g]This population was derived for each HRQoL measure (FACIT-Fatigue Scale and EQ-5D-3L).

based on the EQ-5D-3L health utility index mean scores (A higher score indicates a lower level of fatigue for FACIT-Fatigue, better health state for the EQ-5D-3L). Given these high baseline measurements, improvements in HRQoL were unlikely, and maintenance of the favourable HRQoL of patients in remission would be viewed as a positive outcome.

Table B.2.15. Mean baseline FACIT-Fatigue and EQ-5D-3L health utility index scores by treatment group, QUAZAR AML-001 study (HRQoL-evaluable population)

HRQoL Domain	Oral azacitidine (N=225)	Placebo (N=219)	Overall (N=444)
FACIT-Fatigue scale[c], mean (SD)	[redacted]	[redacted]	[redacted]
EQ-5D-3L health utility index[c], mean (SD)	[redacted]	[redacted]	[redacted]

Abbreviations: EQ-5D-3L = European Quality of Life – 5 dimensions, 3 levels; FACIT = Functional Assessment of Chronic Illness Therapy; HRQoL = health-related quality of life; N = number of patients evaluable; SD = standard deviation.
[c]A higher score indicates a lower level of fatigue for FACIT-Fatigue, better health state for the EQ-5D-3L.
Source: QUAZAR AML-001 CSR (Data on File)[87]

In both treatment groups, scores on the FACIT-Fatigue scale and the EQ-5D-3L health utility index gradually improved over time, indicating that HRQoL was maintained throughout the treatment period.[87, 101] No clinically meaningful differences in FACIT-Fatigue scale were noted between the two treatment groups.[84] A similar trend was observed for the EQ-5D-3L health utility index, with no clinically meaningful differences found.[84] Although statistically significant differences in favour of the placebo group were found at a few time points for the EQ-5D-3L scale, these differences were likely to have occurred by chance, as the comparisons were not adjusted for multiplicity.[87]

Notably, HRQoL assessments were only completed during the Treatment Phase, when treatment-related side effects occurred more frequently in the oral azacitidine group than in the placebo group. Therefore, any long-term benefits in HRQoL associated with improved survival were not captured. Nonetheless, as noted above, the HRQoL of patients in the oral azacitidine group was maintained at a level comparable to that in both the placebo group during the treatment phase. Therefore, treatment with oral azacitidine improved survival without compromising the favourable HRQoL of patients in remission.

B.2.6.3.4 Healthcare resource utilisation

HCRU data collected in the QUAZAR AML-001 study included hospitalisations, medications, clinic visits, medical/diagnostic procedures, and treatment for AEs.[87] At the present time, only data for hospitalisations have been analysed; these data are summarised by treatment for the safety population in **Error! Reference source not found.**. After adjustment for duration of study drug exposure, the results showed that oral azacitidine was associated with significantly fewer hospitalisation events per person-year (0.48 vs. 0.64; p=0.0068) and a lower number of days hospitalised per person-year (7.89 vs. 13.36; p<0.0001) than placebo.[86, 87] These results suggest that maintenance treatment with oral azacitidine will lead to a reduction in HCRU associated with hospitalisations.

Table B.2.16. Summary of hospitalisation data, QUAZAR AML-001 study (safety population)

Hospitalisation parameter	Oral azacitidine (N=236)	Placebo (N=233)
Total person-years exposure, years	363.8	234.9
Number of patients hospitalised, n (%)	108 (45.8)	118 (50.6)
Number of hospital events	173	151
Rate/person-year (2-sided 95% CI)[a]	0.48 [redacted]	0.64 [redacted]
Relative risk (2-sided 95% CI)[b]	0.740 (0.595-0.920)	
Two-sided p-value[b]	0.0068	
Number of days hospitalised	2872	3139
Rate/person-year (2-sided 95% CI)[a]	7.89 [redacted]	13.36 [redacted]
Relative risk (2-sided 95% CI)[b]	0.591 (0.562-0.621)	
Two-sided p-value[b]	<0.0001	

Abbreviations: CI = confidence interval; n = number of patients in the category; N = number of patients evaluable.
[a] The 95% CI for the rate per person-year of exposure is based on the Exact method.
[b] The 95% CI for the relative risk estimate and associated nominal p-value testing that the relative risk is equal to one are based on asymptotic methods.
Source: Oliva et al, 2020[86]; QUAZAR AML-001 CSR (Data on File)[87]

B.2.6.4 Exploratory outcome –MRD status

MRD status was evaluated in this study due to its wide use in clinical practice as a predictor of relapse. MRD is defined as post-chemotherapy persistence of leukemic cells at levels below morphologic detection.[41] A threshold of 0.1% is recommended to differentiate between MRD-positivity and MRD-negativity; patients with MRD levels below 0.1% are considered to have a favourable disease prognosis whereas MRD-positive (levels above the defined threshold) have a higher risk of relapse.[44] Clinical investigations of MRD have clearly shown that many patients with AML who achieve CR after induction chemotherapy have detectable residual disease, and that this is a strong independent prognostic marker of increased relapse risk and shorter survival.[41-43]

At baseline, 103 (43.3%) patients in the oral azacitidine group and 116 (49.6%) patients in the placebo group were identified as MRD-positive, and 133 (55.9%) patients in the oral azacitidine group and 111 (47.4%) patients in the placebo group were identified as MRD-negative (defined as patients who achieved MRD-negative status for at least two consecutive post-baseline assessments).[87]

At the time of the primary analysis, among patients who were MRD-positive at baseline, a higher proportion achieved MRD-negative status at any point during treatment in the oral azacitidine group (38/103 [36.9%]) than in the placebo group (22/116 [19.0%]).[89] Notably, among patients who were MRD-positive at baseline and achieved MRD-negative status during treatment (i.e. MRD responders), a higher proportion of patients achieved MRD negativity greater than six months after randomisation in the oral azacitidine group (9/38 [23.7%]) than in the placebo group (1/22 [4.5%]).[89] In addition, among patients who were MRD-negative at baseline, a higher proportion maintained MRD-negative status during treatment in the oral

azacitidine group (73/133 [54.9%]) than in the placebo group (51/111 [45.9%]).[89] Furthermore, the median duration of MRD negativity was significantly extended with oral azacitidine compared with placebo (11.0 months vs. 5.0 months; HR 0.62 [95% CI 0.48-0.78]).[89]

The results demonstrate that maintenance treatment with oral azacitidine can help patients who are in CR/CRi to achieve or maintain MRD-negative status. Furthermore, oral azacitidine may significantly extend the duration of MRD negativity compared with placebo and it may induce MRD negativity after prolonged periods of MRD positivity. These findings further substantiate the results of subgroup analyses showing that oral azacitidine provides OS and RFS benefits independent of baseline MRD status (see forest plots in Section **Error! Reference source not found.**).

B.2.6.5 Efficacy conclusions

At a median follow-up of 41.2 months (data cut-off date, 15 July 2019), the results of the QUAZAR AML-001 study showed that treatment with oral azacitidine provided a clinically meaningful difference in median OS of 9.9 months. Furthermore, median RFS was 10.2 months among patients treated with oral azacitidine, more than twice as long as patients who received placebo (4.8 months). In the most recent data cut (data cut-off date, 8 September 2020), at a median follow-up of 51.7 months, the OS at 3 years from the time of randomisation favoured oral azacitidine, i.e. it continued to be associated with a significantly longer OS compared with placebo, with a clinically meaningful difference in median OS of 9.9 months.[84]

Median time to relapse was longer in the oral azacitidine group (10.2 months) compared with placebo (4.9 months). Lower relapse rates were observed in the oral azacitidine group than in the placebo group at six months, one year and two years. Patients in the oral azacitidine group remained on study treatment for longer than patients in the placebo group (median time to discontinuation for any reason; 11.4 months in the oral azacitidine group and 6.1 months in the placebo group).[93] Lower treatment discontinuation rates were observed in the oral azacitidine group than in the placebo group at six months, one year and two years.

Notably, oral azacitidine improved survival while preserving HRQoL at a level similar to placebo. In addition, oral azacitidine was associated with significantly fewer hospitalisations than placebo, which can result in reduced HCRU among patients who are in remission.

Clinical investigations of MRD have clearly shown that many patients with AML who achieve CR after induction chemotherapy may have detectable residual disease, and that this is a strong independent prognostic marker of increased relapse risk and shorter survival.[41-43] Among patients who were MRD-positive at baseline and achieved MRD-negative status during treatment (i.e. MRD responders), a higher proportion of patients achieved MRD negativity greater than six months after randomisation in the oral azacitidine group than in the placebo group, suggesting that oral azacitidine can induce MRD negativity after prolonged MRD-positive status.

B.2.7 Subgroup analysis

Subgroup analyses were conducted to determine whether the OS and RFS findings in the full study population were consistent across patient demographic and disease-related subgroups. These included age at induction therapy, sex, race, geographic region, CR/CRi status at randomisation, cytogenetic risk category, receipt of consolidation therapy after induction,

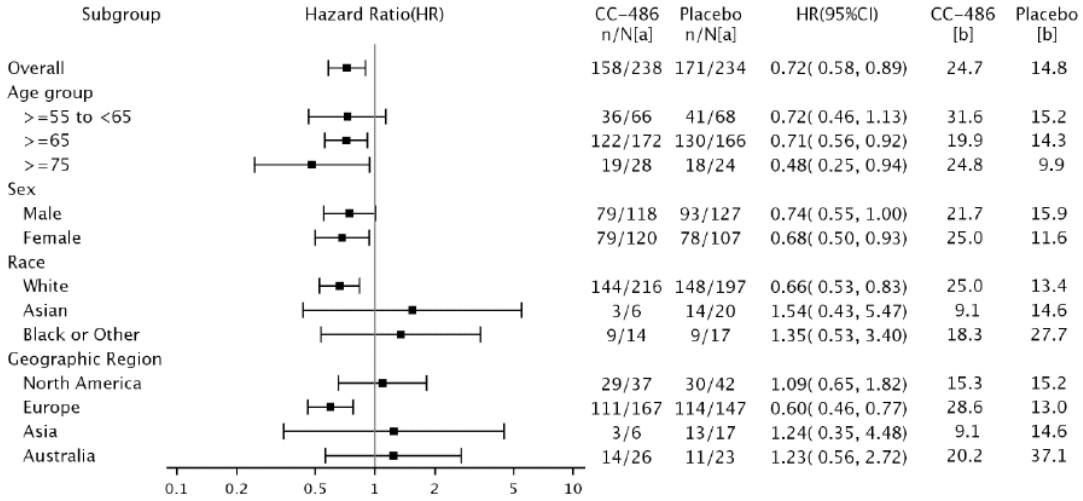
ECOG performance status score, prior MDS or CMML, and MRD status at screening. The results showed that the OS and RFS benefit provided by oral azacitidine in the full study population was consistent across key subgroups. Additional analyses by mutation status and number of consolidation courses also showed survival results consistent with those for the ITT population – in that they all show a positive, considerable benefit in survival.

B.2.7.1 OS

The majority of the predefined demographic subgroup analyses demonstrated meaningful improvement with oral azacitidine over placebo, and the OS results were consistent with the benefit observed in the overall study population (Figure B.2.6). Notably, reduction in risk of death was observed for oral azacitidine compared with placebo across all age group categories. This reduction was clearly evident among patients ≥65 years (HR 0.71 [95% CI, 0.56-0.92]) and ≥75 years (HR 0.48 [95% CI, 0.25-0.94]) which is clinically relevant, because at these ages the patients are not eligible for HSCT and the risk of relapse is very high (as confirmed by two UK AML treating clinicians).[95] The results by race and region indicate some differences between those categorised as White or European compared to other demographics, with a 34% and 40% reduction in risk of death in the oral azacitidine group, respectively.[95]

Among patients from Europe (N = 314), the median OS was [redacted] for the oral azacitidine arm and [redacted] for the placebo arm (HR: 0.60; 95% CI: 0.46, 0.77) and median RFS was [redacted] in the oral azacitidine arm and [redacted] in the placebo arm (HR: 0.56; 95% CI: 0.43, 0.73).[92] The treatment effect was in line, and also surpassing that seen in the ITT population (see Appendix E for further detail).[92] The results for the European subgroup may be more reflective of UK clinical practice, including the influence of European guidance. For example, the current pan-London clinical guidelines[102] have been derived in part from the ELN Consensus Guidelines on the diagnosis and management of AML,[18] the original BCSH AML guideline,[69] and incorporating further details on clinical trials and diagnostic or treatment options relevant to London and the UK.

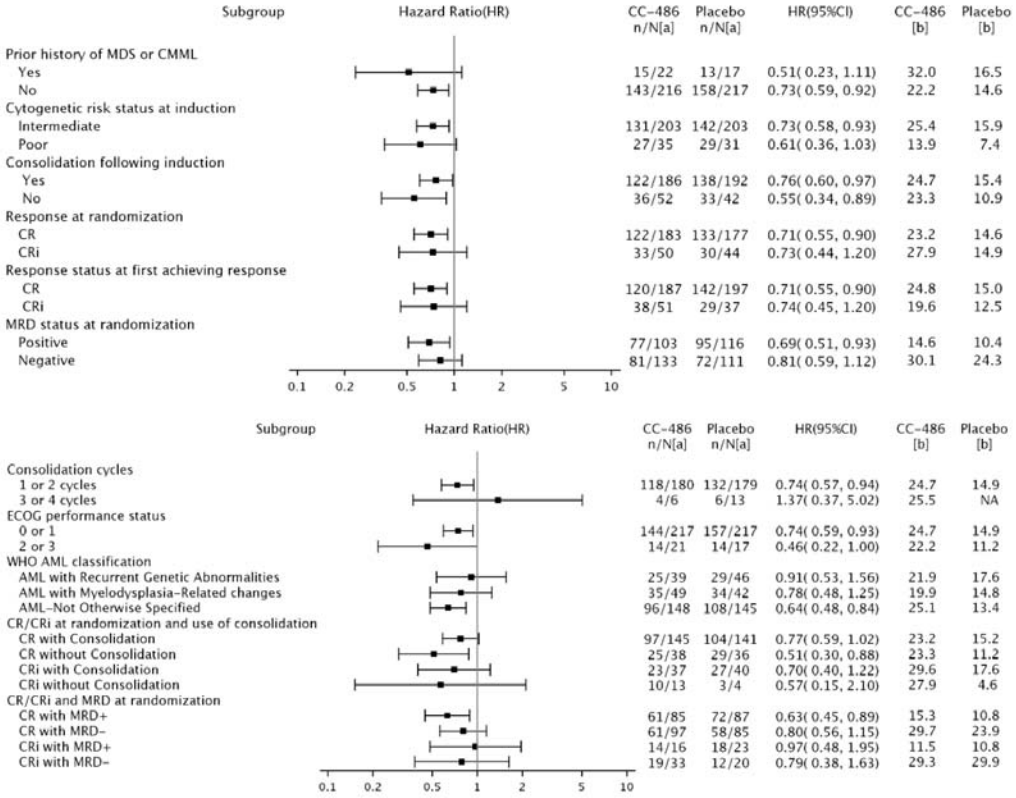
Figure B.2.6. Forest plot of OS by demographic subgroup, QUAZAR AML-001 study (ITT population)



Abbreviations: CC-486 = oral azacitidine; CI = confidence interval; OS = overall survival.
[a] Number of events/number of patients.
[b] Median OS in months.
Source: EMA/308711/2021[95]

Point estimates for the HRs were consistently <1 across all disease characteristic subgroups, with the exception of the small number of patients who received 3 or 4 cycles of consolidation (Figure B.2.7).[87, 95] Of note, an OS benefit was observed with oral azacitidine whether or not patients received consolidation therapy following induction (see Section B.2.7.5). In the subgroup analysis by CR/CRi status at randomisation, an OS benefit was observed with oral azacitidine in both subgroups based on HRs (HR 0.71 [95% CI: 0.55-0.90] for patients in CR and HR 0.73 [95% CI: 0.44-1.20]) for patients in CRi). Oral azacitidine demonstrated a favourable treatment effect compared with placebo in MRD-positive and MRD-negative patients. Based on HRs, the treatment effect was more pronounced in MRD-positive patients (HR 0.69 [95% CI: 0.51-0.93]) than in MRD-negative patients (HR 0.81 [95% CI: 0.59-1.12]). Subgroup analyses for both MRD and CR/CRi status at randomisation were performed. As expected, patients who were MRD negative and in CR at randomisation had the longest survival.[87, 95]

Figure B.2.7. Forest plot of OS by disease-related subgroup, QUAZAR AML-001 study (ITT population

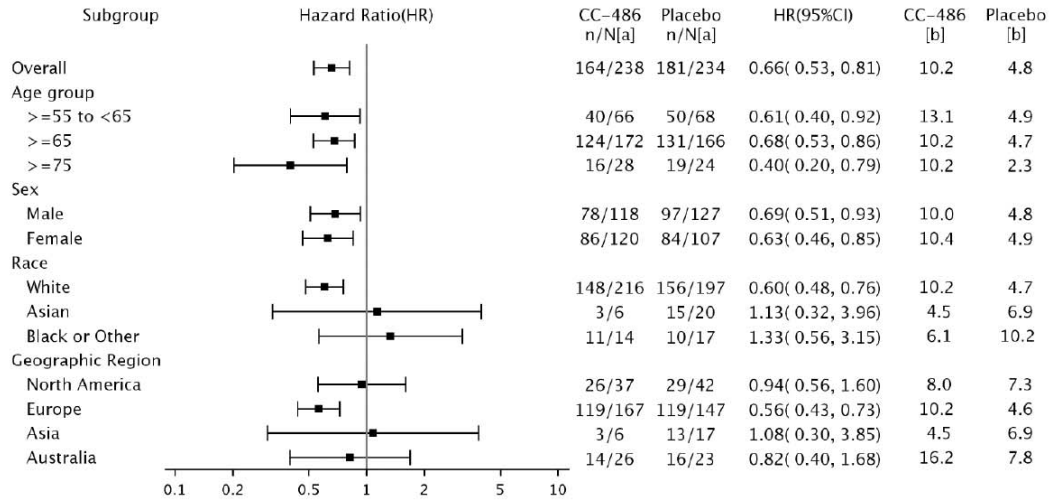


Abbreviations: AML = acute myeloid leukaemia; CI = confidence interval; CMML = chronic myelomonocytic leukaemia; CR = complete remission; CRi = complete remission with incomplete blood count recovery; ECOG = Eastern Cooperative Oncology Group; MDS = myelodysplastic syndromes; MRD = measurable residual disease; OS = overall survival; WHO = World Health Organisation.
[a] Number of events/number of patients.
[b] Median OS in months.
Source: EMA/308711/2021[95]

B.2.7.2 RFS

The HRs for demographic subgroups (see Figure B.2.8) of age and sex were <1 for all subgroup categories, indicating a risk reduction with oral azacitidine of 32% to 60% across age group categories, 31% for males, and 37% for females.[92] These results were consistent with the 34% reduction in risk of relapse or death for the overall ITT population in the unstratified subgroup analysis. In the demographic subgroup analysis by race, subgroup categories of Asian and Black/Other had a small number of patients (N=26 and N=31, respectively) and HRs >1.[92] The HRs were <1 for all geographical regions, except for Asia (N=23).[92] In the European subgroup (n=314), which is likely to reflect UK clinical practice the HR was 0.56. In the subgroups with a small number of patients, the overall result may have been influenced by the outcome for individual patients.[87, 92]

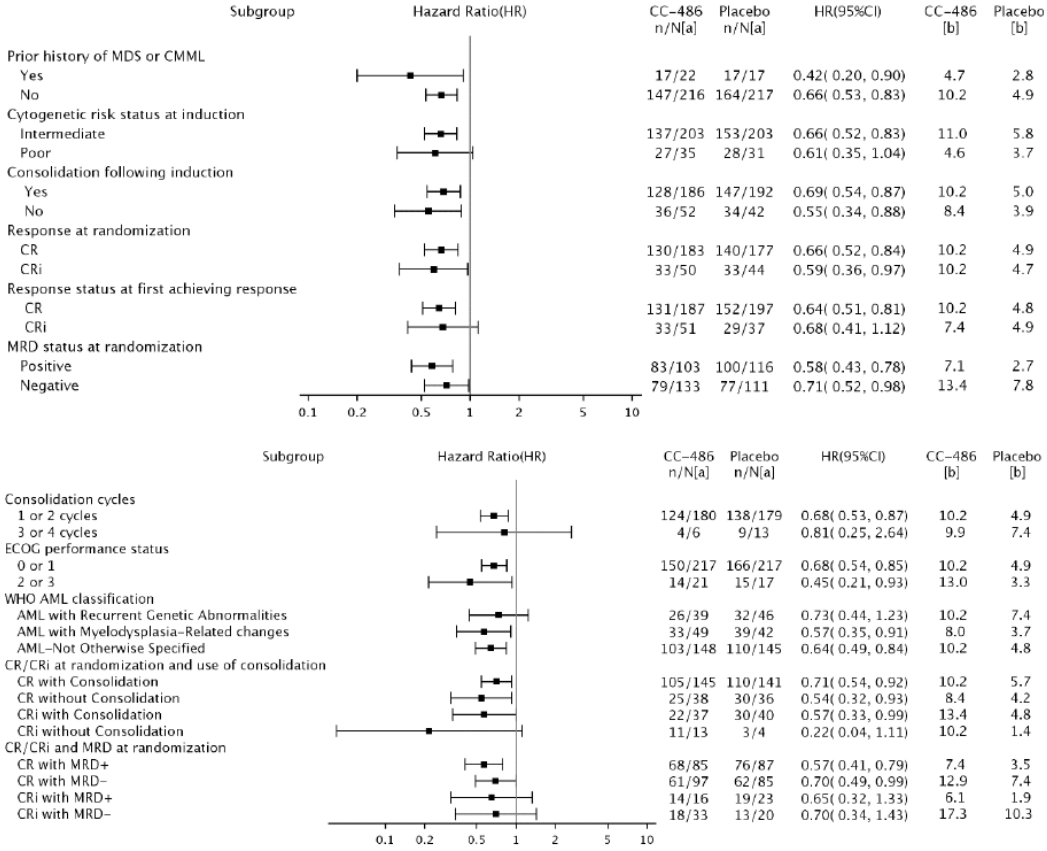
Figure B.2.8. Forest plot of RFS by demographic subgroup, QUAZAR AML-001 study (ITT population)



Abbreviations: CC-486 = oral azacitidine; CI = confidence interval; RFS = relapse-free survival.
[a] Number of events/number of patients.
[b] Median OS in months.
Source: FDA, 2020[92]

Point estimates for the HRs were consistently <1 across all disease characteristic subgroups, including a favourable effect for oral azacitidine compared with placebo (Figure B.2.9). Of note, a benefit with respect to RFS was observed with oral azacitidine whether or not patients received consolidation therapy following induction (see Section B.2.7.5).[92] In the subgroup analysis by CR/CRi status at randomisation, a benefit with respect to RFS was observed with oral azacitidine in both subgroups based on HRs, which indicated a reduction in risk of relapse or death with oral azacitidine of 34% for patients in CR at randomisation and 41% for patients in CRi at randomisation. As expected, patients who were MRD negative prior to randomisation had longer RFS than patients who were MRD positive.[87, 92]

Figure B.2.9. Forest plot of RFS by disease-related subgroup, QUAZAR AML-001 study (ITT population)



Abbreviations: AML = acute myeloid leukaemia; CI = confidence interval; CMML = chronic myelomonocytic leukaemia; CR = complete remission; CRi = complete remission with incomplete blood count recovery; ECOG = Eastern Cooperative Oncology Group; MDS = myelodysplastic syndromes; MRD = measurable residual disease; RFS = relapse-free survival; WHO = World Health Organisation.
[a] Number of events/number of patients.
[b] Median RFS in months.
Source: FDA, 2020[92]

The risk of relapse is particularly high among older patients with AML who achieve CR/CRi with IC. Therefore, an additional subgroup analysis was conducted to evaluate the impact of treatment with oral azacitidine on RFS among patients aged ≥75 years.[103] Despite the small sample size of these patients in the ITT population (N=28 for oral azacitidine; N=24 for placebo), oral azacitidine was associated with a significant RFS benefit compared with placebo (median RFS: 10.2 months vs. 2.3 months; HR 0.40 [95% CI: 0.20-0.79], p=0.0061) for this patient group.[103]

B.2.7.3 Additional *post hoc* analyses (*FLT3-ITD/TKD*)

A *post hoc* analysis on survival outcomes was also conducted on patients with *FLT3*-internal tandem duplication (ITD)/-tyrosine kinase domain (TKD) at diagnosis (n=66) as this is an important prognostic factor (see Section B.1.3.3).[104]

FLT3-ITD/TKD mutations appeared to confer a negative prognosis in the placebo arm, but this was not apparent in the oral azacitidine arm (data cut-off date, 15 July 2019).[104] Median OS

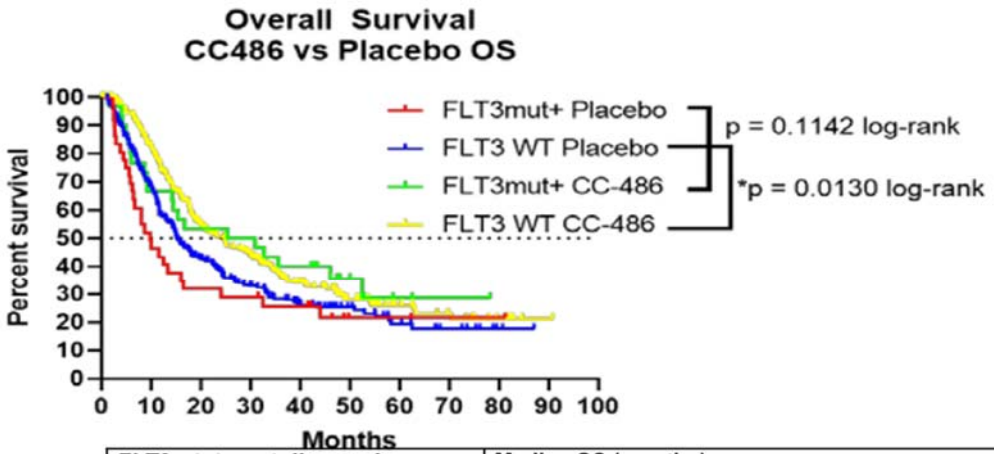
and RFS were prolonged in the oral azacitidine vs. placebo arms, in patients with *FLT3*-ITD/TKD mutations (**Error! Reference source not found.**).[104, 105]

Figure B.2.10. OS and RFS by FLT3 status at diagnosis (data cut-off date, 15 July 2019)

Abbreviations: AZA = oral azacitidine; mut = mutant; FLT3-ITD = fms-like tyrosine 3-internal tandem duplication; FT3-TKD = fms-like tyrosine 3-tyrosine kinase domain; OS = overall survival; RFS = relapse-free survival; wt=wile type
Note: *FLT3*^{wt} includes patients who were negative for *FLT3*-ITD and for *FLT3*-TKD.

Data from the secondary data lock (8 September 2020) confirmed that patients with FLT3 mutations (FLT3^{mut+}; includes FLT3-ITD and FLT3-TKD) in the oral azacitidine group had improved OS compared with placebo with a median OS of 28.2 months and 9.7 months for the oral azacitidine and placebo groups, respectively (Figure B.2.11).[104-106]

Figure B.2.11. OS by FLT3 status at diagnosis (data cut-off date, 8 September 2020)



FLT3 status at diagnosis	Median OS (months)	
	Oral azacitidine (n=30)	Placebo (n=36)
FLT3^{mut}	28.2	9.7
FLT3^{wt}	24.7	15.2

Abbreviations: CC-486 = oral azacitidine; mut = mutant; FLT3 = fms-like tyrosine 3; OS = overall survival; wt = wild-type.
Source: Reid et al., 2021[104]; Dohner et al, 2021 (Oral Presentation)[105]; Dohner et al, 2021[106]

Despite the small sample size of these patients in the ITT population, these results show a positive and significant improvement in survival consistent with results in the ITT population.

B.2.7.4 OS multivariate analysis

OS multivariate analysis confirmed the independent prognostic impact of *FLT3*-ITD/TKD^{mut} (unfavourable vs. *FLT3*^{wt} [HR 1.54; p<0.012]) when controlling for the *FLT3-ITD/TKD* mutation and for the randomised treatment arm (oral azacitidine vs. placebo). Oral azacitidine also significantly improved OS independent of *FLT3* mutation status (HR 0.72; p=0.003).[105] In summary, *FLT3*-ITD/TKD mutations at diagnosis appeared to have a negative prognostic influence in the placebo arm. Treatment benefit with oral azacitidine vs. placebo was observed in patients in remission with *FLT3*-ITD/TKD mutations at AML diagnosis. Multivariate analyses confirmed the independent prognostic influence of *FLT3* mutations, and oral azacitidine showed improvement in OS independent of these mutations.[105]

B.2.7.5 OS and RFS in subgroups defined by number of consolidation courses

Subgroup analyses based on the number of consolidation courses received prior to study entry demonstrated consistent benefits in terms of OS and RFS for oral azacitidine compared with placebo. The comparison was performed between patients who did not receive consolidation courses prior to study entry (20% of the ITT population), patients who received one cycle of consolidation (45%), and patients with ≥2 consolidation courses (35%)[107] Baseline characteristics were generally similar between treatment arms and among the different cohorts.

In patients without consolidation treatment, median OS from time to randomization was 23.3 months and 10.9 months in the oral azacitidine group and placebo group, respectively. Median RFS was 8.4 months and 3.9 months for oral azacitidine and placebo, respectively. In patients with one prior consolidation course, median OS was 21.0 months vs. 14.3 months and median RFS was 10.0 months vs. 4.7 months for oral azacitidine compared and placebo, respectively. In the cohort of patients with ≥2 cycles of consolidation, median OS was 28.6 months vs. 17.6 months and median RFS was 13.0 months vs. 6.1 months for oral azacitidine compared with placebo, respectively (Table B.2.17).[107]

Table B.2.17. OS and RFS from time of randomisation in patients who received no consolidation, 1 consolidation cycle, or ≥2 consolidation cycles in the QUAZAR AML-001 study (ITT population)

Number of consolidation cycles	Oral azacitidine (N=238)	Placebo (N=234)
No consolidation		
Number of patients, n	52	42
Median OS, months (95% CI)	23.3 (13.5-37.5)	10.9 (6.3-15.7)
HR (95% CI)	0.55 (0.34-0.89)	
Median RFS, months (95% CI)	8.4 (7.5-16.2)	3.9 (1.9-4.9)
HR (95% CI)	0.55 (0.34-0.88)	

Number of consolidation cycles	Oral azacitidine (N=238)	Placebo (N=234)
1 consolidation		
Number of patients, n	110	102
Median OS, months (95% CI)	21.0 (16.7-30.5)	14.3 (11.7-18.0)
HR (95% CI)	0.75 (0.55-1.02)	
Median RFS, months (95% CI)	10.0 (7.4-11.7)	4.7 (4.0-7.4)
HR (95% CI)	0.72 (0.53-0.99)	
≥2 consolidation		
Number of patients, n	76	90
Median OS, months (95% CI)	28.6 (17.8-41.3)	17.6 (11.6-28.7)
HR (95% CI)	0.75 (0.50-1.11)	
Median RFS, months (95% CI)	13.0 (7.7-21.2)	6.1 (4.6-7.5)
HR (95% CI)	0.59 (0.41-0.87)	

Abbreviations: CI = confidence interval; HR = hazard ratio; OS = overall survival; RFS = relapse-free survival
Source: Wei et al., 2020[107]

Overall, outcomes were improved in the cohorts with prior consolidation treatment except for the oral azacitidine subgroups with no or one prior consolidation course, where OS was similar with 23.3 months and 21.0 months, respectively. The subgroups defined by consolidation were not prespecified and the sample size of the study was not powered for subgroup analyses; therefore, these results should be interpreted with caution.[107]

B.2.8 Meta-analysis

Efficacy data supporting the use of oral azacitidine for the treatment of AML are provided by a single Phase 3 study (QUAZAR AML-001). Therefore, a meta-analysis was not conducted.

B.2.9 Indirect and mixed treatment comparisons

The comparators included in the final scope for this appraisal which are considered relevant include:

- Established clinical management without oral azacitidine (a 'watch and wait' strategy with BSC)
- Midostaurin in a subgroup of patients with *FLT3*-mutation positive AML[31]

Direct head-to-head data comparing oral azacitidine + BSC to matching placebo + BSC, considered a proxy for 'watch and wait' strategy with BSC, are available from the phase III QUAZAR AML-001 trial;[84, 87] however, the comparative efficacy of oral azacitidine as maintenance treatment has not been assessed in any head-to-head clinical studies with midostaurin in patients with AML and a *FLT3-ITD* and/or *FLT3-TKD* (*FLT3* mutation). Therefore, an indirect treatment comparison (ITC) or mixed treatment comparison (MTC) was

necessary to determine the comparative efficacy of oral azacitidine versus midostaurin in the FLT3 subgroup.

B.2.9.1 Data sources

The SLR conducted to assess the clinical effectiveness for maintenance treatment of patients with AML who have achieved first CR after induction with IC with or without consolidation identified two studies which assessed the clinical efficacy of oral azacitidine and midostaurin. The QUAZAR AML-001 study compared oral azacitidine to placebo + BSC.[84] RATIFY was included for midostaurin in patients with *FLT3*-mutation-positive AML,[82] (see Appendix D for further detail on the SLR). These trials were considered relevant to UK clinical practice and were used to assess the feasibility of conducting an ITC (**Error! Reference source not found.**) to compare oral azacitidine to midostaurin in the FLT3-positive subgroup.

Data from QUAZAR AML-001 was obtained from the CSR, IPD and trial publication.[87, 88] Data from RATIFY was obtained from Stone et al. (2017)[82] to inform the study design/eligibility comparison and the secondary landmark analysis of midostaurin in maintenance from Larson et al. (2021)[83] was used to inform outcome data for the analyses.

Table B.2.18. Summary of the trials used to carry out the indirect treatment comparison

	Oral azacitidine	Placebo[a]	Midostaurin
QUAZAR AML-001[87, 88]	Yes	Yes	
RATIFY[82, 83]		Yes	Yes

[a] Only the QUAZAR AML-001 trial reported details of supportive care administered in the control arm; therefore, to form a connected network, supportive care with placebo was assumed to be equivalent to placebo alone.

QUAZAR-AML-001 is the pivotal Phase 3, multicentre, double-blind, randomised, placebo-controlled, parallel-group study conducted to characterise the efficacy and safety of maintenance treatment with oral azacitidine in adults patients (≥55 years) with AML in CR/CRi after induction therapy with or without consolidation chemotherapy, who were not candidates for HSCT (see Section B.2.3). There were no restrictions on study eligibility with regard to AML mutation (66 of the 472 enrolled patients had a FLT3 mutation).[106] The primary outcome was OS, defined as time from randomisation until death from any cause. The key secondary outcome included RFS which was defined as the time from maintenance therapy randomization (i.e., time zero) to the date of documented relapse after CR/CRi or death from any cause, whichever occurred first. Other secondary outcomes included safety, HRQoL and HCRU.

RATIFY was a Phase 3, multicentre, double-blind, randomised, placebo-controlled study of midostaurin in combination with standard chemotherapy in adult patients (≤59 years) with newly diagnosed *FLT3*-mutation-positive AML.[82] In this study, eligibility for HSCT was not a formal exclusion criterion. The study consisted of a screening phase, a blinded treatment phase, and a follow-up phase. In total, 717 patients from 225 sites across 17 countries were randomized and received treatment with midostaurin (n=360) or placebo (n=357) in combination with intensive induction chemotherapy (cytarabine and daunorubicin) and consolidation chemotherapy (high-dose cytarabine). Patients who remained in CR after completion of consolidation chemotherapy entered a 12-month maintenance phase in which

they received monotherapy with either midostaurin (50 mg; n=120) or placebo (n=85), administered orally twice daily.

The primary efficacy outcome of the RATIFY study was OS, defined as the time from randomisation to death from any cause. The key secondary outcome was event-free survival (EFS), defined as the time from randomisation to relapse, death from any cause, or failure to achieve protocol-specified CR. Time zero for both time-to-event outcomes was defined at randomisation to intensive induction chemotherapy. Additional secondary outcomes included OS censored at the time of HSCT, CR rate, disease-free survival (DFS; defined as the time from protocol-specified CR to relapse or death from any cause), DFS one year after completing planned maintenance treatment, and HSCT rate. For the purpose of this analysis, the secondary outcomes of EFS and DFS are considered synonymous with RFS.

In the full study population, both OS (HR: 0.78, 95%CI: 0.63-0.96; p=0.009) and EFS (HR: 0.78, 95%CI: 0.66-0.93; p=0.002) were significantly longer among patients who received midostaurin in combination with standard chemotherapy than among those who received placebo in combination with standard chemotherapy.

Although the RATIFY study included a 12-month maintenance phase, patients were not re-randomised prior to the start of maintenance treatment, as the study was designed to assess the addition of midostaurin to standard chemotherapy (induction and consolidation) versus chemotherapy alone.

Therefore, for the purposes of the ITC, the following data was used:

- IPD data from the *FLT3* mutation subgroup of the QUAZAR AML-001 trial was matched to the extent possible to the eligibility criteria in the RATIFY trial (see section B2.9.3.1 below). Analyses for OS were conducted using the September 2020 data-cut, with RFS from the July 2019 data-cut.
- Data from a secondary landmark analysis (Larson et al., 2021)[83] of the RATIFY trial was used to align with the outcome definition in QUAZAR AML-001. In this secondary landmark analysis, OS and DFS was measured from the start of the maintenance treatment in the 205 patients that entered the maintenance phase of the RATIFY study.
    - In this study, the contribution of midostaurin maintenance treatment to survival failed to demonstrate statistical significance (see Appendix D for further detail).[83]

B.2.9.2 Feasibility assessment

Selection of the appropriate methodology to determine the comparative efficacy of oral azacitidine versus midostaurin relies upon the availability of study data and between-study heterogeneity. Therefore, a feasibility assessment was conducted to determine the most appropriate method to derive estimates of the comparative efficacy of oral azacitidine and midostaurin in the FLT3-subgroup.

An assessment of the evidence base identified substantial heterogeneity in the study characteristics of the QUAZAR AML-001 and RATIFY trials. Specifically, the studies differed across:

- Trial design: The primary analysis of the RATIFY trial was not prospectively designed to assess the efficacy of midostaurin as a maintenance therapy; rather, the trial was designed to assess the addition of midostaurin to induction and consolidation with standard chemotherapy versus chemotherapy alone.
- Patient time to randomisation: Although the RATIFY trial included a 12-month maintenance therapy phase, patients were not re-randomized prior to the start of maintenance therapy, confounding the contribution of maintenance treatment to overall outcomes.
- Patient age: The inclusion criteria for QUAZAR AML-001 was ≥55 years compared with RATIFY which included patients aged 18-59 years, highlighting the limited overlap between the two populations
- Cytogenetic risk: Favourable cytogenetic risk patients were included in RATIFY but not in QUAZAR AML-001.
- AML mutational status: Patient eligibility for the RATIFY trial was restricted to a specific subgroup of AML patients (i.e., FLT3 mutation-positive AML), whereas patients were included regardless of their mutational status in the QUAZAR AML-001 trial.
- HSCT eligibility: Stem cell transplant eligibility was not a formal exclusion criterion in the RATIFY trial; however, 57% of patients underwent HSCT. The QUAZAR AML-001 trial excluded patients who were eligible for HSCT at study screening; 6% of patients treated with oral AZA underwent HSCT.
- Time zero definitions: Time-to-event outcomes defined in the primary analyses of the included trials were differentially defined, as time zero definitions differed between studies.
    - A comparison of OS outcomes across the primary analyses of the trials would bias results against oral AZA, as patients in the RATIFY trial would benefit from immortal time bias.
    - A comparison of RFS and DFS across the primary analyses of the trials would be inappropriate, as time of randomization to maintenance therapy is not always associated with achievement of CR/CRi.
- History of consolidation therapy: Heterogeneity was observed across history of consolidation between the QUAZAR-AML-001 and RATIFY trials.

Many of these variables are known prognostic factors and potential effect modifiers, with distributions across studies generally favouring the pool of patients from the RATIFY study. Furthermore, the primary analysis in the RATIFY study[82] accounted for patients from the start of induction therapy whereas the QUAZAR-AML-001 trial considered patients from the initiation of maintenance, therefore the secondary landmark analysis from the RATIFY trial had to be used.[83] This feature introduces methodological heterogeneity into the analysis. Patient characteristics such as race, type of AML, ECOG performance, MRD status, bone marrow blasts in the RATIFY trial were available only from the start of induction therapy rather than maintenance, highlighting further between-study heterogeneity. Together, these differences underscore the extent of the clinical and methodological heterogeneity between

studies, leading to distinct patient populations who are not eligible for inclusion in the comparator study. The validity of findings from ITCs rely heavily upon the exchangeability assumption, which suggests that different sets of RCTs used must be, on average, similar in all important factors (e.g. effect modifiers). Due to the extent of the between-study heterogeneity identified in this study, indirect estimates of oral azacitidine and midostaurin are likely limited in their validity and generalisability.

Details on the feasibility assessment are provided in Appendix D.

B.2.9.2.1 Adjusted indirect comparisons

Both a simulated treatment comparison (STC) and a matching adjusted indirect comparison (MAIC) can generate similar effect size estimates and the choice of one method over another may depend on the number of comparators and outcomes.[108] STC may be preferred the analysis has multiple comparators and few outcomes, whereas MAIC may be preferred if there are multiple outcomes and few comparators (MAICs require a detailed assessment of inclusion criteria; as such, scalability to multiple comparisons may be limited). However, recent simulation studies have shown that STCs produce less biased estimates compared with MAICs.[109] Since an MAIC uses a reweighting method, and therefore does not permit extrapolation, bias can only be completely removed when the population of the study with summary-level data are entirely contained within the population of the IPD study. STC is the preferred option when there is minimal overlap between study populations since it can extrapolate beyond the range of the IPD.

Additionally, population adjustment methods may not be needed if the imbalance in treatment effect modifiers is small. The NICE technical support document (TSD) 18 recommends that an STC can be performed when it is likely to produce less biased estimates of treatment differences than achieved through standard methods such as Bucher ITCs and network meta-analyses.[110] However, given the differences in study design and patient populations an STC was assessed as not justifiable based on the criteria from the NICE TSD.[110] A detailed assessment consisting of an evaluation of effect modifier status of relevant variables and evidence of substantial imbalance was performed (see Appendix D for further detail).

In summary, the feasibility assessment determined that adjustment across populations is not possible due to:

- Inadequate data for RATIFY,[82, 83] with no reported baseline characteristics of patients progressing to the maintenance phase.
- Significant differences between study populations resulting in insufficient overlap of population characteristics across studies when using the reported baseline characteristics of the full study population of the RATIFY study.[82]

It is evident from the available evidence that QUAZAR AML-001 and RATIFY differ in terms of likely effect modifiers (age, and *FLT3* mutation status). However, an STC is unlikely to produce less biased effect estimates compared to Bucher ITCs restricted to the FLT3 population. Significant differences exist in AML occurring in patients older than 55 to 60 years of age and these differences likely are not well characterized by a linear relationship.[111] Since there is poor overlap in terms of age between QUAZAR AML-001 and RATIFY, any extrapolation performed is subject to bias and unreasonably large uncertainties. Further, as

mentioned previously, all patients in RATIFY contained a FLT3 mutation so it was deemed most appropriate to restrict the QUAZAR AML-001 trial to just the FLT3 subpopulation.

To this end, despite the limitations of conducting an unadjusted ITC (see Section B.2.9.6), a matched Bucher ITC was conducted to inform directional estimates of comparative efficacy.

The methodology for the Bucher ITC is provided in Appendix D and results are presented in B.2.9.3. Since the primary analysis violated the proportional hazards (PH) assumption, time-varying parametric and spline models were used, percent survival at 6, 12 and 24 months were derived from these models for all treatments in both QUAZAR AML-001 and RATIFY (the methodology is presented in Appendix D and results are presented in Section B.2.9.4).

B.2.9.3 Bucher indirect treatment comparison results

B.2.9.3.1 Primary analysis population

Since access to IPD from oral azacitidine was available, matching was performed to align the inclusion and exclusion criteria between the RATIFY[83] and QUAZAR AML-001 trials[87] (see Appendix D; Section D.1.2.3). Patients from QUAZAR AML-001 were removed from the IPD if they did not satisfy the eligibility criteria used in the RATIFY study e.g. QUAZAR AML-001 included patients without *FLT3* mutations and CRi but the RATIFY study did not. Therefore, IPD for these patients were removed from QUAZAR AML-001 to match the inclusion and exclusion criteria of RATIFY.

Initially, the QUAZAR AML-001 study consisted of 472 patients within the study. After matching the inclusion/exclusion criteria of QUAZAR AML-001 to RATIFY, [redacted] patients were removed from the study and [redacted] patients (oral azacitidine, n=[redacted]; placebo, n=[redacted]) with a *FLT3*-mutation and achieved CR remained within the primary analysis population (see Appendix D; Section D.1.2.3; Figure B.5.3).

B.2.9.3.2 OS

Both prior to matching (n=472) and after matching (n=[redacted]), results demonstrated that the HR is numerically favourable for oral azacitidine compared to midostaurin ([redacted]) and ([redacted]), respectively. Results for OS are provided in Table B.2.19.

Table B.2.19. Results for OS

Scenario	Oral azacitidine vs midostaurin HR (95% CI)
Unmatched	[redacted]
Primary Analysis	
Matched[a]	[redacted]

Abbreviations: CI = confidence interval; HR = hazard ratio; OS = overall survival
[a]416 patients with CRi and no FLT3 mutation were removed from the unmatched population to align with inclusion/exclusion criteria in RATIFY.

The matched results were assessed for suitability for usage in the economic model. This was judged by assessing the PH assumption as described in detail in Section **Error! Reference source not found.**. For OS it was concluded that the PH assumption was likely violated and hence the HR approach was not used. Parametric and spline models were fitted to each trial relaxing the assumption of proportional hazards.

B.2.9.3.3 RFS

Similarly, both prior to matching (n=472) and after matching (n=[redacted]), the HR is [redacted] for oral azacitidine compared to midostaurin ([redacted]) and ([redacted]), respectively. The matched results were used in the model. Results for RFS are provided in Table B.2.20.

Table B.2.20. Results for RFS

Scenario	Oral azacitidine vs midostaurin HR (95% CI)
Unmatched	[redacted]
Primary Analysis	
Matched[a]	[redacted]

Abbreviations: CI = confidence interval; HR = hazard ratio; RFS = relapse-free survival
[a]416 patients with CRi and no FLT3 mutation were removed from the unmatched population to align with inclusion/exclusion criteria in RATIFY.

The matched results were assessed for suitability for usage in the economic model. This was judged by assessing the PH assumption as described in in detail in Section B.3.3.2.4.1. For RFS it was concluded that the PH assumption was likely violated and hence the HR approach was not used. Parametric and spline models were fitted to each trial relaxing the assumption of proportional hazards.

B.2.9.4 Time-varying methods (parametric and spline models)

Based on AIC, BIC, and clinical validity, generalised gamma models were used to determine percent survival for OS and the 1 knot odds linear model for RFS in QUAZAR AML-001 and RATIFY (see Section **Error! Reference source not found.** for OS and Section **Error! Reference source not found.** for RFS, see Appendix D for further detail). Results demonstrated that patients in QUAZAR AML-001 survived and were relapse-free longer than patients in RATIFY (see Appendix D). Similar results were obtained using time-varying spline models where a model with 1 internal knot and an odds linear predictor was used for OS and a model with 1 internal knot and an normal linear predictor was used for RFS. Best fitted time varying spline models were determined using AIC, BIC and clinical validity (see Appendix D for further detail).

B.2.9.5 Conclusion

In the primary Bucher analysis for OS and RFS, oral azacitidine was observed to provide [redacted] compared to midostaurin across both outcomes[redacted]

It is important to note that the trials of oral azacitidine and midostaurin were considerably different from one another in terms of study design, inclusion/exclusion criteria and baseline patients' characteristics. Since the primary analysis violated the PH assumption, time-varying parametric and spline models were used which accounted for violation of the PH assumption (see Appendix D for further detail). This analysis showed similar results to the primary analysis, with increased benefit for oral azacitidine in both OS and RFS.

In summary, the results show numerical benefits for oral azacitidine patients compared to patients in RATIFY for OS and RFS. Despite the limitations (detailed below), this ITC represents the best possible evidence to inform comparative effectiveness.

B.2.9.6 Uncertainties in the indirect and mixed treatment comparisons

The validity of findings from an anchored indirect comparison relies on the assumption of constant efficacy (i.e. the relative effectiveness of a treatment is the same across all included studies) and by extension requires that studies comprising the evidence base are similar (i.e. exchangeable) with respect to all important factors (e.g. effect modifiers). Therefore, significant differences in patient and study characteristics across studies, as identified in this analysis, can limit the validity and generalisability of derived effect estimates and represents a limitation of the comparison. A summary of key differences identified across studies is provided below:

Study design: In contrast with QUAZAR AML-001, the RATIFY study was not prospectively designed to assess the efficacy of midostaurin as a maintenance treatment; rather, the study was designed to assess the addition of midostaurin to standard chemotherapy versus chemotherapy alone. Further, there are differences in the time of randomisation between the QUAZAR AML-001 and RATIFY studies. Patients in QUAZAR AML-001 were randomised to maintenance treatment. In comparison, patients in RATIFY were randomised to induction chemotherapy but not re-randomised at the start of maintenance treatment, making drawing any inferences regarding the effectiveness of midostaurin in the maintenance setting difficult. Together, these distinctions highlight the heterogeneity in study design and severely limit comparability across studies.

Inclusion and exclusion criteria: The included studies differed significantly in their eligibility criteria (including patient age, cytogenetic risk, *FLT3* mutational status, and HSCT eligibility) and thus could not be matched upon. Each of these variables are known prognostic factors and potential effect modifiers, with distributions across studies generally favouring the pool of patients from the RATIFY study; specifically, patients in the RATIFY were younger, a majority had favourable- or intermediate-risk cytogenetic characteristics, and many were HSCT eligible at study screening (HSCT was performed in over 20% of patients during first CR). Therefore, due to differences in study populations, any estimates of comparative efficacy derived from comparing the included studies are subject to bias.

Baseline characteristics: Patient baseline characteristics were reported for all patients at randomisation to induction chemotherapy for RATIFY rather than the subset that received maintenance treatment. In comparison, baseline characteristics were measured at randomisation to maintenance treatment in QUAZAR AML-001. This limits comparability across studies because patient characteristics are reported for different AML populations. As a result, although the studies differed significantly with respect to patient characteristics (e.g. significant differences in median age and cytogenetic risk), no adjustment could be made. Since the patient populations of the included studies are not equivalent, estimates of comparative efficacy are subject to bias.

Small sample size. After matching, the QUAZAR AML-001 included [redacted] which makes it difficult to generalise to the broader population. Additionally, effect estimates derived

for the matched QUAZAR AML-001 population lack the precision observed in the unmatched population.

B.2.10 Adverse reactions

Summary

- Oral azacitidine was well tolerated, with a low rate of discontinuation due to treatment-emergent AEs (TEAEs) (only 13% of patients had ≥1 TEAE leading to discontinuation of treatment with oral azacitidine) and there were [redacted] reported treatment-related deaths
- Rates of SAEs and Grade 3/4 TEAEs were relatively similar between treatment groups
- The most frequently reported Grade 3/4 TEAEs in both groups were neutropenia, thrombocytopenia, and anaemia
- Although GI TEAEs were more common in the oral azacitidine group than in the placebo group, the majority of these events were low in severity and declined in frequency over time
  - The most common gastrointestinal (GI) TEAEs were nausea, vomiting, diarrhoea, and constipation
  - The use of prophylactic antiemetics and anti-diarrhoea medication was not mandated because of the double-blind nature of the study; however, the oral azacitidine SmPC states that patients should be given anti-emetics prior to each dose of oral azacitidine for the first 2 treatment cycles and may be omitted after 2 cycles if there is no nausea and vomiting, to reduce the risk of GI TEAEs
  - Few GI TEAEs led to treatment discontinuation

B.2.10.1 Extent of exposure

A summary of treatment exposure and relative dose intensity (RDI) is provided in Table B.2.21.

The mean treatment duration was [redacted] in the oral azacitidine group and [redacted] in the placebo group.[87] The median average daily dose of oral azacitidine was [redacted] in both treatment groups. The mean number of treatment cycles received was [redacted] in the oral azacitidine group and [redacted] in the placebo group, with a mean cycle length of greater than 28 days in both groups. The mean RDI was [redacted] in the oral azacitidine group and [redacted] in the placebo group, with [redacted] respectively, receiving > 85% to ≤ 100% of planned dose intensity.[87] Furthermore, treatment compliance was high (mean overall compliance was [redacted] in the oral azacitidine and placebo groups, respectively).[87]

Table B.2.21. Summary of treatment exposure, QUAZAR AML-001 study (safety population)

Parameter	Oral azacitidine (N=236)	Placebo (N=233)
Treatment duration[a], months		
Mean (SD)	[redacted]	[redacted]
Median (min, max)	11.6 (0.5, 74.3)	5.7 (0.7, 68.5)
Treatment duration[b], person-years	[redacted]	[redacted]
Average length of cycle[c], days		
Mean (SD)	[redacted]	[redacted]
Median (min, max)	[redacted]	[redacted]
Average number of days dosed per cycle[d]		
Mean (SD)	[redacted]	[redacted]
Median (min, max)	[redacted]	[redacted]
Number of cycles		
Mean (SD)	[redacted]	[redacted]
Median (min, max)	12.0 (1.0, 80.0)	6.0 (1.0, 73.0)
Number of treatment cycles initiated, n (%)		
1 or more	236 (100.0)	233 (100.0)
2 or more	[redacted]	[redacted]
3 or more	[redacted]	[redacted]
4 or more	[redacted]	[redacted]
5 or more	[redacted]	[redacted]
6 or more	[redacted]	[redacted]
12 or more	[redacted]	[redacted]
18 or more	[redacted]	[redacted]
24 or more	[redacted]	[redacted]
30 or more	[redacted]	[redacted]
Relative dose intensity (%)[e]		
Mean (SD)	[redacted]	[redacted]

≤ 75%, n (%)	[redacted]	[redacted]
> 75% to ≤ 85%, n (%)	[redacted]	[redacted]
> 85% to ≤ 100%, n (%)	[redacted]	[redacted]
> 100%, n (%)	[redacted]	[redacted]

Abbreviations: max = maximum; min = minimum; n = number of patients in the category; N = number of patients evaluable; SD = standard deviation.
[a] Treatment duration in months is defined as (treatment end date — first dose date +1)/30.4375. Treatment end date is last dose date + 14 days (the prescribed rest period of each cycle), or the death date, whichever is earlier.
[b] Total person-years of treatment duration is calculated as the sum of treatment duration(days)/365.25 across all patients.
[c] Average cycle length is defined as treatment duration in days/number of cycles.
[d] Average number of days dosed per cycle is defined as total number of days dosed during the entire treatment period/number of cycles.
[e] Relative dose intensity is defined as the ratio of dose intensity to the planned dose intensity (300 mg/day x 14 days/28 days = 150 mg/day for all subjects).
Source: QUAZAR AML-001 CSR (Data on File)[87]; ClinicalTrials.gov[93]

B.2.10.2 Summary of adverse events

When comparing the incidence of TEAEs, it is worthwhile to note that duration of exposure to oral azacitidine (11.6 months) was approximately twice as long as exposure in the placebo group (5.7 months).[87] The proportion of patients who experienced at least one TEAE considered by the investigator to be related to study treatment was higher in the oral azacitidine group than in the placebo group (89.8% vs. 51.5%).[93] The rates of serious TEAEs (oral azacitidine: 33.5%; placebo: 25.3%), Grade 3/4 TEAEs (oral azacitidine: 71.6%; placebo: 63.1%) and TEAEs leading to death (oral azacitidine: 3.8%; placebo: 1.7%) were slightly higher in the oral azacitidine group than in the placebo group.[87] None of the TEAEs leading to death were considered to be related to study treatment.[87] A summary of the TEAEs is presented in Table B.2.22.

Table B.2.22. Summary of ≥1 TEAEs, QUAZAR AML-001 study (safety population)

Category	Oral azacitidine (N=236)	Placebo (N=233)
TEAEs, n (%)	231 (97.9)	225 (96.6)
TEAEs related to study treatment, n (%)	212 (89.8)	120 (51.5)
Serious TEAEs, n (%)	79 (33.5)	59 (25.3)
Treatment-related serious TEAEs, n (%)	22 (9.3)	5 (2.1)
Grade 3/4 TEAEs[a], n (%)	169 (71.6)	147 (63.1)
Treatment-related Grade 3/4 TEAEs[a], n (%)	113 (47.9)	54 (23.2)
TEAEs leading to death, n (%)	9 (3.8)	4 (1.7)
TEAEs leading to dose reduction, n (%)	37 (15.7)	6 (2.6)
TEAEs leading to dose interruption, n (%)	102 (43.2)	40 (17.2)

TEAEs leading to dose reduction and interruption, n (%)	24 (10.2)	3 (1.3)
TEAEs leading to study treatment discontinuation, n (%)	31 (13.1)	10 (4.3)

Abbreviations: n = number of patients in the category; N = number of patients evaluable; TEAE = treatment-emergent adverse event.
[a] Graded using Common Terminology Criteria for Adverse Events version 4.0.
Notes: AML relapse as defined by MedDRA high-level group term leukaemia's is excluded. AEs were evaluated from the first dose date through 28 days after the last dose of study treatment.
Source: ClinicalTrials.gov[93]

B.2.10.3 AEs

The most common TEAEs were GI events, which occurred more frequently in the oral azacitidine group (91.1%) than in the placebo group (61.8%). GI events included nausea, vomiting, diarrhoea, and constipation (Table B.2.23).[87] However, the majority of these events were mild or moderate in severity (Grade 1/2); Grade 3/4 GI TEAEs only occurred in 14.4% of patients in the oral azacitidine group and 5.6% of patients in the placebo group, and included diarrhoea (5.1% vs. 1.3%), vomiting (3.0% vs. 0%), nausea (2.5% vs. 0.4%), and constipation (1.3% vs. 0%).[87, 94, 95] In addition, most GI TEAEs occurred in the first two treatment cycles and the frequency decreased considerably with continued treatment. This finding may have occurred because use of prophylactic anti-emetics and anti-diarrhoea medication was not mandated by protocol, but clinicians could prescribe them as required. In contrast, the SmPC of oral azacitidine stipulates the use of an anti-emetic prior to each dose of oral azacitidine for the first two treatment cycles, with subsequent review. Thus, in real-world practice, earlier initiation of antiemetic treatment may reduce the incidence and severity of nausea and vomiting. Although GI events were the most common TEAEs observed during maintenance treatment with oral azacitidine, a relatively small percentage of patients who experienced these events required dose reduction ([redacted] for oral azacitidine vs. [redacted] for placebo), dose interruption [redacted] or treatment discontinuation [redacted][87]

The most common haematologic TEAEs were neutropenia, thrombocytopenia, and anaemia.[87] The percentage of patients with haematologic AEs within each treatment group were generally consistent over time up to Cycle 12.[87] Haematologic TEAEs were primarily managed with dosing modifications in the oral azacitidine group including treatment interruption in 27% of patients and dose reductions in 8% of patients. Neutropenia was the most frequent TEAE leading to treatment modifications: of the 105 patients who experienced neutropenia in the oral azacitidine group, 45% had treatment interruptions and 12% had dose adjustments (further information on TEAEs leading to dose modifications and treatment discontinuation is provided in Section B.2.10.4).

Grade 3/4 TEAEs were generally higher in the oral azacitidine group, the types of TEAEs were consistent with the safety profile of azacitidine and/or characteristic of disease relapse. The most common Grade 3/4 TEAEs reported with oral azacitidine were neutropenia, thrombocytopenia, anaemia, and febrile neutropenia (Table B.2.23).

Table B.2.23. TEAEs reported in >10% of patients, QUAZAR AML-001 study (safety population)

Event	Oral azacitidine (N=236)		Placebo (N=233)	
	Any grade	Grade 3/4	Any grade	Grade 3/4
TEAEs, n (%)	231 (98)	169 (72)	225 (97)	147 (63)
Nausea	153 (65)	6 (3)	55 (24)	1 (<1)
Vomiting	141 (60)	7 (3)	23 (10)	0 (0)
Diarrhoea	119 (50)	12 (5)	50 (21)	3 (1)
Neutropenia	105 (44)	97 (41)	61 (26)	55 (24)
Constipation	91 (39)	3 (1)	56 (24)	0 (0)
Thrombocytopenia	79 (33)	53 (22)	63 (27)	50 (21)
Fatigue	70 (30)	7 (3)	45 (19)	2 (1)
Anaemia	48 (20)	33 (14)	42 (18)	30 (13)
Asthenia	44 (19)	2 (1)	13 (6)	1 (<1)
Pyrexia	36 (15)	4 (2)	44 (19)	1 (<1)
Arthralgia	32 (14)	2 (1)	24 (10)	1 (<1)
Abdominal pain	31 (13)	2 (1)	16 (7)	0 (0)
Upper respiratory tract infection	31 (13)	1 (<1)	32 (14)	0 (0)
Decreased appetite	30 (13)	2 (1)	15 (6)	2 (1)
Cough	29 (12)	0 (0)	39 (17)	0 (0)
Febrile neutropenia	28 (12)	27 (11)	18 (8)	18 (8)
Back pain	28 (12)	3 (1)	23 (10)	2 (1)
Leukopenia	25 (11)	18 (8)	19 (8)	14 (6)
Pain in extremity	25 (11)	1 (<1)	12 (5)	0 (0)
Dizziness	25 (11)	0 (0)	21 (9)	0 (0)
Headache	23 (10)	0 (0)	26 (11)	1 (<1)
Peripheral oedema	21 (9)	0 (0)	24 (10)	1 (<1)

Abbreviations: n = number of patients in the category; N = number of patients evaluable; TEAE = treatment-emergent adverse event.
Notes: TEAEs were evaluated from the first dose date through 28 days after the last dose of study treatment. Events were coded according to preferred terms from the Medical Dictionary of Regulatory Activities, version 22 and were graded with the use of the National Cancer Institute Common Terminology Criteria for Adverse Events, version 4.0.

Source: Wei et al., 2020[84]

At least one serious TEAE was reported for 33% of the patients in the oral azacitidine group and 25% in the placebo group (Table B.2.24). The most common serious TEAEs were infections, which were reported in 17% of patients in the oral azacitidine group and 8% of patients in the placebo group.[88] The most frequently reported serious TEAEs were febrile neutropenia and pneumonia (Table B.2.24).

Table B.2.24. Serious TEAEs reported in ≥1% of patients in either treatment arm, QUAZAR AML-001 study (safety population)

Event	Oral azacitidine (N=236)	Placebo (N=233)
Serious TEAEs, n (%)	79 (33)	59 (25)
Febrile neutropenia	16 (7)	9 (4)
Pneumonia	9 (4)	7 (3)
Pyrexia	5 (2)	1 (0.4)
Cellulitis	4 (2)	1 (0.4)
Sepsis	4 (2)	5 (2)
Influenza	3 (1)	0 (0)
Diarrhoea	3 (1)	0 (0)
Back pain	3 (1)	0 (0)
Atrial fibrillation	3 (1)	0 (0)
Cholecystitis	3 (1)	2 (1)
Anaemia	2 (1)	3 (1)
Thrombocytopenia	2 (1)	3 (1)

Abbreviations: n = number of patients in the category; N = number of patients evaluable; TEAE = treatment-emergent adverse event.
Notes: Events were coded according to preferred terms from the Medical Dictionary of Regulatory Activities. A patient is counted only once for multiple events within preferred term/system organ class.
Source: Supplementary appendix to Wei et al, 2020[88]

B.2.10.4 AEs leading to dose reduction, dose interruption and treatment discontinuation

AEs leading to dose reduction were reported for 16% of patients in the oral azacitidine group and 3% of patients in the placebo group.[84] The most frequent AEs leading to dose reduction (reported for ≥1% of patients in either treatment arm) were neutropenia (6% vs. 0.4%), diarrhoea (3% vs. 0%), thrombocytopenia (2% vs. 1%), and nausea (2% vs. 0%) for oral azacitidine vs. placebo, respectively.[88]

AEs leading to dose interruption were reported for 43% of patients in the oral azacitidine group and 17% of patients in the placebo group.[84] The most frequent AEs leading to dose interruption (reported for ≥1% of patients in either treatment arm) were neutropenia (20% vs. 6%), thrombocytopenia (8% vs. 2%), nausea (6% vs. 0.4%), diarrhoea (4% vs. 1%), vomiting (4% vs. 0%), febrile neutropenia (2% vs. 0.4%), and alanine aminotransferase increased (2% vs. 1%) for oral azacitidine vs. placebo, respectively.[88]

Discontinuation of study treatment because of AEs was reported for 13% of patients in the oral azacitidine group and 4% of patients in the placebo group.[84] In the oral azacitidine group, AEs leading to treatment discontinuation reported by >1 patient in either treatment arm included nausea (2% vs. 0%), diarrhoea (2% vs. 0%), vomiting (1% vs. 0%), abdominal pain (1% vs. 0%), fatigue (1% vs. 0%), and thrombocytopenia (0.4% vs. 1%) for oral azacitidine vs. placebo, respectively.[88]

B.2.10.5 Deaths

Deaths during treatment were low and most occurred after Cycle 6: [redacted] patients in the oral azacitidine group and [redacted] patients in the placebo group.[87] Overall, the incidence of deaths due to TEAEs were low. AEs led to death in nine patients (4%) in the oral azacitidine group (two died from sepsis, two from cerebral haemorrhage, one from both sepsis and multiorgan failure, and one each from intracranial haemorrhage, cardiogenic shock, aspiration pneumonia, and suicide).[84] AEs led to death in four patients (2%) in the placebo group (two died from multiorgan failure, one from cerebral haemorrhage, and one from general health deterioration).[2] [redacted] leading to death were considered by the investigator to be treatment related.[87]

B.2.10.6 Treatment-emergent AESIs

A summary of treatment-emergent adverse events of special interest (AESIs) of any grade is presented in Table B.2.25. Whilst, AESIs were slightly higher in the oral azacitidine group (97%) than the placebo group (91%),[87] events were largely manageable with dose modifications and standard therapeutic interventions and a few events were fatal (oral azacitidine vs. placebo: [redacted]) or lead to discontinuation of study therapy (oral azacitidine vs. placebo: [redacted]).[87]

Table B.2.25. Summary of treatment-related AESI (any grade), QUAZAR AML-001 study (safety population)

AESI	Oral azacitidine (N=236)	Placebo (N=233)
Myelosuppression, n (%)	[redacted]	[redacted]
Haemorrhagic events, n (%)	[redacted]	[redacted]
Infections, n (%)	[redacted]	[redacted]
Renal failure, n (%)	[redacted]	[redacted]
Hepatic failure, n (%)	[redacted]	[redacted]
Ischaemic colitis, n (%)	[redacted]	[redacted]

AESI	Oral azacitidine (N=236)	Placebo (N=233)
Cardiac events, n (%)	[redacted]	[redacted]
Psychiatric disorder, n (%)	[redacted]	[redacted]
Tumour lysis syndrome, n (%)	[redacted]	[redacted]
Interstitial lung disease, n (%)	[redacted]	[redacted]
Gastrointestinal events, n (%)	[redacted]	[redacted]
Anxiety, confusional state, insomnia, n (%)	[redacted]	[redacted]

Abbreviations: AESI = adverse event of special interest; n = number of patients in the category; N = number of patients evaluable.
Source: QUAZAR AML-001 CSR (Data on File)[87]

B.2.10.7 Safety conclusion

In the QUAZAR AML-001 study oral azacitidine had a manageable safety profile, with a low rate of discontinuation due to TEAEs (only 13% of patients had ≥1 TEAE leading to discontinuation of treatment with oral azacitidine), and [redacted] reported treatment-related deaths.[87]

Rates of SAEs and Grade 3/4 TEAEs were relatively similar between groups; the most frequently reported Grade 3/4 TEAEs in both groups were neutropenia, thrombocytopenia, and anaemia.

Although GI TEAEs were more common in the oral azacitidine group than in the placebo group, the majority of these events were low in severity and declined in frequency over time. The most common GI TEAEs were nausea, vomiting, diarrhoea, and constipation.[84] Few GI TEAEs led to treatment discontinuation. The use of prophylactic antiemetics and anti-diarrhoea medication was not mandated because of the double-blind nature of the study; however, the oral azacitidine SmPC states that patients may be given antiemetics prior to/or during oral azacitidine treatment duration, to reduce the risk of GI TEAEs.

B.2.11 Ongoing studies

There are no further data cuts for the QUAZAR AML-001 study that will provide additional evidence in the next 12 months for the indication being appraised.

B.2.12 Innovation

As most patients with AML experience disease relapse after induction chemotherapy (Section B.1.3.4.3), effective maintenance treatment for patients who attain remission may play a role in preventing disease relapse and prolonging OS. Oral azacitidine addresses a substantial unmet need for a well-tolerated and easily administered AML maintenance treatment that significantly prolongs survival among patients with AML who are in remission after IC, without compromising HRQoL.

B.2.12.1 Oral azacitidine addresses a high unmet need for an effective and well-tolerated maintenance treatment in AML

The main goal of maintenance treatment in AML patients is to delay relapse and prolong survival.[30] Maintenance treatment is not currently standard of care in the UK. Oral azacitidine is the first and only oral HMA specifically indicated for use as maintenance therapy in all patients with AML in CR/CRi, providing significantly prolonged survival without compromising HRQoL. Maintenance therapy with oral azacitidine significantly extended OS by 9.9 months compared with watch and wait strategy with BSC and more than doubled the duration of RFS (median RFS was 10.2 months corresponding to an improvement of 5.3 months) without compromising the favourable HRQoL of patients in remission.[84] Moreover, oral azacitidine has a manageable safety profile, with a low rate of discontinuation due to treatment-emergent adverse events.[87]

B.2.12.2 Oral azacitidine has a unique pharmacokinetic and pharmacodynamic profile that allows for sustained antileukemic activity

The PK profile of oral azacitidine, combined with the dosing regimen, provides the opportunity to deliver oral azacitidine at low systemic doses over a prolonged period of time (14 days of each 28-day cycle). Prolonged exposure to oral azacitidine allows for sustained anti-leukaemic activity by increasing exposure of diseased cells to the drug. [5, 6] Therefore, oral azacitidine is well suited to long-term use in the AML maintenance setting.

B.2.12.3 Oral azacitidine reduces the burden of disease of patients with AML

Oral azacitidine is orally administered and allows patients to take their medication at home, thereby avoiding the inconvenience associated with frequent and costly hospital/clinical visits for treatment with injectable therapies. Oral azacitidine is also associated with fewer hospitalisations compared with placebo which reduces the burden of disease for patients with AML.[87] In the context of the COVID-19 pandemic, these benefits of an oral treatment is expected to translate into a reduction in NHS burden, i.e. preventing patients from requiring visits to hospital for treatment.

B.2.13 Interpretation of clinical effectiveness and safety evidence

B.2.13.1 Findings from the clinical evidence

The pivotal Phase 3, QUAZAR AML-001 study met its primary outcome and demonstrated a statistically significant and clinically meaningful survival benefit (9.9 month improvement in OS at 3 years from the time of randomisation) for oral azacitidine compared to placebo (Section B.2.6). The study results also demonstrated delay of disease relapse (median RFS of 5.4 months) in patients who achieved CR/CRi following IC, compared with placebo (key secondary outcome). These results were supported by sensitivity analyses of OS and RFS demonstrating the robustness and consistency of the primary and key secondary efficacy outcomes (Sections B.2.6.1 and B.2.6.2).[95]

In addition, all subgroup analyses showed that the OS and RFS benefits provided by oral azacitidine were consistent across demographic and disease-related subgroups. Specifically,

the HRs were <1 for and for the European subgroup, which is likely to reflect UK clinical practice (Section B.2.7).

Other secondary efficacy outcomes (time to relapse and time to discontinuation) in addition to HRQoL measurements are also supportive of the demonstrated benefit of oral azacitidine as maintenance treatment for AML (Section **Error! Reference source not found.**).[95]

Oral azacitidine has a manageable safety profile, with a low rate of discontinuation due to TEAEs. GI events, the most common TEAEs among oral azacitidine-treated patients, are typically mild to moderate in severity and decline in frequency over time (Section B.2.10.7).

In the Bucher ITC, prior to matching (n=472) and after matching ([redacted]), OS was in favour of oral azacitidine compared to midostaurin ([redacted]), respectively. In addition, prior to matching (n=472) and after matching ([redacted]) RFS appears was in favour of oral azacitidine compared to midostaurin ([redacted]) and [redacted] respectively. Despite the limitations of the matched Butcher ITC, the results demonstrate that OS and RFS benefits [redacted] with oral azacitidine compared with midostaurin (Section B.2.9.5).

B.2.13.2 Strengths and limitations of the clinical evidence base

Overall, clinical data for oral azacitidine provide an appropriate evidence base for assessment of its clinical and cost-effectiveness for the maintenance treatment of patients with AML.

The strengths of the clinical evidence base are:

- The QUAZAR AML-001 is a robust, high quality, international, multicentre, RCT and was used in the EMA/MHRA marketing authorisation submissions
- The study population is consistent with the population in the NICE scope
- The study included [redacted] patients from the UK and the baseline characteristics of patients in the study are generalisable to the UK as validated by UK clinicians
- The study compared oral azacitidine + BSC compared with BSC alone (placebo + BSC), BSC is standard practice when patients are in remission, and hence the most relevant comparator was used in the study
- The data are sufficiently mature to demonstrate the effects of oral azacitidine on OS and RFS, and provide sufficient certainty around the clinical benefits of oral azacitidine; median follow-up of 51.7 months (data cut-off date 8 September 2020, patients with event [death]: oral azacitidine [redacted]) and 41.2 months (data cut-off date 15 July 2019, patients with event [death]: oral azacitidine 164 [68.9%]; placebo 181 [77.4])[87]
- Subgroup analyses demonstrate consistent effects (OS and RFS) across subgroups defined by demographic and disease-related characteristics, specifically for the European subgroup which is likely to reflect UK clinical practice

The limitations of the clinical evidence base are primarily associated with the Bucher ITC of oral azacitidine versus midostaurin. Significant differences in patients (inclusion and exclusion criteria, baseline patient characteristics, and study characteristics (time of randomisation) across the QUAZAR-AML-001 and RATIFY trials limit the validity and generalizability of derived effect estimates and represents a limitation of the Bucher ITC (see Section B.2.9.6).

B.2.13.3 End-of-life criteria

Oral azacitidine meets the end of life treatment criteria, given that patients who achieve CR/CRi after induction ± consolidation chemotherapy without undergoing maintenance treatment have a short life expectancy (median OS of patients in the placebo group, i.e. BSC of the QUAZAR AML-001 trial was 14.8 months[84]) and there is sufficient evidence from the QUAZAR AML-001 study to indicate that oral azacitidine offers an extension to life of >3 months (prolongs median OS by 9.9 months, compared with placebo + BSC) (Table B.2.26).[84]

Table B.2.26. End-of-life criteria

Criterion	Data available	Reference in submission (section and page number)
The treatment is indicated for patients with a short life expectancy, normally less than 24 months	• In the QUAZAR AML-001 study patients in the placebo group (i.e. BSC) had a median OS of 14.8 months,[84] substantially lower that 24 months. • In the QUAZAR AML-001 trial, FLT3 mutation positive patients in the placebo group (i.e. BSC) had a median OS of 9.7 months,[106] substantially lower that 24 months.	• B.2.6.1; page 59-61 • B.2.7.3; page 75
There is sufficient evidence to indicate that the treatment offers an extension to life, normally of at least an additional 3 months, compared with current treatment	In the QUAZAR AML-001 study, oral azacitidine + BSC prolongs median OS by 9.9 months compared to placebo + BSC.[84]	B.2.6.1; page 59-61

Abbreviations: BSC = best supportive care; CR = complete remission; CRi = complete remission with incomplete blood count recovery; HR = hazard ratio; NHS = National Health Service; OS = overall survival

B.3 Cost effectiveness

B.3.1 Published cost-effectiveness studies

An SLR was conducted to identify cost-effectiveness studies relevant to the decision problem. The search was performed on February 12, 2020 and updated on June 11, 2021. After removing duplicates, a total of 2,695 records were identified from the searches. Title/abstract screening and subsequent full-text screening resulted in the final inclusion of 21 records representing 19 unique studies. Of the 21 identified records, seven were full publications, 10 were conference abstracts and four were HTA reports. The detailed SLR methodology and results can be found in Appendix G. Overall, the SLR did not identify any relevant cost-effectiveness studies from the published literature for the population of interest, i.e. none of the identified studies focused specifically on AML maintenance treatment. Therefore, the cost-effectiveness of oral azacitidine compared with relevant comparators was evaluated using a de novo model further described in Section B.3.2.1.

B.3.2 Economic analysis

As none of the studies identified in the SLR focused on oral azacitidine as maintenance treatment, a *de novo* model was required to assess the cost-effectiveness of oral azacitidine compared with relevant comparators.

The patient population included in this economic evaluation aligns with the decision problem described in Section B.1.1. This is in line with the marketing authorisation for oral azacitidine and with the NICE final scope. The population reflects the patient population in the QUAZAR AML-001 trial which was inclusive of both the categories of FLT3 mutations; tyrosine kinase domain (TKD) and internal tandem duplications (ITD).

The analysis demonstrates the benefits of oral azacitidine compared with relevant treatments for two distinct patient groups:

- The ITT population, compared with watch and wait with BSC (n=472)
- FLT3 population (*FLT3*-ITD and/or *FLT3*-TKD), compared with midostaurin ([redacted])

B.3.2.1 Model structure

A three-state partitioned survival model (PartSA) was developed in Microsoft Excel® to assess the cost-effectiveness of oral azacitidine vs comparators. The model consisted of the following health states: (i) RFS; (ii) Relapse and (iii) Death (Figure B.3.1).In the model, patients accrue costs and utilities for each cycle they spend in each state (excluding death). Time-to-event analysis directly from the trial was used to inform the distribution of patients between health states between oral azacitidine and watch and wait with BSC for the ITT analysis. In the absence of head-to-head evidence for oral azacitidine versus midostaurin, in the FLT3 subgroup, the evidence base is derived from an STC in the form of HRs and time to event analysis (see Section B.2.9.3 and section B.3.3.2 for more details). The model is run over a defined number of cycles (periods of time) allowing an estimate of total costs and quality-adjusted life expectancy for the cohort over the specified time horizon.

Previous NICE submissions within the AML landscape where the interventions were intended to induce complete remission (or potentially cure) have implemented a fixed cure point– a

point defined in terms of months/years that patients are assumed to be in long term remission and follow mortality rates of the general population. This statistical cure point has been set between 3-5 years [NICE TA523][31] and [NICE TA 642][112]. A UK clinical expert suggested that some patients can be classified as being a long-term survivor if they have disease-free survival and 5-years is not an unreasonable assumption. In contrast to AML induction treatments with or without consolidation and/or HSCT where the goal is to achieve remission and potentially cure, the goal of AML maintenance is to avoid disease progression and prolong life but not necessarily be able to cure patients.[18, 113] NICE critiqued the assumption used in both the midostaurin [NICE TA523][31] and gilteritinib [NICE TA 642][112] models that a fraction of the population is functionally cured, as the evidence review group identified literature that suggests that mortality rates after HSCT remain substantially higher than the general population. In light of the points discussed above, it was not deemed appropriate to include a cure point in the base case analysis, however, a cure point of 5-years is explored as a scenario analysis and patients who are assumed to be cured, follow a standardised mortality ratio of 2.0 in line with the midostaurin submission.[31]

The partitioned survival model was selected as this is the preferred modelling approach when the disease can be accurately represented by simple, defined health states. This structure is fully aligned with two of the key objectives of maintenance treatment for AML, namely avoiding disease progression and prolonging life. Previous models identified in the SLR have also adopted this structure. The use of time-to-event analysis to estimate the health state distribution also benefits from the availability of mature survival curves from QUAZAR AML-001,[87] reducing the need for extensive curve extrapolation. UK clinical experts considered this approach to capture the key elements of AML.

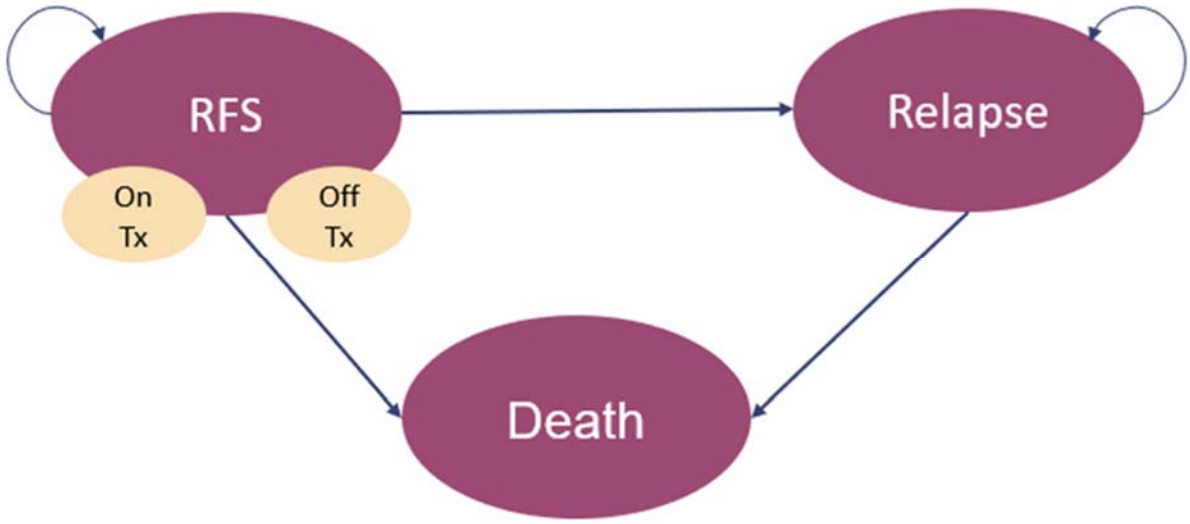
The health states included in the model were defined as:

- Relapse-free: includes patients who are alive and have not relapsed
- Relapse: includes patients who are alive but have relapsed according to IWG 2003 response criteria in AML[84]
- Death: this state is informed by the overall survival curve, which accounts for the number of patients who have died from either AML or other causes.

These are aligned with the primary and secondary outcomes from the QUAZAR AML-001, following the natural history and progression of the disease.

In this model structure, OS was partitioned into RFS (on or off-treatment) and relapse states. In each cycle of the model, the proportion of patients in the relapse state was calculated as the difference between OS and RFS based on the selected curve extrapolations. AEs were modelled as events, rather than as health states, such that costs related to the occurrence of an AE were applied to the proportion of patients estimated to experience the AE. Since there was no active therapeutic agent administered in addition to watch and wait, all patients in this arm were considered to be “off treatment” while in RFS. In this model structure, patients who relapse can-not achieve remission (i.e. move from relapse to RFS), although these patients are not modelled explicitly, they are captured through OS.

Figure B.3.1. Model structure



Abbreviations: RFS = relapse free survival; Tx = treatment

HSCT was modelled as part of subsequent treatment rather than explicitly as a separate heath state. Oral azacitidine is licenced for patients who are not suitable for transplant, and therefore it is unlikely in clinical practice that patients will go on to receive HSCT after oral azacitidine unless they have relapsed. This is supported by data from the QUAZAR AML-001 trial where only a few patients underwent HSCT (6.3%)[87] and the majority of which were post-relapse.

Including HSCT as a health state in the model would require inputs that were not captured in the QUAZAR AML-001 trial nor available from the literature for this population, such as the proportion of patients achieving a successful transplant and the outcomes following the transplant. In models where HSCT is included as a health state, the proportion of patients receiving HSCT tends to be substantially higher and HSCT is often administered during first CR rather than post-relapse. For example, in the RATIFY trial, 59% of midostaurin treated patients underwent HSCT. The HSCT procedure was performed during the first CR in 28.1% of midostaurin treated patients. Although HSCT is deemed as a curative treatment, the effects from subsequent HSCT were assumed to be captured by the RFS and OS data. These effects are expected to be low given subjects in the QUAZAR AML-001 trial who received another therapy (e.g., HSCT) for AML without documented relapse were censored on the date of the last bone marrow assessment, prior to receiving the other therapy. Thus, the efficacy of these subsequent therapies did not contribute to RFS. Moreover, the OS hazard ratio in favour of oral azacitidine was maintained when censoring for HSCT (see full details in Sections B.2.6.1 and B.2.6.2). For all the reasons discussed above, addition of HSCT as a separate health state within the model would therefore add considerable uncertainty within the model, without adding any clarity as to the cost-effectiveness of oral azacitidine. This assumption was supported by clinical opinion. Moreover, this approach aligns with other models in AML.

The analysis was constructed from the perspective of the National Health Service (NHS) and the Personal Social Services (PSS) in England and Wales. A discount rate of 3.5% per annum was applied for costs and benefits in line with the NICE reference case.

A lifetime horizon (i.e., 30 years) was applied to ensure all costs and QALYs were captured. This was considered appropriate given a mean starting age of the cohort was 67.9. Therefore, by the end of the 30-year time horizon, the mean age is 97.9 years and <1% of patients in the model remained alive.

The cycle length selected for the model was 28 days to align with treatment cycles and to provide a reasonable level of granularity to model key clinical events in this disease area. This cycle length aligns with those observed in existing AML models[114, 115] (described in Table B.3.1). Half-cycle correction was applied to the calculation of LYs and QALYs as transitions could occur continuously rather than the start and end of a model cycle.

Table B.3.1. Features of the economic analysis

Factor	Previous appraisals		Current appraisal	
	[TA 523][31]	[TA 399][116]	Chosen values	Justification
Time horizon	54 years (=lifetime)	10 years (=lifetime)	30 years (=lifetime)	Lifetime horizon (i.e., 30 years, as the vast majority of patients have died by the end of year 30 in the model)
Model structure	PartSA	Semi-Markov	PartSA	In line with previous AML models see Appendix G
Health states	Five health states: (i) AML diagnosis (ii) Relapse (iii) Complete response (iv) Steam cell therapy (xi) Death	Four health states: (i) Remission (ii) Relapse (iii) Non-remission (iiii) Death	Three health states : (i) RFS (ii) Relapse (iii) Death	Health states reflective of AML disease pathway and validated by UK clinicians
Cycle length	28 days	4 weeks	28 days	Corresponds to treatment cycle length in the QUAZAR AML-001 trial
Half-cycle correction	Yes	Not stated	Yes	As per NICE reference case
Measurement of health effects	QALYs	QALYs	QALYs	As per NICE reference case
Discount (costs/effects)	3.5%	3.5%	3.5%	As per NICE reference case
Perspective	NHS and personal and social services	NHS and personal and social services	NHS and personal and social services	As per NICE reference case
Treatment waning effect?	No	No	No	In line with previous AML models
Source of utilities	A SLR and TTO study were conducted to identify utility values	Utilities were mapped from trial-based disease specific EORTC QLQC30 data to EQ-5D utility values using published algorithms	QUAZAR AML-001 trial for RFS health state Joshi 2019 for relapse heath state	EQ-5D is the preferred instrument to capture HRQoL as per NICE reference case. Where available, health state utilities were informed from the trial. However, in the absence of trial data, to

Factor	Previous appraisals		Current appraisal	
	[TA 523][31]	[TA 399][116]	Chosen values	Justification
				inform health state utility value for relapse, literature was required.
Source of AE	RATIFY trial	AML-001 trial	QUAZAR AML-001	QUAZAR AML-001 trial

Abbreviations: AE = adverse event; AML = acute myeloid leukaemia; NHS = National Health Service; NICE = National Institute for Health and Care Excellence; OS = overall survival; PartSA = partitioned survival analysis; PFS = progression free survival; QALY = quality-adjusted life year; RFS = relapse free survival; SLR = systematic literature review; TA = technology appraisal; TTO = time trade off

B.3.2.2 Intervention technology and comparators

The pathway and full justification of comparators are presented in Section B.1.3.6.

B.3.2.2.1 Intervention

The intervention is oral azacitidine with BSC. Oral azacitidine is available as 200 mg or 300 mg film-coated tablets to be taken orally with or without food.[1] The recommended starting dose of oral azacitidine is 300 mg once daily (QD) for the first 14 days of every 28-day treatment cycle until disease progression or unacceptable toxicity.[1] Medications included in BSC are detailed in section B.3.5.2.1.2

B.3.2.2.2 Comparator – watch and wait plus BSC

Watch and wait with BSC was chosen as the comparator for the base case analysis with the ITT population. The comparator arm of the QUAZAR AML-001 trial, placebo plus BSC, is used to model this comparator. This represents the standard of care in current clinical practice because there are currently no approved or funded therapies indicated for this population for the independent maintenance treatment of AML in the UK.

B.3.2.2.3 Comparator - midostaurin

For AML patients with mutations in FLT3 , NICE has recommended the use of midostaurin as an option for treating newly diagnosed acute FLT3 mutated AML patients.[31] Midostaurin is an oral, type III, multi-target receptor tyrosine kinases (RTK) inhibitor that acts on FLT3 and multiple other RTKs. For patients in complete response, midostaurin is administered orally at 50mg twice daily as single agent maintenance treatment until relapse for up to 12 cycles of 28 days each.[117]

B.3.3 Clinical parameters and variables

Section B.3.3.1 describes the primary ITT survival analysis conducted for oral azacitidine and no active treatment from the QUAZAR-001 trial.

To inform a comparison of oral azacitidine with midostaurin in patients with FLT-3 mutation, an indirect comparison was conducted (described in Section **Error! Reference source not found.**). The analysis used data from patients with a FLT-3 mutation in the QUAZAR-001 trial, and data from the ITT population of the RATIFY trial (FLT-3 patients) assessing the efficacy of midostaurin in patients with AML. The implementation of this subgroup analysis is described in Section B.3.3.2.

B.3.3.1 Survival modelling of oral azacitidine and SoC

B.3.3.1.1 Data

The modelled baseline patient characteristics presented in Table B.3.2 have been taken from the ITT population of QUAZAR-AML-001 as they were considered to be representative of the patient population in the UK that would be eligible for maintenance treatment with oral azacitidine.

Table B.3.2. Baseline characteristics

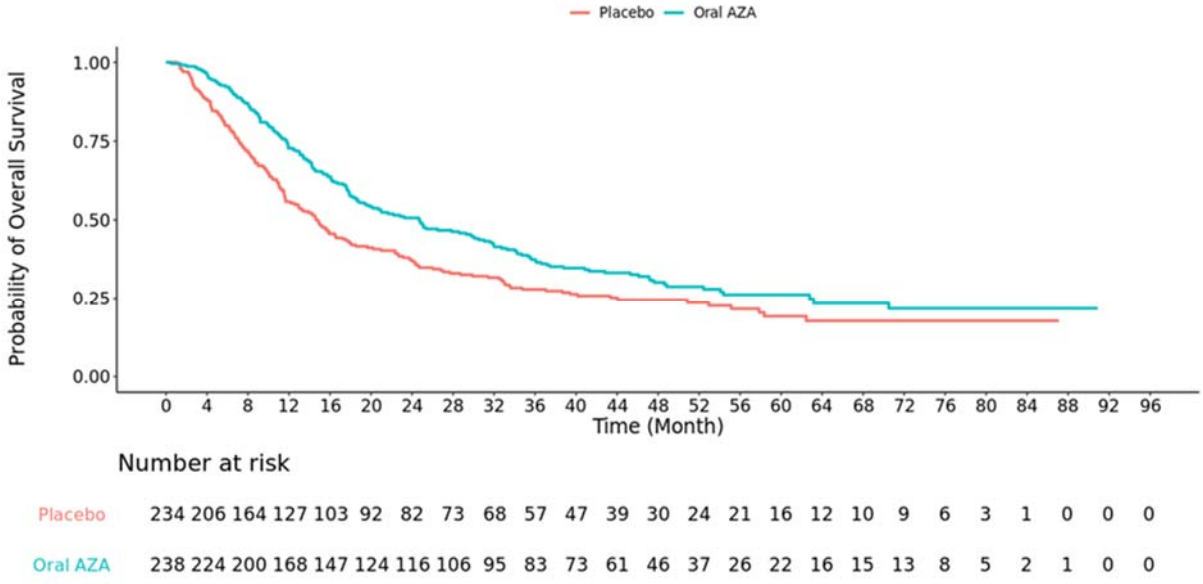
Patient Characteristics	QUAZAR AML-001 (n=472)
Mean age, years (SD)	67.9 (5.66)
Proportion males	52%
Mean weight, kg (SD)	[redacted]
Mean height, cm (SD)	[redacted]

Abbreviations: SD = standard deviation
Source: Wei et al., 2020[84]; QUAZAR AML-001 CSR (Data on File)[87]

The analyses presented in this section used data from the September 2020 data base lock (DBL) of the ITT population of the QUAZAR AML-001 trial to estimate OS with oral azacitidine and with placebo. The September 2020 DBL represented the most recent and mature data set for the estimation of OS as it had been taken during the extension period of the trial. During the extension period of QUAZAR-001, subjects receiving oral azacitidine who were demonstrating clinical benefit could continue to receive oral azacitidine after unblinding, until they met the criteria for study discontinuation or until oral azacitidine became commercially available. Subjects in the placebo group and any subjects who previously discontinued, irrespective of randomization group, were followed up for OS as defined in the clinical trial protocol. Therefore, the September 2020 DBL was used for the estimation of OS as it was considered the most mature and robust.

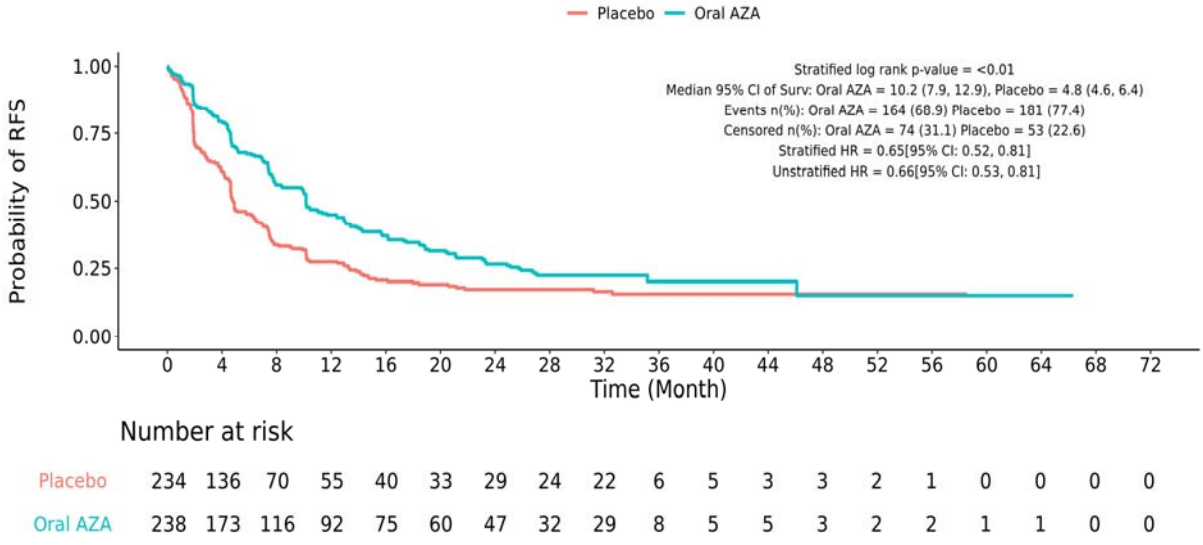
Data from the prior July 2019 DBL was used for the estimation of RFS with oral azacitidine and with no active treatment as the September 2020 DBL had collected [redacted], see Appendix M. The OS and RFS KM curves are presented in Figure B.3.2 and Figure B.3.3. Further details are available in appendix M.

Figure B.3.2. KM curves for OS (ITT population, Sep 2020 data-cut)



Abbreviations: AZA = azacitidine; CI = confidence interval; HR = hazard ratio; ITT = intention-to-treat; KM = Kaplan-Meier; OS = overall survival

Figure B.3.3. KM curves for RFS (ITT population, July 2019 data-cut)



Abbreviations: AZA = azacitidine; CI = confidence interval; HR = hazard ratio; ITT = intention-to-treat; KM = Kaplan-Meier; RFS = relapse-free survival

B.3.3.1.2 Methods

A more detailed overview of the survival analysis methods is outlined in appendix N.1.1. This covers an examination of the observed data to explore the proportionality of the hazards, the fitting of standard parametric survival models (individual and joint models) and the selection of suitable base case and scenario models (by evaluating model fit, model assumptions and the plausibility of extrapolation).

B.3.3.1.3 Overall survival parametrization

In the economic model, all OS models were adjusted for background mortality to ensure that mortality in the modelled population would not be lower than in the general population at any time. UK background mortality was based on the UK National Life Tables, United Kingdom 2017–2019.[118] The survival extrapolation figures as presented in this section (and in appendix N) do not have this adjustment applied. This adjustment is accounted for in the Excel model.

B.3.3.1.3.1 Examination of observed data

A visual inspection of the log-cumulative hazard plot suggests that the two lines are not parallel but reasonably straight (Figure B.3.4). Similarly, the Schoenfeld residual plot displayed a non-horizontal line and the Grambsch-Therneau global Schoenfeld residual test value was statistically significant (p-value 0.0008; Figure B.3.5).

These findings suggest that survival models which assume a proportional hazards relationship between oral azacitidine, and SoC's OS curve may not be appropriate. Although some curvature is present, the lines are relatively straight, suggesting that AFT models would be appropriate.[119] This was confirmed by the quartile-quartile plot which showed no violation of the AFT assumption (Figure B.3.6).

Figure B.3.4. Log-cumulative hazard plot – OS, ITT population

Abbreviations: AZA = azacitidine; ITT = intention-to-treat; OS = overall survival

Figure B.3.5. Schoenfeld residuals plot from Cox PH model – OS, ITT population

Abbreviations: ITT = intention-to-treat; OS = overall survival; PH = proportional hazards

Figure B.3.6. Q-Q plot OS

Abbreviations: AZA = azacytidine; BSC = best supportive care; OS = overall survival; Q-Q = quantile-quantile

B.3.3.1.3.2 Standard parametric models

Parametric curves from both the individual models and joint models are presented in Figure B.3.7 to Figure B.3.10. This information is also presented per model (both treatments) in appendix N.1.2. Note in these figures, KM curves are drawn with a solid line; parametric curves are drawn with a dashed line. Model fit statistics (AIC, BIC) for all parametric distributions are presented in Table B.3.3. The marginal survival gains both pre- and post-extrapolation for each model is presented in Table B.3.4.

Figure B.3.7. Parametric curves fit to the OS outcome in the ITT population, individual models, oral azacitidine

Abbreviations: AZA = azacitidine; ITT = intention-to-treat; OS = overall survival

Figure B.3.8. Parametric curves fit to the OS outcome in the ITT population, individual models, placebo

Abbreviations: AZA = azacitidine; ITT = intention-to-treat; OS = overall survival

Figure B.3.9. Parametric curves fit to the OS outcome in the ITT population, joint models, oral AZA

Abbreviations: AZA = azacitidine; ITT = intention-to-treat; OS = overall survival

Figure B.3.10. Parametric curves fit to the OS outcome in the ITT population, joint models, placebo

Abbreviations: AZA = azacitidine; ITT = intention-to-treat; OS = overall survival

Table B.3.3. Model fit statistics (AIC and BIC) for parametric models of the OS outcome in the ITT population

Parametric Model	AIC	Ranks based on AIC	BIC	Ranks based on BIC
Joint models				
Exponential	[redacted]	[redacted]	[redacted]	[redacted]
Weibull	[redacted]	[redacted]	[redacted]	[redacted]
Log-logistic	[redacted]	[redacted]	[redacted]	[redacted]
Log-normal	[redacted]	[redacted]	[redacted]	[redacted]
Generalised Gamma	[redacted]	[redacted]	[redacted]	[redacted]
Gompertz	[redacted]	[redacted]	[redacted]	[redacted]
Individual models – Oral azacitidine arm				
Exponential	[redacted]	[redacted]	[redacted]	[redacted]
Weibull	[redacted]	[redacted]	[redacted]	[redacted]
Log-logistic	[redacted]	[redacted]	[redacted]	[redacted]
Log-normal	[redacted]	[redacted]	[redacted]	[redacted]
Generalised Gamma	[redacted]	[redacted]	[redacted]	[redacted]
Gompertz	[redacted]	[redacted]	[redacted]	[redacted]
Individual models – Placebo arm				
Exponential	[redacted]	[redacted]	[redacted]	[redacted]
Weibull	[redacted]	[redacted]	[redacted]	[redacted]
Log-logistic	[redacted]	[redacted]	[redacted]	[redacted]
Log-normal	[redacted]	[redacted]	[redacted]	[redacted]
Generalised Gamma	[redacted]	[redacted]	[redacted]	[redacted]
Gompertz	[redacted]	[redacted]	[redacted]	[redacted]

Abbreviations: AIC = Akaike's information criteria; BIC = Bayesian information criterion; ITT= intention-to-treat; OS = overall survival

Table B.3.4. Evaluation of Criterion 5 – estimated rate of OS gain per month by receiving oral azacitidine instead of placebo in the ITT population, before and after the trial cut-off

Models	Pre-extrapolation	Extrapolated tail
KM	0.084	N/A
Joint models		
Exponential	[redacted]	[redacted]
Weibull	[redacted]	[redacted]
Log-Logistic	[redacted]	[redacted]
Log-Normal	[redacted]	[redacted]
Generalised Gamma	[redacted]	[redacted]
Gompertz	[redacted]	[redacted]
Individual models		
Exponential	[redacted]	[redacted]
Weibull	[redacted]	[redacted]
Log-Logistic	[redacted]	[redacted]
Log-Normal	[redacted]	[redacted]
Generalised Gamma	[redacted]	[redacted]
Gompertz	[redacted]	[redacted]

Notes: The rate of survival gain in the pre-extrapolation period is defined as the difference in survival between oral azacitidine and placebo at [redacted] months divided by the number of months in the pre-extrapolation period (ie [redacted] months). The rate of survival gain in the post-extrapolation period is defined as the marginal relative difference in the extrapolated period (post cut-off) divided by the number of months post-cut-off. Negative values represent the rate of survival loss for oral azacitidine (ie, gain for placebo), which in the case of most fitted models indicate a crossing of curves.
Abbreviations: ITT = intention-to-treat; KM = Kaplan–Meier; OS = overall survival

B.3.3.1.3.3 Model selection

Our proposed base case model for OS is the joint generalised gamma, presented in Figure B.3.11 below. The joint generalised gamma distribution has the lowest AIC and BIC values among all distributions, indicating it has the best statistical fit to the observed data. Visual inspection of the joint generalised gamma survival function supports this conclusion, in that the generalised gamma curves most closely fit the data and lead to clinically plausible extrapolations.

Other models, such as the joint and individual Gompertz, as well as the individual generalised gamma, also had good visual fit to the data (during the observed period). However, their extrapolations were considered implausible. The joint Gompertz plateaued and led to an implausibly long right tail. The individual generalised gamma and Gompertz led to a crossing of the curves between placebo and oral azacitidine (see appendix N.1.2). Our expert consultations suggested that this was not considered clinically likely.

It is clear from the KM curve (Figure B.3.2) that the trial hazards (in both arms) are significantly decreasing over time (as illustrated by the gentle plateauing of the curve). This is not to say that the hazards associated with AML decrease over time; it is instead a reflection of patient heterogeneity with respect to hazards/prognosis. Higher hazard (sicker) patients 'leave' the risk set early on, which means that over time the average hazard for the population goes down. This generates decreasing hazards over time at the population level (when it's quite possible, perhaps likely, that all patients are exposed to increasing hazards at the individual level). This

pattern will continue until the population hazard function converges to the hazards of the mildest patients/deepest responders.

The more effective a treatment is, the more gradual the 'curvature' of this transition will be. This can be seen quite clearly in the log-cumulative hazard plot (Figure B.3.4). The curvature of the placebo arm is considerably more pronounced than that of oral AZA. If one tracks this pattern naïve to the clinical dynamics discussed above (as the maximum likelihood survival models do) then a cross-over of these curves is inevitable but spurious.

The joint generalised gamma is an AFT model, and the relatively straight log-cumulative hazard plots (Figure B.3.4) and quantile-quantile plots (Figure B.3.6) indicate that assuming an accelerated failure time relationship between the treatment arms is an appropriate way to model this survival data and preferred over PH models (as suggested in Tremblay et al. 2016[119]). In addition, the joint generalised gamma (as well as the other survival models) also satisfied Criterion 5 of Tremblay et al. 2016[119]. As shown in Table B.3.4, the rate of survival gain in the extrapolated tail is lower than the rate of gain observed in the KM curve.

As indicated in appendix N.1.1, hybrid and cure models were also explored in scenario analyses.

Figure B.3.11. Parametric curves fit to the OS outcome in the ITT population – Generalised gamma distribution, joint model – base case selection

Abbreviations: AZA = azacitidine; ITT = intention-to-treat; OS = overall survival

Figure B.3.12. Parametric curves fit to the OS outcome in the ITT population – Log-normal distribution, individual model

Abbreviations: AZA = azacitidine; ITT = intention-to-treat; OS = overall survival

B.3.3.1.4 Relapse free survival parametrization

B.3.3.1.4.1 Examination of observed data

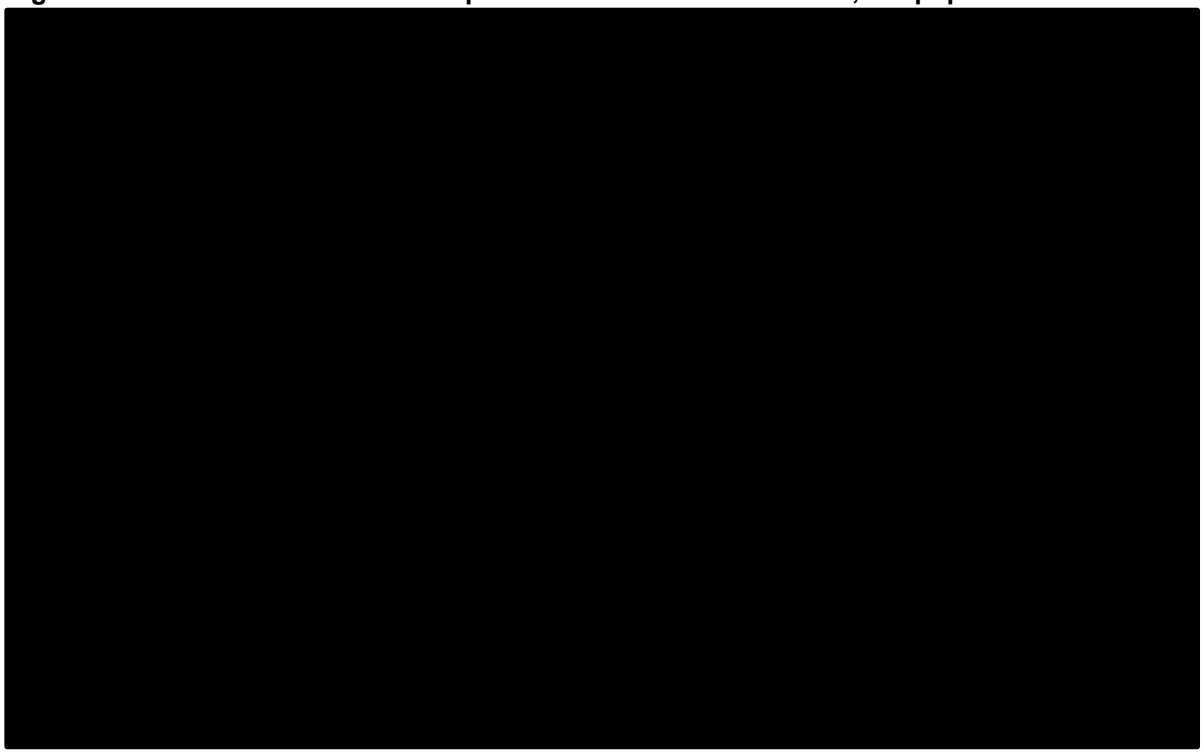
A visual inspection of the log-cumulative hazard plot suggested that the two lines are not parallel but have a relatively straight shape (Figure B.3.13). The Schoenfeld residual plot displayed a non-horizontal line and the Grambsch-Therneau global Schoenfeld residual test value was statistically significant (p-value <0.001; Figure B.3.14).

Given the shape of the KM-estimated hazard functions, the suspected violations of the PH assumption but the relatively straight log-cumulative hazard curves, individual model fits and joint AFT models (log-normal, log-logistic, generalized gamma) may be preferred over PH models.[119] This was confirmed by the quartile-quartile plot which showed no violation of the AFT assumption (Figure B.3.15).

Figure B.3.13. Log-cumulative hazard plot – RFS, ITT population

Abbreviations: AZA = azacitidine; ITT = intention-to-treat; RFS = relapse-free survival

Figure B.3.14. Schoenfeld residuals plot from Cox PH model – RFS, ITT population



Abbreviations: ITT = intention-to-treat; PH = proportional hazards; RFS = relapse-free survival

Figure B.3.15. Q-Q plot RFS

Abbreviations: AZA = azacytidine; BSC = best supportive care; RFS = relapse-free survival; Q-Q = quantile-quantile

B.3.3.1.4.2 Standard parametric models

Parametric curves from both the individual models and joint models are presented in Figure B.3.16 to Figure B.3.19. This information is also presented per model (both treatments) in appendix N.1.3. Note in these figures, KM curves are drawn with a solid line; parametric curves are drawn with a dashed line. Model fit statistics (AIC, BIC) for all parametric distributions are presented in Table B.3.5. The marginal survival gains both pre- and post-extrapolation for each model is presented in Table B.3.6.

Figure B.3.16. Parametric curves fit to the RFS outcome in the ITT population, individual models, oral azacitidine

Abbreviations: AZA = azacitidine; ITT = intention-to-treat; RFS = relapse-free survival

Figure B.3.17. Parametric curves fit to the RFS outcome in the ITT population, individual models, placebo

Abbreviations: AZA = azacitidine; ITT = intention-to-treat; RFS = relapse-free survival

Figure B.3.18. Parametric curves fit to the RFS outcome in the ITT population, joint models, oral azacitidine

Abbreviations: AZA = azacitidine; ITT = intention-to-treat; RFS = relapse-free survival

Figure B.3.19. Parametric curves fit to the RFS outcome in the ITT population, joint models, placebo

Abbreviations: AZA = azacitidine; ITT = intention-to-treat; RFS = relapse-free survival

Table B.3.5. Model fit statistics (AIC and BIC) for parametric models of the RFS outcome in the ITT population

Parametric Model	AIC	Ranks based on AIC	BIC	Ranks based on BIC
Joint models				
Exponential	[redacted]	[redacted]	[redacted]	[redacted]
Weibull	[redacted]	[redacted]	[redacted]	[redacted]
Log-logistic	[redacted]	[redacted]	[redacted]	[redacted]
Log-normal	[redacted]	[redacted]	[redacted]	[redacted]
Generalised Gamma	[redacted]	[redacted]	[redacted]	[redacted]
Gompertz	[redacted]	[redacted]	[redacted]	[redacted]
Individual models – Oral azacitidine arm				
Exponential	[redacted]	[redacted]	[redacted]	[redacted]
Weibull	[redacted]	[redacted]	[redacted]	[redacted]
Log-logistic	[redacted]	[redacted]	[redacted]	[redacted]
Log-normal	[redacted]	[redacted]	[redacted]	[redacted]
Generalised Gamma	[redacted]	[redacted]	[redacted]	[redacted]
Gompertz	[redacted]	[redacted]	[redacted]	[redacted]
Individual models – Placebo arm				
Exponential	[redacted]	[redacted]	[redacted]	[redacted]
Weibull	[redacted]	[redacted]	[redacted]	[redacted]
Log-logistic	[redacted]	[redacted]	[redacted]	[redacted]
Log-normal	[redacted]	[redacted]	[redacted]	[redacted]
Generalised Gamma	[redacted]	[redacted]	[redacted]	[redacted]
Gompertz	[redacted]	[redacted]	[redacted]	[redacted]

Abbreviations: AIC = Akaike's information criteria; BIC = Bayesian information criterion; ITT = intention-to-treat, RFS = relapse-free survival.

Table B.3.6. Evaluation of Criterion 5 – estimated rate of RFS gain per month by receiving oral azacitidine instead of placebo in the ITT population, before and after the trial cut-off

Models	Pre-extrapolation	Extrapolated tail
KM	[redacted]	N/A
Joint models		
Exponential	[redacted]	[redacted]
Weibull	[redacted]	[redacted]
Log-Logistic	[redacted]	[redacted]
Log-Normal	[redacted]	[redacted]
Generalised Gamma	[redacted]	[redacted]
Gompertz	[redacted]	[redacted]
Individual models		
Exponential	[redacted]	[redacted]
Weibull	[redacted]	[redacted]
Log-Logistic	[redacted]	[redacted]
Log-Normal	[redacted]	[redacted]
Generalised Gamma	[redacted]	[redacted]
Gompertz	[redacted]	[redacted]

Notes: The rate of survival gain in the pre-extrapolation period is defined as the difference in relapse-free survival between oral azacitidine and placebo at [redacted] months divided by the number of months in the pre-extrapolation period (ie [redacted] months). The rate of survival gain in the post-extrapolation period is defined as the marginal relative difference in the extrapolated period (post cut-off) divided by the number of months post-cut-off. Negative values represent the rate of survival loss for oral azacitidine (ie, gain for placebo), which in the case of most fitted models indicate a crossing of curves.
Abbreviations: ITT = intention-to-treat, KM = Kaplan–Meier; RFS = relapse-free survival.

B.3.3.1.4.3 Model selection

Our proposed base case model for RFS is the joint log-logistic model. This model exhibits no cross-over of the treatment arms (see appendix N.1.3), has very good visual fit and has higher precision than the individual models, due to the higher statistical power of fitting a single model to both treatment arms. This rests on the assumption that the relative treatment effect can be modeled by an AFT factor, which the (reasonably straight) lines in the log-cumulative hazard plot (Figure B.3.13) supports. From a statistical fit perspective, the log-logistic distribution is the best fitting model in terms of AIC and BIC (Table B.3.5). In addition, the model exhibits much lower marginal survival in the extrapolation vs the observed period (see Table B.3.6) which satisfies Criterion 5 of the Tremblay et al. guidance.[119]

All parametric models fitted (including the joint log-logistic) do not fit well to the 'tail end' of the placebo RFS curve. The curve appears to plateau sharply, and even cross the RFS curve of oral azacitidine. Expert consultations suggested such a cross-over was not clinically plausible, and it is more likely that this is due to statistical noise driven by a low sample size: only around 10% of the original sample is still 'at risk' at that point in the curve (Figure B.3.3).

We include the joint log-normal model as a scenario, the model satisfies Criterion 5 and has the second lowest AIC after the base case log-logistic model (excluding the Gompertz, which has an implausible functional form).

B.3.3.2 Survival modeling of midostaurin and FLT-3 subgroup

B.3.3.2.1 Data and matching

A detailed description of the QUAZAR-001 FLT3 and RATIFY study data is available in Section **Error! Reference source not found.** and in Appendix D1.2.3, including their assessment for differences in effect modifier status as recommended by NICE TSD[110] and outlined in Appendix D.1.2.1. Information on the matching process is provided in detail in Section **Error! Reference source not found.**.

B.3.3.2.2 Methods

A more detailed overview of the indirect treatment comparison between the QUAZAR AML-001 and RATIFY studies is included in Section **Error! Reference source not found.** and appendix D.1.2.3. This covers the definition of the data, the matching and specification of the models used, and the proportional hazards tests performed. The three approaches considered, all based on a matched QUAZAR-001 sample, were hazard ratios from the Bucher ITC, parametric models, and spline models. The sections below describe the rationale for the selection of models used to estimate survival in the FLT3 mutation population.

B.3.3.2.3 Overall survival parametrization

As the case for the ITT population, OS models in the economic model were adjusted for UK background mortality,[118] survival extrapolation figures presented in this section do not have this adjustment applied.

B.3.3.2.3.1 Examination of observed data

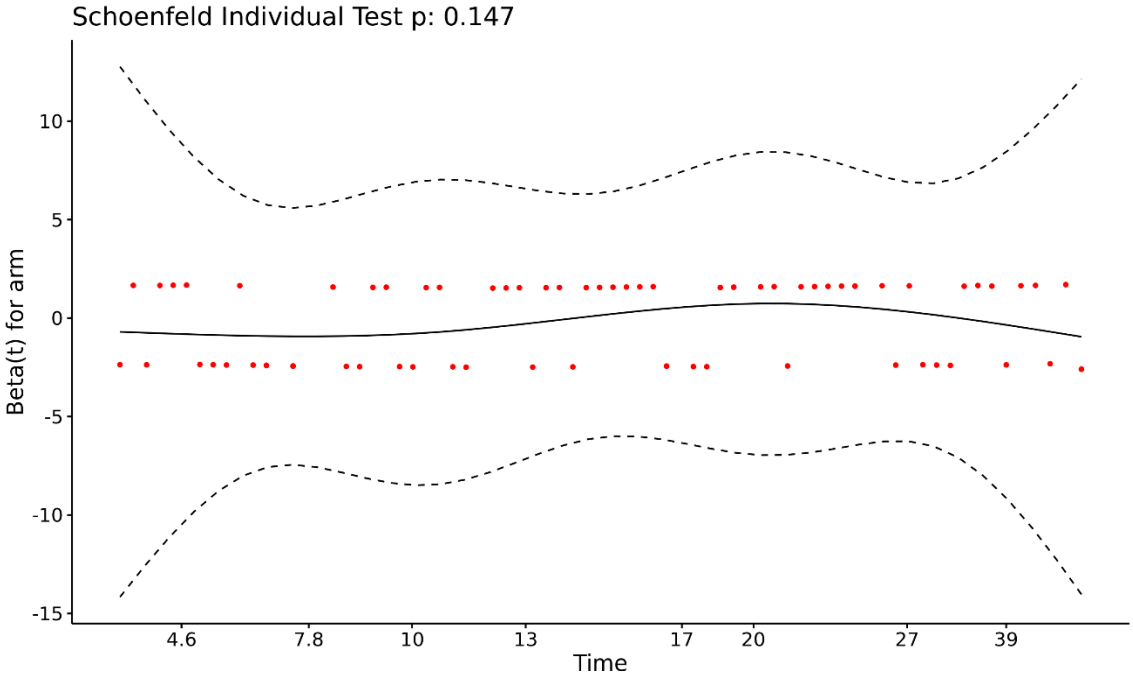
Based on the observed data from the QUAZAR-001 FLT3 subgroup and the RATIFY maintenance group, it was assumed that the PH assumption was likely violated for the OS curves (as outlined in appendix D.1.2.4.1). Although the Schoenfeld residual plot for the QUAZAR AML-001 FLT3 subgroup showed an almost horizontal line (see Figure B.3.20 and a visual inspection of the KM curve did not provide strong evidence of a violation of PH (see Appendix D.1.2.4.1), the global test assessing violation of PH was significant (p-value <0.001). PH was assumed to be violated for the maintenance subgroup within RATIFY based on visual inspection of the KM curves which feature cross-over (see Appendix D.1.2.4.1) and the Schoenfeld residual plot not being straight (Figure B.3.21).

In light of the crossing curves in the RATIFY trial maintenance subgroup and the Schoenfeld residual plot not being straight, and considering the indications of non-proportionality in the QUAZAR-001 trial FLT3 subgroup, proportional hazards models and AFT models were considered less appropriate so individual models were fit to the QUAZAR AML-001 FLT3 IPD and digitized KM data from the RATIFY maintenance subgroup trial. Models included a treatment covariate and treatment effects on ancillary parameters.

Figure B.3.20. Plot of Schoenfeld Residual Over Time for Treatment in QUAZAR AML-001

[redacted]

Figure B.3.21. Plot of Schoenfeld Residual Over Time for Treatment in the RATIFY maintenance subgroup



B.3.3.2.3.2 Parametric models

Parametric curves from individual models are presented in Figure B.3.22 to Figure B.3.25. The fit statistics of all parametric models along with individual plots of the curves can be found in appendix D.1.2.3.4.1.

Figure B.3.22. Parametric curves fit to the OS outcome in the FLT3 population, individual models, oral azacitidine

Abbreviations: AZA = azacitidine; OS = overall survival

Figure B.3.23. Parametric curves fit to the OS outcome in the FLT3 population, individual models, placebo

Abbreviations: AZA = azacitidine; FLT3 = fms-like tyrosine kinase 3; OS = overall survival

Of the 7 parametric models, the generalized gamma and log normal were considered most appropriate for selection as base case or scenarios. The generalized gamma had the best fit to the data based on AIC and the oral azacitidine arm remained apart from the no active treatment arm aligning with the clinical expectations (see Appendix D.1.2.4.4). The lognormal was the third best fitting based on AIC and had a closer fit to the observed median (2.15 vs. 2.07, observed median of 2.35) when compared with the next-best fit, the log logistic curve. The Gompertz was second best fitting based on AIC but was not considered given the observed plateau (see Figure B.3.22 and Figure B.3.23), which was not in line with clinical expectations that there would be no plateau. The remaining distributions, Weibull, exponential and gamma, had poor fit to the data, as judged by the AIC statistics and visual inspection, and were therefore removed not considered appropriate for consideration.

B.3.3.2.3.3 Spline models

The fit statistics and individual plots of all spline models can be found in appendix D.1.2.3.4.2. Parametric spline models are presented in Figure B.3.23, Figure B.3.24, and Figure B.3.25.

Figure B.3.24. Spline models fit to the OS outcome in the FLT3 population, individual models, oral azacitidine

Abbreviations: AZA = azacitidine; FLT3 = fms-like tyrosine kinase 3; OS = overall survival

Figure B.3.25. Spline models fit to the OS outcome in the FLT3 population, individual models, placebo

Abbreviations: AZA = azacitidine; FLT3 = fms-like tyrosine kinase 3; OS = overall survival

All of the extrapolations led to crossing for oral azacitidine and no active treatment which was not expected based on clinical opinion. The 1 knot odds linear predictor was considered as the two curves remain close even after the point of crossover, whereas the remaining models led to divergence after crossover. The 1 knot odds linear predictor also had the second best fit to the data based on AIC and was very close to the best fitting 1 knot normal linear predictor with just [redacted] AIC points difference.

B.3.3.2.3.4 Recommendations and conclusions

Following the assessment of the proportional hazards assumption, which appears violated for the OS endpoint in the maintenance subgroup of the RATIFY trial and also showed signs of non-proportionality in the FLT3 population within the QUAZAR-001 trial, hazard ratios from the Bucher ITC were not used as non-proportionality was expected to apply to the treatment arms of the trials i.e., between the oral azacitidine and the midostaurin arms. Instead, individual models were fitted using parametric and spline-based approaches relaxing the proportional hazards and AFT assumptions.

We recommend using the generalized gamma as the base case model (see Figure B.3.26) with the log-normal and 1 knot odds linear spline models tested using scenario analysis. The generalized gamma had the second-best fitting based on AIC indicating good statistical fit with

the observed data, and the oral azacitidine and no active treatment arms remained apart, in line with clinical expectations. The 1 knot normal linear predictor model had the best fit to the data as judged by AIC, however, the oral azacitidine and no active treatment curves crossed which is not expected based on clinical advice. The log-normal model was also considered plausible with the oral azacitidine and no active treatment arms not crossing but had a worse fit to the data judging bit the AIC, therefore, it was not favoured.

Figure B.3.26. Parametric extrapolation of OS using a generalised gamma model
Abbreviations: AZA = azacitidine; OS = overall survival

B.3.3.2.4 Relapse free survival parametrization

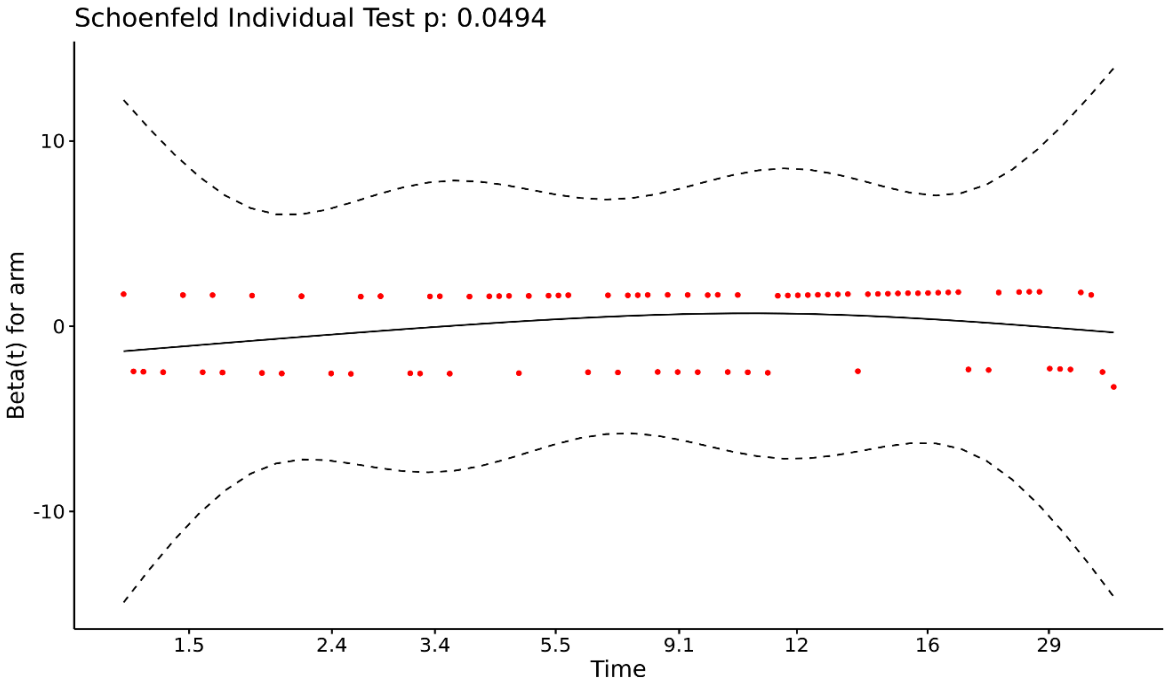
B.3.3.2.4.1 Examination of observed data

Similar to the analysis for OS, it was assumed that PH was violated for RFS for both studies. Although the Schoenfeld residual plot for the FLT3 subgroup within QUAZAR AML-001 showed an almost horizontal line (Figure B.3.27), the global test assessing PH was significant (p=0.0011) and the KM curves (comparing oral azacitidine to placebo; see Appendix D.1.2.3.3) cross-over at multiple time points. PH was also assumed to be violated for the maintenance subgroup within RATIFY since the KM curves (comparing midostaurin to placebo; see Appendix D.1.2.3.3) cross-over, the Schoenfeld residual plot was not a straight line (

Figure B.**3.28**) and the global test assessing PH was significant (p=0.0494).

Figure B.3.27. Plot of Schoenfeld residual over time for treatment in QUAZAR AML-001

Figure B.3.28. Plot of Schoenfeld residual over time for treatment in the maintenance subgroup of RATIFY



Based on the observed data comparing active treatment and placebo arms, for the FLT3 subgroup in QUAZAR-001 and the maintenance group in RATIFY, it was assumed that the proportional hazards and AFT assumptions were violated for the RFS endpoint (as outlined in appendix D.1.2.3.3). Therefore, models were fit to the QUAZAR AML-001 FLT3 IPD and digitized KM data from the maintenance subgroup of the RATIFY trial individually. Models included a treatment covariate and treatment effects on ancillary parameters.

B.3.3.2.4.2 Parametric models

Parametric curves from individual models are presented in Figure B.3.29 and Figure B.3.30. The fit statistics of all parametric models can be found in appendix D.1.2.3.4.1.

Figure B.3.29. Parametric curves fit to the RFS outcome in the FLT3 population, individual models, oral azacitidine

Abbreviations: AZA = azacitidine; FLT3 = fms-like tyrosine kinase 3; RFS = relapse-free survival

Figure B.3.30. Parametric curves fit to the RFS outcome in the FLT3 population, individual models, placebo

Abbreviations: AZA = azacitidine; FLT3 = fms-like tyrosine kinase 3; RFS = relapse-free survival

The generalized gamma was deemed to be a plausible option given there was no plateau. The generalized gamma was also the second-best fitting model. Although the Gompertz was best fitting, the plateau seen in the extrapolations was not considered to be plausible. The log normal also fit reasonably well to the observed data and is considered for usage as base case or scenario. The remaining curves had a poor fit to the data based on AIC and visual assessment and were removed from consideration.

B.3.3.2.4.3 Spline models

The fit statistics of all spline models can be found in appendix D.1.2.3.4.2. Parametric spline models are presented in Figure B.3.31 to Figure B.3.32.

Figure B.3.31. Spline models fit to the RFS outcome in the FLT3 population, oral azacitidine

Abbreviations: AZA = azacitidine; FLT3 = fms-like tyrosine kinase 3; RFS = relapse-free survival

Figure B.3.32. Spline models fit to the RFS outcome in the FLT3 population, placebo

Abbreviations: AZA = azacitidine; FLT3 = fms-like tyrosine kinase 3; RFS = relapse-free survival

RFS spline models with 2 knots were excluded from considerations as they predicted an increase in survival. This was due to small sample size ([redacted] in QUAZAR AML-001 FLT3 subgroup), high rates of censoring, and poor fitting models.

Of the remaining one knot models, the hazard linear predictor led to a divergence of oral azacitidine and no active treatment curves after the point of crossover. Informed by clinical advisor opinion, the one knot odds linear predictor was deemed to be plausible. The one knot normal linear predictor may also be considered.

B.3.3.2.4.4 Recommendations and conclusions

Following the assessment of the proportional hazards and AFT assumption, which were both assumed to be violated for the RFS endpoint, hazard ratios from the Bucher ITC were not used and individual models were fitted using parametric and spline-based approaches.

The 1 knot odds linear model is proposed as the base case model as it had the second-best fit to data as judged by AIC and BIC and was considered most plausible based on clinical opinion. The 1 knot hazard linear predictor model had the best fit to the data as judged by AIC, however, the curves plateaued which is not expected based on clinical advice. Another suitable model was the generalized gamma model which had the second-best fitting based on AIC. The model is recommended for usage in a scenario analysis instead of the base case, this was considered a conservative assumption. The log-normal and 1 knot normal linear predictor models were furthermore considered suitable for the estimation of RFS based on AIC.

Figure B.3.33. Time-varying Spline model for RFS using 1 internal knot and an odds linear predictor

Abbreviations: AZA = azacitidine; OS = overall survival

B.3.3.3 Time on treatment

The SmPC of oral azacitidine recommends discontinuation upon blast counts >15% or unacceptable toxicities and in the QUAZAR AML-001 trial, patients were discontinued from study treatment upon disease relapse or unacceptable toxicities.[87] The time on treatment KM curve of oral azacitidine from the QUAZAR AML-001 trial was used to model drug costs in the base case analysis, see Figure B.3.34. This was preferred to usage of the median or mean which whilst accounting for a reasonable number of cycles, would fail to capture discounting appropriately.

Figure B.3.34. KM curve time on treatment with oral azacitidine

Abbreviations: KM = Kaplan-Meier; RFS = relapse-free survival

Time on treatment with midostaurin was assumed to last 11.08 cycles, informed by published literature.[83] Time on treatment with oral azacitidine in the FLT3 subgroup was based on the time on treatment KM curve from the FLT3 subgroup of the QUAZAR AML-001 trial (Figure B.3.35). To estimate drug costs for midostaurin, a one-off cost was estimated by multiplying the time on treatment (=11.08 cycles) by the drug cost per cycle and included only in the first cycle.

Figure B.3.35. KM curve time on treatment with oral azacitidine FLT3 subgroup

Abbreviations: FLT3 = fms-like tyrosine kinase 3; KM = Kaplan-Meier; RFS = relapse-free survival

B.3.3.4 Subsequent therapy

Following relapse, a proportion of surviving patients were assumed to receive a single line of subsequent therapy. Subsequent therapies were included as a cost input only, with no specific impact on outcomes (survival, quality of life, etc.) as it was assumed that effects would be captured in the OS curve from the trial. Costs for subsequent therapy were applied once, as patients in the model transitioned from the RFS to the relapse health state. Costs for subsequent therapies are described in Section B.3.5.2.1.4.

The proportion of patients receiving a subsequent therapy, and the mix of subsequent therapies, was informed by the QUAZAR AML-001 trial and was validated by clinical advisors. Decitabine was used in a small proportion of patients (<5%) in the QUAZAR AML-001 trial. This treatment was not expected to be used in UK clinical practice and hence not included in the model. Estimates for the FLT3 subgroup treated with oral azacitidine and watch and wait with BSC were based on the QUAZAR-001 study from the FLT3 subgroup and were assumed to also apply to the midostaurin arm. For costing purposes, salvage chemotherapy was assumed to consist of 3+7: daunorubicin and 3+7: cytarabine.

Table B.3.7. Subsequent therapies received in the clinical trials

Treatment	Subsequent therapy		
	Low-Dose Cytarabine	Injectable azacitidine	Salvage chemotherapy
Oral azacitidine	14.3%	8.4%	26.1%
Oral azacitidine (FLT3)	16.7%	6.7%	23.3%
Watch and wait plus BSC	10.7%	15.4%	33.8%
Watch and wait plus BSC (FLT3)	11.1%	8.3%	36.1%
Midostaurin	11.1%	8.3%	36.1%

Abbreviations: BSC = best supportive care. FLT3 = fms-like tyrosine kinase 3Source: QUAZAR AML-001 CSR, Table 14.1.10.3 (Data on File)[87]

B.3.3.5 AEs

The model included Grade 3 or 4 AEs occurring in ≥5% or more of patients in the safety population of the QUAZAR AML-001 trial, including all randomized subjects who received at least one dose of study treatment (n=236 for oral AZA, n=233 for placebo), as well as AEs identified by clinical advisors to have a substantial impact on quality of life. Note, non-UK clinical advisors indicated that leukopenia would be captured within the existing list of AEs and thus, leukopenia was not included as a separate AE in the model to avoid double-counting. AE rates in patients treated with Midostaurin were informed by the ITT population of the RATIFY trial, AEs of grade 3/4 occurring in >10% patients were included based on the maintenance phase, due to the restriction in the RATIFY trial.[31] AE rates for oral azacitidine and watch and wait plus BSC within the FLT3 subgroup were obtained from the FLT3 subgroup of the QUAZAR-001 trial.

Table B.**3.8** summarises the AEs included in the model and the percentage of patients experiencing each AE in each model arm.

Table B.3.8. Percentage of patients experiencing Grade 3 or 4 AEs

Treatment	Adverse events							
	Neutropenia	Thrombocytopenia	Anemia	Febrile neutropenia	Diarrhoea	Vomiting	Nausea	Fatigue
Oral azacitidine	41.1%	22.5%	14.0%	11.4%	5.1%	3.0%	2.5%	3.0%
Oral azacitidine (FLT3)	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
Watch and wait plus BSC	23.6%	21.5%	12.9%	7.7%	1.3%	0.0%	0.4%	0.9%
Watch and wait plus BSC (FLT3)	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
Midostrauin	8.3%	1.7%	0.8%	0.8%	0.8%	0.0%	0.0%	0.0%

Abbreviations: AE = adverse event; FLT3 = fms-like tyrosine kinase 3
Source: Oral azacitidine and Placebo with BSC - QUAZAR AML-001 CSR (Data on File)[87], Midostaurin – NICE TA523[31]

B.3.4 Measurement and valuation of health effects

B.3.4.1 Health-related quality-of-life data from clinical trials

The utility analysis was performed on the ITT population of the QUAZAR AML-001 trial. The EQ-5D-3L questionnaire was used to inform utilities for the RFS: on treatment and RFS: off treatment health states in the model.

In the trial, the EQ-5D-3L questionnaire had been administered to subjects on Day 1 of each 28-day cycle until the study treatment was discontinued, as well as at the end of the study (see Section B.2.6.3.3 for more details on health related quality of life measures assessed). The analysis included 442 subjects in the EQ-5D-3L evaluable population defined as those who had a valid assessment of EQ-5D-3L health utility at baseline and at least one evaluable assessment at post-baseline visits (225 in the Oral azacitidine group and 217 in the placebo group). The EQ-5D-3L health utility scores based on the UK value set[120] were derived, and missing items were handled according to the scoring manual.[121] The score derivation was carried out by two independent data analysts and cross validated for quality assurance purposes. The EQ-5D-3L health utility scores based on the UK value set[120] were derived, and missing items were handled according to the scoring manual.[121] The score derivation was carried out by two independent data analysts and cross validated for quality assurance purposes.

To account for the repeated nature of the data, linear mixed effects models (LMM) with random intercepts for repeated measures were used to derive the EQ-5D-3L utility values in the pre-progression health state. The utility models tested included treatment and adverse event covariates in the specifications. The optimal model was defined as the model which best reflected reality and generated plausible results. The optimal model was selected based on the level of significance and the magnitude of each estimated coefficients and the AIC and BIC statistics. Utilities did not significantly differ between the treatment arms, subjects with any ongoing AEs had slightly lower average health utility scores relative to subjects without ongoing AEs. The model without treatment and adverse event variables was selected for inclusion in the model. The modelling of AEs was based on utility decrements from the literature.

Table B.3.9. QUAZAR AML-001 trial utility values

Source	Health State	Utility Value	SE	Data Source/Notes
QUAZAR AML-001 Trial	RFS: on treatment	[redacted]	[redacted]	QUAZAR AML-001 CSR
	RFS: off treatment	[redacted]	[redacted]	QUAZAR AML-001 CSR

Abbreviations: CSR = clinical study report; RFS = relapse free survival; SE = standard error

As the QUAZAR AML-001 trial did not capture data on health-related quality of life (HRQoL) for patients beyond the treatment period or post relapse, alternative data sources were sought using the published literature (see Section B.3.4.3).

B.3.4.2 Mapping

EQ-5D-3L values were collected directly from QUAZAR AML-001. Hence, no mapping was required.

B.3.4.3 Health-related quality-of-life studies

To inform the utility estimates used in the model, an SLR was conducted to identify and summarise health utility values for adults (≥18 years) with AML who received high intensity first line therapy (induction with or without consolidation) with or without maintenance treatment. The detailed SLR methodology can be found in Appendix H. The search was conducted on February 12, 2020 and updated on June 11, 2021. In total, 2,604 records were identified from the searches after removing duplicates. Title/abstract screening and subsequent full-text screening of records resulted in the final inclusion of 20 records representing 20 unique studies reporting on health utility values and health-related quality of life (HRQoL). Of these, eight unique studies reporting utility values by health state were identified, comprised of four utility elicitation studies and four economic evaluations.

While three sets of utilities were identified, the utility value for relapse calculated based on Joshi 2019 was selected for use in the base case as it was obtained using a composite time trade-off methodology to elicit health state utilities for AML from 210 individuals in the UK general population. Due to the differences between the QUAZAR AML-001 trial and the sample described by Joshi 2019, the difference between RFS and relapse was isolated to avoid the impact of sample differences; the utility for relapse was calculated as the difference between the RFS and relapse utilities in Joshi 2019 (0.38), which was then applied to the RFS utility estimated from the QUAZAR AML-001 trial. This assumes that the difference between RFS and relapse would be similar in the QUAZAR population. The difference between RFS and relapse was considered reasonable by expert clinical advisors. Moreover, rationale was provided that using the same value for on and off treatment is reasonable. Even though there are side effects on treatment, there is still a positive psychological effect of being treated.

The utilities identified are summarised in Table B.3.10. A scenario is tested using the difference from Tremblay 2018 RFS on treatment and relapse (0.28). The Stein 2019 utility values were sourced from an online discrete choice experiment survey was then conducted to capture preferences for the health states from a nationally representative sample of 300 adults in the US.[122] Since the population participating in the study does not align with the NICE reference case, the Stein 2019 values are not tested in the model results but are provided for reference only.

Table B.3.10. Health state utility values from literature sources

Source	Health State	Utility Value	SE	Data Source/Notes
Joshi 2019	RFS: on treatment	0.89	0.15	Joshi et al. 2019[123]
	RFS: off treatment	0.89	0.15	Joshi et al. 2019
	Relapse	0.51	0.46	Joshi et al. 2019
Tremblay 2018[a]	RFS: on treatment	0.81	0.20	Batty et al. 2014,[124] assumption for SE
	RFS: off treatment	0.83	0.20	Leunis et al. 2014,[60] assumption for SE
	Relapse	0.53	0.20	Pan et al. 2010,[125] assumption for SE
Stein 2019	RFS: on treatment	0.87	0.20	Stein et al. 2018[122] and clinical expert opinion, assumption for SE
	RFS: off treatment	0.87	0.20	Stein et al. 2018[122] and clinical expert opinion, assumption for SE
	Relapse	0.62	0.20	Stein et al. 2018[122] and clinical expert opinion, assumption for SE

[a] Note: in the Tremblay 2018 dataset, the health state utility values were assumed to incorporate disutility related to toxicity and adverse events resulting from treatment.
Abbreviations: CSR = clinical study report; RFS = relapse free survival; SE = standard error

B.3.4.4 Utility decrements for adverse events

Utility decrements were included in the model to capture the HRQoL impact of AEs. An overview of AEs included in the model is provided in Section 0. Table B.3.11 shows utility decrement values from the literature for each identified AE. The duration of each AE was informed by clinical advisor opinion that each of the included events would last for approximately 1 week. The total disutility due to AEs that a patient experiences on oral azacitidine, no active therapy or midostaurin was determined based on the percentage of patients experiencing each AE and the disutility of that AE and is shown in Table B.3.12. The QALY decrement related to AEs i.e., the disutility adjusted for the duration, was applied in the first model cycle.

Table B.3.11. Disutility decrement per adverse event

Adverse Event	Reported Utility Decrement (annual)	Duration of Adverse Event (Weeks)[a]	Data Source for Disutility
Neutropenia	0.090	1.0	Nafees et al. 2008[126] and TA642[112]
Thrombocytopenia	0.090	1.0	TA642[112] - assumed same as neutropenia
Anemia	0.119	1.0	TA642[112]
Febrile neutropenia	0.150	1.0	TA642[112]
Diarrhea	0.176	1.0	Stein et al. 2018[122]
Vomiting	0.048	1.0	Nafees et al. 2008[126]
Nausea	0.048	1.0	Nafees et al. 2008[126]
Fatigue	0.115	1.0	TA642[112]

[a] Durations informed by clinical advisor opinion.

Table B.3.12. Average total AE QALY decrement per patient

Maintenance Treatment	Average Total Disutility per Patient
Oral azacitidine	0.026
No active therapy	0.018
Oral azacitidine (FLT3 Subgroup)	0.012
No active therapy (FLT3 Subgroup)	0.012
Midostaurin (FLT3 Subgroup)	0.003

Abbreviations: AE = adverse event; FLT3 = fms-like tyrosine kinase 3; QALY = quality adjusted life year

B.3.4.5 Utility Decrement for Hematopoietic Stem Cell Transplant

A utility decrement for HSCT was included in the model to capture the impact of HSCT on HRQoL. A study in the literature by Matza et al. 2019 was found to report a health state utility for HSCT.[127] Since HSCT is an event in the model rather than a health state, the difference between health states (0.21 difference between transplant and remission) was used to derive a disutility which was then applied for one 28-day cycle. The HSCT disutility was applied in the first model cycle to 6.3% of oral azacitidine treated patients and 13.7% of patients on no active therapy. In the FLT3 population, the same uptake of HSCT was assumed for oral azacitidine as used for ITT. For midostaurin the value based on the Larson et al (2021)[83] study was 5.8%.

B.3.4.6 HRQoL data used in the cost-effectiveness analysis

The utility for pre-progression survival health state was estimated through a utility analysis using the EQ-5D data collected in the QUAZAR AML-001 trial.[87] As no utility information was generated from patients who progressed the utility for post-progression survival health state was obtained from the literature based on Joshi 2019. Utility decrements were applied to account for adverse events and patients who received HSCT.

Table B.3.13. Summary of utility values for cost-effectiveness analysis

State	Utility value, mean	SE	Reference in submission (section and page number)	Justification
RFS :on treatment	[redacted]	[redacted]	Section B.3.4.1	QUAZAR AML-001 CSR
RFS: off treatment	[redacted]	[redacted]	Section B.3.4.1	QUAZAR AML-001 CSR
Relapse	0.51	0.46	Section B.3.4.1	Time trade-off methodology, 210 individuals in the UK general population
HSCT disutility	0.210	NA*	Section B.3.4.5	Treated as disutility since this is an event as opposed to health state in the model

*Total QALY decrement is varied in the PSA assuming 20% variation around the mean
Abbreviations: AE = adverse event; HSCT = Hematopoietic stem cell transplantation; RFS = Relapse-free survival; SE = standard error

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

Healthcare resource use and cost data were identified through relevant databases as outlined in the following sections. An SLR was conducted to identify relevant healthcare resource use and cost data. The details of the methodology and results of the SLR are provided in Appendix G and Appendix I. In addition, clinical experts validated the applicability of the costs and resources used in the model.

Costs were included in the model in 2020 GBP (£) where costs were only available from previous years, they were inflated using the HCHS inflation index[128] up to and including 2015/16 and NHSCII index[128] was used thereafter.

B.3.5.1 Intervention and comparators' costs and resource use

B.3.5.1.1 Drug acquisition costs

Intervention costs

All patients were assumed to initiate treatment in the first model cycle. Drug acquisition costs were applied each cycle in the RFS state for patients on treatment. Drug acquisition costs for oral azacitidine were calculated based on treatment dose, number of administrations per cycle, number of cycles as defined by the treatment protocol, and the unit price.

The recommended starting dose of oral azacitidine is 300 mg orally once daily on days 1 through 14 of repeated 28-day treatment cycles as per the SmPC.[1] The base case analysis is focused exclusively on the 14-day dosing schedule.

Table B.3.14 presents the different dosing schedules possible depending on the recommending dose reductions/dose length frequency adjustments for patients experiencing adverse drug reactions. The list price for oral azacitidine for both 200mg and 300mg packs is set at [redacted].

According to the SmPC[1] of oral azacitidine, dose interruptions and/or reductions may be necessary depending on toxicity levels and adverse drug reactions (see section B.2.10). Depending on the severity of toxicity and severity of adverse drug reaction, a dose reduction of oral azacitidine may be necessary and/or an increase of the time period between doses. In these circumstances if necessary oral azacitidine may be discontinued depending on the physician's discretion[1]. For grade 3 events a maximum of one dose reduction to a daily dose of 200 mg for 14 days in the event of toxicity, and for grade 4 events a maximum of one treatment schedule (frequency) modification from 14 to 7 days of 200 mg in the event of continuing toxicity that did not respond to the initial dose reduction. To pick up the aforementioned discontinuation of treatment we use the KM data directly from QUAZAR AML-001. This is also supplemented by the mean RDI to capture dose changes.

Table B.3.14. Drug dosing schedules

Drug	Dose	Dosing	Schedule	Route	Doses per Cycle	Cycle Length	Total Dose per Administration	Total Dose per Cycle
Oral azacitidine	200 mg	Fixed	Days 1-7	Oral	7	28	200 mg	1,400 mg
	200 mg	Fixed	Days 1-14	Oral	14	28	200 mg	2,800 mg
	300 mg	**Fixed**	**Days 1-14**	**Oral**	**14**	**28**	**300 mg**	**4,200 mg**
	300 mg	Fixed	Days 1-21	Oral	21	28	300 mg	6,300 mg

Abbreviations: mg = milligrams
Source: QUAZAR AML-001 CSR (Data on File)[87]

Table B.3.15 provides the formulations and unit costs for oral azacitidine, along with the calculated costs per cycle. Only the recommended dose was used in the base case analysis. The model assumes a mean relative dose intensity of 86.9% based on the QUAZAR AML-001 trial.[88] (B.2.3.1).

Table B.3.15. Drug acquisition costs for different oral azacitidine dosing

Treatment	Unit Strength	Unit Description	Unit List Price (£)	Cost per Cycle (£)
Oral azacitidine Days 1-7	200 mg	Tablet	[redacted]	[redacted]
Oral azacitidine Days 1-14	200 mg	Tablet	[redacted]	[redacted]
Oral azacitidine Days 1-14	**300 mg**	**Tablet**	[redacted]	[redacted]
Oral azacitidine Days 1-21	300 mg	Tablet	[redacted]	[redacted]

Abbreviations: mg = milligrams.

As part of premedication, the model assumed that 5 days of ondansetron using the dose 8mg twice a day was given as premedication prior to each cycle of oral azacitidine therapy and for midostaurin, it was assumed that patients receive 2.5 days of ondansetron as a conservative approach. Per cycle this cost amounted, £2.21 for use with oral azacitidine and £1.10 for use with midostaurin. Drug cost was retrieved from eMIT 2020[129]. Drug costs relating to BSC were also included for this treatment (B.3.5.2.1.2).

Table B.3.16. Overview of the cost associated with ondansetron tablets for premedication use

Drug	Unit cost (£)	Days of premedication per cycle	Cost per cycle (£)	Source of cost
Ondansetron (oral azacitidine)	0.22	5	£2.20	eMIT 2020. Ondansetron 8mg tablets (10 pack size).
Ondansetron (Midostaurin)	0.22	2.5	£1.10	eMIT 2020. Ondansetron 8mg tablets (10 pack size).

Abbreviations: mg = milligrams.

Comparators – Watch and wait with BSC

No additional drug costs were assigned to the comparator except for those related to BSC (B.3.5.2.1.2).

Comparators – Midostaurin (FLT3 subgroup only)

For patients in complete response, midostaurin is administered orally at 50mg twice daily as single agent maintenance treatment until relapse for up to 12 cycles of 28 days each.[117]

Table B.3.17 provides the formulation and unit costs for midostaurin along with the calculated cost per cycle. Drug costs relating to BSC were also included for this comparator (B.3.5.2.1.2).

Table B.3.17. Drug acquisition costs for midostaurin

Treatment	Unit Strength	Unit Description	Unit List Price (£)	Dose per cycle	Cost per Cycle (£)	Source
Midostaurin	25 mg	Capsule	100.18	28	11,219.88	BNF

Abbreviations: mg = milligrams; BNF = British National Formulary

B.3.5.1.2 Treatment administration costs

Administration costs associated for each treatment were incorporated into the model and were taken from NHS reference costs 2019/2020.[130] Drug administration costs for IV and SC chemotherapies were incurred at each treatment initiation; for oral chemotherapies, the cost was included per cycle. Oral administration of chemotherapy is covered by HRG code SB11Z (Deliver exclusively oral chemotherapy), while for chemotherapy drugs delivered IV HRG code SB13Z (deliver more complex parenteral chemotherapy at first attendance) and for drugs delivered by SC, HRG code SB12Z (deliver simple parenteral chemotherapy at first attendance) was used.

Table B.3.18 below presents the associated administration costs for each treatment formulation and Table B.3.19 provides an overview of the treatment administration details per cycle.

Table B.3.18. Overview of the costs associated with different treatment formulations and the resources used

Description	Unit cost (£)	Source
Same day chemotherapy admission or attendance.	152.28	NHS reference costs 2019-2020.[130] SB97Z: same day chemotherapy admission or attendance.

Description	Unit cost (£)	Source
Chemotherapy admin. fees: SC.	221.35	NHS reference costs 2019-2020.[130] SB12Z: deliver simple parenteral chemotherapy at first attendance (outpatient).
Chemotherapy admin. fees: IV	302.53	NHS reference costs 2019-2020.[130] SB13Z: deliver more complex parenteral chemotherapy at first attendance (outpatient).
Chemotherapy management fees: oral (per administration)	207.79	NHS reference costs 2019-2020.[130] SB11Z: deliver exclusively oral chemotherapy.
Chemotherapy supervision fees (per administration)	136.36	NHS reference costs 2019-2020.[130] 370: non-admitted non-face to face attendance, follow up.
Hospitalisation per day	802.00	NHS reference costs 2019-2020.[130] Non-elective short stay.

Abbreviations: mg = milligrams; IV = intravenous; SC = subcutaneous

Table B.3.19. Number of administration/resource use per cycle

Description	Oral azacitidine	Watch and wait plus BSC	Midostaurin
Chemotherapy management fees: oral (per administration)	1	0	1

B.3.5.2 Health-state unit costs and resource use

Patients incur medical costs associated with the health states; RFS (on treatment and off treatment) and relapse. The costs consisted of disease management in the form of routine healthcare use, and disease management in the form of best supportive care with pharmacological agents. Total costs were inclusive of subsequent therapy costs including HSCT.

B.3.5.2.1 Disease management

B.3.5.2.1.1 Resource use costs

Resource use has been calculated per treatment cycle. The frequency per cycle has been informed from both clinical advisor opinion and the QUAZAR trial.[88] The proportion of patients receiving red blood cell (RBC) and platelet transfusions in relapse was informed by the QUAZAR AML-001 trial; this was validated by the UK clinician. Costs assigned to these resources have been obtained from NHS reference costs 2019-2020.[130] Acquisition costs for all types of treatments have been sourced primarily from eMIT 2020[129] and where necessary supplemented by the online BNF 2021. Treatment administration costs have been sourced from NHS reference costs 2019-2020.[130]

Table B.3.20 provides a summary of resource use per cycle and by treatment arm. For oral azacitidine in the FLT3 subgroup, resource use was assumed to be the same as that of oral azacitidine in ITT population, with the exception of bone marrow aspirate/biopsy. All resource use estimates were guided by UK clinical expert opinion. For red blood cell transfusion, it was

assumed that patients would receive two units of transfusions, once per cycle and for platelet transfusions it was assumed that patients would receive one unit of transfusion, twice per cycle. This was based on UK clinical expert opinion. The associated unit costs are provided in Table B.3.21.

Table B.3.20. Frequency and proportion of patients requiring resources per cycle

Treatment	Resource use							
	Health state	Hematologist visit	Nurse visit	CBC/differentia lab test	Chemistry and liver panel	RBC transfusion	Platelet transfusion	Bone marrow aspirate/biopsy
Oral azacitidine	RFS: on treatment	1.0	2.0	4.0	1.0	0.0%	0.0%	17.5%
	RFS: off treatment	1.0	1.5	1.3	1.0	0.0%	0.0%	13.5%
	Relapse	2.0	2.0	8.0	2.0	22.7%	21.8%	0.0%
Oral azacitidine (FLT3)	RFS: on treatment	1.0	2.0	4.0	1.0	0.0%	0.0%	25.0%
	RFS: off treatment	1.0	1.5	1.3	1.0	0.0%	0.0%	25.0%
	Relapse	2.0	2.0	8.0	2.0	22.7%	21.8%	0.0%
Watch and wait plus BSC	RFS: on treatment							
	RFS: off treatment	1.0	1.5	1.3	1.0	0.0%	0.0%	17.0%
	Relapse	2.0	2.0	8.0	2.0	21.8%	21.8%	0.0%
Watch and wait plus BSC (FLT3)	RFS: on treatment							
	RFS: off treatment	1.0	1.5	1.3	1.0	0.0%	0.0%	17.0%
	Relapse	2.0	2.0	8.0	2.0	21.8%	21.8%	0.0%
Midostaurin	RFS: on treatment	1.0	1.5	3.0	1.0	0.0%	0.0%	25.0%
	RFS: off treatment	1.0	1.0	1.3	0.5	0.0%	0.0%	25.0%
	Relapse	2.0	2.0	8.0	2.0	21.8%	21.8%	0.0%

Abbreviations: RFS = relapse free survival; CBC = complete blood count; FLT3 = fms-like tyrosine kinase 3; RBC = red blood cell; BSC = best supportive care

Table B.3.21. Resource use unit costs

Resource use	Cost input (£) Unit cost	Source
Haematologist visit	166.00	NHS reference costs 2019-2020. Service code 303, clinical haematology outpatients[130]
Nurse visit	99.30	NHS reference costs 2019-2020. Service code NURS-Specialist nursing, cancer related, adult, face to face[130]
CBC/differential lab test	1.00	NHS reference costs 2019-2020. Service code DAPS04-clinical biochemistry[130]
Chemistry and liver panel	1.00	NHS reference costs 2019-2020. Service code DAPS04-clinical biochemistry, (liver results can be found in blood tests) [130]
RBC transfusion	221.46	NHS reference costs 2019-2020. Service code SA44A-single plasma exchange or other intravenous blood transfusion 19 years and over- medical oncology (370) [130]
Platelet transfusion	221.46	NHS reference costs 2019-2020. Service code SA44A-single plasma exchange or other intravenous blood transfusion 19 years and over- medical oncology (370) [130]
Bone marrow aspirate/biopsy	78.09	NHS reference costs 2019-2020. Service code SA33Z, 370 medical oncology- diagnostic bone marrow extraction[130]

Abbreviations: CBC = complete blood count, NHS = National Health Service, RBC = red blood cell

B.3.5.2.1.2 Costs related to best supportive care

BSC costs were included in the model to capture ongoing disease management costs for patients in the relapse health state and includes medications such as antibiotics, antifungals, and hydroxyurea. All patients in the model received BSC regardless of treatment arm, except patients in RFS, this was based on UK clinical expert opinion. The percentage of patients in each health state expected to receive each component of BSC was validated by UK clinical expert opinion. The economic evaluation considered components of BSC listed in Table B.3.22 differences in BSC between patients treated with oral azacitidine, midostaurin and watch and wait.

The dosing regimens for BSC drugs were obtained from their respective SmPC and validated by UK clinical experts. Acquisition costs for all types of treatments have been sourced primarily from eMIT 2020[129] and where necessary supplemented by the online BNF 2021. Treatment administration costs have been sourced from NHS reference costs 2019-2020.[130] (Table B.3.23).

Table B.3.22. Proportion of patients receiving each component of BSC

Treatment	BSC component					
	Health state	Hydroxycarbamide	Ciprofloxacin	Posaconazole	Fluconazole	Tranexamic acid
Oral azacitidine	RFS: on treatment	0%	0%	0%	0%	0%
	RFS: off treatment	0%	0%	0%	0%	0%
	Relapse	15%	30%	15%	15%	15%
Oral azacitidine (FLT3)	RFS: on treatment	0%	0%	0%	0%	0%
	RFS: off treatment	0%	0%	0%	0%	0%
	Relapse	20%	30%	15%	15%	15%
Watch and wait plus BSC	RFS: on treatment					
	RFS: off treatment	0%	0%	0%	0%	0%
	Relapse	15%	30%	15%	15%	15%
Watch and wait plus BSC (FLT3)	RFS: on treatment					
	RFS: off treatment	0%	0%	0%	0%	0%
	Relapse	15%	30%	15%	15%	15%
Midostaurin	RFS: on treatment	0%	0%	0%	0%	0%
	RFS: off treatment	0%	0%	0%	0%	0%
	Relapse	20%	30%	15%	15%	15%

Abbreviations: RFS = relapse-free survival; BSC = best supportive care; FLT3 = fms-like tyrosine kinase 3

Table B.3.23. Drug costs for BSC

Drug	Admin route, dosing, dose, dose per cycle	Dose per tablet	Units per pack	Cost per pack	Cost per unit (£)	Cost per mg (£)	Cost per cycle (£)	Source
Hydroxycarbamide	Oral, weight based, 40mg/kg, 7	500mg	100	9.61	0.10	0.0002	4.00	eMIT 2020[129]
Ciprofloxacin	Oral, fixed, 500mg, 14	500	10	3.08	0.31	0.0006	4.31	eMIT 2020[129]
Posaconazole	Oral, fixed, 400mg,21	100	24	17.32	7.30	0.0731	460.22	eMIT 2020[129]
Fluconazole	Oral, fixed, 200mg, 21	200	7	0.51	0.07	0.0004	1.53	eMIT 2020[129]
Tranexamic acid	Oral, fixed, 1000mg, 21	500	60	7.98	0.13	0.0003	5.59	eMIT 2020[129]

Abbreviations: mg = milligrams

B.3.5.2.1.3 Subsequent therapy resource use and costs

Subsequent treatment has been included in the model as described in Section B.3.3.4. Following relapse, a proportion of surviving patients move on to receive a single line of subsequent therapy. These are included to inform the costs only with no impact on efficacy outcomes as it is assumed that any effects would be captured in the OS curves from the respective clinical trials. The subsequent treatments considered are low dose cytarabine, injectable azacitidine and salvage chemotherapy; for costing purpose, salvage chemotherapy is assumed to include a combination of daunorubicin and cytarabine.

Table B.3.24 below shows the dosing regimen for each of the subsequent treatments used in the model. Low dose cytarabine is based on a fixed dosing regimen with the others based on BSA. Stopping rules/capping of cycles have been included where patients receive subsequent treatments for a certain time-period.

Table B.3.24. Dosing regimen for the different subsequent treatments included in the model

Schedule and dosing details	Low-dose cytarabine	Injectable azacitidine	Salvage chemotherapy	
			Daunorubicin	Cytarabine
Dose (mg for Fixed, mg/kg for Weight, and mg/m2 for BSA)	20	75	60	200
Dosing (Fixed, Weight, or BSA)	Fixed	BSA	BSA	BSA
Schedule	Days 1-10, q12h	Days 1-7	Days 1-3	Days 1-7
Route	Subcutaneous	Subcutaneous	Intravenous	Intravenous
Doses per cycle	20	7	3	7
Cycle length (days)	28	28	28	28
Total administration (mg)	20.00	139.03	111.23	370.76
Dose per cycle (mg)	400	973	334	2,595
Number of cycles	4	3	1	1

Abbreviations: mg = milligrams

The costing details for subsequent treatment are shown in Table B.3.25. It is assumed that there is no vial sharing for any treatments that are given via IV/SC route. The model assumes that there is no discount on the list prices for the subsequent treatments. The drug acquisition costs have been taken from eMIT 2020 database primarily and then where necessary supplemented from BNF online 2021.

Table B.3.25. Overview of the drug costs per cycle for the subsequent treatments considered in the model

Costing details	Low-dose cytarabine	Injectable azacitidine	Salvage chemotherapy	
			Daunorubicin	Cytarabine
Unit strength (mg)	100	100	20	100
Unit description	Vial	Vial	Vial	Vial
Unit costs per pack (£)	4.50	220.00	71.50	4.48
Discount on list price (%)	0%	0%	0%	0%
Number of doses per unit: no vial sharing	1.00	2.00	6.00	4.00
Cost per cycle (£)	90.08	4,494.00	1,287.00	125.33
Source of drug costs	eMIT 2020[129]	Online BNF 2021[131]	Online BNF 2021[131]	eMIT 2020[129]

Abbreviations: mg = milligrams

Table B.3.26 shows units of treatment administration resource use for each subsequent treatment considered in the model. Alongside this, the days of premedication needed prior to subsequent treatment are also included. Costs and dosing of these have been discussed earlier.

Table B.3.26. Subsequent therapy treatment administration frequency

Treatment administration details	Low-dose cytarabine	Injectable azacitidine	Salvage chemotherapy
Chair time per admin. (hours)	0	4	0
Chemotherapy admin. fees: IV and SC (number per cycle)	0	7	7
Chemotherapy management fees: oral (number per cycle)	1	0	0
Chemotherapy supervision fees (number per cycle)	0	1	1
Hospitalization per cycle (days)	0	0	28
Number of days of premedication per cycle- prochlorperazine	28	28	14
Number of days of premedication per cycle- ondansetron	0	0	0
Number of days of premedication per cycle- dexamethasone	0	7	7
Number of days of premedication per cycle-antihistamine chlorphenamine	0	0	7

Treatment administration details	Low-dose cytarabine	Injectable azacitidine	Salvage chemotherapy
Cost per cycle in (£)	208.84	2,297.46	24,144.42

Abbreviations: IV = intravenous; SC = subcutaneous

B.3.5.2.1.4 Stem cell treatment unit costs

Although oral azacitidine is licensed for use in patients who are ineligible or choose not to have a haematopoietic stem cell transplant, a small proportion of patients in the QUAZAR study[88] did go on to receive HSCT. The proportion of patients that received HSCT in the midostaurin treatment arm has been taken from Larson et al, 2021. The proportion of patients undergoing HSCT is informed from the relevant clinical trials QUAZAR AML-001 and RATIFY. NHS reference costs 2019-2020 have been used to estimate the cost of a HSCT.[130]

Table B.3.27. Proportion of patients undergoing stem cell transplant per treatment

Parameter	Treatment	Proportion	Sources
Proportion of patients receiving stem cell transplant	Oral azacitidine	6.3%	Supplementary appendix to Wei et al., 2020[88]
	Watch and wait plus BSC	13.7%	Supplementary appendix to Wei et al., 2020[88]
	Oral azacitidine (FLT3)	6.3%	Assumed same as ITT
	Midostaurin (FLT3)	5.8%	Larson et al 2021[83]
Parameter	**Unit costs (£)**	**Source**	
Stem cell transplant	15,065.00	NHS reference costs 2019-2020. Peripheral blood stem cell transplant, autologous, 19 years and over. Code SA26A[130]	

Abbreviations: CSR = clinical study report; FLT3 = fms-like tyrosine kinase 3; ITT = intention-to-treat; NHS = National Health Service; BSC = best supportive care

B.3.5.3 Adverse events costs and resource use

To represent the safety data of the treatments used, data of adverse event reactions from the respective clinical trials have been incorporated into the model. The model includes all grade 3 or 4 treatment emergent adverse event reactions occurring in ≥5% of patients. For oral azacitidine and watch and wait plus BSC, neutropenia and thrombocytopenia had the highest incidences. For midostaurin, the corresponding grade 3 or 4 treatment emergent adverse event reactions occurring in >10% of patients have been included in line with the midostaurin TA523[31] submission based on the maintenance phase. An overview of these adverse events for all interventions and comparators included in the model have been provided earlier in

Table B.**3.8**.

In order to capture the resource use and costs associated to treat adverse events, both inpatient and outpatient related costs were retrieved and using these the average cost per event is calculated. Costs were informed from NHS reference costs 2019-2020.[130] For instances where specific adverse events could not be found, e.g., adult febrile neutropenia, the cost was assumed to be the same as that associated for acute myeloid leukaemia with CC Score 0-1, this is in line with assumptions from TA52331. Through clinical validation, the duration of these adverse events are assumed to be one week. Table B.3.28 shows the proportion of patients treated as inpatient and outpatient for the respective adverse events based on UK clinical expert opinion. Table B.3.29 shows the costs associated for inpatient treatment, outpatient treatment and the average of both that is used in the model.**Error! Reference source not found.****Table B.3.28. Overview of the percentage of patients treated as inpatient and outpatient for AEs**

AEs	Neutropenia	Thrombocytopenia	Anaemia	Febrile Neutropenia	Diarrhoea	Vomiting	Nausea	Fatigue
Percentage treated as inpatient	0%	10%	10%	100%	5%	5%	0%	5%
Percentage treated as outpatient	100%	90%	90%	0%	95%	95%	100%	95%

Abbreviations: AEs = adverse events

Table B.3.29. Cost per adverse event for inpatient and outpatient treatment and the average cost calculated

AEs	Cost per inpatient stay (£)	Cost per outpatient stay (£)	Average cost per event (£)	Source for inpatient treatment	Source for outpatient treatment
Neutropenia	754.00	394.00	394.00	NHS reference costs 2019-2020. SA25M (total)	NHS reference costs 2019-2020. SA25M (day case)
Thrombocytopenia	363.00	13.80	48.72	NHS reference costs 2019-2020. SA12K (total)	NHS reference costs 2019-2020. SA12K (day case)

AEs	Cost per inpatient stay (£)	Cost per outpatient stay (£)	Average cost per event (£)	Source for inpatient treatment	Source for outpatient treatment
Anaemia	361.00	305.00	310.60	NHS reference costs 2019-2020. SA04L (total)	NHS reference costs 2019-2020. SA04L (day case)
Febrile neutropenia	754.00	394.00	754.00	NHS reference costs 2019-2020. SA25M (total)	NHS reference costs 2019-2020. SA25M (day case)
Diarrhoea	797.00	365.00	386.60	NHS reference costs 2019-2020. FD01J- (total)	NHS reference costs 2019-2020. FD01J- (day case)
Vomiting	797.00	365.00	386.60	NHS reference costs 2019-2020. FD01J (total)	NHS reference costs 2019-2020. FD01J (day case)
Nausea	754.00	394.00	394.00	NHS reference costs 2019-2020. SA25M (total)	NHS reference costs 2019-2020. SA25M (day case)
Fatigue	754.00	394.00	412.00	NHS reference costs 2019-2020. SA25M (total)	NHS reference costs 2019-2020. SA25M (day case)

Abbreviations: AEs = adverse events

B.3.5.4 Miscellaneous unit costs and resource use

B.3.5.4.1 End of life care costs

End of life costs are applied once upon transition from relapse to death and costs have been informed from TA523[31]. Data collected from Nuffield trust, 2014[132] was used and end of life costs included acute hospital care (all hospital contacts, emergency inpatient admissions, non-emergency inpatient admissions, outpatient visits, accident and emergency visits), local authority-funded social care, district nursing care, and general practitioner visit costs. The total cost that the Nuffield trust reported (in 2014) was £13,176 and this was inflated to 2019/2020 prices. The HCHS inflation index[128] had been used up to and including 2015/16 and the NHSCII inflation index[128] had been used thereafter. This gave a total cost of £14,708.43.

B.3.5.5 Summary of base-case analysis inputs

Table B.3.30. Summary of variables applied in the economic model

Variable	Deterministic value (base-case analysis)	Distribution in DSA and PSA	Lower bound for DSA	Upper bound for DSA	Reference to section in submission
General model settings					
Time horizon, years	30	Not included			B.3.2.1
Discounting per year – costs	3.5%	Not included			
Discounting per year – clinical outcomes	3.5%	Not included			
Baseline patient characteristics					
Male proportion	51.9%	Fixed			B.3.3.1.1
Mean age, year	67.9	Fixed			
Body surface area, m2	[redacted]	Normal	[redacted]	[redacted]	
Survival parameters					
RFS, OS oral AZA	Parametric model	Multivariate normal			B.3.3.1
RFS, OS no active treatment	Parametric model	Multivariate normal			
AEs					
AE incidence – oral AZA					B.3.3.5
Neutropenia	41.1%	Beta	32.9%	49.3%	
Thrombocytopenia	22.5%	Beta	18.0%	27.0%	
Anemia	14.0%	Beta	11.2%	16.8%	

Variable	Deterministic value (base-case analysis)	Distribution in DSA and PSA	Lower bound for DSA	Upper bound for DSA	Reference to section in submission
Febrile neutropenia	11.4%	Beta	9.1%	13.7%	
Diarrhea	5.1%	Beta	4.1%	6.1%	
Vomiting	3.0%	Beta	2.4%	3.6%	
Nausea	2.5%	Beta	2.0%	3.0%	
Fatigue	3.0%	Beta	2.4%	3.6%	
AE incidence – Watch and wait plus BSC					
Neutropenia	23.6%	Beta	18.9%	28.3%	
Thrombocytopenia	21.5%	Beta	17.2%	25.8%	
Anemia	12.9%	Beta	10.3%	15.5%	
Febrile neutropenia	7.7%	Beta	6.2%	9.2%	
Diarrhea	1.3%	Beta	1.0%	1.6%	
Vomiting	0.0%	Beta	0.0%	0.0%	
Nausea	0.4%	Beta	0.3%	0.5%	
Fatigue	0.9%	Beta	0.7%	1.1%	
AE incidence – midostaurin					
Neutropenia	8.3%	Beta	30.4%	45.6%	
Thrombocytopenia	1.7%	Beta	21.6%	32.4%	
Anaemia	0.8%	Beta	24.8%	37.2%	
Febrile neutropenia	0.8%	Beta	16.0%	24.0%	
Diarrhea	0.8%	Beta	22.4%	33.6%	

Variable	Deterministic value (base-case analysis)	Distribution in DSA and PSA	Lower bound for DSA	Upper bound for DSA	Reference to section in submission
Vomiting	0.0%	Beta	12.8%	19.2%	
Nausea	0.0%	Beta	32.0%	48.0%	
Fatigue	0.0%	Beta	30.4%	45.6%	
AE incidence – oral AZA (FLT3)					
Neutropenia	[redacted]	Beta	[redacted]	[redacted]	
Thrombocytopenia	[redacted]	Beta	[redacted]	[redacted]	
Anemia	[redacted]	Beta	[redacted]	[redacted]	
Febrile neutropenia	[redacted]	Beta	[redacted]	[redacted]	
Diarrhea	[redacted]	Beta	[redacted]	[redacted]	
Vomiting	[redacted]	Beta	[redacted]	[redacted]	
Nausea	[redacted]	Beta	[redacted]	[redacted]	
Fatigue	[redacted]	Beta	[redacted]	[redacted]	
AE incidence – Watch and wait plus BSC (FLT3)					
Neutropenia	[redacted]	Beta	[redacted]	[redacted]	
Thrombocytopenia	[redacted]	Beta	[redacted]	[redacted]	
Anemia	[redacted]	Beta	[redacted]	[redacted]	
Febrile neutropenia	[redacted]	Beta	[redacted]	[redacted]	
Diarrhea	[redacted]	Beta	[redacted]	[redacted]	
Vomiting	[redacted]	Beta	[redacted]	[redacted]	
Nausea	[redacted]	Beta	[redacted]	[redacted]	

Variable	Deterministic value (base-case analysis)	Distribution in DSA and PSA	Lower bound for DSA	Upper bound for DSA	Reference to section in submission
Fatigue	[redacted]	Beta	[redacted]	[redacted]	
AE duration, weeks					
Neutropenia	1	Not included			B.3.4.4
Thrombocytopenia	1	Not included			
Anemia	1	Not included			
Febrile neutropenia	1	Not included			
Diarrhea	1	Not included			
Vomiting	1	Not included			
Nausea	1	Not included			
Fatigue	1	Not included			
AE costs inpatient, £					
Neutropenia	754.00	Fixed	603.20	904.80	B.3.5.3
Thrombocytopenia	363.00	Fixed	290.40	435.60	
Anaemia	361.00	Fixed	288.80	433.20	
Febrile neutropenia	754.00	Fixed	603.20	904.80	
Diarrhea	797.00	Fixed	637.60	956.40	
Vomiting	797.00	Fixed	637.60	956.40	
Nausea	754.00	Fixed	603.20	904.80	
Fatigue	754.00	Fixed	603.20	904.80	
AE costs outpatient, £					

Variable	Deterministic value (base-case analysis)	Distribution in DSA and PSA	Lower bound for DSA	Upper bound for DSA	Reference to section in submission
Neutropenia	394.00	Fixed	315.20	472.80	B.3.5.3
Thrombocytopenia	13.80	Fixed	11.04	16.56	
Anaemia	305.00	Fixed	244.00	366.00	
Febrile neutropenia	394.00	Fixed	315.20	472.80	
Diarrhea	365.00	Fixed	292.00	438.00	
Vomiting	365.00	Fixed	292.00	438.00	
Nausea	394.00	Fixed	315.20	472.80	
Fatigue	394.00	Fixed	315.20	472.80	
Mortality					
Background mortality	Age- and sex-specific estimates	Fixed			B.3.3.1.3
Treatment costs per cycle, £					
Drug acquisition costs per cycle					
Oral azacitidine (all doses and pack sizes)	[redacted]	Fixed			B.3.5.1.1
Midostaurin	11,219.88	Fixed			
Hydroxyurea	4.00	Fixed			
Amoxicillin	0.40	Fixed			
Ciprofloxacin	4.31	Fixed			
Posaconazole	460.22	Fixed			
Fluconazole	1.53	Fixed			

Variable	Deterministic value (base-case analysis)	Distribution in DSA and PSA	Lower bound for DSA	Upper bound for DSA	Reference to section in submission
Voriconazole	47.46	Fixed			
Tranexamic acid	5.59	Fixed			
Dose intensity					
Oral azacitidine	[redacted]	Beta	[redacted]	[redacted]	B.3.5.1.1
Administration cost					
Same day chemotherapy admission or attendance.	152.28	Gamma	121.82	182.74	B.3.5.1.2
Chemotherapy admin. fees: SC.	221.35	Gamma	177.08	265.62	
Chemotherapy admin. fees: IV	302.53	Gamma	242.02	363.04	
Chemotherapy management fees: oral (per administration)	207.79	Gamma	166.23	249.35	
Chemotherapy supervision fees (per administration)	136.36	Gamma	109.09	163.63	
Hospitalisation per day	802.00	Gamma	641.60	962.40	
Healthcare resource use					
Healthcare resource use per cycle – oral azacitidine					

Variable	Deterministic value (base-case analysis)	Distribution in DSA and PSA	Lower bound for DSA	Upper bound for DSA	Reference to section in submission
Haematologist visit RFS on treatment	1	Fixed			B.3.5.2.1
Haematologist visit RFS off treatment	1	Fixed			
Haematologist visit relapse	2	Fixed			
Nurse visit RFS on treatment	2	Fixed			
Nurse visit RFS off treatment	1.5	Fixed			
Nurse visit relapse	2	Fixed			
CBC/differentia lab test RFS on treatment	4	Fixed			
CBC/differentia lab test RFS off treatment	1.3	Fixed			
CBC/differentia lab test relapse	8	Fixed			
Chemistry and liver panel RFS on treatment	1	Fixed			
Chemistry and liver panel RFS off treatment	1	Fixed			

Variable	Deterministic value (base-case analysis)	Distribution in DSA and PSA	Lower bound for DSA	Upper bound for DSA	Reference to section in submission
Chemistry and liver panel relapse	2	Fixed			
RBC transfusion RFS on treatment	0%	Fixed			
RBC transfusion RFS off treatment	0%	Fixed			
RBC transfusion relapse	21.8%	Fixed			
Platelet transfusion RFS on treatment	0%	Fixed			
Platelet transfusion RFS off treatment	0%	Fixed			
Platelet transfusion relapse	21.8%	Fixed			
Bone marrow aspirate/biopsy RFS on treatment	17.5%	Fixed			
Bone marrow aspirate/biopsy RFS off treatment	13.5%	Fixed			
Bone marrow aspirate/biopsy relapse	0%	Fixed			
Healthcare resource use – Watch and wait plus BSC (ITT and FLT3)					

Variable	Deterministic value (base-case analysis)	Distribution in DSA and PSA	Lower bound for DSA	Upper bound for DSA	Reference to section in submission
Haematologist visit RFS on treatment	0	Fixed			B.3.5.2.1
Haematologist visit RFS off treatment	1	Fixed			
Haematologist visit relapse	2	Fixed			
Nurse visit RFS on treatment	0	Fixed			
Nurse visit RFS off treatment	1.5	Fixed			
Nurse visit relapse	2	Fixed			
CBC/differentia lab test RFS on treatment	0	Fixed			
CBC/differentia lab test RFS off treatment	1.3	Fixed			
CBC/differentia lab test relapse	8	Fixed			
Chemistry and liver panel RFS on treatment	0	Fixed			
Chemistry and liver panel RFS off treatment	1	Fixed			

Variable	Deterministic value (base-case analysis)	Distribution in DSA and PSA	Lower bound for DSA	Upper bound for DSA	Reference to section in submission
Chemistry and liver panel relapse	2	Fixed			
RBC transfusion RFS on treatment	0%	Fixed			
RBC transfusion RFS off treatment	0%	Fixed			
RBC transfusion relapse	21.8%	Fixed			
Platelet transfusion RFS on treatment	0	Fixed			
Platelet transfusion RFS off treatment	0	Fixed			
Platelet transfusion relapse	21.8%	Fixed			
Bone marrow aspirate/biopsy RFS on treatment	0	Fixed			
Bone marrow aspirate/biopsy RFS off treatment	17.5%	Fixed			
Bone marrow aspirate/biopsy relapse	0	Fixed			
Healthcare resource use – midostaurin					

Variable	Deterministic value (base-case analysis)	Distribution in DSA and PSA	Lower bound for DSA	Upper bound for DSA	Reference to section in submission
Haematologist visit RFS on treatment	1	Fixed			B.3.5.2.1
Haematologist visit RFS off treatment	1	Fixed			
Haematologist visit relapse	2	Fixed			
Nurse visit RFS on treatment	1.5	Fixed			
Nurse visit RFS off treatment	1	Fixed			
Nurse visit relapse	2	Fixed			
CBC/differentia lab test RFS on treatment	3	Fixed			
CBC/differentia lab test RFS off treatment	1.3	Fixed			
CBC/differentia lab test relapse	8	Fixed			
Chemistry and liver panel RFS on treatment	1	Fixed			
Chemistry and liver panel RFS off treatment	0.5	Fixed			

Variable	Deterministic value (base-case analysis)	Distribution in DSA and PSA	Lower bound for DSA	Upper bound for DSA	Reference to section in submission
Chemistry and liver panel relapse	2	Fixed			
RBC transfusion RFS on treatment	0%	Fixed			
RBC transfusion RFS off treatment	0%	Fixed			
RBC transfusion relapse	21.8%	Fixed			
Platelet transfusion RFS on treatment	0%	Fixed			
Platelet transfusion RFS off treatment	0%	Fixed			
Platelet transfusion relapse	21.8%	Fixed			
Bone marrow aspirate/biopsy RFS on treatment	25.0%	Fixed			
Bone marrow aspirate/biopsy RFS off treatment	25.0%	Fixed			
Bone marrow aspirate/biopsy relapse	0%	Fixed			
Healthcare resource use per cycle – oral azacytidine (FLT3)					

Variable	Deterministic value (base-case analysis)	Distribution in DSA and PSA	Lower bound for DSA	Upper bound for DSA	Reference to section in submission
Haematologist visit RFS on treatment	1	Fixed			
Haematologist visit RFS off treatment	1	Fixed			
Haematologist visit relapse	2	Fixed			
Nurse visit RFS on treatment	2	Fixed			
Nurse visit RFS off treatment	1.5	Fixed			
Nurse visit relapse	2	Fixed			
CBC/differentia lab test RFS on treatment	4	Fixed			
CBC/differentia lab test RFS off treatment	1.3	Fixed			
CBC/differentia lab test relapse	8	Fixed			
Chemistry and liver panel RFS on treatment	1	Fixed			
Chemistry and liver panel RFS off treatment	1	Fixed			

Variable	Deterministic value (base-case analysis)	Distribution in DSA and PSA	Lower bound for DSA	Upper bound for DSA	Reference to section in submission
Chemistry and liver panel relapse	2	Fixed			
RBC transfusion RFS on treatment	0%	Fixed			
RBC transfusion RFS off treatment	0%	Fixed			
RBC transfusion relapse	21.8%	Fixed			
Platelet transfusion RFS on treatment	0%	Fixed			
Platelet transfusion RFS off treatment	0%	Fixed			
Platelet transfusion relapse	21.8%	Fixed			
Bone marrow aspirate/biopsy RFS on treatment	25.0%	Fixed			
Bone marrow aspirate/biopsy RFS off treatment	25.0%	Fixed			
Bone marrow aspirate/biopsy relapse		**Fixed**			
Healthcare resource use cost, £					
Haematologist visit	166.00	Gamma	132.80	199.20	B.3.5.2.1

Variable	Deterministic value (base-case analysis)	Distribution in DSA and PSA	Lower bound for DSA	Upper bound for DSA	Reference to section in submission
Nurse visit	99.30	Gamma	79.44	119.16	
CBC/differentia lab test	1.00	Gamma	0.80	1.20	
Chemistry and liver panel	1.00	Gamma	0.80	1.20	
RBC transfusion	221.46	Gamma	177.17	265.75	
Platelet transfusion	221.46	Gamma	177.17	265.75	
Bone marrow aspirate/biopsy	78.09	Gamma	62.47	93.71	
Subsequent care post relapse					
Subsequent care unit cost, £					B.3.5.2.1.3
Hydroxyurea	0.10	Fixed			
Amoxicillin	0.02	Fixed			
Ciprofloxacin	0.31	Fixed			
Posaconazole	7.30	Fixed			
Fluconazole	0.07	Fixed			
Voriconazole	1.13	Fixed			
Tranexamic acid	0.13	Fixed			
Subsequent care distribution					
Hydroxyurea in ITT population	15%	Fixed			

Variable	Deterministic value (base-case analysis)	Distribution in DSA and PSA	Lower bound for DSA	Upper bound for DSA	Reference to section in submission
Hydroxyurea in FLT3 population	20%	Fixed			
Amoxicillin	0%	Fixed			
Ciprofloxacin	30%	Fixed			
Posaconazole	15%	Fixed			
Fluconazole	15%	Fixed			
Voriconazole	0%	Fixed			
Tranexamic acid	15%	Fixed			
Subsequent therapy – oral azacytidine					
Low-dose cytarabine	14.3%	Fixed			B.3.3.4
Injectable azacitidine	8.4%	Fixed			
Salvage chemotherapy	26.%	Fixed			
Subsequent therapy – oral azacytidine (FLT3)					
Low-dose cytarabine	16.7%	Fixed			B.3.3.4
Injectable azacitidine	6.7%	Fixed			
Salvage chemotherapy	23.3%	Fixed			
Subsequent therapy – Watch and wait plus BSC					
Low-dose cytarabine	10.7%	Fixed			B.3.3.4
Injectable azacitidine	15.4%	Fixed			
Salvage chemotherapy	33.8%	Fixed			

Variable	Deterministic value (base-case analysis)	Distribution in DSA and PSA	Lower bound for DSA	Upper bound for DSA	Reference to section in submission
Subsequent therapy – Watch and wait plus BSC (FLT3)					
Low-dose cytarabine	11.1%	Fixed			B.3.3.4
Injectable azacitidine	8.3%	Fixed			
Salvage chemotherapy	36.1%	Fixed			
Subsequent therapy – midostraurin					
Low-dose cytarabine	11.1%	Fixed			B.3.3.4
Injectable azacitidine	8.3%	Fixed			
Salvage chemotherapy	36.1%	Fixed			
Peripheral Blood Stem Cell Transplant costs, £					
Cost per HSCT	15,065.00	Gamma	12,052.00	18,078.00	B.3.5.2.1.4
End of life costs, £					
End of life cost per mortality event	14,708.43	Fixed			B.3.5.4.1
Utilities					
Health state utilities					
Pre-progression survival	[redacted]	Beta	[redacted]	[redacted]	B.3.4.6
Post-relapse survival	[redacted]	Beta	[redacted]	[redacted]	
AE disutility					
Neutropenia	0.0897	Beta	0.0718	0.1076	B.3.4.4
Thrombocytopenia	0.0897	Beta	0.0718	0.1076	

Variable	Deterministic value (base-case analysis)	Distribution in DSA and PSA	Lower bound for DSA	Upper bound for DSA	Reference to section in submission
Anemia	0.2975	Beta	0.0952	0.1428	
Febrile neutropenia	0.0375	Beta	0.1200	0.1800	
Diarrhea	0.0070	Beta	0.1408	0.2112	
Vomiting	0.0156	Beta	0.0384	0.0576	
Nausea	0.0156	Beta	0.0384	0.0576	
Fatigue	0.2875	Beta	0.0920	0.1380	

Abbreviations: AE = adverse event; DSA = deterministic sensitivity analysis; FLT3 = fms-like tyrosine kinase 3; HSCT = Hematopoietic stem cell transplantation, OS = overall survival; PSA = probabilistic sensitivity analysis; RFS = relapse-free survival; BSC = best supportive care

B.3.5.6 Assumptions

Table B.3.31 provides an overview of a number of assumptions which should be taken into consideration when assessing the results provided in Section B.3.6

Table B.3.31. Model assumptions

Model feature	Source/assumption	Justification
Efficacy		
Data cut	For OS: September 2020 data cut For RFS: July 2019 data cut	The September 2020 data cut for OS is the most mature, and therefore is associated with less uncertainty than the July DBL. July RFS was deemed most robust and was used in economic modelling. In the extension phase that contributed to the Sept 2020 data cut, patients were followed up for survival, with only limited data available relating to relapse. See Section B.3.3.1 (Varied in scenario analysis)
Subsequent therapy	Efficacy of subsequent therapy was not considered, only costs.	It was assumed that efficacy for subsequent therapies was captured in the existing OS curves for both treatment arms. See Section B.3.3.4
Population	ITT population used for base case analysis	The ITT population included all randomised patients, regardless of whether they received study treatment. See Section B.2.4 and **Error! Reference source not found.** (varied in scenario analysis)
Survival	Choice of survival curves for OS (ITT) : Joint generalised gamma Choice of survival curves for RFS (ITT) : Joint log-logistic Choice of survival curves for OS (FLT3) : generalised gamma Choice of survival curves for RFS (FLT3) : 1 knot odds linear model	The choice of joint generalised gamma for OS for the ITT analysis was based on clinical plausibility, visual fit and lowest AIC and BIC . See Section B.3.3.1.3.3 The choice of joint log-logistic for RFS for the ITT analysis was based on clinical plausibility, visual fit and lowest AIC and BIC. See Section B.3.3.1.4.3 The choice of individual generalised gamma for OS (FLT3 subgroup) was based on clinical plausibility and assessment of AIC and BIC. See Section B.3.3.2.3.4 The choice of 1 knot odds linear model for RFS (FLT3 subgroup) was based on clinical plausibility and assessment of AIC and BIC. See Section B.3.3.2.4.4 (Varied in scenario analysis)
Time on Treatment	Oral azacitidine ITT : KM data QUAZAR AML-001 Oral azacitidine (FLT3) : KM data QUAZAR AML-001 Watch and wait with BSC : Not applicable Midostaurin: mean 11.02 cycles	KM curves accurately account for discontinuation and capture discounting appropriately. See Section B.3.3.3**Error! Reference source not found.**
Model		
Time horizon	30 years (=lifetime)	In line with NICE reference case. See Section B.3.2.1

Model feature	Source/assumption	Justification
HSCT	Modelled as subsequent therapy and not as a health state	Insufficient data available in the QUAZAR AML-001 trial or the literature to model HSCT as a health state and HSCT was not expected to results in a high proportion of cures. See Section B.3.2.1
Utility		
RFS on treatment	EQ-5D data from the QUAZAR AML-001	In line with NICE reference case. All patients in the same health state have the same utility value regardless of treatment arm. See Section B.3.4.1 (Varied in scenario analysis)
RFS off treatment	EQ-5D data from the QUAZAR AML-001	In line with NICE reference case. All patients in the same health state have the same utility value regardless of treatment arm. See Section B.3.4.1 (Varied in scenario analysis)
Relapse	Difference between the RFS and relapse utilities in Joshi 2019	EQ-5D trial data was not available to inform health state utility value for relapse from QUAZAR AML-001 so input from literature was required. Due to the differences between the QUAZAR AML-001 trial and the sample described by Joshi 2019, the difference between RFS and relapse was isolated to avoid the impact of sample differences. See Section B.3.4.3 (Varied in scenario analysis)
AE/HSCT disutility	The disutility associated with grade 3 or 4 AE and HSCT is applied in the first cycle of the model.	AEs are not a driver of the incremental results. This is a simplifying assumption.
Costs		
BSC	Assumed to only include medications such as antibiotics, antifungals, and hydroxyurea. All patients in the model received BSC regardless of treatment arm. BSC for the relapse health state for all treatments assumed to contain the same medication.	Base on UK clinical expert opinion. See Section B.3.5.2.1.2
Health state resource use	For oral azacitidine in the FLT3 subgroup, resource use was assumed to be the same as that of oral azacitidine in ITT population, with the exception of bone marrow aspirate/biopsy	Base on UK clinical expert opinion. See Section B.3.5.2.1.1
Subsequent therapy	Salvage chemotherapy assumed to include a combination of daunorubicin and cytarabine. Midostaurin subsequent therapy proportions were assumed	Efficacy related to subsequent therapy was not modelled, therefore for costing purposes, this was a sufficient simplifying assumption to capture costs relating to salvage chemotherapy. See Section B.3.5.2.1.3

Model feature	Source/assumption	Justification
	to be equal to that of the FLT3 subgroup for watch and wait.	
Cost of midostaurin	The drug and administration cost of midostaurin applied the first cycle of the model as a one-off cost	Time to discontinuation based on KM curves was not available. See Section B.3.5.1.1
Costs of AE	The costs of managing/treating AE of grade 3 or 4 is applied in the first cycle of the model.	AEs are not a driver of the incremental results. This is a simplifying assumption.
Relative dose intensity	Mean value of [redacted]%	Based on QUAZAR AML-001 trial

Abbreviations: AE = adverse event; BSC = best supportive care; CR = complete response; FLT3 = fms-like tyrosine kinase 3; HRQoL = health-related quality of life; ITT = intention-to-treat; OS = overall survival; RFS = relapse-free survival; HSCT = Hematopoietic stem cell transplantation; KM = Kaplan Meier

B.3.6 Base-case results

The economic model results are presented below using PAS price.

B.3.6.1 Base-case incremental cost-effectiveness analysis results

The base case results for oral azacitidine compared to watch and wait strategy with BSC is presented in Table B.3.32. Oral azacitidine generated higher total QALYs, higher LYs and higher costs than the watch and wait strategy with BSC, resulting in an ICER of £49,704.

The comparison between the clinical trial estimates and the estimates included in the cost-effectiveness model are presented in Appendix J with the disaggregated results for the base case. The cost-effectiveness result is predominantly driven by the higher relapse free life years and the higher drug acquisition cost for oral azacitidine compared to watch and wait plus BSC.

Table B.3.32. Base case results with oral azacitidine PAS (discounted)

Technology	Total costs (£)	Total LYG	Total QALYs	Incremental costs (£)	Incremental LYG	Incremental QALYs	ICER versus baseline (£/QALY)
Watch and wait with BSC	[redacted]	2.799	[redacted]	-	-	-	-
Oral azacitidine	[redacted]	3.864	[redacted]	[redacted]	1.06	[redacted]	49,704

Abbreviations: ICER = incremental cost-effectiveness ratio; LYG = life years gained; QALY = quality-adjusted life year; BSC = best supportive care

B.3.7 Sensitivity analyses

B.3.7.1 Probabilistic sensitivity analysis

Probabilistic sensitivity analysis (PSA) was conducted by sampling from a predefined distribution as outlined in Section B.3.5.5, for each model parameter 1,000 times in order to capture the uncertainly in costs and outcomes. The PSA showed a result consistent with the deterministic analysis with oral azacitidine generating more QALYs with a higher cost than the watch and wait strategy with BSC (Table B.3.33). The probabilistic ICER and deterministic ICER differ by -9.20%. Uncertainty can be seen around both costs and QALYs, but predominantly in QALYs (Figure B.3.37). The CEAC shows that the watch and wait strategy with BSC had the highest probability of being cost-effective until a willingness-to-pay threshold £46,000, after which oral azacitidine had the higher probability of being cost-effective. At a £50,000 WTP threshold, oral azacitidine has 60% probability of being cost-effective when compared to the watch and wait strategy with BSC.

Table B.3.33. Base case results with oral azacitidine PAS (Probabilistic)

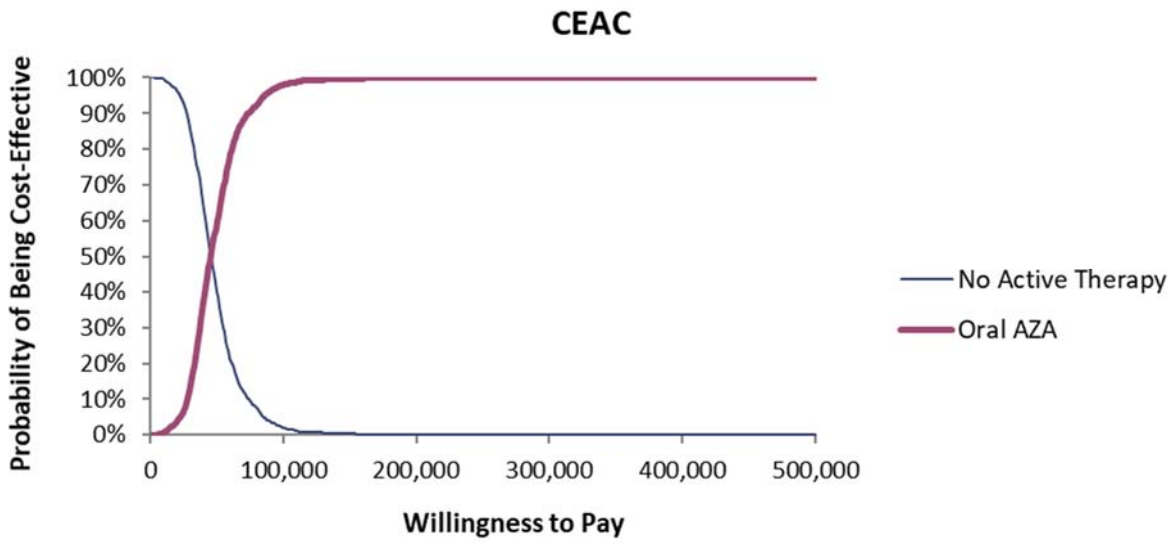
Technology	Total costs (£)	Total LYG	Total QALYs	Incremental costs (£)	Incremental LYG	Incremental QALYs	ICER versus baseline (£/QALY)
Watch and wait plus BSC	[redacted]	2.817	[redacted]	-	-	-	-
Oral azacitidine	[redacted]	3.879	[redacted]	[redacted]	1.06	[redacted]	45,130

Abbreviations: ICER = incremental cost-effectiveness ratio; LYG = life years gained; QALY = quality-adjusted life year; BSC = best supportive care

Figure B.3.36. Scatter plot of incremental costs and QALYs on the cost-effectiveness plane – (PAS) price

Abbreviations: AZA = azacitidine, QALY = quality adjusted life year

Figure B.3.37. Cost-effectiveness acceptability curve – (PAS price)



Abbreviations: AZA = azacitidine; CEAC = cost effectiveness acceptability curve

B.3.7.2 Deterministic sensitivity analysis

To identify key model drivers that impact results, one-way sensitivity analysis (OWSA) was conducted by varying parameters which were expected to have the most uncertainty. The low and high values were based on 95% CI for parameters when available or varied by assuming a SE of 20% around the mean value. The results of the OWSA are presented in the form of tornado plots for incremental costs, QALYs and ICERs. The 10 most influential parameters are presented in the tornado plots. The following key inputs were varied in the OWSA:

- Baseline patient characteristics (i.e. weight & height)
- Treatment administration costs
- Parameters of the parametric curves fitted to OS and RFS
- RDI rate
- Health state utility values
- AE rates and AE disutility
- Disease management costs
- End of life costs

The tornado plots for costs, QALYs and the ICER for PAS price are presented in (Figure B.3.38 Figure B.3.39 and Figure B.3.40). The RDI followed by the treatment administration cost of chemotherapy management had the greatest impact on the incremental costs. The greatest driver of incremental QALYs was the health state utilities for RFS both on and off treatment, followed by the health state utility values for relapse. The most influential driver of the ICER was the health state utility value for RFS both on and off treatment, followed by the RDI rate. The RDI was varied by 20%, however, as this value is a proportion, in the DSA it was capped with an upper bound of 100% and the lower bound was adjusted accordingly to create a balanced range, resulting in a lower bound value of 79.2%.

Figure B.3.38. Tornado plot of deterministic sensitivity analysis: impact on incremental costs – (PAS) price

Abbreviations: AE = adverse event; AZA = azacitidine; incr. = incremental; RDI = relative dose intensity; SCT = stem stell transplant; RBC = red blood-cell count

Figure B.3.39. Tornado plot of deterministic sensitivity analysis: impact on incremental QALYs – (PAS) price

Abbreviations: AZA = azacitidine; QALYs = quality adjusted life years; RBC = red-blood cell count; RFS = relapse-free survival; SCT = stem cell transplant

Figure B.3.40. Tornado plot of deterministic sensitivity analysis: ICER – (PAS) price

Abbreviations: ICER = incremental cost effectiveness ratio; AZA = azacitidine; QALYs = quality adjusted life years; RBC = red-blood cell count; RFS = relapse-free survival; SCT = stem cell transplant

B.3.7.3 Scenario analysis

To test the robustness of the base case results to alternative structural and methodological assumptions, scenario analyses outlined in Table B.3.34 was conducted and the results are provided in Table B.3.35. The base case results were robust to alternative scenarios. Oral azacitidine generated higher QALYs at a higher cost than watch and wait with BSC under all alternative scenarios. Implementing different survival curves for both OS and RFS had minimal impact on the ICER. Discounting had the substantial impact on the ICER, similarly, using the Europe subgroup, resulted in a 20% lower ICER compared to the base case.

Table B.3.34. Descriptions of the scenario analyses conducted

Scenario	Description
Discount rate	The discount rate associated with costs and outcomes were varied between 0 and 6%
Time horizon	The time horizon was varied between 10, 15, 20 and 25 years
Data cut	Use of 2019 OS data cut
Vial sharing	Include
Survival model: extrapolation OS	Scenario 1: Cure modelling with a 5-year cure point
	Scenario 2: Hybrid model
	Joint log-normal model as this model had the second lowest AIC
AE disutility	AE disutility doubled
Population	EU population with 2019 data cut. Same survival curves as base case
Utility values	Utility values based on Joshi 2019 for all health states

Abbreviations: AE = adverse event; AIC = Akaike information criterion; KM = Kaplan-Meier; OS = overall survival; RFS = relapse-free survival;

Table B.3.35. Results from the scenario analyses - PAS price

	PAS price				
	Incremental costs	Incremental QALYs	Incremental LYs	ICER	% change in ICER
Base case	[redacted]	[redacted]	1.06	49,704	-
Discount rate scenarios					
Discount rate. Costs: 0%, QALYs: 0%	[redacted]	[redacted]	1.33	44,595	-11.46%
Discount rate. Costs: 6%, QALYs: 6%	[redacted]	[redacted]	0.93	53,053	+6.31%
Discount rate. Costs: 0%, QALYs: 6%	[redacted]	[redacted]	0.93	60,731	+18.16%
Discount rate. Costs: 6%, QALYs: 0%	[redacted]	[redacted]	1.33	38,958	-27.59%
Time horizon scenarios					
Time horizon: 10 years	[redacted]	[redacted]	0.82	57,605	+13.72%
Time horizon: 15 years	[redacted]	[redacted]	0.97	52,445	+5.23%
Time horizon: 20 years	[redacted]	[redacted]	1.04	50,490	+1.56%
Data cut					
2019 data cut for OS	[redacted]	[redacted]	1.01	50,287	+1.16%
Vial sharing					
Include	[redacted]	[redacted]	1.06	49,786	+0.16%
Survival model : Extrapolation OS					
Cure model	[redacted]	[redacted]	1.49	45,397	-9.49%
Hybrid model	[redacted]	[redacted]	1.00	50,589	+1.75%
Survival model: Extrapolation RFS					
Joint log-normal model	[redacted]	[redacted]	1.06	50,235	+1.06%
Adverse events scenarios					
AE rates doubled	[redacted]	[redacted]	1.06	49,925	+0.44%
Population					
Europe only	[redacted]	[redacted]	1.36	41,320	-20.29%
Utility					
Joshi 2019	[redacted]	[redacted]	1.06	45,271	-8.91%

Abbreviations: AE = adverse event; ICER = incremental cost effectiveness ratio; LY = life year; OS = overall survival; QALY = quality adjusted life year; RFS = relapse-free survival

B.3.7.4 Summary of sensitivity analyses results

The base case probabilistic ICER is closely aligned with the base case deterministic ICER for ITT analysis, indicating that the base results have a low amount of uncertainty. However, from the scatter plot, it is suggestive that there is some degree of uncertainty around QALYs. This is also the case in the OWSA as the health state utility values for RFS both on treatment and

off treatment had the largest impact on the ICER. The RDI has had a substantial impact on the ICER on the cost side, this is due to the RDI being defined as the ratio of dose intensity to planned dose intensity. Lowering the RDI value results in a lower ratio resulting in lower drug costs for oral azacitidine.

B.3.8 Subgroup analysis

The cost-effectiveness results for the FLT3 mutation subgroup, comparing oral azacitidine to midostaurin and watch and wait plus BSC, are presented below. The comparison between the clinical trial estimates and the estimates included in the cost-effectiveness model are presented in Appendix J with the disaggregated results for the FLT3 subgroup.

B.3.8.1 Deterministic results

The deterministic results from the subgroup analysis are provided in Table B.3.36. Results of the base-case analysis indicated that oral azacitidine was associated with an additional discounted 1.23 LYs, 1.18 additional discounted QALYs, and decreased discounted costs of [redacted] over a 30-year lifetime horizon. As such, midostaurin was strictly dominated by oral azacitidine. When comparing against watch and wait with BSC in this subgroup, oral azacitidine was associated with an additional discounted 2.10 LYs, 1.62 additional discounted QALYs, and increased discounted costs of [redacted] over a 30-year lifetime horizon. As such, treatment with oral azacitidine led to cost per QALY gained of £25,010.

Table B.3.36. Deterministic results with oral azacitidine - PAS price: subgroup FLT3 (discounted)

Technology	Total costs (£)	Total LYG	Total QALYs	Incremental costs (£)	Incremental LYG	Incremental QALYs	ICER versus baseline (£/QALY)	Pairwise ICER vs oral azacitidine
Watch and wait plus BSC	[redacted]	2.731	[redacted]	-	-	-	-	25,010
Oral azacitidine	[redacted]	4.828	[redacted]	[redacted]	2.10	[redacted]	25,010	-
Midostaurin	[redacted]	3.600	[redacted]	[redacted]	0.87	[redacted]	300,652	Oral azacitidine is dominant

Abbreviations: FLT3 = fms-like tyrosine kinase 3; ICER = incremental cost-effectiveness ratio; LYG = life years gained; QALY = quality-adjusted life year; BSC = best supportive care

B.3.8.2 Probabilistic sensitivity analysis

The PSA showed a result consistent with the deterministic analysis.

Table B.3.37. Probabilistic results with oral azacitidine - PAS price: subgroup FLT3 (discounted)

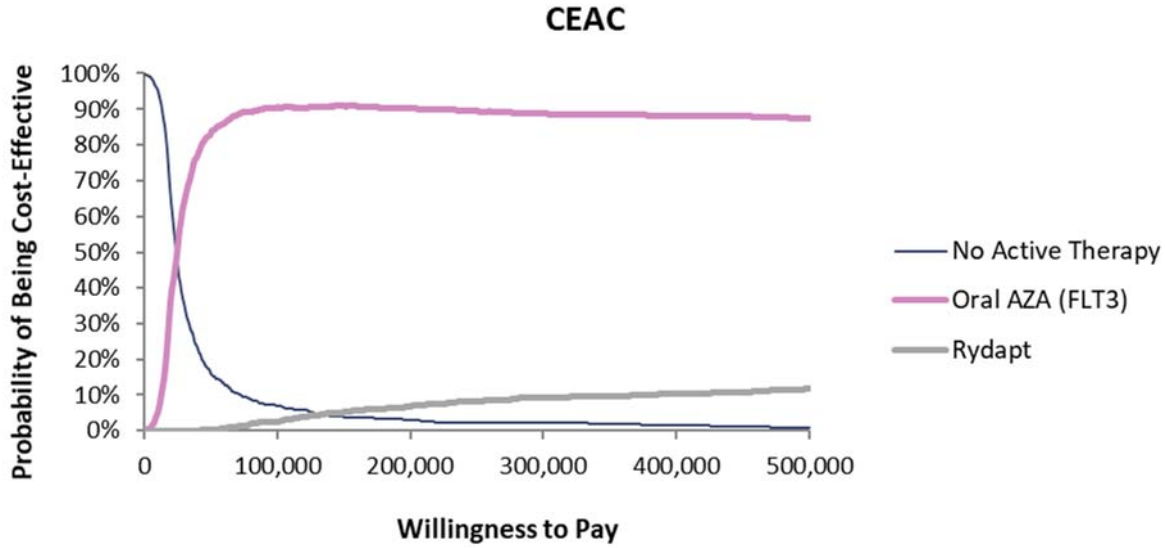
Technology	Total costs (£)	Total LYG	Total QALYs	Incremental costs (£)	Incremental LYG	Incremental QALYs	ICER versus baseline (£/QALY)	Pairwise ICER vs oral azacitidine
Watch and wait plus BSC	[redacted]	2.685	[redacted]	-	-	-	-	24,354
Oral azacitidine	[redacted]	4.757	[redacted]	[redacted]	2.07	[redacted]	24,354	-
Midostaurin	[redacted]	3.560	[redacted]	[redacted]	0.87	[redacted]	272,290	Oral azacitidine is dominant

Abbreviations: FLT3 = fms-like tyrosine kinase 3; ICER = incremental cost-effectiveness ratio; LYG = life years gained; QALY = quality-adjusted life year; BSC = best supportive care

Figure B.3.41. Scatter plot of incremental costs and QALYs on the cost-effectiveness plane FLT-3 subgroup - PAS price

Abbreviations: AZA = azacitidine; FLT3 = fms-like tyrosine kinase 3; QALYs = quality adjusted life years

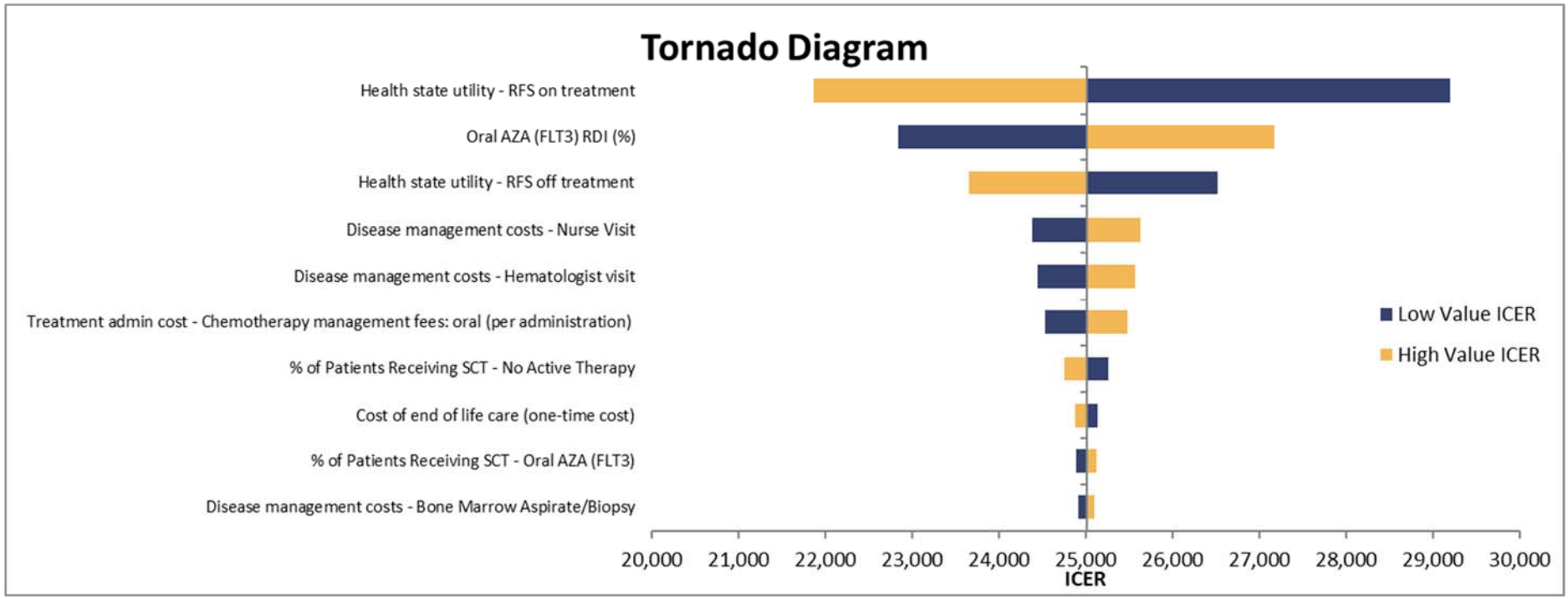
Figure B.3.42. Cost-effectiveness acceptability curves FLT-3 subgroup - PAS price



Abbreviations: CEAC = Cost-effectiveness acceptability curve; FLT3 = fms-like tyrosine kinase 3

B.3.8.3 Deterministic sensitivity analysis

Figure B.3.43. Tornado plot of deterministic sensitivity analysis: impact on ICER FLT-3 subgroup – PAS price



Abbreviations: AZA = azacitidine; FLT3 = fms-like tyrosine kinase 3; ICER = incremental cost effectiveness ratio; RFS = relapse-free survival; SCT = stem cell transpla

B.3.9 Validation

B.3.9.1 Validation of cost-effectiveness analysis

A multi-step approach was undertaken to ensure that the model was both mathematically (technical validation) and clinically valid (plausibility within UK clinical practice). Guidance was sought from two clinical experts to ensure clinical validity by discussing in detail the model structure, inputs and key assumptions

B.3.9.1.1 Internal Validation

A checklist was used to manage quality control across different elements of the health economic model. This included the execution of a number of stress tests on the model by testing the robustness of the model when using extreme values. The checks include but were not limited to those detailed in Table B.3.38.

Table B.3.38. Model validation checklist

Item check	Reason
If lifetime horizon is implemented, check that the overwhelming majority (>99%+) of patients are dead at the end of the model	To confirm that the lifetime horizon is sufficient enough to capture important differences in costs or outcomes between the technologies being compared
Set all utility values equal to 1 and set all disutilities to zero. QALYs should equal to LYs.	To confirm that QALYs are equal to LYs i.e. life-years are not adjusted by QoL
Set all utility/disutility values to zero	To confirm that zero QALYs are accumulated for all included treatments
Set all mortality rates (including background mortality) to zero.	To confirm that there are no deaths in model, and total LYs should equal time horizon of model.
Set all mortality rates (including background mortality) to 1	To confirm that all patients are dead in cycle 1, but still produce (some) expected costs and QALYs (due to half cycle, and one-off costs/disutilities)
If included, set all AE probabilities to zero	To confirm that no AEs occur, and that AE-related costs and disutilities are also estimated to be zero.
Set unit costs for all included treatments to zero	To confirm that estimated treatment costs are zero.
Set the discount rate of benefits and costs to 0%	To confirm that discounted benefits/costs match undiscounted results exactly.

Abbreviations: AE = adverse event; LY = life year; QALY = quality adjusted life year; QoL = quality of life

B.3.9.1.2 External Validation

B.3.9.1.2.1 Inclusion of comparators and clinical trials

The relevant comparators and the associated clinical trials identified from the SLR were verified based on a combination of:

- Published treatment guidelines for the management of AML including the ELN, published in 2017 which is the main guideline used in the UK[18], the British Committee for Standards in Haematology (BSCH) published in 2006[69] and the European Society for Medical Oncology (ESMO) guideline published in 2013[70]
- Guidance from two UK clinical experts with extensive knowledge of the treatments used in practice

B.3.9.1.2.2 Validation of long-term survival extrapolation

As described in Sections B.3.3.1 and B.3.3.2, the selection of survival models for the OS and RFS endpoints were based on fit statistics, AIC and BIC, assessment of hazards and visual inspection. To help guide the choice of distribution further, clinical experts' opinions on clinical plausibility of the extrapolated survival functions were used to inform the final selections.

B.3.10 Interpretation and conclusions of economic evidence

The model was developed to assess the cost effectiveness of oral azacitidine for the treatment of adults with AML who have complete disease remission, or complete remission with incomplete blood count recovery, following induction therapy with or without consolidation treatment who are not eligible for, including those who choose not to proceed to, haematopoietic stem cell transplantation. Comparators in the cost effectiveness analysis were established clinical management without oral azacitidine (a "watch and wait" strategy with best supportive care) and midostaurin. Midostaurin was assessed in a FLT3 mutation subgroup only.

The watch and wait with BSC comparison was based on the head-to-head comparison from QUAZAR AML-001. Results of the base-case analysis showed that oral azacitidine was associated with an additional 1.06 discounted LYs, [redacted] additional discounted QALYs, and an increase in discounted costs of [redacted] compared against watch and wait with BSC over a 30-year lifetime horizon. As such, treatment with oral azacitidine led to cost per QALY gained of £49,704. This shows that Oral azacitidine is cost effective at a willingness to pay per QALY threshold of £50,000.

The comparison with midostaurin was restricted to a FLT3 mutation population and was based on the results of subgroup analysis of the QUAZAR AML-001 trial and the maintenance sample from the RATIFY trial. Results of the base-case analysis indicated that oral azacitidine was associated with an additional discounted 1.23 LYs, [redacted] additional discounted QALYs, and decreased discounted costs of [redacted] over a 30-year lifetime horizon. As such, midostaurin was strictly dominated by oral azacitidine. When comparing against watch and wait with BSC in this subgroup, oral azacitidine was associated with an additional discounted 2.10 LYs, [redacted] additional discounted QALYs, and increased discounted costs of [redacted] over a 30-year lifetime horizon. As such, treatment with oral azacitidine led to cost per QALY gained of £25,010. This shows that Oral azacitidine is cost effective at a willingness to pay per QALY threshold of £50,000 also in the FLT-3 subgroup.

Extensive sensitivity and scenario analyses were conducted to test the robustness of the base-case results. The results of most of the sensitivity and scenario analyses were aligned with the results of the base-case analysis for all comparisons. The DSA indicated that the model outcomes were most sensitive to health state utility values followed by relative dose intensity.

A common limitation in lifetime models is the assumption that the defined survival functions accurately estimate the long-term survival of patients when only short-term clinical data are available. The KM curves from QUAZAR AML-001 that informed the OS and RFS extrapolations in the model were mature helping to mitigate some of this uncertainty. Furthermore, care was taken to select curves that balanced good statistical fit with clinical plausibility and extensive analyses extrapolations were conducted to assess any uncertainty in extrapolation. Since no trial data was collected after a patient had relapsed, the health state utility value for the relapse state was derived from the literature. Whilst consideration of multiple sources was made, there is uncertainty if this value would match those of the QUAZAR AML-001 trial. An attempt was made to reduce this source of uncertainty by combining the literature-based relapse disutility with the trial-based relapse free utility value. The FLT3 subgroup analysis is limited by differences in the QUAZAR AML-001 and RATIFY trials in terms of the trials in terms of study design, inclusion/exclusion criteria and baseline characteristics(Section B.2.9).

B.4 References

1. EMA. *Onureg: EPAR- Product information [online].* [cited 2021 29 November 2021]; Available from: https://www.ema.europa.eu/en/documents/product-information/onureg-epar-product-information_en.pdf.
2. Wei, A., et al., *The QUAZAR AML-001 Maintenance Trial: Results of a Phase III International, Randomized, Double-Blind, Placebo-Controlled Study of CC-486 in Patients with Acute Myeloid Leukemia (AML) in First Remission.* Blood, 2019. **134**(Supplement_2): p. LBA-3-LBA-3.
3. Scott, L.J., *Azacitidine: A Review in Myelodysplastic Syndromes and Acute Myeloid Leukaemia.* Drugs, 2016. **76**(8): p. 889-900.
4. Huls, G., et al., *Azacitidine maintenance after intensive chemotherapy improves DFS in older AML patients.* Blood, 2019. **133**(13): p. 1457-1464.
5. Laille, E., et al., *Pharmacokinetics and Pharmacodynamics with Extended Dosing of CC-486 in Patients with Hematologic Malignancies.* PLoS One, 2015. **10**(8): p. e0135520.
6. Garcia-Manero, G., et al., *Efficacy and safety of extended dosing schedules of CC-486 (oral azacitidine) in patients with lower-risk myelodysplastic syndromes.* Leukemia, 2016. **30**(4): p. 889-96.
7. Roman, E., et al., *Myeloid malignancies in the real-world: Occurrence, progression and survival in the UK's population-based Haematological Malignancy Research Network 2004-15.* Cancer epidemiology, 2016. **42**: p. 186-198.
8. *Department of Health and Social care. The UK Rare Diseases Framework [online].* [cited 2021 29 November 2021]; Available from: https://www.gov.uk/government/publications/uk-rare-diseases-framework.
9. De Kouchkovsky, I. and M. Abdul-Hay, *'Acute myeloid leukemia: a comprehensive review and 2016 update'.* Blood Cancer J, 2016. **6**(7): p. e441.
10. Vosberg, S. and P.A. Greif, *Clonal evolution of acute myeloid leukemia from diagnosis to relapse.* Genes Chromosomes Cancer, 2019. **58**(12): p. 839-849.
11. Shallis, R.M., et al., *Epidemiology of acute myeloid leukemia: Recent progress and enduring challenges.* Blood Rev, 2019. **36**: p. 70-87.
12. Song, X., et al., *Incidence, Survival, and Risk Factors for Adults with Acute Myeloid Leukemia Not Otherwise Specified and Acute Myeloid Leukemia with Recurrent Genetic Abnormalities: Analysis of the Surveillance, Epidemiology, and End Results (SEER) Database, 2001-2013.* Acta Haematol, 2018. **139**(2): p. 115-127.
13. Arber, D.A., et al., *The 2016 revision to the World Health Organization classification of myeloid neoplasms and acute leukemia.* Blood, 2016. **127**(20): p. 2391-2405.
14. CRUK. *Acute myeloid leukaemia (AML) Statistics [online].* [cited 2021 29 November 2021]; Available from: https://www.cancerresearchuk.org/health-professional/cancer-statistics/statistics-by-cancer-type/leukaemia-aml#heading-Zero. .
15. CRUK. *Acute myeloid leukaemia (AML) [online].* [cited 2021 29 November 2021]; Available from: https://www.cancerresearchuk.org/about-cancer/acute-myeloid-leukaemia-aml.
16. Breems, D.A., et al., *Prognostic index for adult patients with acute myeloid leukemia in first relapse.* J Clin Oncol, 2005. **23**(9): p. 1969-78.
17. Rollig, C., et al., *Long-term prognosis of acute myeloid leukemia according to the new genetic risk classification of the European LeukemiaNet recommendations: evaluation of the proposed reporting system.* J Clin Oncol, 2011. **29**(20): p. 2758-65.
18. Dohner, H., et al., *Diagnosis and management of AML in adults: 2017 ELN recommendations from an international expert panel.* Blood, 2017. **129**(4): p. 424-447.

19. Döhner, H., et al., *Cytogenetics and gene mutations influence survival in older patients with acute myeloid leukemia treated with azacitidine or conventional care.* Leukemia, 2018. **32**(12): p. 2546-2557.
20. *Public Health England. National Cancer Intelligence Network Trends in incidence and outcome for haematological cancers in England: 2001-2010 [online]*. [cited 2021 29 November 2021]; Available from: http://www.ncin.org.uk/view?rid=2818.
21. Medinger, M., C. Lengerke, and J. Passweg, *Novel therapeutic options in Acute Myeloid Leukemia.* Leuk Res Rep, 2016. **6**: p. 39-49.
22. Rautenberg, C., et al., *Relapse of Acute Myeloid Leukemia after Allogeneic Stem Cell Transplantation: Prevention, Detection, and Treatment.* Int J Mol Sci, 2019. **20**(1).
23. Bertoli, S., et al., *Improved outcome for AML patients over the years 2000-2014.* Blood Cancer J, 2017. **7**(12): p. 635.
24. Juliusson, G., et al., *Hematopoietic stem cell transplantation rates and long-term survival in acute myeloid and lymphoblastic leukemia: real-world population-based data from the Swedish Acute Leukemia Registry 1997-2006.* Cancer, 2011. **117**(18): p. 4238-46.
25. Ustun, C., H.M. Lazarus, and D. Weisdorf, *To transplant or not: a dilemma for treatment of elderly AML patients in the twenty-first century.* Bone Marrow Transplant, 2013. **48**(12): p. 1497-505.
26. Loke, J., et al., *The role of allogeneic stem cell transplantation in the management of acute myeloid leukaemia: a triumph of hope and experience.* Br J Haematol, 2020. **188**(1): p. 129-146.
27. Storey, S., T.F. Gray, and A.L. Bryant, *Comorbidity, Physical Function, and Quality of Life in Older Adults with Acute Myeloid Leukemia.* Curr Geriatr Rep, 2017. **6**(4): p. 247-254.
28. Aoki, J., et al., *Impact of age on outcomes of allogeneic hematopoietic stem cell transplantation with reduced intensity conditioning in elderly patients with acute myeloid leukemia.* 2016. **91**(3): p. 302-307.
29. Devillier, R., et al., *HLA-Matched Sibling versus Unrelated versus Haploidentical Related Donor Allogeneic Hematopoietic Stem Cell Transplantation for Patients Aged Over 60 Years with Acute Myeloid Leukemia: A Single-Center Donor Comparison.* Biol Blood Marrow Transplant, 2018. **24**(7): p. 1449-1454.
30. Medeiros, B.C., et al., *Optimizing survival outcomes with post-remission therapy in acute myeloid leukemia.* Am J Hematol, 2019. **94**(7): p. 803-811.
31. NICE. *Midostaurin for untreated acute myeloid leukaemia [TA523] [online]*. [cited 2021 29 November 2021]; Available from: https://www.nice.org.uk/guidance/ta523.
32. Patel, J.P., et al., *Prognostic relevance of integrated genetic profiling in acute myeloid leukemia.* N Engl J Med, 2012. **366**(12): p. 1079-89.
33. Dohner, H., D.J. Weisdorf, and C.D. Bloomfield, *Acute Myeloid Leukemia.* N Engl J Med, 2015. **373**(12): p. 1136-52.
34. *American Cancer Society. Key statistics for acute myeloid leukemia (AML) [online]*. [cited 2021 29 November 2021]; Available from: https://www.cancer.org/cancer/acute-myeloid-leukemia/about/key-statistics.html.
35. Sekeres, M.A. and M. Keng. *Acute myeloid leukemia. Cleveland Clinic Center for Continuing Education [online]*. 2014 [cited 2021 29 November 2021]; Available from: https://teachmemedicine.org/cleveland-clinic-acute-myeloid-leukemia/.
36. Ramos, N.R., et al., *Current Approaches in the Treatment of Relapsed and Refractory Acute Myeloid Leukemia.* J Clin Med, 2015. **4**(4): p. 665-95.
37. Kantarjian, H., *Acute myeloid leukemia--major progress over four decades and glimpses into the future.* Am J Hematol, 2016. **91**(1): p. 131-45.

38. Papaemmanuil, E., et al., *Genomic Classification and Prognosis in Acute Myeloid Leukemia.* N Engl J Med, 2016. **374**(23): p. 2209-2221.
39. NCCN. *NCCN Guidelines: Acute Myeloid Leukemia*. [cited 2021 29 November 2021]; Available from: https://www.nccn.org/patients/guidelines/content/PDF/aml-patient.pdf.
40. O'Donnell, M.R., et al., *NCCN Clinical Practice Guidelines Acute myeloid leukemia.* J Natl Compr Canc Netw, 2012. **10**(8): p. 984-1021.
41. Ravandi, F., R.B. Walter, and S.D. Freeman, *Evaluating measurable residual disease in acute myeloid leukemia.* Blood Adv, 2018. **2**(11): p. 1356-1366.
42. Freeman, S.D., et al., *Prognostic relevance of treatment response measured by flow cytometric residual disease detection in older patients with acute myeloid leukemia.* J Clin Oncol, 2013. **31**(32): p. 4123-31.
43. Arber, D.A., et al., *Initial diagnostic workup of acute leukemia: guideline from the College of American Pathologists and the American Society of Hematology.* Arch Pathol Lab Med, 2017. **141**(10): p. 1342-1393.
44. Schuurhuis, G.J., et al., *Minimal/measurable residual disease in AML: a consensus document from the European LeukemiaNet MRD Working Party.* Blood, 2018. **131**(12): p. 1275-1291.
45. Lerch, E., et al., *Prognosis of acute myeloid leukemia in the general population: data from southern Switzerland.* Tumori, 2009. **95**(3): p. 303-10.
46. Chen, Y., et al., *The outcome and prognostic factors of 248 elderly patients with acute myeloid leukemia treated with standard-dose or low-intensity induction therapy.* Medicine (Baltimore), 2016. **95**(30): p. e4182.
47. Rowe, J.M., et al., *Very Poor Survival of Patients with AML Who Relapse after Achieving a First Complete Remission: The Eastern Cooperative Oncology Group Experience.* Blood, 2005. **106**(11): p. 546-546.
48. Datoguia, T.S., et al., *Overall survival of Brazilian acute myeloid leukemia patients according to the European LeukemiaNet prognostic scoring system: a cross-sectional study.* Med Oncol, 2018. **35**(11): p. 141.
49. CRUK. *Acute myeloid leukaemia (AML) Mortality Statistics [online]*. [cited 2021 29 November 2021]; Available from: https://www.cancerresearchuk.org/health-professional/cancer-statistics/statistics-by-cancer-type/leukaemia-aml/mortality#heading-One.
50. Forsythe, A., et al., *Systematic review of health state utility values for acute myeloid leukemia.* Clinicoecon Outcomes Res, 2018. **10:**: p. 83-92.
51. Buckley, S.A., et al., *Quality of life from the perspective of the patient with acute myeloid leukemia.* Cancer, 2018. **124**(1): p. 145-152.
52. Moller, T., et al., *Health related quality of life and impact of infectious comorbidity in outpatient management of patients with acute leukemia.* Leuk Lymphoma, 2012. **53**(10): p. 1896-904.
53. Scott, N.W., et al. *EORTC QLQ-C30 Reference Values [online]*. 2008 [cited 2021 29 November 2021]; Available from: https://www.eortc.org/app/uploads/sites/2/2018/02/reference_values_manual2008.pdf.
54. Alibhai, S.M., et al., *The impact of acute myeloid leukemia and its treatment on quality of life and functional status in older adults.* Crit Rev Oncol Hematol, 2007. **64**(1): p. 19-30.
55. Panju, A.H., et al., *Associations between quality of life, fatigue, and cytokine levels in patients aged 50+ with acute myeloid leukemia.* Support Care Cancer, 2009. **17**(5): p. 539-46.
56. Oliva, E.N., et al., *Quality of life in elderly patients with acute myeloid leukemia: patients may be more accurate than physicians.* Haematologica, 2011. **96**(5): p. 696-702.

57. Mohamedali, H., et al., *Older age is associated with similar quality of life and physical function compared to younger age during intensive chemotherapy for acute myeloid leukemia.* Leuk Res, 2012. **36**(10): p. 1241-8.
58. Leak Bryant, A., et al., *Patient-reported symptoms and quality of life in adults with acute leukemia: a systematic review.* Oncol Nurs Forum, 2015. **42**(2): p. E91-E101.
59. Cheng, M.J., et al., *A Single Center Survey of Health-Related Quality of Life among Acute Myeloid Leukemia Survivors in First Complete Remission.* J Palliat Med, 2017. **20**(11): p. 1267-1273.
60. Leunis, A., et al., *Impaired health-related quality of life in acute myeloid leukemia survivors: a single-center study.* Eur J Haematol, 2014. **93**(3): p. 198-206.
61. Mamolo, C.M., et al., *Cross-Sectional Survey of Symptoms and Health-Related Quality of Life of Adults with De Novo Acute Myeloid Leukemia (AML) in Clinical Practice.* 2017. **130**(Suppl 1): p. 5660-5660.
62. Castejon, N., et al., *Social preferences for health states associated with acute myeloid leukemia for patients undergoing treatment in the United Kingdom.* Health Qual Life Outcomes, 2018. **16**(1): p. 66.
63. Kayastha, N., et al., *The impact of remission status on patients' experiences with acute myeloid leukemia (AML): an exploratory analysis of longitudinal patient-reported outcomes data.* Support Care Cancer, 2018. **26**(5): p. 1437-1445.
64. Bosshard, R., et al., *Systematic reviews of economic burden and health-related quality of life in patients with acute myeloid leukemia.* Cancer Treat Rev, 2018. **69**: p. 224-232.
65. Zeidan, A.M., et al., *Economic burden associated with acute myeloid leukemia treatment.* Expert Rev Hematol, 2016. **9**(1): p. 79-89.
66. Wiese, M. and N. Daver, *Unmet clinical needs and economic burden of disease in the treatment landscape of acute myeloid leukemia.* Am J Manag Care, 2018. **24**(16 Suppl): p. S347-s355.
67. Aly, A., et al., *Economic burden of relapsed/refractory (R/R) acute myeloid leukemia (AML) in the US.* Blood, 2017. **130**(Suppl 1): p. 3386.
68. Kwon, C., et al., *Treatment patterns and health care resources use (HCRU) in patients with acute myeloid leukemia (AML): real world evidence (RWE) from 30 US institutions.* Blood, 2017. **130**(Suppl 1): p. 5655.
69. Milligan, D.W., et al., *Guidelines on the management of acute myeloid leukaemia in adults.* Br J Haematol, 2006. **135**(4): p. 450-74.
70. Fey, M.F. and C. Buske, *Acute myeloblastic leukaemias in adult patients: ESMO Clinical Practice Guidelines for diagnosis, treatment and follow-up.* Ann Oncol, 2013. **24 Suppl 6**: p. vi138-43.
71. Kurosawa, S., et al., *Prognostic factors and outcomes of adult patients with acute myeloid leukemia after first relapse.* Haematologica, 2010. **95**(11): p. 1857-64.
72. Röllig, C., et al., *Intermediate-dose cytarabine plus mitoxantrone versus standard-dose cytarabine plus daunorubicin for acute myeloid leukemia in elderly patients.* Annals of Oncology, 2018. **29**(4): p. 973-978.
73. NICE. *Gemtuzumab ozogamicin for untreated acute myeloid leukaemia [TA545] [online]*. [cited 2021 29 November 2021]; Available from: https://www.nice.org.uk/guidance/ta545.
74. Besse, K., et al., *On Modeling Human Leukocyte Antigen-Identical Sibling Match Probability for Allogeneic Hematopoietic Cell Transplantation: Estimating the Need for an Unrelated Donor Source.* Biol Blood Marrow Transplant, 2016. **22**(3): p. 410-7.
75. Estey, E., et al., *Prospective feasibility analysis of reduced-intensity conditioning (RIC) regimens for hematopoietic stem cell transplantation (HSCT) in elderly patients with acute myeloid leukemia (AML) and high-risk myelodysplastic syndrome (MDS).* Blood, 2006. **109**(4): p. 1395-1400.

76. Thomas, X., et al., *Autologous stem cell transplantation after complete remission and first consolidation in acute myeloid leukemia patients aged 61–70 years: results of the prospective EORTC–GIMEMA AML–13 study.* 2007. **92**(3): p. 389-396.
77. Linker, C.A., et al., *Auto-SCT for AML in second remission: CALGB Study 9620.* Bone Marrow Transplantation, 2009. **44**(6): p. 353-359.
78. Sakamaki, H., et al., *Allogeneic stem cell transplantation versus chemotherapy as post-remission therapy for intermediate or poor risk adult acute myeloid leukemia: results of the JALSG AML97 study.* International Journal of Hematology, 2010. **91**(2): p. 284-292.
79. Rashidi, A., et al., *Maintenance therapy in acute myeloid leukemia: an evidence-based review of randomized trials.* Blood, 2016. **128**(6): p. 763-73.
80. Oliva, E.N., et al., *Azacitidine for Post-Remission Therapy in Elderly Patients with Acute Myeloid Leukemia : Final Results of the Qoless AZA-Amle Randomized Trial.* Blood, 2019. **134**(Supplement_1): p. 117-117.
81. Robles, C., et al., *Low-dose cytarabine maintenance therapy vs observation after remission induction in advanced acute myeloid leukemia: an Eastern Cooperative Oncology Group Trial (E5483).* Leukemia, 2000. **14**(8): p. 1349-53.
82. Stone, R.M., et al., *Midostaurin plus Chemotherapy for Acute Myeloid Leukemia with a FLT3 Mutation.* N Engl J Med, 2017. **377**(5): p. 454-464.
83. Larson, R.A., et al., *Midostaurin reduces relapse in FLT3-mutant acute myeloid leukemia: the Alliance CALGB 10603/RATIFY trial.* Leukemia, 2021. **35**(9): p. 2539-2551.
84. Wei, A.H., et al., *Oral Azacitidine Maintenance Therapy for Acute Myeloid Leukemia in First Remission.* N Engl J Med, 2020. **383**(26): p. 2526-2537.
85. Cutts, K., et al. *Patient Preferences for Maintenance Treatment of Acute Myeloid Leukemia: A discrete choice experiment subanalysis of patients in Germany and Italy [online]*. 2021 [cited 2021 29 November 2021]; Available from: https://library.ehaweb.org/eha/2021/eha2021-virtual-congress/324081/katelyn.cutts.patient.preferences.for.maintenance.therapy.of.acute.myeloid.html?f=listing%3D3%2Abrowseby%3D8%2Asortby%3D1%2Amedia%3D1.
86. Oliva, E.N., et al., *CC-486 Reduces Hospitalization and Associated Estimated Costs in Patients with Acute Myeloid Leukemia (AML) in First Remission after Intensive Chemotherapy: Results from the QUAZAR AML-001 Maintenance Trial.* Blood, 2020. **136**(Supplement 1): p. 14-15.
87. BMS, *Data on file: CC-486-AML-001 (QUAZAR) Clinical Study Report.* 2020.
88. Wei, A.H., et al., *Supplementary Appendix to: Wei AH et al. Oral Azacitidine Maintenance Therapy for Acute Myeloid Leukemia in First Remission. N Engl J Med. 2020; 383(26): 2526-2537.*
89. Roboz, G.J., et al. *692 CC-486 Prolongs Survival for Patients with Acute Myeloid Leukemia (AML) in Remission after Intensive Chemotherapy (IC) Independent of the Presence of Measurable Residual Disease (MRD) at Study Entry: Results from the QUAZAR AML-001 Maintenance Trial [online]*. 2020 [cited 2021 29 November 2021]; Available from: https://ash.confex.com/ash/2020/webprogram/Paper140837.html.
90. Cella, D., et al., *Combining anchor and distribution-based methods to derive minimal clinically important differences on the Functional Assessment of Cancer Therapy (FACT) anemia and fatigue scales.* J Pain Symptom Manage, 2002. **24**(6): p. 547-61.
91. Kvam, A.K., P.M. Fayers, and F. Wisloff, *Responsiveness and minimal important score differences in quality-of-life questionnaires: a comparison of the EORTC QLQ-C30 cancer-specific questionnaire to the generic utility questionnaires EQ-5D and 15D in patients with multiple myeloma.* Eur J Haematol, 2011. **87**(4): p. 330-7.
92. FDA. *NDA Multidisciplinary Review and Evaluation*. [cited 2021 29 November 2021]; Available from:

https://www.accessdata.fda.gov/drugsatfda_docs/nda/2020/214120Orig1s000MultidisciplineR.pdf.
93. ClinicalTrials.gov. *Efficacy of Oral Azacitidine Plus Best Supportive Care as Maintenance Therapy in Subjects With Acute Myeloid Leukemia (AML) in Complete Remission (QUAZAR AML-001) [online]*. [cited 2021 29 November 2021]; Available from: https://clinicaltrials.gov/ct2/show/results/NCT01757535?term=QUAZAR+AML&draw=2&rank=1&view=results.
94. Ravandi, F., et al., *Management of adverse events in patients with acute myeloid leukemia in remission receiving oral azacitidine: experience from the phase 3 randomized QUAZAR AML-001 trial.* J Hematol Oncol, 2021. **14**(1): p. 133.
95. EMA. *CHMP Assessment Report: Oral Azacitidine*. [cited 2021 29 November 2021]; Available from: https://www.ema.europa.eu/en/documents/assessment-report/onureg-epar-public-assessment-report_en.pdf.
96. BMS, *Data on file: CC-486-AML-001 (QUAZAR) Statistical Analysis Plan.* 2020.
97. Wei, A., et al. *871 Long-Term Overall Survival (OS) with Oral Azacitidine (Oral-AZA) in Patients with Acute Myeloid Leukemia (AML) in First Remission after Intensive Chemotherapy (IC): Updated Results from the Phase 3 QUAZAR AML-001 Trial*. 2021 [cited 2021 29 November 2021]; Available from: https://ash.confex.com/ash/2021/webprogram/Paper147501.html.
98. BMS, *Data on file: CC-486-AML-001 (QUAZAR) OS data cut-off date 8 September 2020.* 2021.
99. Janssen, B. and A. Szende, *Population Norms for the EQ-5D*, in *Self-Reported Population Health: An International Perspective based on EQ-5D*, A. Szende, B. Janssen, and J. Cabases, Editors. 2014, Springer
Copyright 2014, The Author(s). Dordrecht (NL). p. 19-30.
100. Montan, I., et al., *General Population Norms for the Functional Assessment of Chronic Illness Therapy (FACIT)-Fatigue Scale.* Value Health, 2018. **21**(11): p. 1313-1321.
101. Roboz, G.J., et al., *Health-related quality of life (HRQoL) in the phase III QUAZAR-AML-001 trial of CC-486 as maintenance therapy for patients with acute myeloid leukemia (AML) in first remission following induction chemotherapy (IC).* Journal of Clinical Oncology, 2020. **38**(15_suppl): p. 7533-7533.
102. NHS. *Pan-London Haemato-Oncology Clinical Guidelines Acute Leukaemias and Myeloid Neoplasms Part 2: Acute Myeloid Leukaemia*. 2020 [cited 2021 29 November 2021]; Available from: https://rmpartners.nhs.uk/wp-content/uploads/2020/01/Pan-London-AML-Guidelines-Jan-2020.pdf.
103. Ravandi, F., et al. *CC-486 maintenance therapy is safe and well tolerated in patients aged ≥75 years with acute myeloid leukemia (AML) in first remission following induction chemotherapy: Results from QUAZAR AML-001*. 2020 [cited 2021 29 November]; Available from: https://library.ehaweb.org/eha/2020/eha25th/294468/farhad.ravandi.cc-486.maintenance.therapy.is.safe.and.well.tolerated.in.html.
104. Reid, J. and S. Bhurke. *QUAZAR AML-001: Update from EHA2021 on survival outcomes according to NPM1 and FLT3 gene mutational status at diagnosis*. 2021 [cited 2021 29 November 2021]; Available from: https://aml-hub.com/medical-information/quazar-aml-001-update-from-eha2021-on-survival-outcomes-according-to-npm1-and-flt3-gene-mutational-status-at-diagnosis.
105. Dohner, H., A. Wei, and G.J. Roboz. *Survival outcomes from the QUAZAR AML-001 trial with oral azacitidine for patients with acute myeloid leukemia in remission by NPM1 and FLT3 gene mutation status at diagnosis [online]*. 2021 [cited 2021 29 November 2021]; Available from: https://library.ehaweb.org/eha/2021/eha2021-virtual-congress/324539/hartmut.dhner.survival.outcomes.from.the.quazar.aml-

001.trial.with.oral.html?f=menu%3D6%2Abrowseby%3D8%2Asortby%3D2%2Amedia%3D3%2Ace_id%3D2035%2Alabel%3D21988%2Aot_id%3D25551%2Amarker%3D1285.
106. Dohner, H., et al. *804 Prognostic Impact of NPM1 and FLT3 Mutations at Diagnosis and Presence of Measurable Residual Disease (MRD) after Intensive Chemotherapy (IC) for Patients with Acute Myeloid Leukemia (AML) in Remission: Outcomes from the QUAZAR AML-001 Trial of Oral Azacitidine (Oral-AZA) Maintenance [online]*. 2021 [cited 2021 29 November]; Available from: https://ash.confex.com/ash/2021/webprogram/Paper147465.html.
107. Wei, A., et al. *1036 CC-486 Improves Overall Survival (OS) and Relapse-Free Survival (RFS) for Patients with Acute Myeloid Leukemia (AML) in First Remission after Intensive Chemotherapy (IC), Regardless of Amount of Consolidation Received: Results from the Phase III QUAZAR AML-001 Maintenance Trial [online]*. 2020 [cited 2021 29 November 2021]; Available from: https://ash.confex.com/ash/2020/webprogram/Paper138498.html.
108. Ishak, K.J., I. Proskorovsky, and A. Benedict, *Simulation and matching-based approaches for indirect comparison of treatments.* Pharmacoeconomics, 2015. **33**(6): p. 537-49.
109. Phillippo, D.M., et al., *Assessing the performance of population adjustment methods for anchored indirect comparisons: A simulation study.* Stat Med, 2020. **39**(30): p. 4885-4911.
110. Phillippo, D., et al. *NICE DSU Technical Support Document 18: Methods for Population-Adjusted Indirecr Comparisons in Submission to NICE. Appendix D: Worked example of MAIC and STC [online]*. [cited 2021 29 November 2021]; Available from: http://nicedsu.org.uk/wp-content/uploads/2017/05/Population-adjustment-TSD-FINAL.pdf.
111. Frohling, S., et al., *Cytogenetics and age are major determinants of outcome in intensively treated acute myeloid leukemia patients older than 60 years: results from AMLSG trial AML HD98-B.* Blood, 2006. **108**(10): p. 3280-8.
112. NICE. *Gilteritinib for treating relapsed or refractory acute myeloid leukaemia [TA642] [online]*. [cited 2021 29 November 2021]; Available from: https://www.nice.org.uk/guidance/ta642/evidence/final-appraisal-determination-committee-papers-pdf-8830117118.
113. Eleni, L.D., Z.C. Nicholas, and S. Alexandros, *Challenges in treating older patients with acute myeloid leukemia.* J Oncol, 2010. **2010**: p. 943823.
114. Tremblay, G., et al., *Cost-effectiveness analysis for midostaurin versus standard of care in acute myeloid leukemia in the United Kingdom.* Cost Effectiveness & Resource Allocation, 2018. **16**: p. 33.
115. Tremblay, G., et al., *Cost-effectiveness of midostaurin in the treatment of newly diagnosed FLT3-mutated acute myeloid leukemia in France.* European Journal of Health Economics, 2020. **22**: p. 22.
116. NICE. *Azacitidine for treating acute myeloid leukaemia with more than 30% bone marrow blasts [TA399] [online]*. [cited 2021 29 November 2021]; Available from: https://www.nice.org.uk/guidance/ta399/resources/azacitidine-for-treating-acute-myeloid-leukaemia-with-more-than-30-bone-marrow-blasts-pdf-82602918570949.
117. EMA. *Rydapt: EPAR - Product information [online]*. [cited 2021 29 November 2021]; Available from: https://www.ema.europa.eu/en/documents/product-information/rydapt-epar-product-information_en.pdf.
118. *Office for National Statistics. National life tables - life expectancy in the UK: 2017 to 2019 [online]*. [cited 2021 29 November 2021]; Available from: https://www.ons.gov.uk/peoplepopulationandcommunity/birthsdeathsandmarriages/lifeexpectancies/bulletins/nationallifetablesunitedkingdom/2017to2019.

119. Tremblay, G. and P.B. Haines, A., *A Criterion-based Approach for the Systematic and Transparent Extrapolation of Clinical Trial Survival Data.* 2015. **2**(2): p. 147-60.
120. Dolan, P., *Modeling valuations for EuroQol health states.* Med Care, 1997. **35**(11): p. 1095-108.
121. *EuroQol Research Foundation. EQ-5D-3L User Guide: basic information on how to use the EQ-5D-3L instrument, Version 6.0 2018 [online].* [cited 2021 29 November 2021]; Available from: https://euroqol.org/publications/user-guides/.
122. Stein, E.M., et al., *Assessing utility values for treatment-related health states of acute myeloid leukemia in the United States.* Health Qual Life Outcomes, 2018. **16**(1): p. 193.
123. Joshi, N., et al., *Health State Utilities for Acute Myeloid Leukaemia: A Time Trade-off Study.* Pharmacoeconomics, 2019. **37**(1): p. 85-92.
124. Batty, N., Y. Yin, and M. Wetzler, *Decitabine is more cost-effective than cytarabine and daunorubicin in elderly acute myeloid leukemia patients.* J Cancer Res Ther, 2014. **2:**: p. 68-73.
125. Pan, F., et al., *Economic analysis of decitabine versus best supportive care in the treatment of intermediate- and high-risk myelodysplastic syndromes from a US payer perspective.* Clin Ther, 2010. **32**(14): p. 2444-56.
126. Nafees, B., et al., *Health state utilities for non small cell lung cancer.* Health Qual Life Outcomes, 2008. **6**: p. 84.
127. Matza, L.S., et al., *Health state utilities associated with treatment options for acute myeloid leukemia (AML).* Journal of Medical Economics, 2019. **22**(6): p. 567-576.
128. Curtis, L.A.a.B., Amanda, *Unit Costs of Health and Social Care 2019. Unit Costs of Health and Social Care . PSSRU*. 2019.
129. NHS. *Drugs and pharmaceutical electronic market information tool (eMIT)*. [cited 2021 29 November 2021]; Available from: https://www.gov.uk/government/publications/drugs-and-pharmaceutical-electronic-market-information-emit.
130. NHS. *National Cost Collection 2019/20 Report [online]*. [cited 2021 29 November 2021]; Available from: https://www.england.nhs.uk/wp-content/uploads/2021/06/National-Cost-Collection-2019-20-Report-FINAL.pdf.
131. NICE. *British National Formulary (BNF) [online]*. [cited 2021 29 November 2021]; Available from: https://bnf.nice.org.uk/drug/.
132. Georghiou, T. and M. Bardsley. *Exploring the cost of care at the end of life. Nuffield Trust [online]*. 2014 [cited 2021 29 November 2021]; Available from: https://www.nuffieldtrust.org.uk/files/2017-01/end-of-life-care-web-final.pdf.

NATIONAL INSTITUTE FOR HEALTH AND CARE EXCELLENCE

Single technology appraisal

Oral azacitidine for maintenance treatment of acute myeloid leukaemia after induction therapy [ID3892]

Clarification questions

January 2022

File name	Version	Contains confidential information	Date
ID3892 Oral azacitidine clarification questions_AIC & CIC redacted 28thJan2022	Final	Yes	28th January 2022

Notes for company

Highlighting in the template

Square brackets and grey highlighting are used in this template to indicate text that should be replaced with your own text or deleted. These are set up as form fields, so to replace the prompt text in [grey highlighting] with your own text, click anywhere within the highlighted text and type. Your text will overwrite the highlighted section.

To delete grey highlighted text, click anywhere within the text and press DELETE.

Section A: Clarification on effectiveness data

1. Literature searches

A1. The search methods in D.1.1.1. of Appendix D report that three joint searches of Medline/Embase/Cochrane were undertaken (these include the original 18.01.20, and two updates 19.02.21 and 11.06.21). The search strategy provided in appendix D Table B.5.1. appears to be for the June update only. This also appears to be the case for the other sections (Appendix G Table B.5.21 & Appendix H Table B.5.28.). Please provide copies of the strategies and results for all dates.

The following section presents the requested search strategies for each of the systematic reviews. For orientation, a summary is provided below in Table 1:

Table 1. Summary of Requested Search Strategies

Systematic Literature Reviews	Search Strategies
Studies of Clinical Evidence	• January 18, 2020 o Table D.1 • February 19, 2021 o Table D.2
Cost-Effectiveness Studies	• February 12, 2020 o Tables G.1, G.2, and G.3
Health-related Quality of Life Studies	• February 12, 2020

	o Table H.1

Appendix D. Identification, selection and synthesis of clinical evidence:

Table D.1. Search strategy, January 18, 2020

#	Searches	Results
1	exp Leukemia, Myeloid, Acute/	93638
2	(acute adj2 myelo* adj2 leu#?emi*).tw,kf.	118115
3	(acute adj2 (nonlympho* or non-lympho*) adj2 leu#?emi*).tw,kf.	6852
4	(acute adj2 granulocytic adj2 leu#?emi*).tw,kf.	338
5	((AML or ANLL) and (leu#?emi* or myelo* or nonlympho* or non-lympho*)).tw,kf. (88638)	88638
6	(AML or ANLL).tw,kf. and exp Leukemia, Myeloid/	43459
7	(acute adj2 basophilic adj2 leu#?emi*).tw,kf.	135
8	(acute adj2 eosinophilic adj2 leu#?emi*).tw,kf.	70
9	(acute adj2 erythroblastic adj2 leu#?emi*).tw,kf.	50
10	(erythroleu#?emi* or erythro-leu#?emi* or erythr?emic myelosis or diguglielmo* or di guglielmo*).tw,kf.	13447
11	((mast-cell* or mastcell*) adj2 leu#?emia*).tw,kf.	965
12	((megakaryocytic or mega-karyocytic) adj2 leu#?emi*).tw,kf.	858
13	((megakaryoblastic or mega-karyoblastic) adj2 leu#?emi*).tw,kf.	2279
14	(acute adj2 monoblastic adj2 leu#?emi*).tw,kf.	801
15	(acute adj2 monocytic adj2 leu#?emi*).tw,kf.	2524
16	(acute adj2 promyelocytic adj2 leu#?emi*).tw,kf.	16842
17	(acute adj2 progranulocytic adj2 leu#?emi*).tw,kf.	32
18	(Schilling-Type adj1 myelo* adj2 leu#?emi*).tw,kf.	0
19	or/1-18	199501
20	exp Animals/ not (exp Animals/ and Humans/)	16978023
21	19 not 20 [ANIMAL-ONLY REMOVED]	146307
22	(comment or editorial or news or newspaper article).pt.	2024048
23	(letter not (letter and randomized controlled trial)).pt.	2159043
24	21 not (22 or 23) [OPINION PIECES REMOVED]	138024
25	systematic review.pt.	128535
26	exp systematic reviews as topic/	27417
27	meta analysis.pt.	110543
28	exp meta-analysis as topic/	60216
29	(meta-analy* or metanaly* or metaanaly* or met analy* or integrative research or integrative review* or integrative overview* or research integration or research overview* or collaborative review*).tw,kf.	432280
30	(systematic review* or systematic overview* or evidence-based review* or evidence-based overview* or (evidence adj3 (review* or overview*)) or meta-	534417

#	Searches	Results
	review* or meta-overview* or meta-synthes* or rapid review* or "review of reviews" or umbrella review? or technology assessment* or HTA or HTAs).tw,kf.	
31	exp Technology assessment, biomedical/	25447
32	(cochrane or health technology assessment or evidence report or systematic reviews).jw.	58467
33	(network adj (MA or MAs)).tw,kf.	29
34	(NMA or NMAs or MTC or MTCs or MAIC or MAICs).tw,kf.	17858
35	indirect* compar*.tw,kf.	6681
36	(indirect treatment* adj1 compar*).tw,kf.	979
37	(mixed treatment* adj1 compar*).tw,kf.	1550
38	(multiple treatment* adj1 compar*).tw,kf.	438
39	(multi-treatment* adj1 compar*).tw,kf.	11
40	simultaneous* compar*.tw,kf.	2400
41	mixed comparison?.tw,kf.	132
42	or/25-41	896206
43	24 and 42 [REVIEWS]	1238
44	limit 43 to yr="2015-current" [REVIEWS - 5 YEARS]	747
45	(controlled clinical trial or randomized controlled trial or pragmatic clinical trial or equivalence trial).pt.	1160850
46	clinical trials as topic/	300899
47	exp Randomized Controlled Trials as Topic/	313900
48	(randomi#ed or randomi#ation? or randomly or RCT or placebo*).tw,kf.	3288971
49	((singl* or doubl* or trebl* or tripl*) adj (mask* or blind* or dumm*)).tw,kf.	671217
50	trial.ti.	799970
51	or/45-50	4170516
52	24 and 51 [RCTS]	10909
53	limit 52 to yr="2005-current" [Limit not valid in DARE; records were retained]	6560
54	44 or 53 [REVIEWS, TRIALS]	7044
55	54 use ppez [MEDLINE RECORDS]	2241
56	exp acute myeloid leukemia/	93638
57	acute leukemia/ and myeloid leukemia/	496
58	(acute adj2 myelo* adj2 leu#?emi*).tw,kw.	119019
59	(acute adj2 (nonlympho* or non-lympho*) adj2 leu#?emi*).tw,kw.	6881
60	(acute adj2 granulocytic adj2 leu#?emi*).tw,kw.	605
61	((AML or ANLL) and (leu#?emi* or myelo* or nonlympho* or non-lympho*)).tw,kw.	89459
62	(AML or ANLL).tw,kw. and exp myeloid leukemia/	43692
63	(acute adj2 basophilic adj2 leu#?emi*).tw,kw.	135
64	(acute adj2 eosinophilic adj2 leu#?emi*).tw,kw.	70

#	Searches	Results
65	(acute adj2 erythroblastic adj2 leu#?emi*).tw,kw.	49
66	(erythroleu#?emi* or erythro-leu#?emi* or erythr?emic myelosis or diguglielmo* or di guglielmo*).tw,kw.	13505
67	((mast-cell* or mastcell*) adj2 leu#?emi*).tw,kw.	1127
68	((megakaryocytic or mega-karyocytic) adj2 leu#?emi*).tw,kw.	872
69	((megakaryoblastic or mega-karyoblastic) adj2 leu#?emi*).tw,kw.	2296
70	(acute adj2 monoblastic adj2 leu#?emi*).tw,kw.	807
71	(acute adj2 monocytic adj2 leu#?emi*).tw,kw.	2532
72	(acute adj2 promyelocytic adj2 leu#?emi*).tw,kw.	16912
73	(acute adj2 progranulocytic adj2 leu#?emi*).tw,kw.	35
74	(Schilling-Type adj2 myelo* adj2 leu#?emi*).tw,kw.	0
75	or/56-74 [AML]	200805
76	exp animal experimentation/ or exp animal model/ or exp animal experiment/ or nonhuman/ or exp vertebrate/	49269157
77	exp human/ or exp human experimentation/ or exp human experiment/	39281090
78	76 not 77	9992874
79	75 not 78 [ANIMAL-ONLY REMOVED]	187753
80	editorial.pt.	1155948
81	letter.pt. not (letter.pt. and randomized controlled trial/)	2158990
82	79 not (80 or 81) [OPINION PIECES REMOVED]	178465
83	meta-analysis/	289403
84	"systematic review"/	351534
85	"meta analysis (topic)"/	41180
86	"systematic review (topic)"/	24415
87	(meta-analy* or metanaly* or metaanaly* or met analy* or integrative research or integrative review* or integrative overview* or research integration or research overview* or collaborative review*).tw,kw.	440872
88	(systematic review* or systematic overview* or evidence-based review* or evidence-based overview* or (evidence adj3 (review* or overview*)) or meta-review* or meta-overview* or meta-synthes* or "review of reviews" or umbrella review? or technology assessment* or HTA or HTAs).tw,kw.	541237
89	biomedical technology assessment/	24330
90	(cochrane or health technology assessment or evidence report).jw.	52052
91	or/83-90	953024
92	82 and 91 [REVIEWS]	1967
93	limit 92 to yr="2015-current" [REVIEWS - 5 YR LIMIT]	1173
94	randomized controlled trial/ or controlled clinical trial/	1360398
95	"clinical trial (topic)"/	107270
96	exp "controlled clinical trial (topic)"/	180248
97	(randomi#ed or randomi#ation or randomly or RCT or placebo*).tw,kw.	3343585

#	Searches	Results
98	((singl* or doubl* or trebl* or tripl*) adj (mask* or blind* or dumm*)).tw,kw.	695716
99	trial.ti.	799970
100	or/94-99	4155555
101	82 and 100 [RCTS, PHASE II/III TRIALS]	14365
102	limit 101 to yr="2005-current" [RCTs - 2005-current]	10356
103	93 or 102 [REVIEWS, TRIALS]	11108
104	conference abstract.pt.	3696512
105	103 not 104	7242
106	103 and 104	3866
107	limit 106 to yr="2018-current" [Limit not valid in DARE; records were retained]	544
108	105 or 107 [MOST RECENT 2 YRS CONFERENCE ABSTRACTS RETAINED]	7786
109	108 use oemezd [EMBASE RECORDS]	3609
110	exp Leukemia, Myeloid, Acute/	93638
111	(acute adj2 myelo* adj2 leu#?emi*).ti,ab,kw.	118957
112	(acute adj2 (nonlympho* or non-lympho*) adj2 leu#?emi*).ti,ab,kw.	6877
113	(acute adj2 granulocytic adj2 leu#?emi*).ti,ab,kw.	605
114	((AML or ANLL) and (leu#?emi* or myelo* or nonlympho* or non-lympho*)).ti,ab,kw.	89410
115	(AML or ANLL).ti,ab,kw. and exp Leukemia, Myeloid/	43690
116	(acute adj2basophilic adj2 leu#?emi*).ti,ab,kw.	0
117	(acute adj2 eosinophilic adj2 leu#?emi*).ti,ab,kw.	70
118	(acute adj2 erythroblastic adj2 leu#?emi*).ti,ab,kw.	49
119	(erythroleu#?emi* or erythro-leu#?emi* or erythr?emic myelosis or diguglielmo* or di guglielmo*).ti,ab,kw.	13505
120	((mast-cell* or mastcell*) adj2 leu#?emi*).ti,ab,kw.	1127
121	((megakaryocytic or mega-karyocytic) adj2 leu#?emi*).ti,ab,kw.	872
122	((megakaryoblastic or mega-karyoblastic) adj2 leu#?emi*).ti,ab,kw.	2296
123	(acute adj1 monoblastic adj2 leu#?emi*).ti,ab,kw.	789
124	(acute adj2 monocytic adj2 leu#?emi*).ti,ab,kw.	2532
125	(acute adj2 promyelocytic adj2 leu#?emi*).ti,ab,kw.	16902
126	(acute adj2 progranulocytic adj2 leu#?emi*).ti,ab,kw.	35
127	(Schilling-Type adj2 myelo* adj2 leu#?emi*).ti,ab,kw.	0
128	or/110-127	200478
129	conference abstract.pt.	3696512
130	128 not 129	168153
131	128 and 129	32325
132	limit 131 to yr="2018-current" [Limit not valid in DARE; records were retained]	5373
133	130 or 132	173526

#	Searches	Results
134	limit 133 to yr="2005-current" [Limit not valid in DARE; records were retained]	91938
135	134 use cctr [TRIALS, 2005-CURRENT]	3379
136	limit 128 to yr="2015-current" [Limit not valid in DARE; records were retained]	54087
137	136 use coch [REVIEWS, 2015-CURRENT]	8
138	135 or 137 [REVIEWS, TRIALS]	3387
139	138 use coch,cctr [COCHRANE DSR, CENTRAL RECORDS]	3387
140	exp Leukemia, Myeloid, Acute/	93638
141	(acute adj2 myelo* adj2 leu#?emi*).tw.	117669
142	(acute adj2 (nonlympho* or non-lympho*) adj2 leu#?emi*).tw.	6850
143	(acute adj2 granulocytic adj2 leu#?emi*).tw.	338
144	((AML or ANLL) and (leu#?emi* or myelo* or nonlympho* or non-lympho*)).tw.	88238
145	(AML or ANLL).tw. and exp Leukemia, Myeloid/	43328
146	(acute adj2 basophilic adj2 leu#?emi*).tw.	135
147	(acute adj2 eosinophilic adj2 leu#?emi*).tw.	70
148	(acute adj2 erythroblastic adj2 leu#?emi*).tw.	49
149	(erythroleu#?emi* or erythro-leu#?emi* or erythr?emic myelosis or diguglielmo* or di guglielmo*).tw.	13408
150	((mast-cell* or mastcell*) adj2 leu#?emi*).tw.	1090
151	((megakaryocytic or mega-karyocytic) adj2 leu#?emi*).tw.	852
152	((megakaryoblastic or mega-karyoblastic) adj2 leu#?emi*).tw.	2265
153	(acute adj2 monoblastic adj2 leu#?emi*).tw.	801
154	(acute adj2 monocytic adj2 leu#?emi*).tw.	2511
155	(acute adj2 promyelocytic adj2 leu#?emi*).tw.	16800
156	(Schilling-Type adj2 myelo* adj2 leu#?emi*).tw.	0
157	(acute adj2 progranulocytic adj2 leu#?emi*).tw.	32
158	or/140-157	199215
159	limit 158 to yr="2015-current" [Limit not valid in DARE; records were retained]	53736
160	159 use dare,clhta [DARE, HTA RECORDS]	85
161	55 or 109 or 139 or 160 [ALL DATABASES]	9322
162	limit 161 to yr="2015-current" [Limit not valid in DARE; records were retained]	5133
163	remove duplicates from 162	3835
164	161 not 162	4189
165	remove duplicates from 164	2866
166	163 or 165 [TOTAL UNIQUE RECORDS]	6701
167	166 use ppez [MEDLINE UNIQUE RECORDS]	2231
168	166 use oemezd [EMBASE UNIQUE RECORDS]	2029
169	166 use coch [DSR UNIQUE RECORDS]	8

#	Searches	Results
170	166 use dare [DARE UNIQUE RECORDS]	50
171	166 use clhta [HTA UNIQUE RECORDS]	35
172	166 use cctr [CENTRAL RECORDS]	2348

Table D.2. Search strategy, February 19, 2021

#	Searches	Results
1	exp Leukemia, Myeloid, Acute/	102975
2	(acute adj2 myelo* adj2 leu#?emi*).tw,kf.	127500
3	(acute adj2 (nonlympho* or non-lympho*) adj2 leu#?emi*).tw,kf.	6892
4	(acute adj2 granulocytic adj2 leu#?emi*).tw,kf.	343
5	((AML or ANLL) and (leu#?emi* or myelo* or nonlympho* or non-lympho*)).tw,kf.	96109
6	(AML or ANLL).tw,kf. and exp Leukemia, Myeloid/	48726
7	(acute adj2 basophilic adj2 leu#?emi*).tw,kf.	138
8	(acute adj2 eosinophilic adj2 leu#?emi*).tw,kf.	71
9	(acute adj2 erythroblastic adj2 leu#?emi*).tw,kf.	50
10	(erythroleu#?emi* or erythro-leu#?emi* or erythr?emic myelosis or diguglielmo* or di guglielmo*).tw,kf.	13634
11	((mast-cell* or mastcell*) adj2 leu#?emia*).tw,kf.	994
12	((megakaryocytic or mega-karyocytic) adj2 leu#?emi*).tw,kf.	878
13	((megakaryoblastic or mega-karyoblastic) adj2 leu#?emi*).tw,kf.	2367
14	(acute adj2 monoblastic adj2 leu#?emi*).tw,kf.	819
15	(acute adj2 monocytic adj2 leu#?emi*).tw,kf.	2666
16	(acute adj2 promyelocytic adj2 leu#?emi*).tw,kf.	17764
17	(acute adj2 progranulocytic adj2 leu#?emi*).tw,kf.	32
18	(Schilling-Type adj1 myelo* adj2 leu#?emi*).tw,kf.	0
19	or/1-18	212945
20	exp Animals/ not (exp Animals/ and Humans/)	18178116
21	19 not 20 [ANIMAL-ONLY REMOVED]	156984
22	(comment or editorial or news or newspaper article).pt.	2177594
23	(letter not (letter and randomized controlled trial)).pt.	2303912
24	21 not (22 or 23) [OPINION PIECES REMOVED]	147846
25	systematic review.pt.	154476
26	exp systematic reviews as topic/	30791
27	meta analysis.pt.	127435
28	exp meta-analysis as topic/	66065
29	(meta-analy* or metanaly* or metaanaly* or met analy* or integrative research or integrative review* or integrative overview* or research integration or research overview* or collaborative review*).tw,kf.	485741
30	(systematic review* or systematic overview* or evidence-based review* or evidence-based overview* or (evidence adj3 (review* or overview*)) or meta-review* or meta-overview* or meta-synthes* or rapid review* or "review of reviews" or umbrella review? or technology assessment* or HTA or HTAs).tw,kf.	577594
31	exp Technology assessment, biomedical/	26445

#	Searches	Results
32	(cochrane or health technology assessment or evidence report or systematic reviews).jw.	63425
33	(network adj (MA or MAs)).tw,kf.	37
34	(NMA or NMAs or MTC or MTCs or MAIC or MAICs).tw,kf.	19969
35	indirect* compar*.tw,kf.	7061
36	(indirect treatment* adj1 compar*).tw,kf.	1169
37	(mixed treatment* adj1 compar*).tw,kf.	1517
38	(multiple treatment* adj1 compar*).tw,kf.	474
39	(multi-treatment* adj1 compar*).tw,kf.	12
40	simultaneous* compar*.tw,kf.	2619
41	mixed comparison?.tw,kf.	144
42	or/25-41	975714
43	24 and 42 [REVIEWS]	1377
44	limit 43 to yr="2015-current" [REVIEWS - 5 YEARS]	891
45	(controlled clinical trial or randomized controlled trial or pragmatic clinical trial or equivalence trial).pt.	1211925
46	clinical trials as topic/	309568
47	exp Randomized Controlled Trials as Topic/	349328
48	(randomi#ed or randomi#ation? or randomly or RCT or placebo*).tw,kf.	3605010
49	((singl* or doubl* or trebl* or tripl*) adj (mask* or blind* or dumm*)).tw,kf.	719773
50	trial.ti.	902940
51	or/45-50	4526731
52	24 and 51 [RCTS]	11940
53	limit 52 to yr="2005-current"	7617
54	44 or 53 [REVIEWS, TRIALS]	8206
55	54 use ppez [MEDLINE RECORDS]	2530
56	exp acute myeloid leukemia/	102975
57	acute leukemia/ and myeloid leukemia/	515
58	(acute adj2 myelo* adj2 leu#?emi*).tw,kw.	128438
59	(acute adj2 (nonlympho* or non-lympho*) adj2 leu#?emi*).tw,kw.	6921
60	(acute adj2 granulocytic adj2 leu#?emi*).tw,kw.	611
61	((AML or ANLL) and (leu#?emi* or myelo* or nonlympho* or non-lympho*)).tw,kw.	96992
62	(AML or ANLL).tw,kw. and exp myeloid leukemia/	49010
63	(acute adj2 basophilic adj2 leu#?emi*).tw,kw.	138
64	(acute adj2 eosinophilic adj2 leu#?emi*).tw,kw.	71
65	(acute adj2 erythroblastic adj2 leu#?emi*).tw,kw.	49
66	(erythroleu#?emi* or erythro-leu#?emi* or erythr?emic myelosis or diguglielmo* or di guglielmo*).tw,kw.	13696
67	((mast-cell* or mastcell*) adj2 leu#?emi*).tw,kw.	1156
68	((megakaryocytic or mega-karyocytic) adj2 leu#?emi*).tw,kw.	892
69	((megakaryoblastic or mega-karyoblastic) adj2 leu#?emi*).tw,kw.	2381
70	(acute adj2 monoblastic adj2 leu#?emi*).tw,kw.	825
71	(acute adj2 monocytic adj2 leu#?emi*).tw,kw.	2673
72	(acute adj2 promyelocytic adj2 leu#?emi*).tw,kw.	17833

#	Searches	Results
73	(acute adj2 progranulocytic adj2 leu#?emi*).tw,kw.	35
74	(Schilling-Type adj2 myelo* adj2 leu#?emi*).tw,kw.	0
75	or/56-74 [AML]	214254
76	exp animal experimentation/ or exp animal model/ or exp animal experiment/ or nonhuman/ or exp vertebrate/	52097339
77	exp human/ or exp human experimentation/ or exp human experiment/	41727917
78	76 not 77	10371135
79	75 not 78 [ANIMAL-ONLY REMOVED]	200962
80	editorial.pt.	1248574
81	letter.pt. not (letter.pt. and randomized controlled trial/)	2303596
82	79 not (80 or 81) [OPINION PIECES REMOVED]	190908
83	meta-analysis/	336166
84	"systematic review"/	430443
85	"meta analysis (topic)"/	44732
86	"systematic review (topic)"/	26071
87	(meta-analy* or metanaly* or metaanaly* or met analy* or integrative research or integrative review* or integrative overview* or research integration or research overview* or collaborative review*).tw,kw.	495088
88	(systematic review* or systematic overview* or evidence-based review* or evidence-based overview* or (evidence adj3 (review* or overview*)) or meta-review* or meta-overview* or meta-synthes* or "review of reviews" or umbrella review? or technology assessment* or HTA or HTAs).tw,kw.	584633
89	biomedical technology assessment/	25330
90	(cochrane or health technology assessment or evidence report).jw.	54143
91	or/83-90	1044160
92	82 and 91 [REVIEWS]	2177
93	limit 92 to yr="2015-current" [REVIEWS - 5 YR LIMIT]	1378
94	randomized controlled trial/ or controlled clinical trial/	1451431
95	"clinical trial (topic)"/	111114
96	exp "controlled clinical trial (topic)"/	205199
97	(randomi#ed or randomi#ation or randomly or RCT or placebo*).tw,kw.	3667059
98	((singl* or doubl* or trebl* or tripl*) adj (mask* or blind* or dumm*)).tw,kw.	748188
99	trial.ti.	902940
100	or/94-99	4516287
101	82 and 100 [RCTS, PHASE II/III TRIALS]	15476
102	limit 101 to yr="2005-current" [RCTs - 2005-current]	11424
103	93 or 102 [REVIEWS, TRIALS]	12316
104	conference abstract.pt.	4042951
105	103 not 104	8274
106	103 and 104	4042
107	limit 106 to yr="2018-current"	719
108	105 or 107 [MOST RECENT 2 YRS CONFERENCE ABSTRACTS RETAINED]	8993
109	108 use oemezd [EMBASE RECORDS]	4232
110	exp Leukemia, Myeloid, Acute/	102975
111	(acute adj2 myelo* adj2 leu#?emi*).ti,ab,kw.	128393

#	Searches	Results
112	(acute adj2 (nonlympho* or non-lympho*) adj2 leu#?emi*).ti,ab,kw.	6918
113	(acute adj2 granulocytic adj2 leu#?emi*).ti,ab,kw.	611
114	((AML or ANLL) and (leu#?emi* or myelo* or nonlympho* or non-lympho*)).ti,ab,kw.	96952
115	(AML or ANLL).ti,ab,kw. and exp Leukemia, Myeloid/	49009
116	(acute adj2basophilic adj2 leu#?emi*).ti,ab,kw.	0
117	(acute adj2 eosinophilic adj2 leu#?emi*).ti,ab,kw.	71
118	(acute adj2 erythroblastic adj2 leu#?emi*).ti,ab,kw.	49
119	(erythroleu#?emi* or erythro-leu#?emi* or erythr?emic myelosis or diguglielmo* or di guglielmo*).ti,ab,kw.	13696
120	((mast-cell* or mastcell*) adj2 leu#?emi*).ti,ab,kw.	1156
121	((megakaryocytic or mega-karyocytic) adj2 leu#?emi*).ti,ab,kw.	892
122	((megakaryoblastic or mega-karyoblastic) adj2 leu#?emi*).ti,ab,kw.	2381
123	(acute adj1 monoblastic adj2 leu#?emi*).ti,ab,kw.	807
124	(acute adj2 monocytic adj2 leu#?emi*).ti,ab,kw.	2673
125	(acute adj2 promyelocytic adj2 leu#?emi*).ti,ab,kw.	17827
126	(acute adj2 progranulocytic adj2 leu#?emi*).ti,ab,kw.	35
127	(Schilling-Type adj2 myelo* adj2 leu#?emi*).ti,ab,kw.	0
128	or/110-127	213936
129	conference abstract.pt.	4042951
130	128 not 129	179441
131	128 and 129	34495
132	limit 131 to yr="2018-current"	7519
133	130 or 132	186960
134	limit 133 to yr="2005-current"	104962
135	134 use cctr [TRIALS, 2005-CURRENT]	3922
136	limit 128 to yr="2015-current"	66954
137	136 use coch [REVIEWS, 2015-CURRENT]	10
138	135 or 137 [REVIEWS, TRIALS]	3932
139	138 use coch,cctr [COCHRANE DSR, CENTRAL RECORDS]	3932
140	55 or 109 or 139 [ALL DATABASES]	10694
141	limit 140 to yr="2020 -Current"	1145
142	remove duplicates from 141	787
143	("20200101" or "20200102" or "20200103" or "20200104" or "20200105" or "20200106" or "20200107" or "20200108" or "20200109" or "20200110" or "20200111" or "20200112" or "20200113" or "20200114" or "20200115" or "20200116" or "20200117").up.	73530
144	("20200101" or "20200102" or "20200103" or "20200104" or "20200105" or "20200106" or "20200107" or "20200108" or "20200109" or "20200110" or "20200111" or "20200112" or "20200113" or "20200114" or "20200115" or "20200116" or "20200117").dc.	110897
145	("20200101" or "20200102" or "20200103" or "20200104" or "20200105" or "20200106" or "20200107" or "20200108" or "20200109" or "20200110" or "20200111" or "20200112" or "20200113" or "20200114" or "20200115" or "20200116" or "20200117").dt.	59170
146	143 or 144 or 145	239924
147	142 not 146 [All databases - update results 18 Jan 2020 - Current]	773

Appendix G. Published cost-effectiveness studies:

Table G.1. MEDLINE database search strategy, February 12, 2020

#	Searches	Results
1	exp Leukemia, Myeloid, Acute/	54278
2	(acute adj2 myelo* adj2 leu#?emi*).tw,kf.	44998
3	(acute adj2 (nonlympho* or non-lympho*) adj2 leu#?emi*).tw,kf.	3175
4	(acute adj2 granulocytic adj2 leu#?emi*).tw,kf.	168
5	((AML or ANLL) and (leu#?emi* or myelo* or nonlympho* or non-lympho*)).tw,kf.	29756
6	(AML or ANLL).tw,kf. and exp Leukemia, Myeloid/	22738
7	(acute adj2 basophilic adj2 leu#?emi*).tw,kf.	66
8	(acute adj2 eosinophilic adj2 leu#?emi*).tw,kf.	35
9	(acute adj2 erythroblastic adj2 leu#?emi*).tw,kf.	19
10	(erythroleu#?emi* or erythro-leu#?emi* or erythr?emic myelosis or diguglielmo* or di guglielmo*).tw,kf.	6546
11	((mast-cell* or mastcell*) adj2 leu#?emia*).tw,kf.	397
12	((megakaryocytic or mega-karyocytic) adj2 leu#?emi*).tw,kf.	392
13	((megakaryoblastic or mega-karyoblastic) adj2 leu#?emi*).tw,kf.	1006
14	(acute adj2 monoblastic adj2 leu#?emi*).tw,kf.	366
15	(acute adj2 monocytic adj2 leu#?emi*).tw,kf.	1134
16	(acute adj2 promyelocytic adj2 leu#?emi*).tw,kf.	7077
17	(acute adj2 progranulocytic adj2 leu#?emi*).tw,kf.	14
18	(Schilling-Type adj1 myelo* adj2 leu#?emi*).tw,kf.	0
19	or/1-18 [AML-Medline]	83434
20	Economics/ or exp "Costs and Cost Analysis"/ or Economics, Nursing/ or Economics, Medical/ or Economics, Pharmaceutical/ or exp Economics, Hospital/ or Economics, Dental/ or exp "Fees and Charges"/ or exp Budgets/ or exp models, economic/ or markov chains/ or monte carlo method/ or exp Decision Theory/	354942
21	budget*.ti,ab,kw,kf. or (economic* or cost or costs or costly or costing or price or prices or pricing or pharmacoeconomic* or pharmaco-economic* or expenditure or expenditures or expense or expenses or financial or finance or finances or financed).ti,kw,kf. or economic model*.ab,kw,kf.	247678
22	(economic* or cost or costs or costly or costing or price or prices or pricing or pharmacoeconomic* or pharmaco-economic* or expenditure or expenditures or expense or expenses or financial or finance or finances or financed).ab. /freq=2	280135
23	(cost* adj2 (effective* or utilit* or benefit* or minimi* or analy* or outcome or outcomes)).ab,kw,kf. or (value adj2 (money or monetary)).ti,ab,kw,kf.	158430
24	(markov or monte carlo or (decision* adj2 (tree* or analy* or model*))).ti,ab,kw,kf.	85758
25	or/20-24 [Filter-Econ-CADTH-Medline]	711894
26	economics, pharmaceutical/ or exp economics, medical/ or exp economics, hospital/ or economics, nursing/ or Cost allocation/ or Cost control/ or Cost savings/ or Cost of illness/ or exp "Fees and Charges"/ or exp Budgets/ or exp "Costs and cost analysis"/ or Cost-benefit analysis/ or Models, economic/ or Markov chains/ or Monte Carlo method/ or Decision tree/ or Direct service costs/ or Drug costs/ or Health expenditures/	326881
27	(pharmacoeconomic$ or (pharmaco adj economic$) or health economic$ or economic aspect$ or economic evaluati$ or cost utili$ analys$ or (cost$ or (cost$ adj2 (effective$ or utili$ or benefit$ or minimi$ or stud$ or effic$ or effect$))) or	1186414

#	Searches	Results
	(economic$ and (evaluat$ or analys$ or model$)) or (economic$ or cost$ or pric$ or pharmacoeconomic$) or budget$ or expenditure$).mp.	
28	cost of illness.mp.	27350
29	(cba or cea or cua or cost minimi?ation analys$).mp.	50999
30	((decision adj2 (tree$ or analys$ or model$)) or markov$ or (monte adj carlo) or (cost$ adj3 estimate$) or (unit adj3 cost$)).mp.	122218
31	(cost effectiveness analys$ or cost benefit analys$).mp.	85107
32	or/26-31 [Filter-Econ-NICE-Medline]	1327189
33	19 and (or/25,32) [Medline results]	1183
34	(2005$ or 2006$ or 2007$ or 2008$ or 2009$ or 201$ or 202$).yr,dp.	14365582
35	33 and 34 [Medline Results--Economic Filter, 2005-]	711
36	limit 33 to yr="2005 -Current"	711
37	remove duplicates from 36	708

Table G.2. EMBASE database search strategy, February 12, 2020

#	Searches	Results
1	exp *acute myeloid leukemia/	18834
2	acute leukemia/ and myeloid leukemia/	347
3	(acute adj2 myelo* adj2 leu#?emi*).tw,kw.	66433
4	(acute adj2 (nonlympho* or non-lympho*) adj2 leu#?emi*).tw,kw.	1845
5	(acute adj2 granulocytic adj2 leu#?emi*).tw,kw.	28
6	((AML or ANLL) adj7 (leu#?emi* or myelo* or nonlympho* or non-lympho*)).tw,kw.	43576
7	(AML or ANLL).tw,kw. and exp myeloid leukemia/	19898
8	(acute adj2 basophilic adj2 leu#?emi*).tw,kw.	60
9	(acute adj2 eosinophilic adj2 leu#?emi*).tw,kw.	21
10	(acute adj2 erythroblastic adj2 leu#?emi*).tw,kw.	24
11	(erythroleu#?emi* or erythro-leu#?emi* or erythr?emic myelosis or diguglielmo* or diguglielmo*).tw,kw.	5024
12	((mast-cell* or mastcell*) adj2 leu#?emia*).tw,kw.	562
13	((megakaryocytic or mega-karyocytic) adj2 leu#?emi*).tw,kw.	430
14	((megakaryoblastic or mega-karyoblastic) adj2 leu#?emi*).tw,kw.	1181
15	(acute adj2 monoblastic adj2 leu#?emi*).tw,kw.	325
16	(acute adj2 monocytic adj2 leu#?emi*).tw,kw.	1136
17	(acute adj2 promyelocytic adj2 leu#?emi*).tw,kw.	9024
18	(acute adj2 progranulocytic adj2 leu#?emi*).tw,kw.	19
19	(Schilling-Type adj1 myelo* adj2 leu#?emi*).tw,kw.	0
20	or/1-19 [AML-EMBASE]	89503
21	*Economics/ or *Cost/ or exp *Health Economics/ or *Budget/ or *Statistical Model/ or *Probability/ or *monte carlo method/ or *Decision Theory/ or *Decision Tree/	280751
22	budget*.ti,ab,kw. or (economic* or cost or costs or costly or costing or price or prices or pricing or pharmacoeconomic* or pharmaco-economic* or expenditure or expenditures or expense or expenses or financial or finance or finances or financed).ti,kw.	277614
23	(economic* or cost or costs or costly or costing or price or prices or pricing or pharmacoeconomic* or pharmaco-economic* or expenditure or expenditures or expense or expenses or financial or finance or finances or financed).ab. /freq=2	376257

#	Searches	Results
24	(cost* adj2 (effective* or utilit* or benefit* or minimi* or analy* or outcome or outcomes)).ab,kw. or (value adj2 (money or monetary)).ti,ab,kw.	217820
25	economic model*.ab,kw. or markov.ti,ab,kw. or monte carlo.ti,ab,kw. Or (decision* adj2 (tree* or analy* or model*)).ti,ab,kw.	102343
26	or/21-25 [Filter-Econ-CADTH-EMBASE]	799808
27	exp *pharmacoeconomics/ or exp *socioeconomics/ or exp *economic aspect/ or exp *health economics/ or *cost of illness/ or *cost minimization analysis/ or *cost effectiveness analysis/ or *cost benefit analysis/ or exp *economic evaluation/ or exp *economics/ or "cost control"/ or *cost utility analysis/ [EMTREE NICE Econ Filter]	429879
28	(pharmacoeconomic$ or (pharmaco adj economic$) or health economic$ or economic aspect$ or economic evaluati$ or cost utili$ analys$ or (cost$ or (cost$ adj2 (effective$ or utili$ or benefit$ or minimi$ or stud$ or effic$ or effect$))) or (economic$ and (evaluat$ or analys$ or model$)) or (economic$ or cost$ or pric$ or pharmacoeconomic$) or budget$ or expenditure$).mp.	1484118
29	cost of illness.mp.	20164
30	(cba or cea or cua or cost minimi?ation analys$).mp.	46984
31	*Models, economic/ or *Markov chains/ or *Monte Carlo method/ or *Decision tree/	9361
32	((decision adj2 (tree$ or analys$ or model$)) or markov$ or (monte adj carlo) or (cost$ adj3 estimate$) or (unit adj3 cost$)).mp.	141193
33	Direct service costs/ or Drug costs/ or Health expenditures/	241132
34	(cost effectiveness analys$ or cost benefit analys$).mp.	217778
35	or/27-34 [Filter: Economic-NICE-Embase]	1739842
36	20 and (or/26,35) [Embase results]	2314
37	36 not (conference abstract or conference review).pt. [EM exc conf abs-2005-]	1132
38	(2005$ or 2006$ or 2007$ or 2008$ or 2009$ or 201$ or 202$).yr,dp.	18867312
39	(2017$ or 2018$ or 2019$ or 2020$).yr,dp.	4913755
40	37 and 38 [Embase results 2005- excluding conference abstracts]	770
41	36 and 39 and (conference abstract or conference review).pt. [EM conf abs-2005-]	413

Table G.3. NHS Economic Evaluation database search strategy, February 12, 2020

#	Searches	Results
1	(acute adj2 myelo* adj2 leu#?emi*).tw.	28
2	(acute adj2 (nonlympho* or non-lympho*) adj2 leu#?emi*).tw.	2
3	(acute adj2 granulocytic adj2 leu#?emi*).tw.	0
4	((AML or ANLL) and (leu#?emi* or myelo* or nonlympho* or non-lympho*)).tw.	13
5	(acute adj2 basophilic adj2 leu#?emi*).tw.	0
6	(acute adj2 eosinophilic adj2 leu#?emi*).tw.	0
7	(acute adj2 erythroblastic adj2 leu#?emi*).tw.	0
8	(erythroleu#?emi* or erythro-leu#?emi* or erythr?emic myelosis or diguglielmo* or di guglielmo*).tw.	0
9	((mast-cell* or mastcell*) adj2 leu#?emia*).tw.	0
10	((megakaryocytic or mega-karyocytic) adj2 leu#?emi*).tw.	0
11	((megakaryoblastic or mega-karyoblastic) adj2 leu#?emi*).tw.	0
12	(acute adj2 monoblastic adj2 leu#?emi*).tw.	0
13	(acute adj2 monocytic adj2 leu#?emi*).tw.	0

#	Searches	Results
14	(acute adj2 promyelocytic adj2 leu#?emi*).tw.	3
15	(acute adj2 progranulocytic adj2 leu#?emi*).tw.	0
16	(Schilling-Type adj2 leu#?emi*).tw.	0
17	or/1-16 [AML-EED]	34

Appendix H. Health-related quality of life studies:

Table H.1. MEDLINE, EMBASE, and Cochrane Central Register of Controlled Trials database search strategy, February 12, 2020

#	Searches	Results
1	exp Leukemia, Myeloid, Acute/	106857
2	(acute adj2 myelo* adj2 leu#?emi*).tw,kf.	130053
3	(acute adj2 (nonlympho* or non-lympho*) adj2 leu#?emi*).tw,kf.	6847
4	(acute adj2 granulocytic adj2 leu#?emi*).tw,kf.	337
5	((AML or ANLL) and (leu#?emi* or myelo* or nonlympho* or non-lympho*)).tw,kf.	98579
6	(AML or ANLL).tw,kf. and exp Leukemia, Myeloid/	51086
7	(acute adj2 basophilic adj2 leu#?emi*).tw,kf.	139
8	(acute adj2 eosinophilic adj2 leu#?emi*).tw,kf.	70
9	(acute adj2 erythroblastic adj2 leu#?emi*).tw,kf.	50
10	(erythroleu#?emi* or erythro-leu#?emi* or erythr?emic myelosis or diguglielmo* or di guglielmo*).tw,kf.	13629
11	((mast-cell* or mastcell*) adj2 leu#?emia*).tw,kf.	1004
12	((megakaryocytic or mega-karyocytic) adj2 leu#?emi*).tw,kf.	881
13	((megakaryoblastic or mega-karyoblastic) adj2 leu#?emi*).tw,kf.	2400
14	(acute adj2 monoblastic adj2 leu#?emi*).tw,kf.	829
15	(acute adj2 monocytic adj2 leu#?emi*).tw,kf.	2695
16	(acute adj2 promyelocytic adj2 leu#?emi*).tw,kf.	17963
17	(acute adj2 progranulocytic adj2 leu#?emi*).tw,kf.	32
18	(Schilling-Type adj1 myelo* adj2 leu#?emi*).tw,kf.	0
19	or/1-18 [AML-Medline]	216972
20	"Value of Life"/ or Quality of Life/ or Quality-Adjusted Life Years/ or exp health status indicators/	1258664
21	quality of life.ti,kf,kw. or ((instrument or instruments) adj3 quality of life).ab. or quality adjusted life.ti,ab,kf,kw. or (qaly* or qald* or qale* or qtime* or life year or life years).ti,ab,kf,kw. or disability adjusted life.ti,ab,kf,kw. or daly*.ti,ab,kf,kw.	364612
22	(sf36 or sf 36 or short form 36 or shortform 36 or short form36 or shortform36 or sf thirtysix or sfthirtysix or sfthirty six or sf thirty six or shortform thirtysix or shortform thirty six or short form thirtysix or short form thirty six or (sf6 or sf 6 or short form 6 or shortform 6 or sf six or sfsix or shortform six or short form six or shortform6 or short form6) or (sf8 or sf 8 or sf eight or sfeight or shortform 8 or shortform 8 or shortform8 or short form8 or shortform eight or short form eight) or (sf12 or sf 12 or short form 12 or shortform 12 or short form12 or shortform12 or sf twelve or sftwelve or shortform twelve or short form twelve) or (sf16 or sf 16 or short form 16 or shortform 16 or short form16 or shortform16 or sf sixteen or sfsixteen or shortform sixteen or short form sixteen) or (sf20 or sf 20 or short form 20 or shortform 20 or short form20 or shortform20 or sf twenty or sftwenty or shortform twenty or short form twenty)).ti,ab,kf,kw.	112698

#	Searches	Results
23	(hql or hqol or h qol or hrqol or hr qol or (hye or hyes) or (health* adj2 year* adj2 equivalent*) or (pqol or qls) or (quality of wellbeing or quality of well being or index of wellbeing or index of well being or qwb) or nottingham health profile* or sickness impact profile).ti,ab,kf,kw.	65723
24	((health adj3 (utilit* or status)) or (utilit* adj3 (valu* or measur* or health or life or estimat* or elicit* or disease or score* or weight)) or (preference* adj3 (valu* or measur* or health or life or estimat* or elicit* or disease or score* or instrument or instruments)) or disutilit* or rosser or willingness to pay or standard gamble* or (time trade off or time tradeoff) or tto or (hui or hui1 or hui2 or hui3) or (eq or euroqol or euro qol or eq5d or eq 5d or euroqual or euro qual) or duke health profile or functional status questionnaire or dartmouth coop functional health assessment*).ti,ab,kf,kw.	312239
25	or/20-24 [Filter-Utilities-QoL-CADTH-Medline]	1622974
26	(aqol or "assessment of quality of life").ti,ab,kw,kf.	6382
27	(facit or facitf or facit-f).ti,ab,kw,kf.	4539
28	(fatigue? adj2 (scale? or score?)).ti,ab,kw,kf.	20956
29	"European Organization for Research and Treatment of Cancer Quality of Life".ti,ab,kw,kf.	3453
30	(EORTC QLQ-C30 or eortc qlq c30).ti,ab,kw,kf.	12973
31	"Functional Assessment of Cancer Therap$ ".ti,ab,kw,kf.	7864
32	((fact adj3 (assess$ or questionnai$ or questionai$ or survey? or tool?)) or (factg or fact g)).ti,ab,kw,kf.	6391
33	((Rotterdam Symptom adj1 (Checklist? check list? or questionai$ or questionnai$ or survey?)) or RSCL?).ti,ab,kw,kf.	356
34	Functional Assessment of Chronic Illness Therapy.ti,ab,kw,kf.	3129
35	AML-QOL.ti,ab,kw,kf.	7
36	(sf-36 or sf-6D).ti,ab,kw,kf.	71190
37	(symptom? distress adj2 (scale? or instrument? or survey or questionai$ or questionnai$)).ti,ab,kw,kf.	765
38	(quality of life adj3 (measur$ or survey? or questionn$ or questionai$)).ti,ab,kw,kf. [Not in CADTH filter]	114047
39	or/26-38 [additional utility terms per protocol; and specific to cancer]	215086
40	(2005$ or 2006$ or 2007$ or 2008$ or 2009$ or 201$ or 202$).dp,yr.	39084071
41	19 and (or/25,39)	2913
42	and/40-41 [Results-Medline-Utilities-2005-]	2564
43	42 use ppez [MEDLINE results]	604
44	exp acute myeloid leukemia/	106857
45	acute leukemia/ and myeloid leukemia/	521
46	(acute adj2 myelo* adj2 leu#?emi*).tw,kw.	131012
47	(acute adj2 (nonlympho* or non-lympho*) adj2 leu#?emi*).tw,kw.	6876
48	(acute adj2 granulocytic adj2 leu#?emi*).tw,kw.	605
49	((AML or ANLL) and (leu#?emi* or myelo* or nonlympho* or non-lympho*)).tw,kw.	99488
50	(AML or ANLL).tw,kw. and exp myeloid leukemia/	51391
51	(acute adj2 basophilic adj2 leu#?emi*).tw,kw.	139
52	(acute adj2 eosinophilic adj2 leu#?emi*).tw,kw.	70
53	(acute adj2 erythroblastic adj2 leu#?emi*).tw,kw.	49
54	(erythroleu#?emi* or erythro-leu#?emi* or erythr?emic myelosis or diguglielmo* or di guglielmo*).tw,kw.	13691
55	((mast-cell* or mastcell*) adj2 leu#?emia*).tw,kw.	1031

#	Searches	Results
56	((megakaryocytic or mega-karyocytic) adj2 leu#?emi*).tw,kw.	894
57	((megakaryoblastic or mega-karyoblastic) adj2 leu#?emi*).tw,kw.	2416
58	(acute adj2 monoblastic adj2 leu#?emi*).tw,kw.	835
59	(acute adj2 monocytic adj2 leu#?emi*).tw,kw.	2702
60	(acute adj2 promyelocytic adj2 leu#?emi*).tw,kw.	18033
61	(acute adj2 progranulocytic adj2 leu#?emi*).tw,kw.	35
62	(Schilling-Type adj1 myelo* adj2 leu#?emi*).tw,kw.	0
63	or/44-62 [AML-EMBASE]	218173
64	socioeconomics/ or exp Quality of Life/ or Quality-Adjusted Life Year/ or nottingham health profile/ or sickness impact profile/ or health status indicator/ [EMTREE]	946487
65	quality of life.ti,kw. or ((instrument or instruments) adj3 quality of life).ab. or quality adjusted life.ti,ab,kw. or (qaly* or qald* or qale* or qtime* or life year or life years).ti,ab,kw. or disability adjusted life.ti,ab,kw. or daly*.ti,ab,kw.	361282
66	(sf36 or sf 36 or short form 36 or shortform 36 or short form36 or shortform36 or sf thirtysix or sfthirtysix or sfthirty six or sf thirty six or shortform thirtysix or shortform thirty six or short form thirtysix or short form thirty six or (sf6 or sf 6 or short form 6 or shortform 6 or sf six or sfsix or shortform six or short form six or shortform6 or short form6) or (sf8 or sf 8 or sf eight or sfeight or shortform 8 or shortform 8 or shortform8 or short form8 or shortform eight or short form eight) or (sf12 or sf 12 or short form 12 or shortform 12 or short form12 or shortform12 or sf twelve or sftwelve or shortform twelve or short form twelve) or (sf16 or sf 16 or short form 16 or shortform 16 or short form16 or shortform16 or sf sixteen or sfsixteen or shortform sixteen or short form sixteen) or (sf20 or sf 20 or short form 20 or shortform 20 or short form20 or shortform20 or sf twenty or sftwenty or shortform twenty or short form twenty)).ti,ab,kw.	112589
67	(hql or hqol or h qol or hrqol or hr qol or (hye or hyes) or (health* adj2 year* adj2 equivalent*) or (pqol or qls) or (quality of wellbeing or quality of well being or index of wellbeing or index of well being or qwb) or nottingham health profile* or sickness impact profile).ti,ab,kw.	65563
68	((health adj3 (utilit* or status)) or (utilit* adj3 (valu* or measur* or health or life or estimat* or elicit* or disease or score* or weight)) or (preference* adj3 (valu* or measur* or health or life or estimat* or elicit* or disease or score* or instrument or instruments)) or disutilit* or rosser or willingness to pay or standard gamble* or (time trade off or time tradeoff) or tto or (hui or hui1 or hui2 or hui3) or (eq or euroqol or euro qol or eq5d or eq 5d or euroqual or euro qual) or duke health profile or functional status questionnaire or dartmouth coop functional health assessment*).ti,ab,kw.	310946
69	or/64-68 [Filter: CADTH: Health Utilities/Quality of Life – OVID Embase]	1314376
70	(aqol or "assessment of quality of life").ti,ab,kw.	6378
71	(facit or facitf or facit-f).ti,ab,kw.	4535
72	(fatigue? adj2 (scale? or score?)).ti,ab,kw.	20932
73	"European Organization for Research and Treatment of Cancer Quality of Life".ti,ab,kw.	3447
74	(EORTC QLQ-C30 or eortc qlq c30).ti,ab,kw.	12966
75	"Functional Assessment of Cancer Therap$ ".ti,ab,kw.	7854
76	((fact adj3 (assess$ or questionnai$ or questionai$ or survey? or tool?)) or (factg or fact-g)).ti,ab,kw.	6379
77	((Rotterdam Symptom adj1 (Checklist? check list? or questionai$ or questionnai$ or survey?)) or RSCL?).ti,ab,kw.	356
78	Functional Assessment of Chronic Illness Therapy.ti,ab,kw.	3129

#	Searches	Results
79	AML-QOL.ti,ab,kw.	7
80	(sf-36 or sf-6D).ti,ab,kw.	71107
81	(symptom? distress adj2 (scale? or instrument? or survey or questionai$ or questionnai$)).ti,ab,kw.	764
82	(quality of life adj3 (measur$ or survey? or questionn$ or questionai$)).ti,ab,kw. [Not in CADTH filter]	113937
83	"quality of life index"/	2880
84	"quality of life assessment"/	11001
85	or/70-84 [HRQoL--Additional terms per protcol; additional EMTREE & Instruments--MF]	223837
86	63 and 69 [AML & CADTH FILTER-EMBASE]	2556
87	(63 and 85) not 86 [AML & ADDITIONAL UTIL HRQOL TERMS]	86
88	or/86-87	2642
89	88 not (CONFERENCE ABSTRACT or CONFERENCE REVIEW).pt.	1844
90	limit 89 to yr="2005 -Current" [RESULTS -EMBASE-AML-UTIL-HRQOL]	1542
91	90 use oemezd [EMBASE results]	910
92	exp Leukemia, Myeloid, Acute/	106857
93	(AML or ANLL).tw,kw. and exp Leukemia, Myeloid/	51391
94	(acute adj2 myelo* adj2 leu#?emi*).tw,kw.	131012
95	(acute adj2 (nonlympho* or non-lympho*) adj2 leu#?emi*).tw,kw.	6876
96	(acute adj2 granulocytic adj2 leu#?emi*).tw,kw.	605
97	((AML or ANLL) and (leu#?emi* or myelo* or nonlympho* or non-lympho*)).tw,kw.	99488
98	(AML or ANLL).tw,kw. and exp Leukemia, Myeloid/	51391
99	(acute adj2 basophilic adj2 leu#?emi*).tw,kw.	139
100	(acute adj2 eosinophilic adj2 leu#?emi*).tw,kw.	70
101	(acute adj2 erythroblastic adj2 leu#?emi*).tw,kw.	49
102	(erythroleu#?emi* or erythro-leu#?emi* or erythr?emic myelosis or diguglielmo* or di guglielmo*).tw,kw.	13691
103	((mast-cell* or mastcell*) adj2 leu#?emia*).tw,kw.	1031
104	((megakaryocytic or mega-karyocytic) adj2 leu#?emi*).tw,kw.	894
105	((megakaryoblastic or mega-karyoblastic) adj2 leu#?emi*).tw,kw.	2416
106	(acute adj2 monoblastic adj2 leu#?emi*).tw,kw.	835
107	(acute adj2 monocytic adj2 leu#?emi*).tw,kw.	2702
108	(acute adj2 promyelocytic adj2 leu#?emi*).tw,kw.	18033
109	(acute adj2 progranulocytic adj2 leu#?emi*).tw,kw.	35
110	(Schilling-Type adj1 myelo* adj2 leu#?emi*).tw,kw.	0
111	or/92-110 [AML-central]	217935
112	"Value of Life"/ or Quality of Life/ or Quality-Adjusted Life Years/ or exp health status indicators/ [MeSH]	1258664
113	quality of life.ti,kw. or ((instrument or instruments) adj3 quality of life).ab. or quality adjusted life.ti,ab,kw. or (qaly* or qald* or qale* or qtime* or life year or life years).ti,ab,kw. or disability adjusted life.ti,ab,kw. or daly*.ti,ab,kw.	361282

#	Searches	Results
114	(sf36 or sf 36 or short form 36 or shortform 36 or short form36 or shortform36 or sf thirtysix or sfthirtysix or sfthirty six or sf thirty six or shortform thirtysix or shortform thirty six or short form thirtysix or short form thirty six or (sf6 or sf 6 or short form 6 or shortform 6 or sf six or sfsix or shortform six or short form six or shortform6 or short form6) or (sf8 or sf 8 or sf eight or sfeight or shortform 8 or shortform 8 or shortform8 or short form8 or shortform eight or short form eight) or (sf12 or sf 12 or short form 12 or shortform 12 or short form12 or shortform12 or sf twelve or sftwelve or shortform twelve or short form twelve) or (sf16 or sf 16 or short form 16 or shortform 16 or short form16 or shortform16 or sf sixteen or sfsixteen or shortform sixteen or short form sixteen) or (sf20 or sf 20 or short form 20 or shortform 20 or short form20 or shortform20 or sf twenty or sftwenty or shortform twenty or short form twenty)).ti,ab,kw.	112589
115	(hql or hqol or h qol or hrqol or hr qol or (hye or hyes) or (health* adj2 year* adj2 equivalent*) or (pqol or qls) or (quality of wellbeing or quality of well being or index of wellbeing or index of well being or qwb) or nottingham health profile* or sickness impact profile).ti,ab,kw.	65563
116	((health adj3 (utilit* or status)) or (utilit* adj3 (valu* or measur* or health or life or estimat* or elicit* or disease or score* or weight)) or (preference* adj3 (valu* or measur* or health or life or estimat* or elicit* or disease or score* or instrument or instruments)) or disutilit* or rosser or willingness to pay or standard gamble* or (time trade off or time tradeoff) or tto or (hui or hui1 or hui2 or hui3) or (eq or euroqol or euro qol or eq5d or eq 5d or euroqual or euro qual) or duke health profile or functional status questionnaire or dartmouth coop functional health assessment*).ti,ab,kw.	310946
117	or/112-116 [Filter: CADTH: Health Utilities/Quality of Life]	1620856
118	(aqol or "assessment of quality of life").ti,ab,kw.	6378
119	(facit or facitf or facit-f).ti,ab,kw.	4535
120	(fatigue? adj2 (scale? or score?)).ti,ab,kw.	20932
121	"European Organization for Research and Treatment of Cancer Quality of Life".ti,ab,kw.	3447
122	(EORTC QLQ-C30 or eortc qlq c30).ti,ab,kw.	12966
123	"Functional Assessment of Cancer Therap$ ".ti,ab,kw.	7854
124	((fact adj3 (assess$ or questionnai$ or questionai$ or survey? or tool?)) or (factg or fact-g)).ti,ab,kw.	6379
125	((Rotterdam Symptom adj1 (Checklist? check list? or questionai$ or questionnai$ or survey?)) or RSCL?).ti,ab,kw.	356
126	Functional Assessment of Chronic Illness Therapy.ti,ab,kw.	3129
127	AML-QOL.ti,ab,kw.	7
128	(sf-36 or sf-6D).ti,ab,kw.	71107
129	(symptom? distress adj2 (scale? or instrument? or survey or questionai$ or questionnai$)).ti,ab,kw.	764
130	(quality of life adj3 (measur$ or survey? or questionn$ or questionai$)).ti,ab,kw. [Not in CADTH filter]	113937
131	or/118-130 [Additional HRQoL terms]	214871
132	111 and (or/117,131)	2944
133	("conference 4th pediatric allergy and asthma meeting paam berlin germany 15 17 october 2015" or conference abstract or conference abstract placebo controlled partly blinded crossover study in 12 sle patients or conference proceeding or "conference review").pt.	4134489
134	132 not 133	2146
135	limit 134 to yr="2005 -Current"	1792

#	Searches	Results
136	135 use cctr [CENTRAL results]	269
137	("202001*" or "20200201" or "20200202" or "20200203" or "20200204" or "20200205" or "20200206" or "20200207" or "20200208" or "20200209" or "20200210" or "20200211").dt.	148726
138	43 and ("2020*" or "2021*").dt.	66
139	138 not 137	53
140	limit 139 to yr="2020 -Current" [MEDLINE results - Feb 11, 2020 - Current]	52
141	("202001*" or "20200201" or "20200202" or "20200203" or "20200204" or "20200205" or "20200206" or "20200207" or "20200208" or "20200209" or "20200210" or "20200211").dc.	208571
142	91 and ("2020*" or "2021*").dc.	197
143	142 not 141	186
144	limit 143 to yr="2020 -Current" [Embase results - Feb 11, 2020 - Current]	179
145	("202001*" or "20200201" or "20200202" or "20200203" or "20200204" or "20200205" or "20200206" or "20200207" or "20200208" or "20200209" or "20200210" or "20200211").up.	39856
146	136 and ("2020*" or "2021*").up.	114
147	146 not 145	103
148	limit 147 to yr="2020 -Current" [CENTRAL results - Feb 11, 2020 - Current]	46
149	140 or 144 or 148	277
150	remove duplicates from 149 [All Results – deduplicated - Feb 11, 2020 - Current]	236

A2. Please confirm if any other searches were conducted for adverse events other than those reported in Appendix D.

The systematic review of clinical evidence identified safety and efficacy data for maintenance treatments for patients with acute myeloid leukemia (AML) who have achieved complete remission (CR) or complete remission with incomplete platelet recovery (CRi) after intensive induction chemotherapy, with or without consolidation, and are ineligible for stem cell transplant (SCT). No additional searches were conducted to identify safety data associated with maintenance treatments in the population of interest.

A3. Grey literature searches:

- Please confirm that the number of included studies for resources listed in Table B.5.2. (Appendix D, clinical evidence) are for all searches (original 2020

and two 2021 updates), and if not, please provide full details of hits per search.

The number of included studies outlined in Table B.5.2 (Appendix D, clinical evidence) represents the total identified studies from all grey literature searches including the following systematic review dates: January 18, 2020; February 19, 2021; June 11, 2021. For full details of included studies per search, please see Table 2 below:

Table 2. Number of included studies, clinical evidence

Conference Name	Number of Included Studies by Search Date			
	January 18, 2020	February 19, 2021	June 11, 2021	Total
ClinicalTrials.gov	0	0	0	0
FDA Database	0	0	0	0
ASCO	0	0	0	0
ASH	1	2	0	3
EBMT	0	0	0	0
EHA	0	0	3	3
SOHO	0	1	0	1

Abbreviations: ASCO = American Society of Clinical Oncology; ASH = American Society of Hematology; EBMT = European Society for Blood and Marrow Transplantation; EHA = European Hematology Association; FDA = Food and Drug Administration; SOHO = Society of Hematologic Oncology.

- Please confirm that the number of included studies for resources listed in Table B.5.22. (Appendix G, cost effectiveness) are for all searches (original 2020 and 2021 update), if not please provide full details of hits per search.

The number of included studies outlined in Table B.5.22. (Appendix G, cost-effectiveness) represents the total identified studies from all grey literature searches including the following systematic review dates: February 12, 2020 and June 11, 2021. For full details of included studies per search, please see Table 3 below:

Table 3. Number of included studies, cost-effectiveness

Conference Name	Number of Included Studies by Search Date		
	February 12, 2020	June 11, 2021	Total
CADTH	1	1	2
NICE	1	1	2
ASCO	0	0	0

ASH	0	0	0
EHA	0	0	0
ISPOR EU	0	0	0
ISPOR US	0	0	0

Abbreviations: ASCO = American Society of Clinical Oncology; ASH = American Society of Hematology; CADTH = Canadian Agency for Drugs and Technologies in Health; EHA = European Hematology Association; ISPOR EU = International Society for Pharmacoeconomics and Outcomes Europe; ISPOR US = International Society for Pharmacoeconomics and Outcomes United States; NICE = National Institute for Health and Care Excellence.

A4. Please provide the date range and dates searched for the ScHARRHUD search reported in section H.1.1.2 (Appendix H, HRQoL).

No date restrictions were applied to the supplementary search of the ScHARRHUD database reported in section H.1.1.2 (Appendix H, HRQoL). The ScHARRHUD database was searched on June 21, 2021.

2. Decision problem

A5. Priority question: the final NICE scope specified that the comparator treatment should be "Established clinical management without oral azacitidine

(which may include a watch and wait strategy with best supportive care, low dose cytarabine or subcutaneous azacitidine)".

a. **Please specify how best supportive care was determined, with reference to relevant NICE clinical guidelines.**

In the relevant NICE clinical guidelines, no specific definition of best supportive care could be identified. Therefore, a targeted search was conducted to identify relevant definitions for best supportive care. Relevant AML guidelines were also searched for definitions of best supportive care.

- The targeted search identified a peer-reviewed systematic literature review which aimed at defining, among others, the concepts and definitions of "supportive care" and "best supportive care". The authors found that the terms "supportive care" and "best supportive care" are commonly used to describe treatment for symptom control and improvement of quality of life of patients.[1]

- The European LeukemiaNet (ELN) guidelines do not provide a definition of best supportive care, but recommend "best supportive care including hydroxyurea for patients who cannot tolerate any antileukemic therapy, or who do not wish any therapy". Based on this recommendation, best supportive care does not include active antileukemic treatment.[2]

In the QUAZAR AML-001 trial, best supportive care may have been used in combination with study treatment as deemed necessary. Best supportive care in both treatment groups included, but was not limited to, red blood cell and platelet transfusions, use of an erythropoiesis stimulating agent, antibiotic, antiviral, and antifungal therapy, nutritional support, and Granulocyte colony-stimulating factors (G-CSFs) for subjects experiencing neutropenic infections.[3]

Based on the ELN guidelines, the definition provided in the systematic literature review, and the definition of best supportive care in the QUAZAR AML-001 trial, we refer to best supportive care as supportive treatment without anti-leukemic activity.

b. Please justify the use of placebo as a comparator, given that the final NICE scope states that best supportive care should be used as a comparator.

BMS considers the "watch and wait" comparator to be represented by the placebo arm within the QUAZAR study, with best supportive care common to both randomised treatment groups. Of note, costs for best supportive care are also captured in the submitted economic model.

Throughout the treatment period of the QUAZAR AML-001 trial, patients in both the placebo and oral azacitidine treatment groups were permitted to receive best supportive care, which may have included red blood cell (RBC) and platelet transfusions; use of an erythropoiesis stimulating agent (ESA); antibiotic, antiviral, and/or antifungal therapy; nutritional support; and/or granulocyte colony-stimulating factor (G-CSF) for patients experiencing neutropenic infections.[3] The inclusion of best supportive care in the study design minimised the risk of providing patients with inadequate care and is consistent with current practice ("watch and wait" strategy) for many patients with AML who are in remission after induction/consolidation therapy.[3, 4]

Concomitant medication (defined as non-study medications that started after the date of randomisation but before the end of the study treatment period, or those that started on or before the date of randomisation and ended or remained ongoing during the study treatment period) use was reported by [redacted] of subjects in the ITT population, and the percentages and types of concomitant medications received were comparable between treatment groups. The most frequently reported (> 50%) concomitant medications were those for alimentary tract and metabolism ([redacted]), anti-infectives for systemic use ([redacted]), the nervous system ([redacted]), and the cardiovascular system ([redacted]).[3]

Midostaurin has been approved by the European Medicines Agency (EMA) in 2017 as single-agent maintenance therapy after use in combination with chemotherapy during induction and consolidation for adult patients with newly diagnosed FLT3 mutation-positive AML.[5] However, only about one-third of patients with AML have an

FLT3 mutation.[6, 7] As oral azacitidine is a mutation-agnostic therapy, use of midostaurin as a comparator would be inappropriate for most patients. Notably, a post hoc analysis on survival outcomes of patients with FLT3-positive mutation showed a positive and significant improvement of survival consistent with the results in the ITT population (28.2 months versus 9.7 months; p=0.114, September 2020 data cut-off).[8, 9]

c. Please justify the use of placebo control, in light of the Declaration of Helsinki guidance on use of placebo controls when established therapy is an option.

At the time of study design of the QUAZAR AML-001 trial (2011), no therapies were approved in multiple regions for use in the AML maintenance setting. Furthermore, there is currently no standard of care for maintenance therapy in AML, and in routine clinical practice, many patients are unlikely to receive further active treatment after achieving remission.[4, 10-12] Therefore, placebo was determined to be the appropriate comparator for oral azacitidine in the QUAZAR AML 001 trial and its selection was agreed upon with regulatory agencies (eg, the US FDA; this was an FDA special protocol assessment trial).

d. Please provide evidence demonstrating that low-dose cytabarine and subcutaneous azacitidine are not part of best supportive care.

As explained in response to A5.b. best supportive care does not include active antileukemic treatment.

Low dose cytarabine, and subcutaneous azacitidine are chemotherapeutic agents, with cytotoxic properties, that have licensed indications to treat acute myeloid leukaemia.[13, 14] BMS do not consider them to be “best supportive care” but rather examples of active treatments that target the underlying leukaemia.

The Pan-London AML guidelines refer to low dose cytarabine or azacitidine as potential treatment options for patients who are not fit for intensive chemotherapy.[15] Similarly, the ELN 2017 guidelines refers to both treatments as ‘selected conventional care regimens’ for patients not considered to be candidates for

intensive chemotherapy. This guidance also clearly distinguishes them from 'best supportive care', describing BSC as follows: "…for patients who cannot tolerate any antileukemic therapy, or who do not wish any therapy".[2]

In the QUAZAR phase 3 study, subsequent therapies received by patients included both low dose cytarabine and subcutaneous azacitidine (along with other chemotherapeutic regimens and HSCT).[3]

e. **Please provide references of the consultation on UK clinical practice for AML maintenance with the two UK AML clinicians.**

A summary report of the consultations with the two UK AML clinicians is provided attached to this response document.

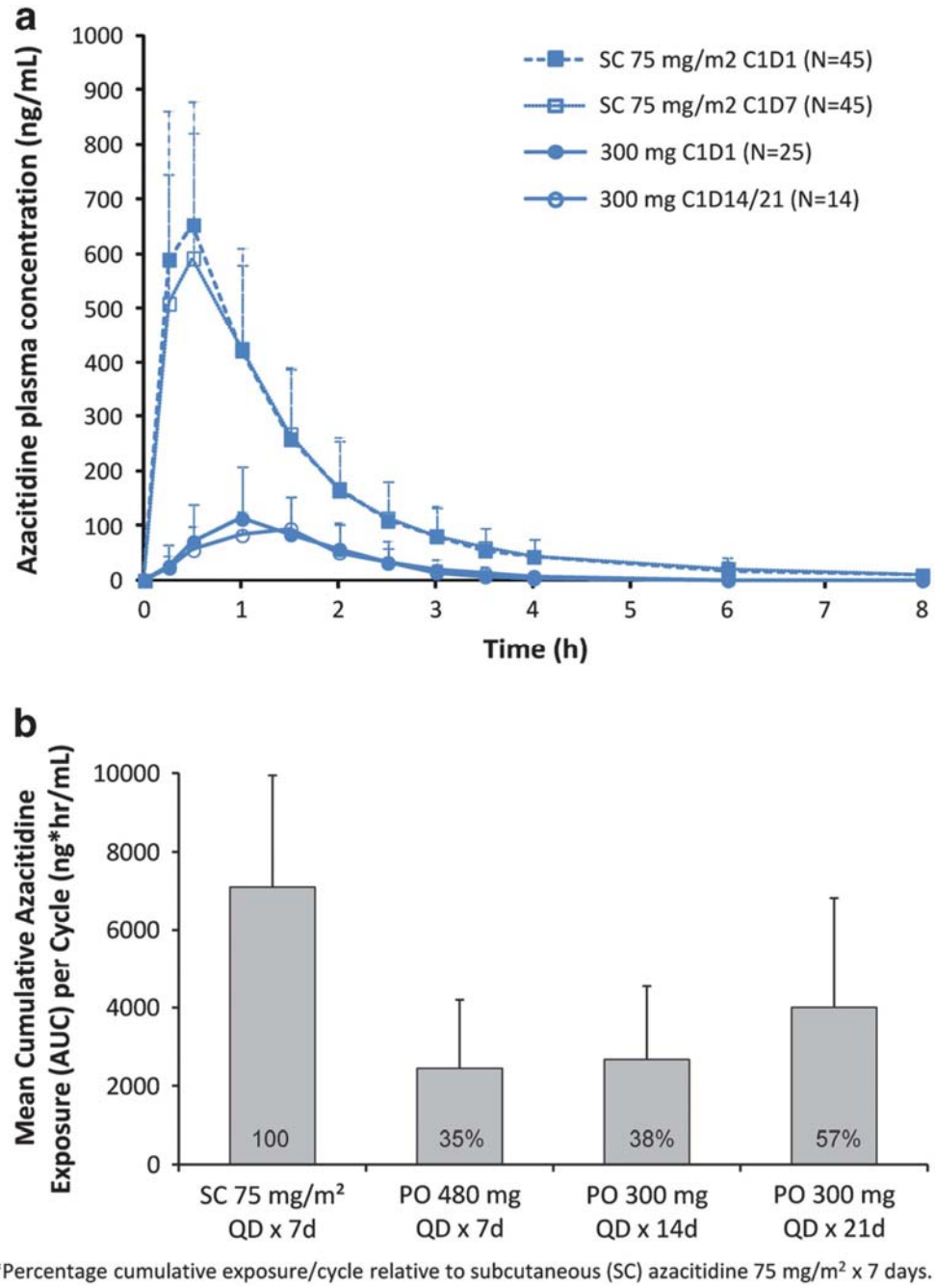
A5. Please comment on the bioavailability of oral azacitidine in comparison with the bioavailability of subcutaneous azacitidine.

Although azacitidine is the active pharmaceutical ingredient in both oral azacitidine and subcutaneous azacitidine, the two formulations are not bioequivalent.

The relative bioavailability of azacitidine after oral (300 mg dose) relative to subcutaneous (75 mg/m^2 dose) administration was approximately 11.5% based on AUC.[16]

A multicentre, open-label study investigated the pharmacokinetic and pharmacodynamic profiles of oral azacitidine in patients with lower-risk myelodysplastic syndromes (LR-MDS).[17, 18] The relative bioavailability of oral azacitidine (300mg QD, 14 or 21 day dosing regimens) compared to subcutaneous azacitidine 75mg/m2 (7 day regimen) were compared as part of this study, and demonstrated a clear difference in both the AUC (Figure 1a), and cumulative azacitidine exposure over the treatment cycles (Figure 1b).[17]

Figure 1 Azacitidine plasma concentration (a) and cumulative azacitidine exposure (b)



(a) Mean (+s.d.) plasma concentration-vs-time profiles following SC azacitidine administration on days 1 and 7, and CC-486 300mg once daily on days 1 and 14; and (b) Cumulative azacitidine exposure per cycle with extended CC-486 dosing regimens relative to azacitidine exposure with subcutaneous (SC) azacitidine 75mg/m2 administered for 7 days.
Source: Garcia-Manero et al. 2016.[17]

A6. In Table B.2.8 of the company submission (CS), the company describes a modified intention to treat population. Please specify how the intention to treat population was modified.

The mITT population included all subjects who met the inclusion/exclusion criteria, experienced no protocol violations during the study, and received at least 1 cycle of treatment. In contrast, the ITT population included all randomised subjects, independently on whether they received study treatment or not.[3]

A7. Section B.1.3.6.2 of the CS states that, "Less than half of all allogenic haematopoietic stem cell transplantation (alloHSCT)-eligible patients have a human leukocyte antigen (HLA) matched sibling available as a donor; availability is even lower for older patients". Please clarify if those eligible for oral azacitidine maintenance would include those for whom no HSCT is available, in addition to those who are not eligible for, including those who choose not to proceed to HSCT.

The sponsor considers that the option for oral azacitidine use as a maintenance therapy would include those patients who cannot proceed to transplant due to no HSCT being available.

A8. 10% (47/472) of patients in the QUAZAR AML-001 study received 'subsequent HSCT' including 6.3% of patients on the azacitidine arm and 13.7% in the comparator arm.

a. Please clarify if by 'subsequent', HSCT was administered during the treatment period.

All 47 patients from the QUAZAR trial who received a subsequent HSCT, did so after treatment discontinuation. The majority of them (41) had relapsed on study drug (32 on placebo and 9 on oral azacitidine) and received HSCT as salvage therapy. The remaining 6 patients, all of which were in the oral azacitidine arm, were transplanted while still in CR1.[19]

b. If so, please clarify how it was that AML patients who were recruited on the basis of being ineligible for HSCT, received HSCT during the trial.

Multiple factors play a role to determine the HSCT eligibility of an AML patient in first CR. It is recommended to take a decision individually for each patient based on an assessment of their risk of relapse if treated with chemotherapy alone, in the context of the mortality risk associated with transplant. This risk is based on factors such as age, fitness, and donor source.[20] It is not uncommon for transplant eligibility of a patient to change over time.

In the literature it is recommended that patients achieving second CR after relapse should proceed quickly to transplant if they are fit enough and a donor is available.[20]

In Table 4 the reasons for transplant ineligibility of the ITT population of the QUAZAR AML-001 trial are provided.

Table 4. Disease baseline characteristics – transplant ineligibility

Parameter	Oral azacitidine (N=238)	Placebo (N=234)	Total (N=472)
Reason ineligible for transplant[a] - n (%)			
Age	154 (64.7)	152 (65.0)	306 (64.8)
Comorbidities	52 (21.8)	50 (21.4)	102 (21.6)
Performance Status	14 (5.9)	9 (3.8)	23 (4.9)
Not acceptable or available donor	37 (15.5)	35 (15.0)	72 (15.3)
Subject decision	19 (8.0)	32 (13.7)	51 (10.8)
Unfavorable cytogenetics	6 (2.5)	10 (4.3)	16 (3.4)
Other	28 (11.8)	21 (9.0)	49 (10.4)

[a] A subject may have had more than 1 reason.
Source: Supplementary appendix to Wei et al. 2020.[21]

In 87% (n=41) of patients that received HSCT, the decision for transplant followed relapse. In 6 patients, representing 2.5% of patients randomised to receive oral azacitidine, a decision was made after randomisation by the treating clinician to proceed with transplant in CR1.[19] The sponsor considers this to be reflective of real-world practice, and to restrict this potentially curative option from the patient pathway to be an unethical scenario.

The protocol required that treatment with oral azacitidine or placebo ceased in the event of subsequent AML therapy, including transplant, however the patients were followed up for survival.

c. Please explain why and how more patients in the comparator arm than in the oral azacitidine arm became eligible for HSCT.

More patients relapsed in the comparator arm than in the treatment arm being one of the main reasons for subsequently undergoing HSCT (please also see response to question A8.a.).[19]

d. Please clarify if the AML patients who received HSCT did not discontinue oral azacitidine maintenance therapy (continued the trial).

Subjects discontinued oral azacitidine prior to undergoing HSCT.

e. As the trial eligibility criteria for patient inclusion specified transplant ineligibility, please clarify if AML patients receiving HSCT was a protocol deviation, or if azacitidine or best supportive care are to improve patient outcome and thus make them eligible for HSCT.

HSCT is the optimal available treatment modality in AML after achieving CR, therefore this was not classified as a protocol deviation or violation. HSCT was viewed as a benefit for patients but the treatment protocol was not designed to make them eligible to undergo transplantation.

f. If azacitidine is to improve the outcome of HSCT-ineligible AML patients, thus rendering them eligible for HSCT, please update Figure B.1.5 from the CS with the proposed treatment pathway.

In 2.5% (n=6) of patients randomised to receive oral azacitidine, HSCT occurred in CR1.[19] The study was not designed to assess the impact of oral azacitidine in achieving transplant eligibility. Whilst it is theoretically possible that oral azacitidine benefitted this small subgroup of patients, there are insufficient data to recommend the treatment pathway described above. This is also beyond the licensed indication for oral azacitidine. Therefore, no updates are required to Figure B.1.5 from the CS.

A9. Regarding the treatment duration in the intervention and control groups in the QUAZAR study:

a. In Table B.2.21, the company notes that the mean treatment duration was [redacted] months for oral azacitidine and [redacted] months for placebo. Can the company please comment on the reasons and implications of this difference?

Per protocol, treatment was discontinued when patients stopped benefitting from the study treatment. In the majority of the cases it was disease relapse that led to treatment discontinuation, and as the difference in RFS demonstrates, this occurred later in the oral azacitidine arm compared to the placebo arm.[22] Please see Table 5 below.

Table 5. Summary of Time to Discontinuation from Treatment Due to Disease Relapse (ITT population)

Parameter	**CC-486 (N=238)**	**Placebo (N=234)**
Subjects with treatment discontinued due to disease relapse – n (%)	143 (60.1)	180 (76.9)
Subjects with treatment discontinued due to adverse event – n (%)	29 (12.2)	11 (4.7)
Subjects with treatment discontinued due to eligibility for bone marrow or stem cell transplant – n (%)	6 (2.5)	0
Subjects with treatment discontinued due to withdrawal of consent/lost to follow-up/protocol violation/other – n (%)	14 (5.9)	15 (6.4)
Treatment discontinued due to death – n (%)	1 (0.4)	2 (0.9)
Censored – n (%)	45 (18.9)	26 (11.1)
Median time to treatment discontinued due to disease relapse (months) (95% CI)[a]	[redacted]	[redacted]
6-month treatment discontinuation due to disease relapse rate estimate (95% CI)[b]	[redacted]	[redacted]
1-year treatment discontinuation due to disease relapse rate estimate (95% CI)[b]	[redacted]	[redacted]
2-year treatment discontinuation due to disease relapse rate estimate (95% CI)[b]	[redacted]	[redacted]

CI = Confidence Interval; ITT = intent-to-treat.
[a] Unstratified Kaplan-Meier analysis

[b] Estimates of treatment discontinuation due to disease relapse rate is based on the cumulative incidence function from a competing risk analysis with treatment discontinuation due to other reasons as competing risk.
Time to discontinuation from treatment is defined as the interval (in months) from the date of randomization to the date of discontinuation from study drug.
Source: QUAZAR AML-001 CSR (Data on File); Wei et al. 2020; Supplementary appendix to Wei et al. 2020; FDA 2020.[3, 21, 22]

b. Can the company also provide treatment duration for ITT and mITT populations?

Treatment duration is based on subjects who took the study drug. In the ITT population [redacted] patients had less than 1 cycle of treatment, therefore conducting this analysis in the ITT population will have a negligible impact on the treatment duration of the mITT, therefore it has not been conducted. The treatment duration for the mITT population is provided in Table 6 below.

Table 6. Treatment duration in the mITT population of the QUAZAR AML-001 trial

Parameter	Oral azacitidine (N=223)	Placebo (N=217)
Treatment duration (months)		
Mean (min, max)	[redacted]	[redacted]
Median	[redacted]	[redacted]

Source: QUAZAR AML-001 CSR (Data on File).[3]

A10. The CS states that midostaurin was recommended by NICE in 2018 for the maintenance treatment of a small subgroup of FLT3-mutation-positive patients who are in remission after previously being treated with midostaurin in combination with chemotherapy agents during induction and consolidation chemotherapy, and most patients with a FLT3 mutation undergo hematopoietic stem cell transplant (HSCT), which would disqualify this patient population from receiving oral azacitidine altogether.

a. Please clarify if only the FLT3-mutation-positive patients who are not eligible for HSCT would be eligible for oral azacitidine maintenance therapy.

Oral azacitidine is indicated for maintenance treatment in adult patients with AML who achieved complete remission (CR) or complete remission with incomplete blood count recovery (CRi) following induction therapy with or without consolidation treatment and who are not candidates for, including those who choose not to

proceed to, haematopoietic stem cell transplantation (HSCT). Patients fulfilling these criteria can receive oral azacitidine independently of their FLT3 mutational status. In other words, both FLT3-positive and FLT3-negative patients who are not candidates for, including those who choose not to proceed to HSCT, are eligible to receive oral azacitidine. Table 7 provides a comparative overview of the eligibility criteria for oral azacitidine and midostaurin in the maintenance setting.

Table 7. Oral azacitidine versus midostaurin eligibility criteria for maintenance treatment

Oral azacitidine:[23*]	**Midostaurin**:[5]
Eligibility requirements for maintenance	
FLT3-mutation-positive patients with AML, who achieved CR or CRi following induction therapy with or without consolidation treatment and who are not candidates for, including those who choose not to proceed to, hematopoietic stem cell transplantation (HSCT).	Midostaurin maintenance is indicated in FLT3-mutation positive patients that received midostaurin in combination with induction chemotherapy (and in combination with any subsequent consolidation chemotherapy), for newly diagnosed AML, and achieved a complete remission.
FLT3-mutation-positive patients in remission but are candidates for and proceeding to transplant are not eligible for oral azacitidine maintenance.	There are no further restrictions to maintenance with midostaurin (monotherapy) with regards to the patient potentially receiving HSCT, unlike with oral azacitidine.

* adapted to focus on FLT3-mutation positive patients.

b. Please clarify if midostaurin is a comparator only for the FLT3-mutation-positive patients who are not eligible for HSCT.

Yes, midostaurin is a comparator only for those patients who have FLT3-positive mutation and who are not eligible for HSCT. Patients who are not eligible for HSCT can receive oral azacitidine, independently of their FLT3 mutational status; whereas patients who have FLT3-positive mutation can receive midostaurin independently of being eligible for HSCT or not (please see response to question A10.a.). Thus, the "intersection" of oral azacitidine and midostaurin are FLT3-mutation-positive patients who are not eligible for HSCT.

Of note, the limitations of the ITC (section B.2.9.6 of the company evidence submission) explained that due to differences in study populations of the QUAZAR AML-001 trial (oral azacitdine versus placebo) and the Ratify trial (midostaurin versus placebo) any estimates of comparative efficacy derived from comparing the included studies are subject to bias.

c. If so, why was midostaurin only suitable for the FLT3 subgroup?

Midostaurin is a multitargeted kinase inhibitor showing efficacy in FLT3-mutation positive AML.[24] Midostaurin is recommended by NICE "within its marketing authorisation as an option in adults for treating newly diagnosed acute FLT3-mutation-positive myeloid leukaemia with standard daunorubicin and cytarabine as induction therapy, with high-dose cytarabine as consolidation therapy, and alone after complete response as maintenance therapy".[25]

Midostaurin was only suitable for the FLT3 subgroup since it is recommended only for this subgroup of AML patients.

A11. Section B.1.3.8 of the CS states that "*maintenance treatment with oral azacitidine represents a new potential therapeutic standard for adult patients with AML in first remission*" (page 29). Please clarify if the company's interpretation is that oral azacitidine will be indicated for the maintenance treatment of AML patients who have achieved remission following induction, having never previously experienced an AML disease relapse following treatment in the past.

According to the licensed indication,[23] there is no restriction on the prescription of oral azacitidine in CR1 or subsequent remissions, however, the QUAZAR-AML-001 trial recruited patients in first CR/CRi[22] and so reflects the efficacy and safety profile of this treatment in first remission.

3. Treatment pathway

A12. Priority question: Figure B.1.5. of the CS suggests that oral azacitidine maintenance therapy could succeed consolidation chemotherapy. The majority of patients in the QUAZAR AML-100 trial received one cycle of

consolidation therapy, and approximately 20% of patients did not receive consolidation therapy.

a. **Please comment on whether consolidation therapy after induction is standard of care in UK clinical practice, and how many cycles of consolidation therapy are usually given in UK clinical practice.**

In the absence of a global standard on optimal number of consolidation cycles in older patients, the protocol of this international trial did not control for the clinical decision making with regards to the provision, or number of cycles of consolidation chemotherapy to be given after achievement of remission.

We raised this question with 2 clinical experts based in the UK, one expert was an investigator for the QUAZAR-AML-001 trial. Both experts confirmed that consolidation therapy is standard of care in UK clinical practice, although there is some uncertainty with regards to the optimal number of cycles. The aim of consolidation therapy in clinical practice is to reduce the risk of relapse of patients who are in CR after induction therapy.

Decisions relating to the number of consolidation chemotherapy cycles can be impacted by the patients' response to induction therapy, including tolerance and fitness. It was acknowledged that, where possible, consolidation would be offered, although both clinicians highlighted there would still be patients who receive no consolidation.

We asked the AML clinicians to provide an estimation on the percentages of patients, that match the QUAZAR-AML-001 eligibility criteria, who would receive 0, 1, 2, 3, 4, or 5 cycles of consolidation. Their responses are provided in Table *8*. The responses from the two AML clinicians show some divergences in their proportions, specifically for no consolidation cycle ([redacted]% versus [redacted]%) and 2 cycles of consolidation therapy ([redacted]% versus [redacted]%), however, both clinicians agreed that the majority of patients would receive consolidation therapy ([redacted]% and

[redacted]%, respectively). In the QUAZAR AML-001 trial about 20% of patients did not receive
consolidation therapy which is likely to be reflective of UK clinical practice as the clinicians estimated that [redacted]% of patients would not receive consolidation therapy.

Table 8. Estimated proportions of patients receiving N number of consolidation therapy cycles in UK clinical practice

Number of consolidation cycles	**Proportion of patients receiving number (N) of consolidation cycles in UK clinical practice**	
	Advisor 1	Advisor 2
Patients receiving consolidation	[redacted]	[redacted]
N = 1	[redacted]	[redacted]
N = 2	[redacted]	[redacted]
N = 3	[redacted]	[redacted]
N = 4	[redacted]	[redacted]
N = 5	[redacted]	[redacted]
N = 0	[redacted]	[redacted]

Abbreviations: UK = United Kingdom

The QUAZAR AML-001 study design did not mandate specific consolidation regimens. As patients were randomised post-consolidation, the rationale for the choice of consolidation agents and the number of consolidation cycles were determined by the treating physician.

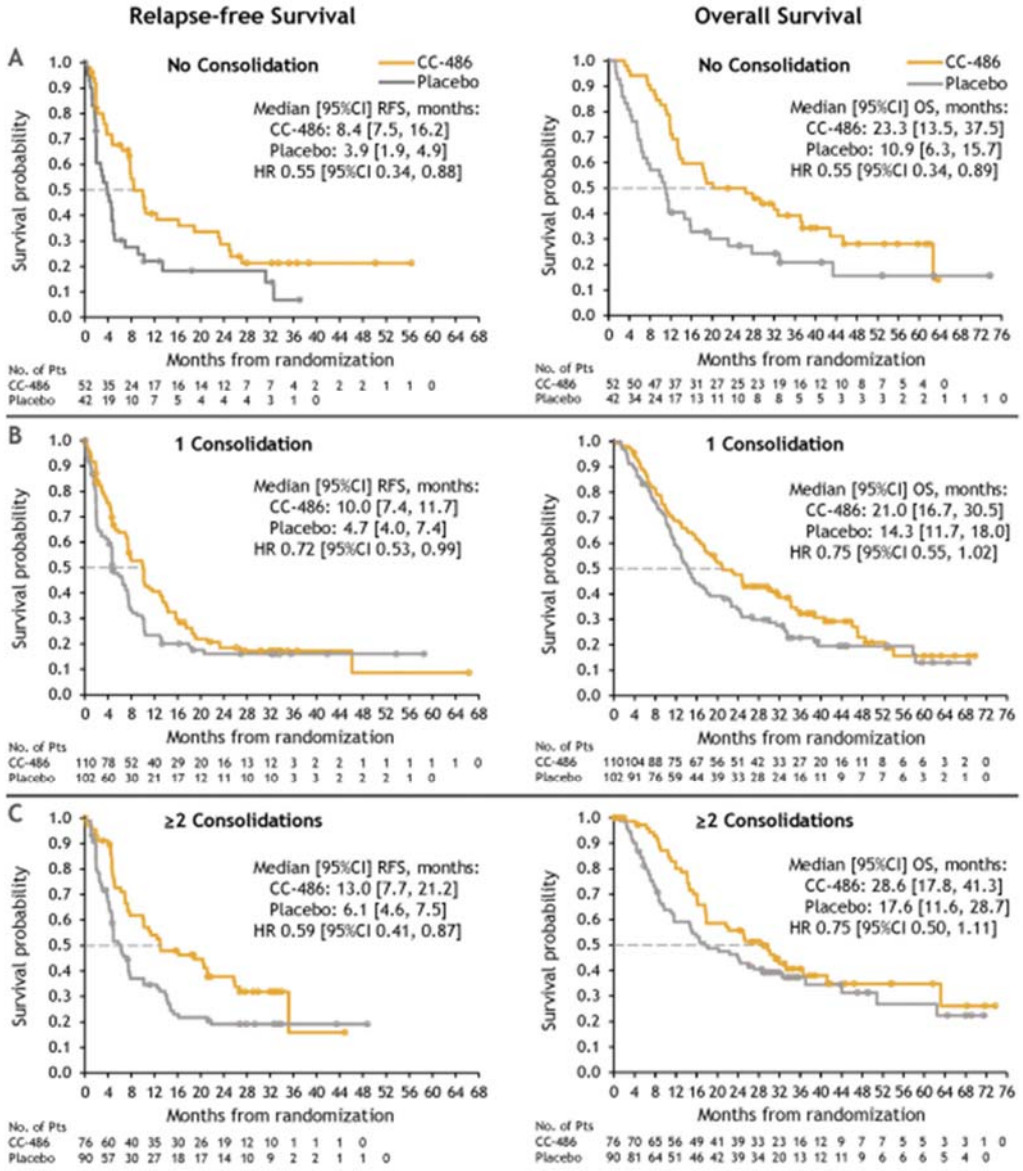
b. Please comment on the potential implications of treating patients with only one cycle of consolidation therapy and not giving consolidation therapy to 20% of patients in the QUAZAR AML-001 trial.

In section B2.7.5 of the CS results of subgroup analyses defined by number of consolidation courses have been presented.

The subgroups analysed were for 0,1 and ≥ 2 cycles of consolidation. Compared to placebo, there was a consistent OS and RFS prolongation with oral azacitidine with each consolidation-based cohort. The KM curves of these subgroup analyses are

shown in Figure 2. Overall, administration of consolidation therapy was associated with treatment benefits in both the oral azacitidine group and the placebo group.[26]

Figure 2 RFS and OS from time to randomisation in patients who received no consolidation, 1 consolidation cycles, or ≥2 consolidation cycles



Abbreviations: CI = confidence interval; HR = hazard ratio; No. = number; OS = overall survival; RFS = relapse-free survival.
Source: Wei et al. 2020.[26]

c. As post-remission therapy consists of consolidation and maintenance, please clarify if the proposed clinical pathway would suggest that consolidation chemotherapies would solely be administered pre-

maintenance with oral azacitidine as in the QUAZAR AML-001 trial, or concomitantly with oral azacitidine.

In accordance with the QUAZAR AML-001 trial design, oral azacitidine maintenance will be administered once the appropriate amount of consolidation therapy has been given to a patient. Consolidation chemotherapy should be administered pre-maintenance, and not concomitantly with oral azacitidine.

d. Please clarify what the recovery time between consolidation therapy and maintenance therapy with oral azacitidine would be in clinical practice.

The design of the study was to start maintenance therapy with oral azacitidine within 120 days or 4 months of achieving CR/CRi with intensive induction chemotherapy.[22] We received feedback on this question from a UK clinician. Each consolidation cycle was estimated to take around [redacted], and the expert explained that patients would start maintenance therapy as soon as blood count recovery was achieved which would approximately be around [redacted] after the last consolidation cycle.
Both experts stated that the 4-months timeframe of the QUAZAR AML-001 trial would typically allow patients to receive 2 cycles of consolidation, including sufficient recovery time before maintenance treatment with oral azacitidine started.

e. Please specify which consolidation therapies that would render patients ineligible to receive oral azacitidine.

Following intensive chemotherapy and achieving a CR/CRi, there is no evidence for specific consolidation therapies to render patients' ineligible to oral azacitidine, which was confirmed by the clinical advisors consulted as part of these responses.

A13. Priority question: In the company submission, it is stated that the enrolment period for the QUAZAR study had to be done within four months of achieving CR or CRi.

a. Please justify the choice of four months.

The eligibility criteria for the QUAZAR AML-001 trial included patients who achieved CR/CRi with induction chemotherapy and then underwent consolidation chemotherapy, the latter of which may involve up to four cycles of treatment.[2, 4] Cycles of consolidation chemotherapy, reported in literature, are typically around

28 days in duration (with treatment occurring within the first 3 to 5 days).[24, 27, 28] Consultation with a UK clinical expert advised that each consolidation cycle would take around 6 weeks. The maximum period between achievement of first CR/CRi and randomisation was selected as 4 months (± 7 days) to incorporate an appropriate amount of time for patients to complete and recover from consolidation regimens before initiating oral azacitidine or placebo.

b. Please specify whether a maximum of four months would suffice for other options, such as consolidation therapy, to succeed.

In the QUAZAR AML-001 trial the following number of cycles of consolidation occurred per treatment arm (Table 9):

Table 9. Disease baseline characteristics – consolidation therapies received

Parameter	Oral azacitidine (N=238)	Placebo (N=234)	Total (N=472)
Received consolidation therapy following induction therapy (n, %)			
Yes	186 (78)	192 (82)	378 (80)
1 Cycle	110 (46)	102 (44)	212 (45)
2 Cycles	70 (29)	77 (33)	147 (31)
3 Cycles	6 (2.5)	13 (6)	19 (4)
4 Cycles	0	0	0
No	52 (22)	42 (18)	94 (20)

Source: Supplementary appendix to Wei et al. 2020.[21]

Based on the clinical experts' advice, a 4-months timeframe would typically allow for two cycles of consolidation therapy, including achievement of blood count recovery.

c. Please provide additional details about how long patients took to enrol in the QUAZAR study. For example, how many enrolled after 1, 2, 3, and 4 months and what was the average enrolment time.

In Table. *10*. details on the time since first CR/CRi to randomisation are presented providing clarification on how long patients took to enrol in the QUAZAR AML-001 trial.

Table. 10 Disease baseline characteristics – Time since first CR/CRi to randomization

Parameter	Oral azacitidine (N=238)	Placebo (N=234)	Total (N=472)
Time Since First Achieving CR/CRi to Randomisation (days)			
Mean (SD)	[redacted]	[redacted]	[redacted]
Median	84.5	86.0	85.0
IQ range (Q1-Q3)	[redacted]	[redacted]	[redacted]

Abbreviations: IQ = interquartile
Source: QUAZAR AML-001 CSR (data on file); Wei et al. 2020.[3, 22]

A14. Priority question: Please specify whether all patients in the QUAZAR trial who relapsed and who were on placebo, continued to receive placebo.

For patients with subsequent evidence of AML relapse with blasts ≥ 5% either in the peripheral blood or bone marrow, and provided the blasts were no greater than 15% in the blood or bone marrow, escalation of the dosing regimen (dose and/or schedule of oral azacitidine/placebo) could be implemented, provided it was in the best interest of the patient to do so as judged by the Investigator. These dose and schedule adjustments were pre-defined in the study protocol.[3]

In all cases patients were discontinued from study treatment (oral azacitidine/placebo) following AML relapse when they had > 15% blasts in the bone marrow or peripheral blood, which was attributable to relapse following CR/CRi, and not attributable to any other cause (eg, bone marrow regeneration after consolidation therapy).[3]

Of note, cross-over from the placebo arm to the oral azacitidine arm was not permitted in the QUAZAR AML-001 trial.[3]

A15. In Tables B.2.11 (OS), B.2.11, B.2.12 (RFS), Table B.2.13 (TTR), and Table B.2.14 (TTD) of the CS, the number of patients censored is substantially higher in the intervention group compared with the placebo group. Can the company please comment on the reasons and implications of this difference for all outcomes?

For the OS analysis, patients were censored most frequently, because they were alive at data cut-off, and similarly for RFS, patients were censored most frequently

because they were alive without documented relapse at time of data-cut off. A summary of censoring data for OS and RFS are provided in Table 11.[3, 22]

Due to the relatively high number of patients censored alive at study closure, the extension phase of the QUAZAR AML-001 (September 2020 data cut-off date) allowed to collect robust and mature OS data of oral azacitidine versus placebo, confirming the sustained benefit with oral azacitidine over the long term.

Table 11. Summary of Censoring for OS and RFS – ITT population (July 2019 data cut-off)

Parameter	**Oral azacitidine (N=238)**	**Placebo (N=234)**	**Total (N=472)**
Overall survival			
Died (n (%))	158 (66.4)	171 (73.1)	
Censored	80 (33.6)	63 (26.9)	143 (30.3)
Reason for censoring (n (%))			
Lost to follow-up	[redacted]	[redacted]	[redacted]
Withdrew consent	[redacted]	[redacted]	[redacted]
Alive at study closure	[redacted]	[redacted]	[redacted]
Relapse-free survival			
Events (n (%))	164 (68.9)	181 (77.4)	345 (73.1)
Documented relapse[a]	[redacted]	[redacted]	[redacted]
Death without documented relapse	[redacted]	[redacted]	[redacted]
Censored	74 (31.1)	53 (22.6)	127 (26.9)
Reason for censoring (n (%))[b]			
No Documented Relapse Or Death	[redacted]	[redacted]	[redacted]
Event After Follow-Up Therapy	[redacted]	[redacted]	[redacted]
Event Out Of Window	[redacted]	[redacted]	[redacted]

[a] Documented relapse is defined as at least 5% blast in the bone marrow blast or reappearance of peripheral blast.
[b] Percentages are based upon number of subjects censored.
Source: Wei et al. 2020; BMS, 2020 (data on file).[3, 22]

4. Systematic literature review (SLR)

A16. Adverse event (AE) data for the QUAZAR AML-100 study was reported in Section B.2.10 of the CS.

a. Please discuss the implications of the extent of exposure to study medication on AEs.

When examining the incidence of TEAEs, it is important to note that duration of exposure to study treatment in the oral azacitidine group (11.6 months) was approximately twice as long as exposure in the placebo group (5.7 months).[29] Since subjects on oral azacitidine have a longer exposure, the chances of having a TEAE would be greater.

b. Please provide the follow-up time period.

TEAEs included adverse events that started between the first dose date and up to 28 days after the last dose date of study treatment or until the date of the last study visit (whichever was longer).[3]

c. Please provide the scale used to judge the severity of treatment-emergent adverse events (TEAEs).

TEAEs were graded using NCI-CTCAE (National Cancer Institute - Common Terminology Criteria for Adverse Events) version 4.0.[21]

AEs that are not defined in the CTCAE were evaluated for severity according to the following scale:

- Grade 1 = Mild – transient or mild discomfort; no limitation in activity; no medical intervention/therapy required
- Grade 2 = Moderate – mild to moderate limitation in activity, some assistance may be needed; no or minimal medical intervention/therapy required
- Grade 3 = Severe – marked limitation in activity, some assistance usually required; medical intervention/therapy required, hospitalization is possible

- Grade 4 = Life threatening – extreme limitation in activity, significant assistance
- required; significant medical intervention/therapy required, hospitalization or hospice care probable
- Grade 5 = Death - the event results in death

d. Please provide list of TEAEs by severity, by system class.

TEAEs of severity Grade 3 or 4 are summarised for the safety population by organ class in Table 12.[3]

Table 12. Treatment-emergent Adverse Events with Severity of Grade 3 or 4 by System Organ Class and Preferred Term Reported for ≥ 2% of Subjects in the CC-486 group Excluding AML Relapse (Safety Population)

System Organ Class Preferred Term[a]	Oral azacitidine (N=236)			Placebo (N=233)		
	Grade 3 n (%)	Grade 4 n (%)	Grade 3 or 4 n (%)	Grade 3 n (%)	Grade 4 n (%)	Grade 3 or 4 n (%)
Subjects with at least 1 Grade 3 or 4 TEAE[b]	[redacted]	[redacted]	169 (71.6)	[redacted]	[redacted]	147 (63.1)
Blood and lymphatic system disorders	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
Neutropenia	[redacted]	[redacted]	97 (41.1)	[redacted]	[redacted]	55 (23.6)
Thrombocytopenia	[redacted]	[redacted]	53 (22.5)	[redacted]	[redacted]	50 (21.5)
Anaemia	[redacted]	[redacted]	33 (14.0)	[redacted]	[redacted]	30 (12.9)
Febrile neutropenia	[redacted]	[redacted]	27 (11.4)	[redacted]	[redacted]	18 (7.7)
Leukopenia	[redacted]	[redacted]	18 (7.6)	[redacted]	[redacted]	14 (6.0)
Infections and infestations	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
Pneumonia	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
Gastrointestinal disorders	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
Diarrhoea	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
Vomiting	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
Nausea	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
Metabolism and nutrition disorders	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
Hypokalaemia	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
General disorders and administration site conditions	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
Fatigue	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
Investigations	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
Blood uric acid increased	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
Vascular disorders	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
Hypertension	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]

Eye disorders	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
Cataract	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
Nervous system disorders	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
Syncope	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]

MedDRA = Medical Dictionary for Regulatory Activities; TEAE = treatment-emergent adverse event

[a] Coded using MedDRA version 22.0. A subject with multiple TEAEs within a preferred term/system organ class is counted once for that preferred term/system organ class in each severity grade and once in the combined severity grade grouping.

[b] Graded using Common Terminology Criteria for Adverse Events (CTCAE) version 4.0.

Notes: Acute myeloid leukemia relapse as defined by MedDRA high-level group term leukemias are excluded. Treatment-emergent adverse events include adverse events that started between first dose date and the date 28 days after the last dose date of study treatment.

Source: QUAZAR AML-001 CSR (Data on File); FDA, 2020; ClinicalTrials.gov; EMA/308711/2021.[3, 29-31]

e. Please discuss the impact of AEs that lead to dose reduction and transitory discontinuations of treatment.

The summary of TEAEs (≥1) captured in the safety population of QUAZAR AML-001 trial are included in Table B.2.22 of the CS. The ITT analysis captures the efficacy measures including this cohort of patients, and so the impact of dose reduction and transitory discontinuations are represented in the OS and RFS data.

The information for prescribers provided in the summary of product characteristics includes recommendations relating to dose adjustments for specific haematological and gastrointestinal adverse events.[23]

The impact of adverse events that led to dose reduction and transitory discontinuations are captured within the RDI calculation.

A17. Table B.3.8 of the company submission shows different rates of AE occurrence for the regular treatment population and FLT3 patients.

a. Please explain why AE rates may differ between treatment populations.

The differences in AE rates between the FLT3 and regular treatment population are reflective of the data collected in the QUAZAR-AML-001 study. It may be that the differences are a result of analysis of smaller subgroups, rather than a true difference between these cohorts, however we are unable to point to a definitive biological reason why AE rate may differ between populations.

b. Please include in this explanation why AE rates may also differ for the 'watch and wait' treatment population

As alluded to above (A17.a.), we are not able to provide a reason why the AE rate may differ between the intervention and the control arm. However, the aim was to use the same methodology as for the intervention arm.

c. Please conduct a scenario analysis applying the AE rates of the regular population to the FLT3 treatment population.

Scenario results from assuming equivalent AE rates for oral azacitidine and watch and wait + BSC for the FLT3 subgroup as the ITT population are provided in Table 13 and the comparison with the base case provided in Table 14.

Table 13. Scenario results: Equivalent adverse event rates - FLT3 population

Technology	Total costs (£)	Total LYG	Total QALYs	Incremental costs (£)	Incremental LYG	Incremental QALYs	ICER versus baseline (£/QALY)	Pairwise ICER vs oral azacitidine
No active therapy	[redacted]	2.731	[redacted]	-	-	-	-	24,621
Oral azacitidine	[redacted]	4.828	[redacted]	[redacted]	2.10	[redacted]	24,621	-
Midostaurin	[redacted]	3.600	[redacted]	[redacted]	0.87	[redacted]	291,526	Oral azacitidine is dominant

Abbreviations: FLT3 = fms-like tyrosine kinase 3; ICER = incremental cost-effectiveness ratio; LYG = life years gained; QALY = quality-adjusted life year

Table 14. Difference in ICER : Scenario results: Equivalent adverse event rates

Model	ICER (£/QALY)
Base case:	Midostaurin is strictly dominated by oral azacitidine
Equivalent AE rates	Midostaurin is strictly dominated by oral azacitidine
Difference	N/A

Abbreviations: ICER = incremental cost-effectiveness ratio, AE = adverse events

5. Trials and data analysis

A18. During induction treatment, it is expected that AML patients will have regular blood transfusions. Red blood cell (RBC) and platelet infusions, use of an erythropoiesis stimulating agent (ESA), antibiotics, etcetera were allowed as concomitant best supportive care medications in the QUAZAR AML-001 trial. Please supply a list of concomitant medications with patient numbers by arm.

All patients who were eligible to enrol in the QUAZAR AML-001 trial were already in first CR/CRi following induction therapy with or without consolidation chemotherapy. Therefore, the best supportive care medications captured in the QUAZAR AML-001 trial did not include the induction phase. In the QUAZAR AML-001 trial, concomitant medications were defined as non-study medications started after the date of randomisation but before the end of the study treatment period (ie, maintenance), or medications started on or before the date of randomisation that ended or remained ongoing during the study treatment period.[3] A list of concomitant medications with patient numbers by treatment arm is provided in Table 14.1.9.1.1 of the clinical study report (data on file).[3]

A19. The company submission states that “Patients with subsequent evidence of AML relapse (≥5% and ≤15% blasts in the peripheral blood or bone marrow) had the

option to continue treatment with an extended dose schedule to 300 mg QD for 21 days, provided it was in the best interest of the patient to do so as judged by the INV"

a. Please justify the choice of 21 days.

The proposed mechanism of action of oral azacitidine is to expose potential leukemic cells extended drug exposure over the treatment cycle. Both 14 and 21 days of treatment with oral azacitidine were studied in initial Phase I/II studies in patients with AML/MDS/CMML, and the phase 3 study design was informed by this early phase work.[16, 17]

b. Please provide details about how the INV made their judgment about what was in the best interest of the patient.

This was based on the opinion of the treating physician and the patients desire to continue treatment with an extended dose schedule. In part this would depend on the overall condition of the patient and their performance status, the ability to and ease of obtaining other therapy options for the patient, and the preferred option of the patient.

A20. Only 8% of patients in the QUAZAR AML-001 study were ECOG 2-3 performance status. Please specify the parameters that disqualified patients in the trial from being eligible for HSCT.

Performance status is an important consideration when assessing a patient's eligibility for intensive induction chemotherapy.[2] The population under investigation within QUAZAR-AML-001 represents a cohort that were deemed fit enough to undergo intensive induction chemotherapy, and that achieved CR/CRi. This may be

reflected in the lower proportion of patients recruited with ECOG 2-3 performance status compared to ECOG 0-1.

The reasons for transplant ineligibility of patients at baseline in the QUAZAR AML-001 trial are provided in Table 15.

Table 15. Disease baseline characteristics – transplant ineligibility

Parameter	Oral azacitidine (N=238)	Placebo (N=234)	Total (N=472)
Reason ineligible for transplant[a] - n (%)			
Age	154 (64.7)	152 (65.0)	306 (64.8)
Comorbidities	52 (21.8)	50 (21.4)	102 (21.6)
Performance Status	14 (5.9)	9 (3.8)	23 (4.9)
Not acceptable or available donor	37 (15.5)	35 (15.0)	72 (15.3)
Subject decision	19 (8.0)	32 (13.7)	51 (10.8)
Unfavorable cytogenetics	6 (2.5)	10 (4.3)	16 (3.4)
Other	28 (11.8)	21 (9.0)	49 (10.4)

[a] A subject may have had more than 1 reason.
Source: Supplementary appendix to Wei et al. 2020.[21]

A21. Table B.2.7. of the CS details "subsequent AML therapies reported for ≥ 10% of subjects in either treatment group" for patients in the QUAZAR AML-001 study.

a. Please clarify if by 'subsequent', these therapies were administered during the maintenance treatment period.

Subsequent AML therapies did not occur with concomitant AML maintenance treatment in the QUAZAR-AML-001 study.

b. If so, owing to the trial's eligibility criteria for concomitant medications excluding cytotoxic chemotherapeutic agents or experimental agents, please justify this protocol violation.

Please note, this question is no longer applicable given the response to question A21.a. above.

c. Could the company please shed more light on the AML therapies used by patients in this trial.

An overview of subsequent AML therapies is provided in the below Table 16.

Table 16. Subsequent AML therapies used after the treatment phase of QUAZAR AML-001 trial (ITT population)

ATC1 Dictionary Level Preferred Name[a]	Oral azacitidine (N=238)		Placebo (N=234)		Total (N=472)	
	N	%	N	%	N	%
Subjects with at least one subsequent AML Therapy	137	57.6	170	72.6	307	65.0
Intensive chemotherapy	69	29.0	88	37.6	157	33.3
Low-intensity therapy	94	39.5	110	47.0	204	43.2
Other	15	6.3	19	8.1	34	7.2
Missing	0	0	1	0.4	1	0.2
Antineoplastic and immunomodulating agents	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
Cytarabine	83	34.9	92	39.3	175	37.1
Fludarabine	32	13.4	48	20.5	80	16.9
Azacitidine	31	13.0	47	20.1	78	16.5
Hydroxycarbamide	28	11.8	34	14.5	62	13.1
Mitoxantrone	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
Idarubicin	20	8.4	33	14.1	53	11.2
Decitabine	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
Etoposide	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
Mercaptopurine	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
Venetoclax	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
Cyclophosphamide	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
Busulfan	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
Cladribine	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]

Tioguanine	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
Amsacrine	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
Granulocyte colony stimulating factor	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
Melphalan	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
Methotrexate	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
Antithymocyte immunoglobulin	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
Clofarabine	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
Filgrastim	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
Sorafenib	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
Thiotepa	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
Alemtuzumab	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
Daunorubicin	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
Daunorubicin hydrochloride	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
Enasidenib	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
Fludarabine phosphate	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
Idasanutlin	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
Imatinib	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
Ivosidenib	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
Lenalidomide	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
Mitoxantrone hydrochloride	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
Other antineoplastic agents	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
Quizartinib	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
Romidepsin	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
Selinexor	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]

Sorafenib tosilate	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
Tretinoin	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
Vinblastine	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
Carmustine	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
Ciclosporin	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
Dasatinib monohydrate	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
Doxorubicin	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
Gemtuzumab ozogamicin	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
Quadecitabine	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
Midostaurin	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
Plerixafor	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
Treosulfan	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
Various	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
Investigational drug	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
Radiotherapy	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
Blood and blood forming organs	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
Other blood products	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
Eltrombopag	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
Dermatologicals	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
Tretinoin	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
Systemic hormonal preparations, excl. sex hormones and insulins	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
Dexamethasone	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]

Subsequent AML therapy is defined as any therapy collected on the CRF for subsequent therapy for AML.

[a] Coded using WHO Drug dictionary version March 2019. A subject with multiple occurrences of a drug class or drug preferred name is counted only once in the specific ATC classification or preferred name, respectively.

Source: BMS, 2020 (Data on file); FDA, 2020.[3, 30]

d. Is it reasonable that the company used the 2019 data cut-off point (instead of the 2020 data cut-off point) to inform relapse-free survival?

For the 2019 data cut-off point relapse-free survival (RFS) data was collected according to the pre-defined study methodology. This was not the case for the 2020 data cut-off point which relied on data from the Extension Phase (EP) of the QUAZAR AML-001 study.

In the EP, the two study arms were unblinded. Any patient who provided additional consent could enter the EP and was followed-up for survival for at least another 12 months until death, withdrawal of consent, study closure or loss to follow-up. While patients on oral azacitidine could continue study treatment, patients on placebo discontinued treatment. Cross-over from the placebo group to the oral azacitidine group was not permitted in the EP.

The objective of the EP was to follow-up on survival of trial participant. Since it was not required in the EP to collect bone marrow and peripheral blood samples.

While there were some isolated bone marrow or peripheral blood samples recorded after the July 2019 database lock (and unblinding of the study), [redacted]

Since bone marrow and peripheral blood samples were not collected routinely and also not according to the methodology as required in the treatment phase of the trial, we consider that reliable RFS data can only be drawn from the July 2019 data cut-off point.

A22. Only 9% of patients included in the QUAZAR AML-001 study had secondary AML.As secondary AML is associated with poor response to chemotherapy and poor outcomes, please perform overall survival (OS) subgroup analysis for type of AML.

The subgroup analysis for the secondary AML population demonstrates improved OS for the oral azacitidine group compared with the placebo group with a median OS of [redacted]months and [redacted]months, respectively (HR: [redacted]). The Kaplan-Meier curves of this subgroup are provided in Figure 3. However, due to the small number of patients with secondary AML, the data should be interpreted with caution.

Figure 3 Kaplan-Meier Plot of Overall Survival - ITT Population with Secondary AML at Baseline (September 2020 data cut-off)

A23. Priority question. Generalisability to the UK clinical practice setting. [redacted] patients (out of a total of 472) in the QUAZAR AML-001 study were recruited from UK study sites.

a. **Please discuss the generalisability of the study baseline disease characteristics to what is expected to be the UK's AML (ineligible for HSCT following induction) population.**

BMS considers that the baseline disease demographics of the QUAZAR-AML-001 trial, which includes a majority of patients from Europe (65%), align to the UK's AML (ineligible for HSCT following induction) population with some caveats:[22]

Age:

The limitation of age to ≥ 55 years for inclusion in the study will of course not account for patients below this age cut off. The recruitment and randomisation to treatment arms have, nonetheless, resulted in a representative median age (range) of 68 (55-86) years for the incident AML population.[22]

Cytogenetic Risk:

The study included patients with intermediate and poor risk cytogenetics (according to National Comprehensive Cancer Network 2011 guidelines).[22] Patients with favourable risk cytogenetics are less likely to proceed to HSCT[2] in first CR (and therefore would be within scope), but as a group are not represented in the QUAZAR data.

We sought feedback from the two UK AML experts regarding the disease and baseline characteristics. Their responses are shown in Table 17 and Table 18. Overall, the percentages for UK clinical practice as provided by the experts align well with the ITT population and the EU subgroup.

Table 17. Disease baseline characteristics across ITT, EU subgroup, and UK subgroup compared to UK clinical practice

Parameter	ITT (N=472)	EU (N=314)	UK ([redacted])	UK clinical practice[a]

Initial AML classification, n (%)				
AML with recurrent genetic abnormalities	85 (18)	[redacted]	[redacted]	[redacted]
AML with myelodysplasia - related changes	91 (19)	[redacted]	[redacted]	[redacted]
Therapy-related myeloid neoplasms	2 (0.4)	[redacted]	[redacted]	[redacted]
AML not otherwise specified	293 (62)	[redacted]	[redacted]	[redacted]
Missing	1 (0.2)	[redacted]	[redacted]	[redacted]
Type of AML, n (%)				
Primary (de novo)	429 (91)	[redacted]	[redacted]	[redacted]
Secondary	43 (9)	[redacted]	[redacted]	[redacted]
Time since original AML diagnosis (months) to randomisation				
Median (range)	4.2 (1.4–10.9)	[redacted]	[redacted]	[redacted]
Prior history of MDS/CMML, n (%)				
Primary	37 (8)	[redacted]	[redacted]	[redacted]
Secondary	0	[redacted]	[redacted]	[redacted]
Missing	2 (0.4)	[redacted]	[redacted]	[redacted]
ECOG performance status, n (%)				
Grade 0	227 (48)	[redacted]	[redacted]	[redacted]
Grade 1	207 (44)	[redacted]	[redacted]	[redacted]
Grade 2–3	38 (8)	[redacted]	[redacted]	[redacted]
Cytogenetic risk category defined by NCCN at diagnosis, n (%)				
Intermediate	406 (86)	[redacted]	[redacted]	[redacted]
Poor	66 (14)	[redacted]	[redacted]	[redacted]
Reason ineligible for transplant[b], n (%)				
Age	306 (65)	[redacted]	[redacted]	[redacted]
Comorbidities	102 (22)	[redacted]	[redacted]	[redacted]
Performance Status	23 (5)	[redacted]	[redacted]	[redacted]
Not acceptable or available donor	72 (15)	[redacted]	[redacted]	[redacted]
Patient decision	51 (11)	[redacted]	[redacted]	[redacted]
Unfavourable cytogenetics	16 (3)	[redacted]	[redacted]	[redacted]
Other	49 (10)	[redacted]	[redacted]	[redacted]
Received consolidation therapy following induction therapy, n (%)				
Yes	378 (80)	[redacted]	[redacted]	[redacted]
1 Cycle	212 (45)	[redacted]	[redacted]	[redacted]

2 Cycles	147 (31)	[redacted]	[redacted]	[redacted]
3 Cycles	19 (4)	[redacted]	[redacted]	[redacted]
4 Cycles	0	[redacted]	[redacted]	[redacted]
No	94 (20)	[redacted]	[redacted]	[redacted]
MRD status at randomisation[a], n (%)				
Negative	244 (52)	[redacted]	[redacted]	[redacted]
Positive	219 (46)	[redacted]	[redacted]	[redacted]
Missing	9 (2)	[redacted]	[redacted]	[redacted]
Response achieved after induction therapy (with or without consolidation therapy), n (%)				
CR	384 (81)	[redacted]	[redacted]	[redacted]
CRi	88 (19)	[redacted]	[redacted]	[redacted]
CR/CRi status at randomization, n (%)				
CR	360 (76)	[redacted]	[redacted]	[redacted]
CRi	94 (20)	[redacted]	[redacted]	[redacted]
Not in CR/CRi	16 (3)	[redacted]	[redacted]	[redacted]
Missing	2 (0.4)	[redacted]	[redacted]	[redacted]
Time from start of induction therapy to randomisation, months				
Median (range)	4.0 (1.3–15.1)	[redacted]	[redacted]	[redacted]
Time from induction therapy to first achieving CR/CRi, days				
Median (range)	35.0 (13.0–455.0)	[redacted]	[redacted]	[redacted]
Time since first achieving CR/CRi to randomisation, days				
Median (range)	85.0 (7.0–263.0)	[redacted]	[redacted]	[redacted]
Bone marrow blasts, %				
Median (range)	2.0 (0.0–6.5)	[redacted]	[redacted]	[redacted]
Peripheral blood blasts, %				
Median (range)	0.0 (0.0–2.0)	[redacted]	[redacted]	[redacted]

[a] Note: The presented percentages were provided by two UK clinical experts in AML. When the experts provided different numbers, these were summarised as a range. E.g., one expert said 50% and the other stated 40%, then this is reported as 40-50%.
[b] Note: Only one expert provided feedback on the percentages.
Source: BMS, internal region analysis; Supplementary appendix to Wei et al., 2020.[21]

Table 18. Demographic characteristics across ITT, EU subgroup, and UK subgroup compared to UK clinical practice

Parameter	ITT (N=472)	EU (N=234)	UK ([redacted])	UK clinical practice[a]
Age (years)				
Mean (SD)	67.9 (5.66)	[redacted]	[redacted]	[redacted]

Median (min, max)	68.0 (55, 86)	[redacted]	[redacted]	[redacted]
Age Category – n (%)				
≥ 55 to < 65 years	134 (28.4)	[redacted]	[redacted]	[redacted]
≥ 65 to < 75 years	286 (60.6)	[redacted]	[redacted]	[redacted]
≥ 75 years	52 (11.0)	[redacted]	[redacted]	[redacted]
≥ 85 years	1 (0.2)	[redacted]	[redacted]	[redacted]
Sex – n (%)				
Male	245 (51.9)	[redacted]	[redacted]	[redacted]
Female	227 (48.1)	[redacted]	[redacted]	[redacted]
Race – n (%)				
White	413 (87.5)	[redacted]	[redacted]	[redacted]

[a] Note: Data presented was provided by one of the two UK clinical experts in AML. The other expert stated that the range of percentages across the ITT, EU subgroup and UK subgroup seem representative for the UK.
Source: BMS, internal region analysis; Supplementary appendix to Wei et al., 2020.[21]

b. Are the patient baseline characteristics from the QUAZAR trial ITT population representative for UK population? It is noted that Number of UK patients is low, would the Europe subset not be more suitable?

BMS considers the data from [redacted] patients recruited from UK sites to be an important contribution to the European and ITT datasets, but due to low numbers not sufficiently robust for modelling. By contrast, the subset of patients from Europe, including those patients from the UK, represent 65% of the total recruitment, and are more suitable for analysis – this has been included as a scenario in the model.

The decision was taken to model the ITT population in the first instance to retain the benefits of the initial randomisation, along with the statistical power associated with the study design.

Whilst we maintain that the ITT population is representative of the UK population (and sought input for clinical experts – please see response to A23.a, there are some rationales that support the Europe subset as an alternative base case analysis:

- Guidelines:
    - The national guidelines for AML were published in 2006,[32] and to some degree have been superseded by more recent guidelines such as the European LeukemiaNet (ELN) 2017,[2] and European Society of Medical Oncology (ESMO) Guidelines 2021.[33] Both of ELN and ESMO guidance

documents have author representation from the UK. In addition, there is direct evidence from local guidelines of the relevance of clinical practice in the rest of Europe, for example, the pan-London AML guidelines indicate that they have been in part derived from the ELN 2017 guidelines.[15] This implies some alignment in the diagnostic and treatment pathway between the UK and the rest of Europe.

- Healthcare provision
    - One clinical expert highlighted broad differences in the healthcare environment in America vs Europe, and the potential impact this may have on patient management. Issues such as insurance and funding of procedures that may be more applicable in America, were considered less of a concern in Europe.

In conclusion, the EU subgroup is a relevant population to analyse regarding the UK context.

c. Please provide the disease baseline characteristics of these patients by study arm.

Disease baseline characteristics specific to UK patients in the QUAZAR trial are provided by treatment arm in the following Table 19 and

Table *20*.

Table 19. Baseline demographics, QUAZAR AML-001 study

Parameter	UK Population			EU Population		
	Oral azacitidine (N=[redacted])	Placebo (N=[redacted])	Total (N=[redacted])	Oral azacitidine (N=167)	Placebo (N=147)	Total (N=314)
Age (years)						
Median (range)	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
Age category, n (%)						
≥55 to <65 years	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]

≥65 to <75 years	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
≥75 years	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
≥85 years	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
Sex, n (%)						
Male	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
Female	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
Race, n (%)						
White	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
Black or African-American	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
Asian	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
Other	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
Missing	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
Ethnicity, n (%)						
Hispanic/Latino	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
Non-Hispanic/Latino	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
Unknown	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
Geographical region, n (%)						
Europe	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]

Source: BMS, internal region analysis.

Table 20. Baseline disease characteristics, QUAZAR AML-001 study

Parameter	UK Population			EU Population		
	Oral azacitidine (N=[redacted])	Placebo (N=[redacted])	Total (N=[redacted])	Oral azacitidine (N=167)	Placebo (N=147)	Total (N=314)
Initial AML classification, n (%)						
AML with recurrent genetic abnormalities	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
AML with myelodysplasia - related changes	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
Therapy-related myeloid neoplasms	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
AML not otherwise specified	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
Missing	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
Type of AML, n (%)						
Primary (de novo)	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]

		[redacted]	[redacted]			
Secondary	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
Time since original AML diagnosis (months) to randomisation						
Median (range)	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
Prior history of MDS/CMML, n (%)						
Primary	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
Secondary	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
Missing	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
ECOG performance status, n (%)						
Grade 0	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
Grade 1	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
Grade 2–3	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
Cytogenetic risk category defined by NCCN at diagnosis, n (%)						
Intermediate	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
Poor	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
MRD status at randomisation, n (%)						
Negative	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
Positive	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
Missing	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
Reason ineligible for transplant, n (%)						
Age	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
Comorbidities	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
Performance Status	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
Not acceptable or available donor	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
Patient decision	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
Unfavourable cytogenetics	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
Other	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
Received consolidation therapy following induction therapy						
Yes	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
1 Cycle	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
2 Cycles	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]

3 Cycles	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
4 Cycles	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
No	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
Response achieved after induction therapy (with or without consolidation therapy), n (%)						
CR	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
CRi	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
CR/CRi status at randomisation, n (%)						
CR	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
CRi	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
Not in CR/CRi	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
Missing	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
Time from start of induction therapy to randomisation, months						
Median (range)	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
Time from induction therapy to first achieving CR/CRi, days						
Median (range)	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
Time since first achieving CR/CRi to randomisation, days						
Median (range)	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
Bone marrow blasts, %						
Median (range)	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
Peripheral blood blasts, %						
Median (range)	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]

Source: BMS, internal region analysis.

d. Is the control arm of the QUAZAR representative of UK clinical practice (including consolidation therapy and the definition of BSC)?

As explained in the responses to A5, placebo plus best supportive care was considered as appropriate comparator treatment in the QUAZAR AML-001 trial. The ITT population of the QUAZAR AML-001 trial is considered representative of the AML population in UK clinical practice with the outlined caveats in response to A23.a. We therefore consider that both the intervention as well as the control arm are reflective of the UK AML population.

e. **In the QUAZAR trial bone marrow biopsies were performed every 3 cycles. Please state whether this is standard practice in the UK, and if not, how this feature of the trial affected its generalisability to the UK setting.**

This is not standard practice in the UK, but facilitated the sensitivity required to capture registrational quality relapse data whilst supporting the care of patients in the controlled study environment**.** The product licence does not mandate bone marrow biopsies, nor specify testing intervals to assess for relapse.[23]

The purpose of the bone marrow biopsy in the QUAZAR-AML-001 trial was to identify relapse. After achieving remission from AML, and as demonstrated in the QUAZAR study, relapse occurs early and in most patients. Whilst we do not expect bone marrows every 3 months to be practical outside of a trial setting, the signs, and symptoms of relapse, along with diagnostic tools including blood tests, bone marrow biopsies and more advanced MRD monitoring techniques should facilitate informed clinical decision making regarding the initial prescription and subsequent continuation of oral azacitidine.

Consultation with 2 clinical experts in AML confirmed that regular bone marrow samples during maintenance treatment is not in line with standard or care, nor their view of how relapse would be initially identified in the UK setting should oral azacitidine be a treatment option. Both clinicians suggest that regular blood test monitoring, that typically occurs in current clinical practice, yields the signs of relapse that would prompt further investigation and action. In addition, where specific mutations exist e.g. NPM1, there may be additional MRD monitoring techniques available that can also provide evidence of early relapse (this is not standardised across the UK).

f. **Please provide scenario analyses using the Europe subset of the trial data. Please explain differences in results and potential issues with both the ITT and Europe subset and their generalisability to UK NHS practice.**

The details of this scenario are included in the response to question B4.

Quality of life questions

A24. Priority question: HRQoL and fatigue were measured on day 1 of each 28-day cycle. Given that oral azacitidine was given in the first 14 days of each 28-day cycle (and patients are off treatment for the remaining 14 days thereafter), treatment related AEs may not be captured in the HRQoL measurement. Please justify the impact and direction of potential bias measuring HRQoL and fatigue on the first day of each cycle.

In theory, assessing HRQoL at the start of each treatment cycle is less likely to capture the effect of treatment-related symptomatic AEs on HRQoL, especially if AEs are short-lived or when treatment cycles are long. Therefore, detrimental effects on HRQoL caused by AEs may be more likely to be underestimated for oral azacitidine (vs. placebo/SOC). Despite this, it is believed that the impact would be marginal. The negative impact of AEs is not anticipated to have a long-lasting effect in most cases, as dose would likely be modified to address the issue. Those AEs with longer-lasting effects would be captured by the HRQoL instrument on day 1 of each 28-day cycle.

To mitigate any risk that treatment-related AEs were not fully captured in the HRQoL measurement from the QUAZAR AML-001 trial, AE disutilities were applied to the health state utility values in the base case. For example, a disutility of 0.115 is applied over the duration of a week to account for patients with grade 3 or 4 fatigue. This approach ensures that the HRQoL impact for patients who experienced fatigue and other treatment-related AEs between measurement intervals (the first day of each cycle) would still be captured.

6. *Indirect treatment comparisons (ITC)*

A25. Priority question: Section B.2.9.1 of the CS states that two studies were identified from the SLR for the indirect comparison. Appendix D is cited for further details of this SLR. However, the SLR inclusion and exclusion criteria (Table B.5.3 in Appendix D) lists the only intervention as "Oral azacitidine", which would also imply the exclusion of any studies of any treatment not compared to oral azacitidine.

a. **Please clarify if this SLR reported in Appendix D is the one use to obtain studies for the indirect comparison.**

The SLR eligibility criteria outlined in Appendix D were used to identify all trials assessing the efficacy and safety of maintenance therapies in AML. Studies included in the SLR were then assessed for the feasibility to be included in an ITC versus Oral azacitidine. Eligibility criteria for the ITC were stricter than the eligibility criteria for the SLR in order to align with the QUAZAR AML-001 trial.

b. **Please clarify the eligibility criteria used to identify studies relevant for the indirect comparison.**

The SLR focused on randomized controlled trials of adult patients (≥18 years) with de novo AML or AML secondary to prior myelodysplastic disease who are in CR or CRi receiving any maintenance therapy. Maintenance therapy was defined as treatment with lower intensity than, and administered after induction therapy, with or without consolidation.[34] Studies included in the SLR were not limited to Oral azacitidine trials only, with studies examining treatments listed as comparators also eligible for inclusion (ie, the SLR included studies assessing oral azacitidine or treatments listed as comparators).

The comparators included in the final scope of this technology appraisal included midostaurin in a subgroup of patients with *FLT3*-mutation positive AML and established clinical management without oral azacitidine (a 'watch and wait' strategy with best supportive care). As the efficacy of the 'watch and wait' approach was evaluated in the QUAZAR AML-001 trial, evidence informing this comparator was not

explored further. However, as the RATIFY trial was the only study informing the efficacy of midostaurin in the SLR, all associated publications were reviewed for maintenance-specific data. Following an assessment, it was determined that Larson et al. 2021[35] was the only publication reporting maintenance-specific data for midostaurin, resulting in its inclusion. All other studies were excluded from evaluation.

c. **Please provide the feasibility assessments for the studies identified in the SLR as they were assessed according to the aforementioned eligibility criteria.**

As mentioned in the response above (A25b), with the exception of the RATIFY trial (Larson et al. 2021),[35] all studies identified in the SLR were excluded from an assessment of feasibility for an ITC as the final scope of this technology appraisal included the following comparators: midostaurin and 'watch and wait'. A detailed feasibility assessment of oral azacitidine and RATIFY is provided in Document B, Appendix D.1.2.2 of the CS.

d. **Please clarify how the RATIFY trial, which did not include oral azacitidine, was selected based on this feasibility assessment.**

The RATIFY trial included midostaurin as a maintenance therapy in adults with AML who have achieved CR following induction therapy. Midostaurin is a treatment included as a comparator in the SLR inclusion criteria.

e. **Please confirm that there were no other randomised control trials or comparative studies by which an indirect comparison with midostaurin could have been achieved.**

The RATIFY trial was the only study identified in the SLR that examined midostaurin as a maintenance treatment in AML.

A26. Priority question: A feasibility assessment of the RATIFY trial for an indirect comparison with QUAZAR AML-001 was conducted in Section B.2.9.2 of the CS.

a. The ERG notes that several sources of incomparability between QUAZAR AML-001 and RATIFY trials were identified in the assessment. Could the company please discuss how each of these sources of heterogeneity affect the validity of the ITC results.

An assessment of the evidence base identified substantial heterogeneity in the study characteristics of the QUAZAR AML-001 and RATIFY trials. Specifically, the studies differed across numerous characteristics including study design, patient eligibility criteria, baseline characteristics, and outcome definitions. The following sections highlight key differences across variables that compromise the validity of indirect comparisons of oral azacitidine and midostaurin.

Study Design:

The study design of the QUAZAR AML-001 and RATIFY trials shared some similarities; however, the trials differed substantially across several important characteristics. Although both trials included treatment with maintenance therapy, only QUAZAR AML-001 was prospectively designed to evaluate the efficacy of maintenance therapy in comparison with placebo. Specifically, patients included in the QUAZAR AML-001 trial were randomized to either maintenance therapy with oral AZA or placebo after achieving CR/CRi following intensive induction chemotherapy, with or without consolidation chemotherapy. In contrast, patients in the primary analysis of RATIFY trial were randomized to receive treatment with midostaurin or placebo in combination with intensive induction and consolidation chemotherapy. Therefore, the RATIFY trial was designed to assess the addition of midostaurin to standard chemotherapy versus chemotherapy alone as part of induction and consolidation. Although the RATIFY trial included a maintenance therapy phase, the 205 patients who entered the maintenance phase were not re-randomized prior to the start of maintenance therapy. The lack of randomization obscures the efficacy of midostaurin as a maintenance therapy due to selection bias, since systematic

baseline differences between groups were not controlled for (eg, prior midostaurin therapy in induction and consolidation for the treatment group), resulting in potentially biased estimates of efficacy.

Inclusion Criteria:

An assessment of the inclusion criteria of the QUAZAR AML-001 and RATIFY studies identified several key differences across trials. Although both trials limited study entry to adult patients with AML, the RATIFY trial included substantially younger patients (≤59 years) in comparison with the QUAZAR AML-001 trial (≥55 years). While all patients in the RATIFY trial were ≤59 years, fewer than 10% of patients in QUAZAR AML-001 were ≤59 years (n = 42). In AML, age is a well-known prognostic factor and a potential effect modifier, with younger patients achieving substantially higher five-year survival rates in comparison with older adults.[36] Substantial differences across studies in age, a key variable in predicting patient prognosis, may violate the exchangeability assumption and bias any indirect estimates against oral AZA.

In addition, unlike the QUAZAR-AML-001 trial, patients included in the RATIFY trial had not achieved CR/CRi at study entry. The RATIFY trial included induction and consolidation chemotherapy as part of the study treatment plan, with entry into the 12-month maintenance phase limited to patients who were in CR after receiving consolidation chemotherapy. Due to this design, all patients entering the maintenance phase of the RATIFY trial had received consolidation chemotherapy, whereas patients in QUAZAR AML-001 trial were allowed to commence maintenance therapy with or without consolidation chemotherapy; in the QUAZAR AML-001 study, 80% of patients received at least one course of consolidation chemotherapy before trial entry. Previous studies have shown that a history of consolidation chemotherapy may be associated with a more favorable disease prognosis.[37] Notably, a post-hoc analysis of the QUAZAR AML-001 trial demonstrated that use of consolidation was generally associated with nominal improvements in OS and RFS within both the oral AZA and placebo treatment arms.[26] Such findings suggest that differences across this variable (ie, history of

consolidation chemotherapy) may modify treatment effect and bias indirect estimates of comparative efficacy against oral AZA.

Exclusion Criteria:

Several key differences were identified when comparing the exclusion criteria across the included trials. At study screening, the RATIFY trial excluded patients with FLT3 mutation-negative AML and thereby limited the study population to patients with the following FLT3 mutational subtypes: the FLT3 internal tandem duplication (ITD), associated with a poor prognosis owing to a high relapse rate, and the FLT3 tyrosine kinase domain (TKD) point mutation, of which the effect on prognosis is uncertain.[38-40] In the QUAZAR AML-001 trial, patients were included regardless of their mutational status. Although, FLT3 mutation-positive patients were included in the QUAZAR AML-001 study, this patient subgroup (n = 66, 14% of total study population) was not the focus of the trial and represents a key difference between studies.

The QUAZAR AML-001 trial excluded patients with favorable-risk cytogenetic characteristics and included patients with intermediate-risk or poor-risk cytogenetic characteristics. In contrast, patients with favorable-risk cytogenetic characteristics were included in the RATIFY trial, along with patients stratified to other cytogenetic risk categories. In AML, cytogenetic risk category is a prognostic factor and a potential treatment effect modifier, with intermediate-risk and poor-risk cytogenetic characteristics associated with a poor disease prognosis.[2] Therefore, the inclusion of patients with favorable risk cytogenetics in the RATIFY trial limits the comparability between studies, with any indirect comparisons potentially biasing results against oral AZA.

The QUAZAR AML-001 trial excluded patients who were eligible for allogeneic bone marrow transplant or HSCT at screening, whereas eligibility for HSCT was not a formal exclusion criterion in the RATIFY trial. Notably, HSCT was performed at some point during the disease course in 59% of patients treated with midostaurin in the full study population of RATIFY. Among patients treated with midostaurin maintenance

therapy, 5.8% received HSCT, with all procedures occurring during first CR; it is unknown how many were treated with HSCT in subsequent lines of therapy. In QUAZAR, 6.3% of patients treated with oral AZA underwent subsequent HSCT, the majority of which occurred post-relapse. While reasons for receipt of HSCT may vary (eg, genetic features, level of fitness, availability of suitable donors etc.), in the RATIFY trial, transplants were mainly motivated by favorable results in high-risk patients with activating FLT3 mutations (ie, patients with an ITD mutation subtype and a high-allelic ratio).[35] The selection of patients with an adverse risk profile to receive transplants in first CR and subsequently be excluded from maintenance therapy may introduce bias that cannot be adjusted for with the available data, as this may result in a patient cohort with a more favorable risk profile in the maintenance phase. Notably, this is reflected in the reported characteristics of patients proceeding to maintenance therapy versus those who did not, respectively: proportion of patients with an FLT3-ITD mutation and a high allelic burden was 21% and 33%. Therefore, the differential HSCT criterion across trials may represent an additional source of bias.

Baseline Characteristics:

Patient baseline characteristics varied substantially between the intention-to-treat (ITT) patient populations of the QUAZAR AML-001 and RATIFY studies. As mentioned above, the RATIFY trial was not prospectively designed to determine the independent effect of midostaurin as maintenance therapy, as patients were randomized to induction chemotherapy and were not re-randomized following completion of consolidation chemotherapy. As such, patient characteristics of the RATIFY trial were predominantly reported at baseline (ie, at randomization to induction), with reporting of characteristics for patients who entered the 12-month maintenance therapy phase limited to only a few variables (age, sex, *FLT3* mutational subtype, and cytogenetic risk). In contrast, the QUAZAR AML-001 trial reported patient characteristics at randomization to maintenance therapy. This represents a key distinction and limits comparability between studies, as summary-level data is reported for different patient populations; that is, the RATIFY trial reports summary-level data for patients with untreated AML (with insufficient data reported

for patients commencing maintenance therapy), and the QUAZAR AML-001 trial reports summary-level data for patients with AML who have achieved CR/CRi following induction chemotherapy (with or without consolidation chemotherapy). In addition, the trials differed significantly in their eligibility criteria for patient age and cytogenetic risk. These differences are reflected in reported patient characteristics in the full study population of QUAZAR AML-001 and the subset of patients who entered the 12-month maintenance phase of the RATIFY trial, respectively: median age was 68.0 years (range: 55.0-86.0 years) and 49.0 years (range: 19.0-60.0 years); proportion of patients stratified to cytogenetic risk categories was 0% versus 57.0% for favorable-risk and 14.0% versus 17.9% for poor-risk characteristics. A similar trend is observed when comparing the full study populations of the QUAZAR AML-001 and RATIFY trials. In AML, both age and cytogenetic risk are known prognostic factors and potential effect modifiers, with younger patients and favorable cytogenetic characteristics leading to a better disease prognosis. Therefore, differences across these variables underscore the heterogeneity between study populations and favor the RATIFY trial, potentially biasing estimates of comparative efficacy against oral AZA.

Outcome Definitions:

In comparison to outcomes reported in the primary analysis of the RATIFY trial, time-to-event outcomes in the landmark analyses were similarly defined to those reported in QUAZAR AML-001 trial (see Table B.5.9 of the company submission [CS]; however, several limitations persist: 1) patients in the RATIFY trial were not re-randomized prior to the start of maintenance therapy, 2) the dataset was not statistically powered to isolate the clinical benefit gained from the maintenance phase of the trial. This limits comparability to survival outcomes reported in the QUAZAR AML-001 trial because results derived from a lack of randomization may incorporate bias (eg, selection bias), and the lack of sufficient statistical power adds uncertainty around the reported effect size of survival outcomes (eg, inflated effect size estimation and low reproducibility), which, in turn, leads to uncertainty when deriving estimates of comparative efficacy between oral AZA and midostaurin, such

that only unusually large differences in outcomes between therapies would be deemed statistically significant.

b. Please provide justification that the anchored population adjustment delivers effect estimates that are applicable to the final scope population.

Anchored Bucher ITCs were performed to compare Oral AZA with midostaurin. Midostaurin is indicated for the treatment of adult patients with AML who have a *FLT3* mutation. In order to compare Oral AZA to midostaurin, the inclusion and exclusion criteria were aligned between the two trials.

Patients from QUAZAR AML-001 were removed from the IPD if they did not satisfy the eligibility criteria used in the RATIFY trial. Specifically, QUAZAR AML-001 included patients without *FLT3* mutations and with CRi but the RATIFY trial did not. Given the substantial differences between the trial populations and lack of baseline characteristics reported for the RATIFY maintenance subgroup, the populations were matched as closely as possible. Matching on age was not feasible as sample size would be greatly reduced (only [redacted] individuals would remain in the QUAZAR AML-001 trial as RATIFY primarily included younger patients). Furthermore, many patients in the RATIFY trial were HSCT eligible at study screening (HSCT was performed in over 20% of patients during first CR). Since patients in QUAZAR AML-001 were ineligible for transplant at study screening, matching on HSCT eligibility was not feasible.

Therefore, the anchored Bucher ITC provides an estimate of Oral AZA versus midostaurin for patients with AML and a *FLT3* mutation who have achieved first CR after induction with intensive chemotherapy with or without consolidation and who are ineligible for transplant.

c. Please provide an assessment of the effect of randomisation in the RATIFY trial not occurring at the start of the period of analysis i.e., start of the maintenance phase.

Patients in the primary analysis of RATIFY trial were randomized to receive treatment with midostaurin or placebo in combination with intensive induction and consolidation chemotherapy. Therefore, the RATIFY trial was designed to assess the addition of midostaurin to standard chemotherapy versus chemotherapy alone as part of induction and consolidation.

Although the RATIFY trial included a maintenance therapy phase, the 205 patients who entered the maintenance phase were not re-randomized prior to the start of maintenance therapy. The lack of randomization obscures the efficacy of midostaurin as a maintenance therapy due to selection bias, since systematic baseline differences between groups were not controlled for (eg, prior midostaurin therapy in induction and consolidation for the treatment group), resulting in potentially biased estimates of efficacy.

Furthermore, patients in the RATIFY trial who remained in remission after consolidation were eligible to enter the maintenance phase. In contrast, time-to-event outcome estimates reported in the QUAZAR AML-001 trial were designed to be conservative, as patients enrolled in the trial may have achieved CR/CRi up to four months (±7 days) prior to randomization. In the QUAZAR AML-001 study, patient recruitment occurred after patients received intensive induction chemotherapy (with or without consolidation), thereby introducing a time lag between achievement of first CR/CRi and randomization. The mean time from achievement of first CR/CRi to randomization was [redacted]. Since patients in the RATIFY trial are eligible to enter the maintenance phase soon after completion of consolidation therapy, survival may appear greater in RATIFY patients and a comparison of time-to-event outcomes (eg, OS and RFS) between the two trials may bias results against the QUAZAR AML-001 study.

A27. Priority question: As per NICE TSD 18, please provide justification that the anchored population adjustment is less likely to produce biased estimates when compared to standard indirect comparisons.

Anchored Bucher ITCs were performed to compare Oral AZA with Midostaurin because no head-to-head evidence comparing these treatments exists. Indirect treatment comparison methods using IPD can play an important role in the generation of clinical evidence in the absence of RCTs since it can be leveraged to match the inclusion/exclusion criteria of comparator trials. Anchored indirect comparisons improve on naïve comparisons by relying on comparing relative effects between two treatments, anchored through a common comparator. Naïve comparisons require much stronger assumptions; all prognostic variables and all effect modifiers are assumed to be accounted for and correctly specified in unanchored ITCs (NICE DSU TSD 18, Sections 3.1.1 and 4.1.4). Thus, anchored comparisons are preferred as they only rely on the assumption that trials are similar with respect to effect modifiers and thus derived effect estimates constitute higher-grade evidence compared to unanchored comparisons (NICE DSU TSD 18, Section 3.1.1).[41]

Population adjustment methods (eg, MAIC, STC) were not used to compare Oral AZA with Midostaurin. Since MAIC uses a reweighting method, and therefore does not permit extrapolation, bias can only be completely removed when the population of the study with summary-level data are entirely contained within the population of the IPD study. Additionally, population adjustment methods may not be needed if the imbalance in treatment effect modifiers is small. The NICE DSU TSD 18 Section 4.2.3 recommends that population adjustment methods should only be performed if they are likely to produce less biased estimates of treatment differences than those that could be achieved through standard methods.[41] Therefore, STCs would be the preferred option when there is minimal overlap between study populations since it can extrapolate beyond the range of the IPD.[42] STCs can also be used to predict outcomes wherein it would be impossible for a MAIC to do so. The means and standard deviations of age vary greatly between the QUAZAR AML-001 and RATIFY populations and this could potentially be addressed through the ability of a STC to extrapolate beyond the data range. However, as the estimate gets further from the

area of central support in the QUAZAR AML-001 trial, the variance in the estimate will grow and validity of point estimates rely on unverifiable assumptions regarding the form of the extrapolation. The mean age in RATIFY (42.7, standard deviation = 7.93) was lower than the minimum age for inclusion in QUAZAR and assuming it was approximately normally distributed would suggest only [redacted] overlap [redacted] in the two trials. Adjustment was therefore not considered reasonable given the need to rely on extensive extrapolation and the resulting unreasonable uncertainty.

Further, all patients in RATIFY contained a *FLT3* mutation so it was deemed most appropriate to restrict the QUAZAR AML-001 trial to this subpopulation. Sex was already very similar between the two trial populations so an STC was deemed unnecessary to adjust for this covariate.

Lastly, stem cell transplant was not considered for matching since adjusting for post-treatment variables breaks randomization and because not enough patients received HSCT in QUAZAR to allow for reasonable estimate of that effect. Further, no standard method exists to address for imbalances in post-treatment variables since differential arm adjustment further creates issues with from a randomization standpoint.

Given the substantial differences between the QUAZAR AML-001 and RATIFY trials in terms of study eligibility criteria and baseline patient characteristics, population adjustment methods would produce effect estimates with a high uncertainty. Therefore, anchored Bucher ITCs in the matched population were used to compare Oral AZA and midostaurin. Anchored comparisons relax the assumption of balanced prognostic variables between the trials that unanchored comparisons require and therefore provide a less biased effect estimate.

Section B: Clarification on cost-effectiveness data

7. Model structure

B1. Priority question: HSCT was implicitly included in the modelling through the survival analysis of the QUAZAR AML-001 ITT population (of which a proportion of patients received HSCT at some point). In addition, costs and disutilities associated with undergoing HSCT were included in the modelling. The ERG is concerned that this way of handling HSCT in the model may cause biases, one because survival analysis of OS and RFS may be biased, and two because no benefit in health-related quality of life post HSCT is captured in the model.

a. **Survival analyses may be biased when HSCT patients are included in the population as their hazard rates over time for OS and RFS would be expected to differ compared with those patients not receiving HSCT (indeed the QUAZAR AML-001 KM curves show long tails). Ideally, survival analyses would be performed by censoring patients who received HSCT. Please provide these survival analyses for both treatment arms (separately or using joint models, as found most appropriate following guidance in NICE DSU TSD 14) and for both OS and RFS. Please report on all the steps outlined in NICE DSU TSD 14.**

The following section focuses on the intention to treat (ITT) population of the QUAZAR AML-001 trial across all recorded time points censored for hematopoietic stem cell transplant (HSCT). This population was included to explore the potential impact of HSCT on modifying survival estimates. In the QUAZAR AML-001 trial, 6.3% of patients treated with oral AZA and 13.7% of patients treated with placebo received HSCT, the majority of these procedures occurred post-relapse. For both overall survival (OS) and relapse-free survival (RFS), analyses censored for HSCT suggested alignment with the assessment for the ITT population, with joint generalized gamma providing the optimal fit for OS and joint log-logistic providing the optimal fit for RFS. For additional details, please see Appendix B.1.

b. **Please provide a scenario analysis or a revised base-case in which HSCT is modelled as a health state. In this analysis please use the results from the survival analyses as above for patients who do not receive HSCT, and the rates of patients receiving HSCT per treatment arm in QUAZAR AML-001. The post-HSCT OS estimates can either be based on QUAZAR AML-001 or other sources.**

Insufficient data were collected to allow modelling of hematopoietic stem cell transplant (HSCT). While patients who underwent HSCT were followed for survival, further parameters around HSCT and following treatments were not recorded, including proportion achieving successful HSCT. These parameters would be required to model HSCT as a separate health state and were not available from the QUAZAR trial dataset. BMS are unaware of any published literature that reported HSCT data in patients who were initially in CR/CRi and ineligible for HSCT following induction therapy with or without consolidation. Addition of HSCT as a separate health state within the model would therefore add considerable uncertainty within the model, without adding any clarity as to the cost-effectiveness of oral azacitidine. This approach was supported by clinical opinion. Moreover, this approach aligns with other models in AML.[43-45]

c. **Further bias is introduced by not capturing the benefit in health-related quality of life that patients may have in the long-term following HSCT. Please justify the exclusion of such a benefit.**

HSCT was not expected to result in a high proportion of cures. In the QUAZAR AML-001 trial, HSCT was primarily conducted following relapse, as a salvage treatment. Although HSCT was conducted with curative intent, the effectiveness of treatment in this relapsed patient group was expected to have little impact on survival. Thus, incorporation of the effectiveness within the overall survival cohort data is therefore considered to capture the natural history of the disease, including those patients who, for the most part, received HSCT as a salvage treatment.

More specifically, in the QUAZAR AML-001 trial, a small proportion of patients treated with oral azacitidine received HSCT post-relapse (6.3%).[3] In models where HSCT is included as a health state, the proportion of patients receiving HSCT tends to be substantially higher and HSCT is often administered during first CR rather than post-relapse. For example, in the RATIFY trial 59% of RYDAPT treated patients underwent HSCT.[24] Among these patients, 47.6% received HSCT during the first CR.[24] Similarly, in the ALFA-0701 trial, 23.7% of MYLOTARG treated patients underwent HSCT.[46] Among these patients, 53.1% received HSCT during the first CR.[46] In contrast, subjects in the QUAZAR AML-001 trial who received another therapy (e.g., HSCT) for AML without documented relapse were censored on the date of the last bone marrow assessment, prior to receiving the other therapy.[3] Thus, the efficacy of these subsequent therapies did not contribute to RFS.[3] The OS hazard ratio [redacted] was maintained when censoring for HSCT (HR: [redacted]; QUAZAR AML-001, Table 14.2.1.5.4).[3] A scenario analysis was conducted where the health state utility value for relapse free survival was based on a weighted average where the proportion of patients on HSCT (6.3%) are assumed to have a utility value equal to 1 and the remaining patients who do not undertake HSCT are assumed to have the base case RFS health state value of [redacted]. The results for the ITT population are provided in Table 21 and a comparison against the base case ICER is provided in Table 22. Results for the FLT3 subgroup are provided in Table 23 and a comparison against the base case ICER is presented in Table 24.

Table 21. Scenario results: RFS utility (weighted average) – ITT population

Technology	Total costs (£)	Total LYG	Total QALYs	Incremental costs (£)	Incremental LYG	Incremental QALYs	ICER versus baseline (£/QALY)
Watch and wait +BSC	[redacted]	2.799	[redacted]	-	-	-	-
Oral azacitidine	[redacted]	3.864	[redacted]	[redacted]	1.06	[redacted]	47,998

Abbreviations: ICER = incremental cost-effectiveness ratio; LYG = life years gained; QALY = quality-adjusted life year; BSC = best supportive care

Table 22. Difference in ICER: Scenario results: RFS utility (weighted average) – ITT population

Model	ICER (£/QALY)
Base case	48,660
Scenario: Utility for RFS: [redacted]	47,998
Difference	£-662 (1.36%)

Abbreviations: ICER = incremental cost-effectiveness ratio, RFS = relapse free survival

Table 23. Scenario results: RFS utility (weighted average) – FLT3 population

Technology	Total costs (£)	Total LYG	Total QALYs	Incremental costs (£)	Incremental LYG	Incremental QALYs	ICER versus baseline (£/QALY)	Pairwise ICER vs oral azacitidine
No active therapy	[redacted]	2.731	[redacted]	-	-	-	-	24,227
Oral azacitidine	[redacted]	4.828	[redacted]	[redacted]	2.10	[redacted]	24,227	-
Midostaurin	[redacted]	3.600	[redacted]	[redacted]	0.87	[redacted]	286,432	Oral azacitidine is dominant

Abbreviations: FLT3 = fms-like tyrosine kinase 3; ICER = incremental cost-effectiveness ratio; LYG = life years gained; QALY = quality-adjusted life year

Table 24. Difference in ICER: Difference in ICER : Scenario results : RFS utility (weighted average) - FLT3 subgroup

Model	ICER (£/QALY)
Base case	Midostaurin is strictly dominated by oral azacitidine
Scenario : Utility for RFS : [redacted]	Midostaurin is strictly dominated by oral azacitidine
Difference in ICER	NA

Abbreviations: ICER = incremental cost-effectiveness ratio, RFS = relapse free survival

d. Please incorporate a utility benefit post HSCT in the current company's base-case using appropriate data sources, and the scenario analysis as requested in b).

Given the reasons provided in responses B1 b and B1 c around data availability and impact of HSCT, BMS do not believe it is appropriate to incorporate a utility benefit post HSCT.

B2. The NICE Decision Support Unit (DSU) technical support document (TSD) 19 recommended the use of state transition models (STMs) alongside partitioned survival models (PSMs) to verify the plausibility of PSM extrapolations and to explore key clinical uncertainties in the extrapolation period.

a. Please justify the use of a partitioned survival approach given the issues highlighted in NICE DSU TSD 19, particularly regarding the extrapolation of PFS and OS while assuming structural independence between these endpoints.
b. If deemed necessary, please use state transition modelling to assist in verifying the plausibility of the PSM extrapolations and to address uncertainties in the extrapolation period (NICE DSU TSD 19, recommendation 11).

B2.a. A Markov model was not considered optimal for the analysis for a number of reasons:

- Whilst parametrising the “relapse free to death” and “relapse free to relapse” may have been feasible, there would be issues estimating the “relapse to death” transition. One issue is that not all patients with relapse can inform inputs to this transition. Only those who relapsed and died during the trial can provide information for estimating this transition. Furthermore, any estimates for this transition would be disproportionally driven by patients who relapsed earlier vs later (due to them generating more follow-up information). Both of these aspects in combination create a sample selection bias towards those who relapse earlier vs later (and are hence the more severe patients within the trial). This adds upward bias to the estimated hazards for this transition. From an incremental perspective, this penalises the more effective therapy as it’s sample of relapses is further skewed towards the more severe end, given limited follow-up.

- The hazard profiles observed in the trial are not constant and are instead time varying. To evaluate a STM in the presence of time varying hazards (for any transition post-baseline) requires a semi-Markov approach and likely an individual simulation-based analysis. This would add additional complexity and computational overhead. Whilst not a reason to abandon the approach in isolation, such costs should carry tangible benefits to the framework, which is not clear.

- The STM and PartSA methods should only potentially differ within the extrapolation period (due to their different structural assumptions). The trial data is quite mature, so this leaves only the tail end of the survival extrapolations for differences to emerge. Therefore, we would expect these differences to be small, and almost certainly fall within the range of the survival outcomes already included in the model across the different options provided.

In the SLR conducted of prior economic evaluations (see appendix G of the company submission), PSMs have been extensively used in economic evaluations in an AML setting. The PSM approach allows for modelling of overall survival (OS) and relapse free survival (RFS) based on observed events, coupled with mature survival data from QUAZAR AML-001 trial, this facilitates the replication of within-trial data to accurately reflect disease progression and the long term expected survival profile of patients treated with oral azacitidine.

We acknowledge the limitations associated with a PSM approach as detailed in the NICE DSU TSD 19. However, the NICE DSU TSD 19[47] does not provide explicit guidance on model selection (i.e PSM vs STM), instead recommends that when using a PSM, the modelling method should be (i) clearly stated, (ii) model choice should be rationalised on the bases of theoretical and practical considerations, (iii) the main structural model assumptions reported and (iv) specific limitations on extrapolation should be highlighted. In essence, it eludes that a PSM is a reasonable approach if a rationale for its use is provided. The rationale is provided in the main submission and reinforced in the paragraph above. Every effort has been made to validate the model structure and extrapolations with clinical experts providing input on the appropriateness of the model structure and extrapolations to ensure both reflect the treatment and disease pathway. Further details on clinical validation of extrapolations are detailed in the response to Question B6c and in the clinical expert summary report.

B2.b. Given that a Markov model was not deemed feasible, it was not considered appropriate to develop and use alongside the PSM.

Intervention and comparator

B3. Priority question: The final scope issued by NICE mentions the following treatments as comparators: midostaurin and established clinical management without oral azacitidine (which may include a "watch and wait" strategy with best supportive care, low dose cytarabine or subcutaneous azacitidine).

a. Please provide an updated economic model and scenario analyses including low dose cytarabine or subcutaneous azacitidine as comparators

(also considering the response to clarification questions A8, A10). Please provide the results of a fully incremental analysis (and updated economic model used for this analysis) with all comparators listed in the scope as comparators modelled separately.

With regards to the use of low dose cytarabine and subcutaneous azacitidine, it is important to note that neither of these treatment options are recommended by NICE for the patient population eligible for maintenance treatment with oral azacitidine, nor is their use mentioned or endorsed as maintenance treatments in either the ELN (2017) or BSCH (2006) guidelines.[2, 32]

The established BMS opinion is that neither of these treatments are legitimate comparators for this appraisal. Both treatments have historically been investigated as AML maintenance options in randomised clinical trials, but neither injectable azacitidine nor low dose cytarabine have demonstrated an overall survival benefit versus comparators in the maintenance setting:

- Injectable azacitidine vs observation/no maintenance (HOVON 97 Trial); this RCT demonstrated a significant improvement in disease-free survival (DFS) after maintenance with injectable azacitidine versus observation/no maintenance (64% vs 42% at 1 year; p=0.04). This study did not show a significant OS benefit (84% vs 70% at 1 year, p=0.69).[48]

- Injectable azacitidine vs BSC (QOLESS AZA-AMLE Trial); this small study (27 patients randomised per treatment arm) did not identify statistically significant differences in DFS or OS between injectable azacitidine and BSC.[49]

- Low-dose cytarabine maintenance therapy vs observation (E5483); this trial reported statistically significant improvements for median DFS (7.4 months vs 3.3 months, p=0.084), but not for median OS (10.8 months vs 7.0 months, p=0.492).[50]

To authenticate our position, we sought expert clinical advice from two UK AML clinicians, who unequivocally confirmed that these treatments are not used in UK clinical practice for AML maintenance. The clinical experts could only provide very limited examples where these treatments could be used in situations resembling

maintenance treatment, such as those patients whose disease was in partial remission, or patients who showed signs of early relapse.

This does not align with the definition of maintenance treatment considered in this appraisal, therefore, low dose cytarabine and subcutaneous azacitidine have been disregarded as relevant comparators to oral azacitidine.

For further details of the clinical expert comments, please refer to the clinical expert summary report.

b. Subsequent treatments include subcutaneous azacitidine and low dose cytarabine as well as salvage chemotherapy. Please provide justification for this and confirm that these treatments are used as subsequent treatments in the English NHS, potentially supporting that with expert opinion.

Subsequent therapies were based on the QUAZAR-AML-001 trial and validated by UK clinical experts. Both UK clinical expert opinions confirmed the use of subcutaneous azacitidine, low dose cytarabine and salvage chemotherapy as subsequent therapy in the treatment pathway in England. Further details of this clinical expert opinion are included in the clinical expert summary report regarding specific treatments and usage in the UK setting for ITT and FLT-3 mutation populations.

Subsequent AML therapies did not occur with concomitant AML maintenance treatment in the QUAZAR-AML-001 study. The study protocol did not influence the clinical decision making regarding the selection of subsequent AML therapies. Data collected, indicated that both intensive and low intensity therapies were given to patients as subsequent AML therapies, which included azacitidine and low dose cytarabine (please see response to question A21.c. above).

Population

B4. Priority question: Please provide scenario analyses using the Europe subset of the QUAZAR AML-001 trial data. Please explain differences in results and potential issues with both the ITT and Europe subset and their generalisability to English NHS practice.

Scenario analysis was only conducted based on the Europe subgroup comparing oral azacitidine with watch and wait + BSC in the ITT population. It was not possible to conduct an analysis in the FLT3 subgroup with only a Europe sample due to sample size limitations restricting the comparison with midostaurin further. A joint generalised gamma distribution was used to model OS and a joint log-logistic distribution used to model RFS. This is further outlined in appendix B.8. Results from the scenario analyses using the Europe subset are provided in Table 25 and disaggregated QALYs by health state and costs by resource use are provided in Table 26 and Table 27, respectively. A comparison between the base case ICER and the ICER based on the Europe subgroup is provided in Table 28. The Europe subset yielded both [redacted] incremental costs and QALYs than the ITT population resulting in a lower ICER of £40,444 compared to the ICER of £48,660 using the ITT population. The [redacted] incremental QALYs in the Europe subset is predominantly driven by the [redacted] QALYs in the RFS: off treatment health state for the watch and wait with BSC treatment arm compared to the ITT population, [redacted] QALYs, respectively. There was a 16.66% decline in the ICER when using the Europe subgroup compared to the base case ITT population. The discussion around the generalizability of the ITT population and Europe subset are provided in the response to question A23.

Table 25. Cost-effectiveness results with oral azacitidine PAS (discounted) - Europe subset

Technology	Total costs (£)	Total LYG	Total QALYs	Incremental costs (£)	Incremental LYG	Incremental QALYs	ICER versus baseline (£/QALY)
Watch and wait with BSC	[redacted]	2.633	[redacted]	-	-	-	-
Oral azacitidine	[redacted]	3.992	[redacted]	[redacted]	1.36	[redacted]	40,444

Abbreviations: ICER = incremental cost-effectiveness ratio; LYG = life years gained; QALY = quality-adjusted life year; BSC = best supportive care

Table 26. Summary of QALY gain by health state - Europe subset

Health state	Total QALYS Oral azacitidine	Total QALYs Watch and wait +BSC	Increment
Total RFS	[redacted]	[redacted]	[redacted]
RFS: On Treatment	[redacted]	[redacted]	[redacted]
RFS: Off Treatment	[redacted]	[redacted]	[redacted]
Relapse	[redacted]	[redacted]	[redacted]
Adverse Event Disutility*	[redacted]	[redacted]	[redacted]
HSCT Disutility*	[redacted]	[redacted]	[redacted]
Total	[redacted]	[redacted]	[redacted]

* These are not health states but components of the generated QALYs . Abbreviations: RFS = relapse free survival; QALYs = quality adjusted life years, HSCT = haematopoietic stem cell transplantation; BSC = best supportive care

Table 27. Summary of predictive resource use by category (PAS price) - Europe subset

Resource	Oral azacitidine	Watch and wait +BSC	Increment
Total RFS	[redacted]	[redacted]	[redacted]
RFS: On Treatment	[redacted]	[redacted]	[redacted]
Drug Costs	[redacted]	[redacted]	[redacted]
Treatment Admin. Costs	[redacted]	[redacted]	[redacted]
Disease Management Costs	[redacted]	[redacted]	[redacted]
Adverse Event Costs	[redacted]	[redacted]	[redacted]
RFS: Off Treatment	[redacted]	[redacted]	[redacted]
Disease Management Costs	[redacted]	[redacted]	[redacted]
Relapse	[redacted]	[redacted]	[redacted]
Disease Management Costs	[redacted]	[redacted]	[redacted]
Subsequent Therapy Costs	[redacted]	[redacted]	[redacted]
SCT Costs	[redacted]	[redacted]	[redacted]
End of Life Costs	[redacted]	[redacted]	[redacted]
Total	[redacted]	[redacted]	[redacted]

Abbreviations: RFS = relapse free survival; HSCT = haematopoietic stem cell transplantation; BSC = best supportive care

Table 28. Scenario analysis: Impact on ICER - Europe subgroup

Model	ICER (£)
Base case	48,660
Scenario: Europe subgroup	40,096
Difference (%)	8,216 (16.88%)

Abbreviations: ICER = incremental cost-effectiveness ratio

B5. Priority question: The majority of patients in the QUAZAR AML-001 trial received only one cycle of consolidation therapy, and approximately 20% of the patients received no consolidation was given. Please perform survival analyses for OS and RFS using the subgroup of patients that received at least one cycle of consolidation therapy following the guidance in NICE DSU TSD 14 and provide an updated model and scenario analysis.

BMS has provided this analysis at the request of the ERG, however, BMS consulted with experts in AML who advised that at time of recruitment into the QUAZAR-AML-001 trial, as well as in current UK clinical practice, there are patients that complete induction chemotherapy (and achieve remission) but may not receive consolidation chemotherapy (as detailed in response to A12a). Therefore, BMS does not consider that excluding these patients from the analysis is appropriate or in accordance with clinical practice in England.

The following survival analysis focuses on the subgroup of the intention to treat (ITT) population of QUAZAR AML-001 who received at least one course of consolidation chemotherapy, hereafter referred to as the ERG consolidation subgroup. In the QUAZAR AML-001 trial, 78% (n=186) of patients treated with oral azacitidine and 82% (n=192) of patients treated with placebo received at least one course of consolidation chemotherapy. For both overall survival (OS) and relapse-free survival (RFS), analyses restricted to the ERG consolidation subgroup suggested alignment with the assessment for the ITT population, with joint generalized gamma providing the optimal fit for OS and joint log-logistic providing the optimal fit for RFS. Selection of curves was based on the criteria described in the NICE DSU TSD 14.[51] Assessments for OS and RFS are presented below.

Overall Survival:

In general, the joint generalized gamma provided among the best statistical fit based on AIC/BIC for the ITT population restricted to the ERG consolidation subgroup (Table 29). Visually the oral azacitidine arm remained apart from the no active treatment arm which corresponds to clinical expectations and the extrapolated tails were clinically plausible. In addition, joint models were generally preferred over

individual models as they have a higher precision due to the higher statistical power of fitting a single model to both treatment arms. Furthermore, given that it was not deemed plausible for curves to cross in both the ITT and FLT-3 populations by UK clinical experts, the use of joint models would be most appropriate. Overall, this aligns with the assessment for the ITT population where the joint generalized gamma was also determined to have the optimal fit based on AIC/BIC, visual inspection, and clinical plausibility (Section B.3.3, Document B of the company submission).

Table 29. Model fit statistics (AIC and BIC) for parametric models of the OS outcome, ITT population restricted to the ERG consolidation subgroup

Parametric Model	AIC	Ranks based on AIC	BIC	Ranks based on BIC
Joint models				
Exponential	[redacted]	[redacted]	[redacted]	[redacted]
Weibull	[redacted]	[redacted]	[redacted]	[redacted]
Log-logistic	[redacted]	[redacted]	[redacted]	[redacted]
Log-normal	[redacted]	[redacted]	[redacted]	[redacted]
Generalized Gamma	[redacted]	[redacted]	[redacted]	[redacted]
Gompertz	[redacted]	[redacted]	[redacted]	[redacted]
Individual models – Oral azacitidine arm				
Exponential	[redacted]	[redacted]	[redacted]	[redacted]
Weibull	[redacted]	[redacted]	[redacted]	[redacted]
Log-logistic	[redacted]	[redacted]	[redacted]	[redacted]
Log-normal	[redacted]	[redacted]	[redacted]	[redacted]
Generalized Gamma	[redacted]	[redacted]	[redacted]	[redacted]
Gompertz	[redacted]	[redacted]	[redacted]	[redacted]
Individual models – Placebo arm				
Exponential	[redacted]	[redacted]	[redacted]	[redacted]
Weibull	[redacted]	[redacted]	[redacted]	[redacted]
Log-logistic	[redacted]	[redacted]	[redacted]	[redacted]
Log-normal	[redacted]	[redacted]	[redacted]	[redacted]
Generalized Gamma	[redacted]	[redacted]	[redacted]	[redacted]
Gompertz	[redacted]	[redacted]	[redacted]	[redacted]

Abbreviations: AIC, Akaike's information criteria; BIC, Bayesian information criterion; ITT, intention-to-treat; OS, overall survival.

Relapse-free Survival:

The selected model for the ITT population restricted to the ERG consolidation subgroup was the joint log-logistic model. This model has good visual fit, and higher precision than the individual models, due to the higher statistical power of fitting a

single model to both treatment arms. From a statistical fit perspective, the log-logistic distribution is the best fitting joint model in terms of AIC and BIC (Table 30). Overall, this aligns with the assessment for the ITT population where joint log-logistic was also determined to have the optimal fit based on AIC/BIC, visual inspection, and clinical plausibility. Furthermore, given that it was not deemed plausible for curves to cross in both the ITT and FLT-3 population by UK clinical experts, joint models were deemed most appropriate (Section B.3.3, Document B of the company submission).

Table 30. Model fit statistics (AIC and BIC) for parametric models of the RFS outcome, ITT population restricted to the ERG consolidation subgroup

Parametric Model	AIC	Ranks based on AIC	BIC	Ranks based on BIC
Joint models				
Exponential	[redacted]	[redacted]	[redacted]	[redacted]
Weibull	[redacted]	[redacted]	[redacted]	[redacted]
Log-logistic	[redacted]	[redacted]	[redacted]	[redacted]
Log-normal	[redacted]	[redacted]	[redacted]	[redacted]
Generalized Gamma	[redacted]	[redacted]	[redacted]	[redacted]
Gompertz	[redacted]	[redacted]	[redacted]	[redacted]
Individual models – Oral azacitidine arm				
Exponential	[redacted]	[redacted]	[redacted]	[redacted]
Weibull	[redacted]	[redacted]	[redacted]	[redacted]
Log-logistic	[redacted]	[redacted]	[redacted]	[redacted]
Log-normal	[redacted]	[redacted]	[redacted]	[redacted]
Generalized Gamma	[redacted]	[redacted]	[redacted]	[redacted]
Gompertz	[redacted]	[redacted]	[redacted]	[redacted]
Individual models – Placebo arm				
Exponential	[redacted]	[redacted]	[redacted]	[redacted]
Weibull	[redacted]	[redacted]	[redacted]	[redacted]
Log-logistic	[redacted]	[redacted]	[redacted]	[redacted]
Log-normal	[redacted]	[redacted]	[redacted]	[redacted]
Generalized Gamma	[redacted]	[redacted]	[redacted]	[redacted]
Gompertz	[redacted]	[redacted]	[redacted]	[redacted]

Abbreviations: AIC, Akaike's information criteria; BIC, Bayesian information criterion; ITT, intention-to-treat; RFS, relapse-free survival.

Cost effectiveness results are provided in Table 31 and a comparison with the base case ICER presented in Table 32. The ERG consolidation subgroup analysis resulted in a 10.1% higher ICER than the base case ITT population.

Table 31. Scenario results: ERG consolidation subgroup (discounted)

Technology	Total costs (£)	Total LYG	Total QALYs	Incremental costs (£)	Incremental LYG	Incremental QALYs	ICER versus baseline (£/QALY)
Watch and wait with BSC	[redacted]	2.955	[redacted]	-	-	-	-
Oral azacitidine	[redacted]	3.821	[redacted]	[redacted]	0.87	[redacted]	53,574

Abbreviations: ICER = incremental cost-effectiveness ratio; LYG = life years gained; QALY = quality-adjusted life year; BSC = best supportive care

Table 32. Difference in ICER: ERG consolidation subgroup

Model	ICER versus baseline (£/QALY)
Base case	48,660
Scenario: ERG consolidation subgroup	53,574
Difference (%)	£+4,914 (10.1%)

Abbreviations: ICER = incremental cost-effectiveness ratio

Treatment Effectiveness

B6. Priority question: The company used the joint generalised gamma distribution and the joint log-logistic distribution to respectively model OS and RFS in its base-case. According to the company, these distributions have the best statistical (based on AIC and BIC) and visual fit to the observed data. However, Figures B.3.10 and B.3.19 in the CS, as well as Table B.5.35 in Appendix J indicate that the modelled OS and RFS of watch and wait + BSC are an underestimation of the OS and RFS observed in the comparator arm of the QUAZAR AML-001 trial. In addition, the company stated that expert consultations suggested that crossing of curves (e.g. using the individual generalised gamma and Gompertz) was not considered clinically likely.

a. **Tables B.2.11 and B.2.12 in the CS show that the OS and RFS difference between oral azacitidine and placebo in the QUAZAR AML-001 trial decreases over time (e.g., the OS survival difference is [redacted] at year 1, [redacted] at 2 years and [redacted] at 3 years). In addition, the proportion of patients receiving HSCT is higher in the placebo arm, which may impact the (difference in) hazard rates over time for OS and RFS. Given the above, please justify why crossing of the survival curves of the two treatment strategies is considered unlikely.**

The observed data does suggest a path towards convergence of the curves (and potentially a cross-over of the curves at some point during the extrapolation, depending on how tightly you fit to the data). However, we outlined in our ITT OS selection chapter within the submission why we consider this an unlikely outcome.

UK clinical experts suggested that although there is a possibility that the survival curves may cross, it is assumed that the likelihood of this happening is low, given the level of toxicity that patients experience is not high enough to explain a cross-over. Furthermore, it was suggested that only a crossing at a relative late time point would be plausible. Subsequent to the ERG clarification questions, additional clinical advice from the same UK clinical experts was sought regarding the plausibility of the curves crossing, both clinicians reiterated and strengthened their position that the curves crossing was not a possibility.

Instead, the convergence can be explained by heterogeneity between patients within the trial. There is strong evidence to support the presence of heterogeneity given the plateauing observed in the OS hazards. Sicker patients drop out of the risk set sooner, making the average hazard at the population level reduce (when individual hazards are actually more likely to be increasing). The treatment which is least effective (BSC) will more quickly present events in the most severe patients (as can be seen from the KM curves) and hence a sharper reduction in hazards, appearing to converge towards a more effective therapy such as oral azacitidine. But there is no strong clinical rationale as to why that trend would continue indefinitely, particularly to the point where survival for oral azacitidine becomes worse than BSC (despite the large incremental benefits observed in the trial).

Whilst this explanation is speculative given the lack of concrete observed data, it provides a data generating process which aligns with our clinical validations.

b. **Please also provide smoothed hazard plots per treatment arm over time, with patient numbers at risk and add the company's modelled hazard rates per treatment arm with the company's selected model.**

Data for modeled vs non-parametric smoothing of hazard plots were generated from fitted parametric (R package [flexsurv]) and smoothed hazard (R package [muhaz])[52] model objects respectively. Hazards for parametric models were estimated over the required time-points using the predict function. The [muhaz] package fits a non-parametric kernel smoothed hazard using a weighted estimate of data within a given distance (bandwidth) of time t. The current analysis made use of default settings which is to allow bandwidths to vary locally across the bandwidth grid, include left and right boundary corrections. These are intended to reduce variance by increasing bandwidth size at left and right endpoints where events and numbers at risk are sparse respectively (Hans-Georg 1994). Numbers at risk are included to contextualize estimates and are calculated using the [survival] package.[53]

Modeled versus smoothed hazards were broadly similar (Figure 4) with the [redacted]. The observed crossing is expected to be an

artifact of noise given clinical expectation of hazards crossing being considered implausible combined with the brief nature and occurrence as numbers of at risk are decreasing. We view this as more consistent with the modeled waning treatment effect on the hazard scale. The sharp increase in hazards near the end of follow-up is likely an artifact of small numbers at risk.

Figure 4.Modeled vs smoothed hazards for OS

Smoothed hazards show a similar pattern in RFS as OS (Figure 5), with a [redacted]. While the signal for crossing of hazard functions is slightly stronger here than for OS, the difference in hazard is small and occurs as numbers at risk have decreased considerably suggesting the modeled converging hazards are an adequate fit.

Figure 5. Modelled vs smoothed hazards for RFS

While the parametric models do not provide perfect fit to smoothed hazards, this is likely appropriate given the risk of over-fitting associated with the latter. The survival models included in the CS are thus a reasonable compromise between observed hazards and clinical plausibility.

c. **To examine the validity of the extrapolation beyond the data, please provide supporting evidence that the extrapolations are consistent with relevant external data (for example the comparator arm) and/or expert opinion. In case of expert opinion, please provide a full description of the methods and results of the expert consultation conducted.**

UK clinical expert insight was sought to validate extrapolations beyond the trial period.

The Clinical experts were selected according to the following criteria:

1. Currently treating patients with AML in the UK
2. Greater than 10 years' experience treating patients with AML

Notably, one expert had experience as an investigator in the QUAZAR study. Both of the experts approached participated in the clinical elicitation exercises. The experts were presented with clinical data in the form of survival analysis and an overview of various draft model inputs. This data presented during the elicitations were consistent with the submission unless the clinical experts guided towards other sources or values. Expert opinions were collected over the course of three virtual meetings with slide sharing. The first and second meetings were held with each clinical expert individually and the third included the two experts together. Each meeting was approximately one hour in duration. Open questions were asked and details of the questions and responses can be found in the clinical expert summary report.

UK clinical experts suggested that it would be unlikely for survival curves to cross therefore experts suggested using a modelling approach that would not result in curves crossing. Joint models were selected as these prevent curves crossing. For OS, the clinician found the Gompertz model to be more optimistic than the other curves and suggested not to use the Gompertz model. For RFS, they experts found the Gompertz model to provide overly optimistic extrapolations for early years and pessimistic in later years. The experts were indifferent between extrapolations provided by the log-normal, log-logistic and generalized gamma models as they yielded similar results. Detailed external validation of extrapolation was not possible due to the lack of external data available at the time of the responses.

B7. In line with the OS and RFS modelling in the FLT-3 subgroup, please describe whether the use of spline-based models for OS and RFS was explored. Please provide these analyses including 1 and 2 knot models (with default knot location) using the hazard, odds as wells as normal scales (resulting in 6 models). Please elaborate on the appropriateness of these spline models and provide an updated economic model as well as scenario analyses enabling the use of these spline models.

There are a number of reasons why splines were not used for the ITT analysis, and have therefore not been included in the economic model:

- Amongst the set of existing models which have been fitted, there are several survival modelling options which fit very well to the observed data (as shown by both model fit statistics [AIC/BIC] and visual inspection), provide clinically plausible extrapolations, and have been validated with a clinical audience.
- Splines, as a vehicle for extrapolation, have a tendency to overfit to patterns in the observed data which are spurious, which is a particular concern given the nature of the observed data in QUAZAR AML-001 (see response to B6a).
- Splines can add value in fitting to complex hazard functions and faithfully representing the observed data – however we feel this is already achieved with our existing strata of parametric models.
- For the FLT3 population, parametric models may not have accurately fit the data for OS and RFS. Thus, splines were explored given our lower levels of confidence in the more standard approaches and small sample size ([redacted] in the QUAZAR AML-001 FLT3 and CR only subgroup).

B8. Appendix N.1.4 shows the OS and RFS survival models for the EU subgroup of the QUAZAR AML-001 trial (N=314). The company aligned the survival models for this subgroup with the QUAZAR AML-001 ITT population, and joint generalized gamma and joint log-logistic models were used for OS and RFS respectively. Instead of aligning with the ITT population, please perform survival analyses for OS and RFS in the EU subgroup following the guidance in NICE DSU TSD 14.

The following section focuses on the intention to treat (ITT) population of the QUAZAR AML-001 trial across all recorded time points restricted to European (EU) patients only. This population was included to explore the influence of restricting the patient pool from the QUAZAR AML-001 trial to the EU patient subgroup. For both overall survival (OS) and relapse-free survival (RFS), analyses restricted to the EU subgroup suggested alignment with the assessment for the ITT population, with joint generalized gamma providing the optimal fit for OS and joint log-logistic providing the optimal fit for RFS. For additional details, please see Appendix B.8.

B9. No treatment waning was assumed in the company's base-case analysis.

a. Please justify the assumption of no treatment waning.

b. Please provide a hazard ratio plot with numbers of patients at risk over time to justify this assumption.

c. Please provide an updated economic model where you explore treatment waning in scenario analyses.

B9.a. Treatment waning is typically tested when there is a large unobserved period. The individual models showed some indication of treatment waning. We put forward the reasons for this in the response to question B6a. However, based on clinical expert opinion, extrapolations based on joint curves were deemed more clinically appropriate than individual curves, hence joint curves were implemented in the base case. We propose to explore the impact of waning with the use of the individual curves (see response to question B9c).

B9.b. Models for OS and RFS in the base case belong to the class of accelerated failure time (AFT) models. The treatment effect that is estimated for these models is interpreted as an acceleration factor and is assumed constant across time. This approach is potentially beneficial when the proportional hazards assumption is inappropriate since AFT models allow the hazard ratio to vary over time. The [flexsurv] package does not automatically produce estimates of hazard ratios with their 95% confidence intervals over time and therefore these were generated through repeated sampling of the coefficients from their variance-covariance matrix using the MASS::mvrnorm() function. For each iteration, hazards for treatment and control were estimated using built-in functions of the [flexsurv] package (eg, hgengamma()) and the hazard ratio at each time was defined as $HR_t = h_{Onureg_t}/h_{BSC_t}$. Confidence bands and central estimates were summarized as their respective quantiles. Numbers at risk were derived as outlined above.

Data for OS were modeled using a joint generalized-gamma model. Data for RFS were modeled using joint a joint log-logistic model.

Cost-effectiveness model data consisted of the modeled survivor functions after all post-processing (e.g., incorporation of background mortality). This was included since estimated hazard ratios would be expected to differ from those used in the model. This allows accurate capture of true waning treatment effects. Time-horizons for these models are limited to 360 months to align with the CEM model.

As expected, given the use of AFT models in the base case, the hazard ratio varies over time, exhibiting a waning effect (Figure 6). This pattern is consistent with the modeled and smoothed hazard plots, and exhibits a strength of AFT models in their ability to accommodate time-varying hazards in a parsimonious approach. The cost-effectiveness model then increases the waning aspect of this treatment effect further (Figure 7) by incorporating general population mortality including a hard step at 150 months where the HR is deterministically set to 1.

Figure 6. OS hazard ratio over time (modeled only)

Figure 7. OS hazard ratio over time (modeled vs CEM)

Treatment waning is more aggressive for RFS but otherwise is comparable to that of OS (Figure 8). Calculations within the model do not noticeably change the parametric model estimates. This is expected given RFS is only constrained by OS in the model to preserve the logic of the partitioned survival structure and prevent crossover (Figure 9).

Figure 8. RFS hazard ratio over time (modeled only)

Figure 9. RFS hazard ratio over time (modeled vs CEM)

These analyses provide evidence that the choice of survival models for OS and RFS are acceptable and naturally incorporate a waning treatment effect. This treatment effect is additionally attenuated within the cost-effectiveness model

B9.c. A scenario analysis was conducted using alternative models to better understand the impact of treatment waning. Independent models were used for this purpose which featured a [redacted]. It should be noted that expert opinion had suggested that converging and crossing survival curves were not considered clinically likely hence this may lack clinical plausibility. For OS, independent log-normal models were used as they appeared to have reasonable visual fit to the data and the second-best AIC (see Figure 10).

Figure 10. Parametric curves fit to the OS outcome in the ITT population – Log-normal distribution, individual model

Abbreviations: AZA = azacitidine; ITT = intention-to-treat; OS = overall survival

For RFS, independent log-logistic models were used as these models had the lowest AIC of the individual models (excluding the Gompertz, which has an implausible functional form). Due to the merging of the curves, both OS and RFS models satisfy Criterion 5 of the Tremblay et al.[54] guidance showing lower marginal survival in the extrapolation vs the observed period.

Figure 11. Parametric curves fit to the RFS outcome in the ITT population – Log-logistic distribution, individual model

Abbreviations: AZA = azacitidine; ITT = intention-to-treat; RFS = relapse-free survival

The results of the scenario analysis are provided in Table 33 and the comparison to the base case ICER provided in Table 34 . The use of independent curves resulted in an increase in the ICER by £5,347. This is predominantly driven by the change in incremental QALYs with both treatments generating similar QALYs in the relapse state when using independent curves. In the base case, QALYs for the relapse health state were [redacted] for oral azacitidine and [redacted] for watch and wait + BSC arm, with the use of independent curves, this resulted in [redacted] QALYs for oral azacitidine and [redacted] QALYs for watch and wait + BSC. It should be noted that although there is a slight increase in the ICER when using the independent curves, UK clinical expert opinion has emphasised that the crossing of curves is not possible and independent curves should not be used.

Table 33. Scenario results: waning (independent survival curves) - ITT population

Technology	Total costs (£)	Total LYG	Total QALYs	Incremental costs (£)	Incremental LYG	Incremental QALYs	ICER versus baseline (£/QALY)
Watch and wait with BSC	[redacted]	2.633	[redacted]	-	-	-	-
Oral azacitidine	[redacted]	3.434	[redacted]	[redacted]	0.80	[redacted]	54,017

Abbreviations: ICER = incremental cost-effectiveness ratio; LYG = life years gained; QALY = quality-adjusted life year; BSC = best supportive care

Table 34. Difference in ICER: waning (independent survival curves) - ITT population

Model	ICER (£/QALY)
Base case	£48,660
Scenario : Additional waning implemented	£54,017
Difference (%)	+£5,357 (11%)

Abbreviations: ICER = incremental cost-effectiveness ratio

B10. Patients in the QUAZAR AML-001 trial discontinued from oral azacitidine upon relapse. However, in the economic model the time-on-treatment curve and RFS curve seem to cross, i.e., the modelled proportion of patients on treatment is higher than the proportion of patients that are relapse-free.

a. Please justify the plausibility of crossing of these curves.

For illustration of this observation, Figure 12 is provided which includes both the KM data and the parametric survival curves used by the model.

Figure 12. Time on treatment and RFS

Treatment duration in months is defined as (treatment end date — first dose date +1)/30.4375, where treatment end date is last dose date + 14 days (the prescribed rest period of each cycle), or the death date, whichever is earlier. Thus, the end date for time on treatment may extend beyond the date of relapse, causing the proportion of patients on treatment to appear higher than the proportion of patients that are relapse-free.

Patients receiving dose extension would still be on treatment and in the relapse state.

As can be seen in Figure 12, during the first part of the time horizon there was a small deviation between time on treatment and relapse free survival. A small deviation is also plausible since estimated KM survival curves are plotted as step functions rather than smooth curves estimated based on parametric distribution for RFS.

b. Please correct this logical inconsistency and provide an updated economic model.

The logic in the model prevents this inconsistency. The calculations are set-up to check that the number of patients in "RFS on treatment" is the minimum value between the number of patients in RFS versus the number of patients from the ToT KM curve at each cycle. Drug costs are calculated as the number of patients in RFS on treatment multiplied by the per-patient drug cost. This has been quality checked as part of the model validations using the TECH-VAR checklist (see question B25). As there are no corrections required, an updated economic model relating to this question has not been provided.

B11. The company used the time-varying generalised gamma distribution and the time-varying spline-based 1 knot odds linear model to model OS and RFS respectively in the FLT-3 subgroup. The Gompertz was the second-best fitting model for OS based on AIC but was not considered given the observed plateau, which was not in line with clinical expectations. Models leading to crossing of curves were not considered clinically likely.

a. In contrast to what was shown in Figure B.3.22 and Figure B.3.23 of the CS, no plateauing of the curves was observed in the economic model when selecting the Gompertz distribution for OS in the FLT-3 subgroup. Please justify this.

The reason for the discrepancy is that economic model ensures that the mortality hazard for the modelled cohort is at least that of the general population from UK Life Tables (adjusted for age and gender). Removing this restriction, results in the

observed plateau mentioned in Figure B.3.22 and B.3.23 in the company submission.

b. Please provide further justification on why the spline-based 1 knot odds linear model was considered more appropriate for the modelling of RFS than standard parametric models.

For the FLT-3 RFS curve selection, the 1 knot odds linear model was deemed an appropriate choice since the tail of the curves collapsed towards zero. Of the models considered the 1 knot odds model also provided a good fit to the observed data. Of the standard parametric models, the generalised gamma and lognormal were also considered to be reasonable selections for RFS in this population.

c. To examine the validity of the extrapolation beyond the data, please provide supporting evidence that the extrapolations are consistent with relevant external data (for example the comparator arm) and/or expert opinion. In case of expert opinion, please provide a full description of the methods and results of the expert consultation conducted.

Please refer to the response to question B6.C for a summary of the clinical expert opinion elicitation procedure. For RFS the experts were unsure regarding curve choice. However, the experts concluded that the generalized gamma model is optimistic, and that the spline model using 1 knot odds linear predictor is most plausible.

d. Please provide a more detailed justification on why crossing of curves was not considered clinically likely.

Subsequent to the ERG clarification questions, additional clinical advice from the same UK clinical experts was sought regarding the plausibility of the curves crossing for the FLT3 subgroup, both clinicians reiterated and strengthened their position that the curves crossing was not a possibility and stated that the same rationale provided for the ITT population is also applicable to the FLT3 subgroup (see response to Q6a). The experts suggested that given the level of toxicity that patients experience it was highly unlikely for a cross-over to take place.

e. Although various spline-based models were considered for OS and RFS in the FLT-3 subgroup, cell D/E 117 on the efficacy sheet of the economic model only allows the selection of the 1 internal knot normal linear predictor and the 1 internal knot odds linear predictor for the modelling of RFS. Please provide an updated model that allows the selection of all spline-based models that were considered for the modelling of RFS.

Other spline models (1 knot hazard and 2 knots models) were not included in the cost-effectiveness mode (CEM) as they predicted nonsensical survival extrapolations. Natural cubic splines are not globally monotone. Datasets of reasonable size are required to impose monotonically decreasing survival estimates, especially in regions where data are sparse (e.g., tails of survival data).[55] 1 knot hazard and 2 knots models predicted an increase in survival due to small sample size ([redacted] in QUAZAR AML-001 *FLT3* and CR only subgroup), high rates of censoring ([redacted]), and poor fit to the data. In Figure 13, the 1 internal knot and hazard linear predictor model for RFS, RFS in the placebo group of both the QUAZAR AML-001 and RATIFY trials gradually increases beyond the trial data. RFS increases more dramatically in the 2 knot spline models (Figure 14, Figure 15 and Figure 16).

Figure 13. Time-varying Spline Model for Relapse-Free Survival Using 1 Internal Knot and a Hazard Linear Predictor

Figure 14. Time-varying Spline Model for Relapse-Free Survival Using 2 Internal Knots and Hazard Linear Predictor

Figure 15. Time-varying Spline Model for Relapse-Free Survival Using 2 Internal Knots and Odds Linear Predictor

Figure 16. Time-varying Spline Model for Relapse-Free Survival Using 2 Internal Knots and Normal Linear Predictor

Adverse events

B12. According to section B.3.3.5 of the CS, the model included grade 3 and 4 AEs occurring in ≥5% or more of patients in the safety population of the QUAZAR AML-001 trial, as well as AEs identified by clinical advisors to have a substantial impact on quality of life. For midostaurin, AEs of grade 3 and 4 occurring in >10% of patients in the maintenance phase of the RATIFY trial were included. However, also many low grade TEAEs that occurred in the safety population of the QUAZAR AML-001 trial are reported in Table B.2.23 of the CS.

a. Please justify using different cut-off points for oral azacytidine (≥5% and midostaurin (>10%).

In the economic model, different cut-off points were used to inform the rates of Grade 3 or 4 adverse events (AEs) for oral azacitidine (≥5%) and midostaurin (≥10%). The decision to use the lower threshold for oral azacitidine was driven by the interest to be conservative, while the higher threshold for midostaurin was driven by a lack of published evidence. Although the NICE technology appraisal of midostaurin (NICE TA523) reports rates of Grade 3 or 4 AEs occurring in ≥5% or more of patients in the treatment arm of the RATIFY trial, information contained within the presented table was redacted. In addition, the redacted safety data were specific to the entire treatment period and incorporated AEs experienced by patients during all phases of the RATIFY study (i.e., induction, consolidation, and maintenance). Therefore, even if available, this data would be limited in its comparability to the maintenance-specific safety data informing oral azacitidine. While recognizing the limitations of using different thresholds in the model as AEs are potentially underestimated for midostaurin, this conservative approach was deemed most appropriate due to the lack of available safety data from the RATIFY trial.

b. For oral azacitidine and midostaurin, please provide an updated model and scenario analyses including all grade 3 and 4 AEs that occur in at least 2% of the corresponding QUAZAR AML-001 and RATIFY trials.

Data are available for grade 3 and 4 AEs that occur in at least 2% of the QUAZAR AML-001 trial population. However, no data are available for grade 3 and 4 AEs that

occur in at least 2% of the RATIFY trial population during the maintenance phase. Therefore, the option to include these additional AEs has been added to the model for the QUAZAR ITT population. As data was not available for grade 3 and 4 AEs that occur in at least 2% in the RATIFY trial population, to include this for only oral azacitidine and not for midostaurin would lead to biased results, therefore this scenario was not conducted in the FLT3 subgroup. Instead, this scenario was conducted in the ITT population using grade 3 and 4 AEs that occur in at least 2% of QUAZAR AML-001. Scenario results are provided in Table 35 and a comparison with the base case ICER is presented in Table 36. This inclusion resulted in a 0.07% increase in the ICER compared to the base case.

Table 35. Scenario results: ITT population with AE incidence of at least 2% (discounted)

Technology	Total costs (£)	Total LYG	Total QALYs	Incremental costs (£)	Incremental LYG	Incremental QALYs	ICER versus baseline (£/QALY)
Watch and wait with BSC	[redacted]	2.799	[redacted]	-	-	-	-
Oral azacitidine	[redacted]	3.864	[redacted]	[redacted]	1.06	[redacted]	48,694

Abbreviations: ICER = incremental cost-effectiveness ratio; LYG = life years gained; QALY = quality-adjusted life year; BSC = best supportive care

Table 36. Difference in ICER: Grade 3 and 4 adverse events with an incidence of at least 2% - ITT population

Model	ICER (£/QALY)
Base case	48,660
Scenario: Grade 3/4 AE with incidence of at least 2%	48,694
Difference (%)	£+34 (0.07%)

Abbreviations: ICER = incremental cost-effectiveness ratio, AE = adverse events

c. For oral azacitidine and midostaurin, please provide an updated model and scenario analyses also including low grade (grade 1 and 2) AEs that occur in at least 5% of the corresponding QUAZAR AML-001 and RATIFY trials.

Similarly, data are available for grade 1 and 2 AEs that occur in at least 5% of the QUAZAR AML-001 population. No data are available for these AEs in the RATIFY trial during the maintenance phase. Therefore, the option to include low grade AEs has been added to the model for the QUAZAR ITT population only and so the scenario was undertaken in the ITT population. Scenario results are provided in Table 37 and a comparison with the base case ICER is presented in Table 38

Table 37. Scenario results with oral azacitidine - PAS price: ITT population with grade 1 and 2 adverse events (discounted)

Technology	Total costs (£)	Total LYG	Total QALYs	Incremental costs (£)	Incremental LYG	Incremental QALYs	ICER versus baseline (£/QALY)
Watch and wait with BSC	[redacted]	2.799	[redacted]	-	-	-	-
Oral azacitidine	[redacted]	3.864	[redacted]	38,160	[redacted]	[redacted]	49,791

Abbreviations: ICER = incremental cost-effectiveness ratio; LYG = life years gained; QALY = quality-adjusted life year; BSC = best supportive care

Table 38. Difference in ICER : Scenario results: Grade 1 and 2 adverse events with an incidence of at least 5%

Model	ICER (£/QALY)
Base case	48,660
Scenario : Grade 1/2 AE with at least 5% incidence	49,791
Difference (%)	£+1,131 (2.3%)

Abbreviations: ICER = incremental cost-effectiveness ratio, AE = adverse events

B13. Based on clinical advisor opinion, it was stated in section B.3.4.4 of the CS that the duration of AEs is assumed to be 1 week. The ERG did not find detailed information on the elicitation of the clinical advisor opinion. Please provide information on the methods used to elicit clinical advisor opinion and the results of the elicitation.

Please refer to the response to question B6.c. for a summary of the clinical expert opinion elicitation procedure. Please also see the full details in the clinical expert summary report shared.

B14. Based on clinical advisor opinion, it was stated in section B.3.4.4 of the CS that the duration of AEs is assumed to be 1 week. Furthermore, the percentage of patients that experienced an AE at least once during the trial follow-up (as reported

in Table B.3.8 of the CS) was used to model the frequency per AE. This implies that AEs were assumed to occur a maximum of once per patient for a duration of 1 week.

a. Please justify the plausibility of assuming a 1-week duration for all AEs, also considering the severity of the AEs (grade 3 or 4).

b. Please perform a scenario analysis assuming an average AE duration of 4 weeks (one model cycle) and provide an updated model including this scenario.

c. Given that oral azaciditine is given in intervals of 14 days and treatment related AEs are expected to occur in the on-treatment phase, please justify the plausibility of modelling AEs only once per patient.

d. Please perform scenario analyses assuming that AEs occur in every on-treatment interval for the percentage of patients as reported in Table 3.8 of

e. the CS for a duration of 1 week and a duration of 4 weeks and provide an updated model including these scenarios.

B14a. Subsequent to the ERG clarification questions, additional clinical expert opinion was sought from the same two clinicians regarding this question. Feedback from the UK clinical experts indicated that the assumption of 1-week duration for all AE was not unreasonable as some AE may have lower duration whilst others may have higher. Explicit references were made to febrile, diarrhoea, vomiting and neutropenia as having a duration equal to or lower than 1 week and anaemia and thrombocytopenia typically having duration longer than 1 week. Furthermore, the clinicians noted that in clinical practice, the strategy is to predict and prevent AE recurrence hence modelling one event per patient may be a reasonable simplifying assumption.

B14b. Results from assuming an average duration of 4 weeks for adverse events is presented in Table 39 for the ITT population and in Table 40 for the FLT3 subgroup. A comparison with the base case ICER is presented in Table 41.

Table 39. Scenario results with oral azacitidine - PAS price: ITT population with 4-week duration of adverse events

Technology	Total costs (£)	Total LYG	Total QALYs	Incremental costs (£)	Incremental LYG	Incremental QALYs	ICER versus baseline (£/QALY)
Watch and wait with BSC	[redacted]	2.799	[redacted]	-	-	-	-
Oral azacitidine	[redacted]	3.864	[redacted]	[redacted]	1.06	[redacted]	48,787

Abbreviations: ICER = incremental cost-effectiveness ratio; LYG = life years gained; QALY = quality-adjusted life year; BSC = best supportive care

Table 40. Scenario results with oral azacitidine - PAS price: FLT3 subgroup with 4-week duration of adverse events

Technology	Total costs (£)	Total LYG	Total QALYs	Incremental costs (£)	Incremental LYG	Incremental QALYs	ICER versus baseline (£/QALY)	Pairwise ICER vs oral azacitidine
No active therapy	[redacted]	2.731	[redacted]	-	-	-	-	24,532
Oral azacitidine	[redacted]	4.828	[redacted]	[redacted]	2.10	[redacted]	24,532	-
Midostaurin	[redacted]	3.600	[redacted]	[redacted]	0.87	[redacted]	290,619	Oral azacitidine is dominant

Abbreviations: ICER = incremental cost-effectiveness ratio; LYG = life years gained; QALY = quality-adjusted life year; BSC = best supportive care

Table 41. Difference in ICER: Scenario results: 4-week duration of adverse events

Model	**ICER (£/QALY)**
Base case (ITT population)	48,660
AE duration 4 weeks:	48,787
Difference (%)	£127 (0.26%)
Base case (FLT3 subgroup)	Oral azacitidine is dominant
AE duration 4 weeks:	Oral azacitidine is dominant
Difference (%)	NA

Abbreviations: ICER = incremental cost-effectiveness ratio, AE = adverse events

B14c. In the model, treatment emergent adverse events (TEAEs) included AEs that occurred between the first dose and up to 28 days following the last dose of study treatment. The approach of using the prevalence of AEs (ie, modelling AEs once per patient) as a simplifying assumption is common in partitioned survival models and is consistent with the methods used in previous AML submissions [56, 57]. The cost and disutilities of AEs are front-loaded in the model given the prevalence of AEs is reflective of the entire starting population in the QUAZAR-AML trial, and thus they are not impacted by discounting or reduced survival over time. This is considered conservative.

B14d. The percentages used in the model represent the overall proportion of patients in the QUAZAR-AML trial population experiencing the given Grade 3 or 4 TEAE. They are not rates, and therefore, they do not reflect the average occurrence of AEs per cycle. Assuming that AEs occur in every on-treatment interval for the percentage of patients as reported in Table 3.8 of the CS is not a valid application of these proportions and would be expected to over-estimate AE cost and disutilities in the model. For this reason, the suggested scenario analysis was not conducted.

Quality of life

B15. Priority question. Health state utility values are, according to CS Figure B.3.43 key drivers of the cost-effectiveness results.

a. **Please provide, per measurement timepoint, separately for oral azacitidine and SoC:**
    a. **the total number of EQ-5D-3L responses**
    b. **the estimated mean utility values and standard error**
    c. **a breakdown of how many patients were relapse-free and were relapsed and the respective utility scores**
    d. **a breakdown of how many patients were on and off treatment and the respective utility scores**
    e. **the extent of missing data observed**
b. **Please explain, with appropriate justifications, how missing data were handled and the implications of this approach.**
c. **Please clarify what the likely causes of missing data were and what the potential impact of these missing data on the estimation of the utility scores would be, separately for patients who had completely and partially missing utility data.**
d. **Please recalculate the utility estimates while imputing missing values (for the patients with completely missing utility data and patients with**

partially missing utility data) using multiple imputation (incorporating potential explanatory variables and using at least 10 imputations).

   a. **Please provide in detail, the methods used to impute and pool the utility data**

   b. **Please elaborate on the plausibility of the imputed utility values**

   c. **Please provide an updated economic model as well as scenario analysis incorporating these newly calculated utility values**

e. **Please compare patient characteristics of patients with complete utility measurements and patients with missing utility measurements for both treatment groups separately and for the whole trial population combined (independent of treatment groups) and comment on potential differences.**

f. **Please rerun the analyses performed to obtain the utility values (i.e. original approach from the CS) for oral azacitidine (stratified for patients being on and off treatment) and SoC separately.**

g. **Please provide an updated economic model as well as a scenario analysis incorporating the estimated utility values in response to sub-questions e and f (i.e. utility values estimated stratified for patients being on and off treatment with and without imputation).**

B15.aa. The total number of EQ-5D-3L responses per timepoint by treatment arm are provided in Appendix B.15 Table 78

B15.ab. The estimated mean utility values with their standard errors per timepoint by treatment arm are provided in Appendix B.15 Table 78

B15.ac. Per the study design, study treatment would be discontinued if patients experienced a relapse (ie., bone marrow blast [BMB] >5%) and utility (or HRQoL) data would not be collected after the end of treatment assessment visit was completed. However, if patients experienced a relapse with BMB between >5% - ≤15%, they may be further treated with dose escalation (from 14 days to 21 days per cycle) at physician's discretion. Therefore, per study design, utility data following a relapse were collected only in some of the patients for a limited number of visits. In addition, BMB was measured every 3 cycles and therefore did not perfectly align with the timing of utility assessment (ie, Day 1 of every cycle), making the estimation of mean utility values by relapse status across visits difficult to do, unless the last BMB value carried forwarded approach was used. Therefore, utility values post relapse should be interpreted with these limitations in mind. The results of mean utility values by relapse status (ie, BMB>5%; yes/no) across visits are presented in Appendix B.15 Table 79. Patients with a relapse in both treatment groups had more assessment visits with decrement in utility value from baseline or lower observed mean utility value than those without a relapse. Again, sample size was too small to yield consistent and reliable utility estimates for patients with a relapse, in addition to those limitations highlighted above.

B15.ad. It is not possible to provide this information as EQ-5D and other HRQoL instruments were collected during treatment phase (Day 1 of each treatment cycle) and the end of treatment visit; therefore there is no "off treatment" assessment.

B15.ae. Missing data per timepoint by treatment arm are provided as the completion rate in Appendix B.15.Table 80

B15.b. According to the recommendation by the Setting International Standards in Analyzing Patient-Reported Outcomes and Quality of Life Endpoints Data Consortium (SISAQOL),[58] the extent of missing data should be quantified using two measures: 1) completion rate (also known as variable denominator rate which is defined as the number of ITT subjects submitting a valid HRQoL assessment at a given timepoint over the number of ITT subjects who are expected to provide HRQoL assessment at that timepoint), as presented in Appendix B.15 Table 79 and

Appendix B.15 Table 80 available data rate (also known as fixed denominator rate, defined as the number of ITT subjects submitting a valid HRQoL assessment at a given timepoint over the number of ITT subjects). The amount of missing data summarized by the available data rate was mostly caused by study design (ie, stopping HRQoL data collection after a given intercurrent event; eg, treatment discontinuation due to any cause per study protocol). Thus, the extent of missing data for this study should be based on the completion rates (ie, data displayed in Appendix B.15 Table 80). As shown in Appendix B.15 Table 80, the extent of missing data was small (<10%; in most visits ≤5%) and very similar between treatment arms across visits. In addition, a lot of them were "intermittent missingness" (ie, patients missed one or two HRQoL assessments in between visits), which can be treated missing completely at random as the reasons behind intermittent missingness are often not treatment related.[59] Therefore, imputing such a small amount of missing data (which were likely to be missing completely at random) should have little or no impact on the health utility values, as currently observed for both treatment arms, while patients were still relapse free. Due to these reasons, no imputation of missing data was performed.

B15.c. For those subjects who were eligible for HRQoL assessment at a given scheduled visit but did not complete the assessment, specific causes for these missing data were not unclear as data were not collected in the study. As mentioned above, the extent of missing data was very small and similar between treatment arms, there should be little or no impact by this type of missing data. For subjects becoming not eligible for HRQoL assessment after experiencing those pre-specified events (treatment discontinuation due to AML relapse, AEs, etc.) at a given time point, AML relapse was the most likely reason for both treatment arms (143/236 [oral azacitidine] vs. 180/233 [placebo]).[22] Only 29 [oral azacitidine] vs 11 [placebo] subjects discontinued due to AEs as primary reason.[22] As AML relapse usually leads to worsening in HRQoL and a greater percentage of patients in the placebo arm with treatment discontinuation due to AML relapse, we would expect the placebo

arm would be more likely to have worse HRQoL results than what was currently observed if HRQoL data were continued to be collected after these events.

B15.da. This request is not possible as missing data imputation was not performed due to the rationale provided in response to B15c.

B15.db. This request is not possible as missing data imputation was not performed due to the rationale provided in response to B15c.

B15.dc. This request is not possible as missing data imputation was not performed due to the rationale provided in response to B15c.

B15.e. Per the ERG request, patients with complete utility measurements were defined as those without missing any utility measurement across all eligible assessment visits. This is a very stringent criterion and not a commonly-used approach to assess differences in baseline demographic and disease characteristics between those missing and not missing HRQoL data in longitudinal studies with many repeated HRQoL assessment visits. For example, a patient who was eligible for 10 HRQoL assessments but completed 9 of them would be categorized in the subgroup with missing utility measurement. Typically, a much less stringent criterion is usually used in HRQoL analysis by categorizing patients into HRQoL-evaluable and non-evaluable populations. Evaluable patients are often defined as patients who has non-missing baseline visit and at least one non-missing post-baseline HRQoL (or utility) assessment visit. With that being said, the analysis per the ERQ request, as well as the less stringent one, was all performed accordingly and the results are presented in Appendix B.15 Table 81, Table 82 and Table 83. No marked

differences in demographic and disease characteristics between subgroups, regardless of criterion used.

B15.f. This analysis cannot be performed as we do not have the off-treatment utilities as mentioned. Utilities and other HRQoL measures were not collected after treatment discontinuation.

B15.g. This request is not possible as missing data imputation was not performed due to the rationale provided in response to B15c.

B16. As described in Appendix O, linear mixed effects models with random intercepts were used to derive EQ-5D-3L utility values in the pre-progression health state. To determine relevant covariates, four different models were fitted and the best fitting model was selected based on the level of significance, the magnitude of coefficients and AIC and BIC statistics. AIC and BIC statistics, however, are not reported.

a. Please provide the AIC and BIC statistics for the four different models.

Please find the AIC and BIC statistics for the four models in Table 42.

Table 42. Utility model AIC and BIC statistics

Parameters	Model 1: Intercepts Only	Model 2: Treatment Arm	Model 3: Ongoing AEs	Model 4: Treatment and Ongoing AEs
AIC	[redacted]	[redacted]	[redacted]	[redacted]
BIC	[redacted]	[redacted]	[redacted]	[redacted]

b. Given that the number of (serious) AEs in the QUAZAR AML-001 trial was considerably higher for oral azacitidine than for the placebo arm, please

discuss the plausibility of the not statistically significant treatment coefficient in model 2.

Although the number (%) of subjects with AEs were greater with the oral azacitidine, the difference was not so substantial between treatment arms (<10%). In addition, not all AEs had a meaningful impact on subjects' health utility value, and the impact of AEs on health utility should be short-lasting as dose would be modified or treatment would be stopped if symptomatic AEs still can't be addressed effectively. These are the likely reasons why there was no significant difference in health utility value between treatment arms.

c. Although the ongoing AEs covariate in model 3 was statistically significant ($p<0.001$), the intercept only model (model 1) was selected and AE utility decrements were informed from the literature. Please justify why the statistically significant AE covariate was ignored and why evidence from the literature was preferred over evidence from the QUAZAR AML-001 trial.

The intercept only model (model 1) was selected over the model inclusive of ongoing AEs (model 3) since the AE covariate, whilst indicating a significant impact on utility, does not provide a means of connecting the disutility to specific adverse events. In other words, the combined impact of all AEs was estimated to be [redacted] but that is not to say this was the same for each event type e.g. thrombocytopenia and fatigue.

However, it should be acknowledged that model 3 may still be the optimal choice for usage in the economic model. The estimate of the intercept can be interpreted as the health state utility value for those not experiencing adverse events. Therefore, the additional adverse event disutility adjustment taken, from the literature values, avoids any potential double counting. Therefore, the supporting model provided includes this change and all subsequent analysis utilise the intercept of model 3 and AE disutility values from the literature. The AE coefficient is not directly used.
In relation to question B18, it should be noted that this change does inflate the RFS utility value from [redacted] to [redacted] hence moving further from the age adjusted UK population norm.

d. Please conduct a scenario analysis using model 3 (intercept + ongoing AEs covariate) and provide an updated model file including this scenario.

We acknowledge that model 3 (intercept + ongoing AEs covariate) should have been used as the base case, therefore we have updated the RFS utility value from [redacted] to [redacted]. The updated base case results are provided in Table 43, with the disaggregated results for QALYs, costs and resource use are provided in Table 44, Table 45 and Table 46. The updated results from the FLT3 subgroup are provided in Table 47 with the disaggregated results for QALYs, costs and resource use are provided in Table 48, Table 49 and Table 50.

Table 43. Base case results with oral azacitidine PAS (discounted)

Technology	Total costs (£)	Total LYG	Total QALYs	Incremental costs (£)	Incremental LYG	Incremental QALYs	ICER versus baseline (£/QALY)
Watch and wait with BSC	[redacted]	2.799	[redacted]	-	-	-	-
Oral azacitidine	[redacted]	3.864	[redacted]	[redacted]	1.06	[redacted]	48,660

Abbreviations: ICER = incremental cost-effectiveness ratio; LYG = life years gained; QALY = quality-adjusted life year; BSC = best supportive care

Table 44.Summary of QALY gain by health state

Health state	Total QALYS Oral azacitidine	Total QALYs Watch and wait +BSC	Increment
Total RFS	[redacted]	[redacted]	[redacted]
RFS: On Treatment	[redacted]	[redacted]	[redacted]
RFS: Off Treatment	[redacted]	[redacted]	[redacted]
Relapse	[redacted]	[redacted]	[redacted]
Adverse Event Disutility*	[redacted]	[redacted]	[redacted]
HSCT Disutility*	[redacted]	[redacted]	[redacted]
Total	[redacted]	[redacted]	[redacted]

* These are not health states but are components of the generated QALYs.
Abbreviations: RFS = relapse free survival; QALYs = quality adjusted life years, HSCT = haematopoietic stem cell transplantation; BSC = best supportive care

Table 45. Summary of costs by health state – PAS price

Health state	Total costs Oral azacitidine	Total costs Watch and wait +BSC	Increment
Total RFS	[redacted]	[redacted]	[redacted]
RFS: On Treatment	[redacted]	[redacted]	[redacted]
RFS: Off Treatment	[redacted]	[redacted]	[redacted]
Relapse	[redacted]	[redacted]	[redacted]

Table 46. Summary of predictive resource use by category – PAS price

Resource	Oral azacitidine	Watch and wait +BSC	Increment
Total RFS	[redacted]	[redacted]	[redacted]
RFS: On Treatment	[redacted]	[redacted]	[redacted]
Drug Costs	[redacted]	[redacted]	[redacted]
Treatment Admin. Costs	[redacted]	[redacted]	[redacted]
Disease Management Costs	[redacted]	[redacted]	[redacted]
Adverse Event Costs	[redacted]	[redacted]	[redacted]
RFS: Off Treatment	[redacted]	[redacted]	[redacted]
Disease Management Costs	[redacted]	[redacted]	[redacted]
Relapse	[redacted]	[redacted]	[redacted]
Disease Management Costs	[redacted]	[redacted]	[redacted]
Subsequent Therapy Costs	[redacted]	[redacted]	[redacted]
SCT Costs	[redacted]	[redacted]	[redacted]
End of Life Costs	[redacted]	[redacted]	[redacted]
Total	[redacted]	[redacted]	[redacted]

Abbreviations: RFS = relapse free survival; HSCT = haematopoietic stem cell transplantation; BSC = best supportive care

Table 47. Deterministic results with oral azacitidine - PAS price: subgroup FLT3 (discounted)

Technology	Total costs (£)	Total LYG	Total QALYs	Incremental costs (£)	Incremental LYG	Incremental QALYs	ICER versus baseline (£/QALY)	Pairwise ICER vs oral azacitidine
No active therapy	[redacted]	2.731	[redacted]	-	-	-	-	24,532
Oral azacitidine	[redacted]	4.828	[redacted]	[redacted]	2.10	[redacted]	24,532	-
Midostaurin	[redacted]	3.600	[redacted]	[redacted]	0.87	[redacted]	291,902	Oral azacitidine is dominant

Abbreviations: FLT3 = fms-like tyrosine kinase 3; ICER = incremental cost-effectiveness ratio; LYG = life years gained; QALY = quality-adjusted life year

Table 48. Summary of QALY gain by health state – FLT3 subgroup

Health state	Total QALYs Watch and wait +BSC	Total QALYS Oral azacitidine	Total QALYs Midostaurin	Increment
Total RFS	[redacted]	[redacted]	[redacted]	[redacted]
RFS: On Treatment	[redacted]	[redacted]	[redacted]	[redacted]
RFS: Off Treatment	[redacted]	[redacted]	[redacted]	[redacted]
Relapse	[redacted]	[redacted]	[redacted]	[redacted]
Adverse Event Disutility*	[redacted]	[redacted]	[redacted]	[redacted]
HSCT Disutility*	[redacted]	[redacted]	[redacted]	[redacted]
Total	[redacted]	[redacted]	[redacted]	[redacted]

* These are not health states but are components of the generated QALYs.
Abbreviations: RFS = relapse free survival; QALYs = quality adjusted life years, HSCT = haematopoietic stem cell transplantation; BSC = best supportive care

Table 49. Summary of costs by health state – PAS price – FLT3 subgroup

Health state	Total costs Watch and wait +BSC	Total costs Oral azacitidine	Total Costs Midostaurin	Increment
Total RFS	[redacted]	[redacted]	[redacted]	[redacted]
RFS: On Treatment	[redacted]	[redacted]	[redacted]	[redacted]
RFS: Off Treatment	[redacted]	[redacted]	[redacted]	[redacted]
Relapse	[redacted]	[redacted]	[redacted]	[redacted]

Abbreviations: RFS = relapse free survival; BSC = best supportive care

Table 50. Summary of predictive resource use by category – PAS price – FLT3 subgroup

Resource	Watch and wait +BSC	Oral azacitidine	Midostaurin	Increment
Total RFS	[redacted]	[redacted]	[redacted]	[redacted]
RFS: On Treatment	[redacted]	[redacted]	[redacted]	[redacted]

Resource	Watch and wait +BSC	Oral azacitidine	Midostaurin	Increment
Drug Costs	[redacted]	[redacted]	[redacted]	[redacted]
Treatment Admin. Costs	[redacted]	[redacted]	[redacted]	[redacted]
Disease Management Costs	[redacted]	[redacted]	[redacted]	[redacted]
Adverse Event Costs	[redacted]	[redacted]	[redacted]	[redacted]
RFS: Off Treatment	[redacted]	[redacted]	[redacted]	[redacted]
Disease Management Costs	[redacted]	[redacted]	[redacted]	[redacted]
Relapse	[redacted]	[redacted]	[redacted]	[redacted]
Disease Management Costs	[redacted]	[redacted]	[redacted]	[redacted]
Subsequent Therapy Costs	[redacted]	[redacted]	[redacted]	[redacted]
SCT Costs	[redacted]	[redacted]	[redacted]	[redacted]
End of Life Costs	[redacted]	[redacted]	[redacted]	[redacted]
Total	[redacted]	[redacted]	[redacted]	[redacted]

Abbreviations: RFS = relapse free survival; HSCT = haematopoietic stem cell transplantation; BSC = best supportive care

B17. HSCT was modelled as part of a subsequent treatment in terms of costs and a one model cycle disutility, without modelling the post-HSCT benefits. As HSCT is expected to have a positive impact on HRQoL, please conduct a scenario analysis

applying a post-HSCT utility increment for the patients assumed to receive HSCT as a subsequent treatment and provide an updated model including this scenario.

Given the reasons provided in responses B1.b. and B1.c., we do not think it is appropriate to incorporate a utility benefit post HSCT.

B18. Patients in the QUAZAR AML-001 trial had a median age of 68 years at baseline. The modelled utility for the pre-progression health state (RFS on and off treatment) was [redacted], which is higher than the UK general population norm for this age group (0.785 for 65-74 years, Szende et al. 2014). Please provide an updated economic model and scenario analysis capping the maximum pre-progression health state utility value based on the UK general population norm.

Due to differences between real world and trial based elicitation of health related quality of life, some difference between utility outcomes can be expected. Moreover, some differences can be expected when comparing populations from clinical trial and the real world due to differences in sample characteristics. Whilst the health state utility value of [redacted], this is adjusted downwards over time in line with guidance from the NICE decision support unit. Of the models fitted to the trial-based EQ-5D data, the model used had the lowest utility value which is a conservative assumption. Moreover, where possible data is sourced from the key clinical trial taking account of the position this takes in the NICE evidence hierarchy. The results from the scenario analysis with a RFS utility of 0.785 is provided in Table 51 for the ITT population and Table 52. A comparison with the base case ICER is presented in Table 53.

Table 51. Scenario results: ITT population with RFS utility of 0.785

Technology	Total costs (£)	Total LYG	Total QALYs	Incremental costs (£)	Incremental LYG	Incremental QALYs	ICER versus baseline (£/QALY)
Watch and wait with BSC	[redacted]	2.799	[redacted]	-	-	-	-
Oral azacitidine	[redacted]	3.864	[redacted]	[redacted]	1.06	[redacted]	51,934

Abbreviations: ICER = incremental cost-effectiveness ratio; LYG = life years gained; QALY = quality-adjusted life year; BSC = best supportive care

Table 52. Scenario results: RFS utility of 0.785 - FLT3 subgroup

Technology	Total costs (£)	Total LYG	Total QALYs	Incremental costs (£)	Incremental LYG	Incremental QALYs	ICER versus baseline (£/QALY)	Pairwise ICER vs oral azacitidine
No active therapy	[redacted]	2.731	[redacted]	-	-	-	-	26,027
Oral azacitidine	[redacted]	4.828	[redacted]	[redacted]	2.10	[redacted]	26,027	-
Midostaurin	[redacted]	3.600	[redacted]	[redacted]	0.87	[redacted]	319,827	Oral azacitidine is dominant

Abbreviations: FLT3 = fms-like tyrosine kinase 3; ICER = incremental cost-effectiveness ratio; LYG = life years gained; QALY = quality-adjusted life year

Table 53. Difference in ICER : Scenario results: RFS utility of 0.785

Model	ICER (£/QALY)
Base case (ITT population)	48,660
RFS utility 0.785:	51,934
Difference (%)	£+3,274 (6.7%)
Base case (FLT3 subgroup)	Oral azacitidine is dominant
RFS utility 0.785:	Oral azacitidine is dominant
Difference (%)	NA

Abbreviations: ICER = incremental cost-effectiveness ratio, AE = adverse events , RFS = relapse free survival

B19. As discussed in A30 HRQoL and fatigue (using the FACIT-Fatigue Scale) were measured on day 1 of each 28-day cycle. Please provide an updated economic model and scenario analysis correcting for the resulting bias.

In theory, assessing HRQoL at the start of each treatment cycle is less likely to capture the effect of treatment-related symptomatic AEs on HRQoL, especially if AEs are short-lived or when treatment cycles are long. Therefore, detrimental effects on HRQoL caused by AEs may be more likely to be underestimated for oral azacitidine (vs. placebo/SOC). Despite this, it is believed that the impact would be marginal. The negative impact of AEs is not anticipated to have a long-lasting effect in most cases, as dose would likely be modified to address the issue. Those AEs with longer-lasting effects would be captured by the HRQoL instrument on day 1 of each 28-day cycle.

To mitigate any risk that treatment-related AEs were not fully captured in the HRQoL measurement from the QUAZAR AML-001 trial, AE disutilities were applied to the health state utility values in the base case. For example, a disutility of 0.115 is applied over the duration of a week to account for patients with grade 3 or 4 fatigue. This approach ensures that the HRQoL impact for patients who experienced fatigue and other treatment-related AEs between measurement intervals (the first day of each cycle) would still be captured.

For these reasons, we have elected not to update the economic model.

Costs and resource use

B20. According to Table B.3.30 of the CS, a relative dose intensity (RDI) of [redacted] was assumed for oral azacitidine.

a. Please explain whether a zero drug waste was assumed and justify the plausibility of this assumption.

b. Please justify why no RDI was assumed for midostaurin, and if applicable, provide an updated model and scenario analysis also including an RDI for midostaurin.

c. Does the dose intensity estimate include dose escalation to 21-day course as observed in the QUAZAR trial? If not, please provide an updated model

d. including the proportion of patients that received a longer course and the duration for which they received it.

B20.a. Drug wastage relating to discontinuation of treatment was not accounted as inspection of the time-on-treatment curve does not indicate that a sudden discontinuation is common with oral azacitidine. Moreover, the average compliance rate was high. Furthermore, we anticipate the inclusion of drug wastage to have minimal impact on the results.

B20.b. The economic model did not incorporate relative dose intensity (RDI) for midostaurin due to the lack of published evidence specific to the maintenance phase of the RATIFY study. Although the NICE technology appraisal of midostaurin (NICE TA523)[25] reports a median RDI of 95%, this represents the exposure to midostaurin across all treatment phases of the RATIFY trial, including induction and consolidation. In contrast to maintenance therapies, induction and consolidation regimens are shorter in duration and higher in intensity, commonly resulting in high rates of toxicity that may require more frequent dose modifications. Due to these considerations, using the reported RDI for midostaurin in the NICE technology appraisal was deemed inappropriate and limited in its comparability to the maintenance-specific value assumed for oral azacitidine. However, despite the lack of available maintenance data

from the RATIFY study, we have provided an updated economic model that incorporates RDI for midostaurin, assuming a value of 95%. Although this may potentially underestimate the observed value in the maintenance phase of the RATIFY trial, it was included in the model as a conservative estimate. Results of a scenario analysis exploring inclusion of this parameter are presented in Table 54 and a comparison with the base case ICER presented in Table 55.

Table 54. Scenario results: Midostaurin with 95% relative dose intensity: FLT3 subgroup

Technology	Total costs (£)	Total LYG	Total QALYs	Incremental costs (£)	Incremental LYG	Incremental QALYs	ICER versus baseline (£/QALY)	Pairwise ICER vs oral azacitidine
No active therapy	[redacted]	2.731	[redacted]	-	-	-	-	26,027
Oral azacitidine	[redacted]	4.828	[redacted]	[redacted]	2.10	[redacted]	26,027	-
Midostaurin	[redacted]	3.600	[redacted]	[redacted]	0.87	[redacted]	304,793	Oral azacitidine is dominant

Abbreviations: FLT3 = fms-like tyrosine kinase 3; ICER = incremental cost-effectiveness ratio; LYG = life years gained; QALY = quality-adjusted life year

Table 55. Difference in ICER: Scenario results: Inclusion of relative dose intensity for midostaurin

Model	ICER (£/QALY)
Base case	Midostaurin is strictly dominated by oral azacitidine
RDI 95% for midostaurin	Midostaurin is strictly dominated by oral azacitidine
Difference in ICER	NA

Abbreviations: ICER = incremental cost-effectiveness ratio, RDI = relative dose intensity

B20.c. We confirm that the does intensity estimate does include escalation to 21-day course as observed in the QUAZAR trial

B21. Please could you confirm if any of the drug costs in the model would fall under a primary care setting?

Drugs in the model that can be prescribed under a primary care setting are ondansetron, ciprofloxacin, fluconazole and tranexamic acid. Hydroxycarbamide and posaconazole are “grey” areas as sometimes their prescribing can be transferred to a primary care setting. For reference, the budget impact analysis assumes that hydroxycarbamide, posaconazole, ondansetron, ciprofloxacin, fluconazole, and tranexamic acid have a VAT of 10% to reflect the charges for drugs delivered by Homecare and this also includes initial supply of drugs in secondary care as per the budget impact analysis submission template.

B22. Please compile a table which lists all the treatments that have been modelled in your base case results and all other analyses, making sure to include:

a. all pre-medication treatments (including for comparators and subsequent treatments), the intervention, comparators, and subsequent treatments (and any concomitant medications)

b. the strength (per ml for injections if applicable), form/mode of administration, pack size, list price (and source) for each treatment included in the table.

B22.a./b. A table detailing a list of all treatments that have been modelled is provided in Table 56. A typological error was made in table B.3.25 of the company submission and in the model sheet Disease Management cell I200 where the unit strength of cytarabine as salvage chemotherapy was stated as 100mg, this should have been stated as 500mg and has been corrected in the table below. As the reported mg in the model is not used to calculate the unit price, the typographical error has no impact on results.

Table 56. List of treatments included in the model

Drug name (type)	Admin route	Dose per tablet	Units per pack	Cost per pack (£) (list price)	Source
Intervention					
Oral azacitidine	Oral	300mg	14	[redacted]	BMS data on file
FLT3 comparator					
Midostaurin	Oral	25mg	56	5609.94	BNF
Premedication					
Ondansetron	Oral	8mg	10	0.93	eMIT 2020
Best supportive care					
Hydroxycarbamide	Oral	500mg	100	9.61	eMIT 2020
Ciprofloxacin	Oral	500mg	10	3.08	eMIT 2020
Posaconazole	Oral	100mg	24	175.32	eMIT 2020
Fluconazole	Oral	200mg	7	0.51	eMIT 2020

Tranexamic acid	Oral	500mg	60	7.98	eMIT 2020
Subsequent therapy					
Low-dose cytarabine	Subcutaneous	100mg	5	22.52	eMIT 2020
Injectable azacitidine	Subcutaneous	100mg	1	220	BNF
Subsequent therapy : Salvage chemotherapy					
Daunorubicin	Intravenous	20mg	10	715	BNF
Cytarabine	Intravenous	500mg	5	22.38	eMIT 2020

Scenario and sensitivity analyses

B23. Compared to the deterministic analysis (ICER of £49,704 per QALY gained), the PSA based on 1,000 iterations resulted in a considerably lower ICER (£45,130 per QALY gained).

a. Please use convergence plots to show the stability of the PSA results (costs and effects) based on 1,000 iterations?

The convergence plot is provided in Figure 17.

Figure 17. Convergence diagnostic - average net monetary benefit

b. Please rerun the PSA on (at least) 5,000 iterations.

Results from the PSA with 5000 iterations is provided in Table 57.

Table 57. Base case results with oral azacitidine PAS (Probabilistic) - 5000 iterations

Technology	Total costs (£)	Total LYG	Total QALYs	Incremental costs (£)	Incremental LYG	Incremental QALYs	ICER versus baseline (£/QALY)
Watch and wait +BSC	[redacted]	2.815	[redacted]	-	-	-	-
Oral azacitidine	[redacted]	3.877	[redacted]	[redacted]	1.062	[redacted]	48,147

Abbreviations: ICER = incremental cost-effectiveness ratio; LYG = life years gained; QALY = quality-adjusted life year; BSC = best supportive care

B24. The scenario analyses using the 2019 data cut for OS and doubling AE disutilities could not be reproduced by the ERG (i.e. resulted in a different ICERs than reported by the company). Please provide details on which cells were changed in the economic model for these scenario analyses and how these cells were changed. If applicable, also provide a corrected model file.

We were able to reproduce the results from both the scenarios. However, a typographically error was made in Table B.3.34 of the company submission which stated AE disutility doubled, this should have been AE rates doubled as detailed in Table B.3.35. Details on how to reproduce the scenarios to output results seen in Table B.3.35 of the company submission are provided below:

Scenario : 2019 data cut for OS

- Cell D/E 28 on the Efficacy sheet : dropdown used to select July 2019 datacut
- Cell D/E 41 on the Efficacy sheet : dropwdown used to select Lognormal curve fit as the lognormal had the lowest BIC.

Scenario : AE rates doubled

- Adverse event rates in sheet Adverse events, Cells D24:K26 should be multiplied by 2.

Results based on the updated base case as mentioned in response to B16 for the 2019 data cut is provided in Table 58 and the comparison with the base case provided in Table 59. The results with doubling adverse event rates is provided in Table 60 and a comparison with the base case ICER is presented in Table 61.

Table 58. Scenario results: 2019 data cut for overall survival

Technology	Total costs (£)	Total LYG	Total QALYs	Incremental costs (£)	Incremental LYG	Incremental QALYs	ICER versus baseline (£/QALY)
Watch and wait +BSC	[redacted]	2.343	[redacted]	-	-	-	-
Oral azacitidine	[redacted]	3.354	[redacted]	[redacted]	1.01	[redacted]	49,248

Abbreviations: ICER = incremental cost-effectiveness ratio; LYG = life years gained; QALY = quality-adjusted life year; BSC = best supportive care

Table 59. Difference in ICER: Scenario results: 2019 data cut for overall survival

Model	ICER (£/QALY)
Base case	48,660
Scenario: 2019 data cut for OS	49,248
Difference (%)	£+588 (1.12%)

Abbreviations: ICER = incremental cost-effectiveness ratio, OS = overall survival

Table 60. Scenario results: Adverse event rates doubled

Technology	Total costs (£)	Total LYG	Total QALYs	Incremental costs (£)	Incremental LYG	Incremental QALYs	ICER versus baseline (£/QALY)
Watch and wait +BSC	[redacted]	2.799	[redacted]	-	-	-	-
Oral azacitidine	[redacted]	3.864	[redacted]	[redacted]	1.06	[redacted]	48,875

Abbreviations: ICER = incremental cost-effectiveness ratio; LYG = life years gained; QALY = quality-adjusted life year; BSC = best supportive care

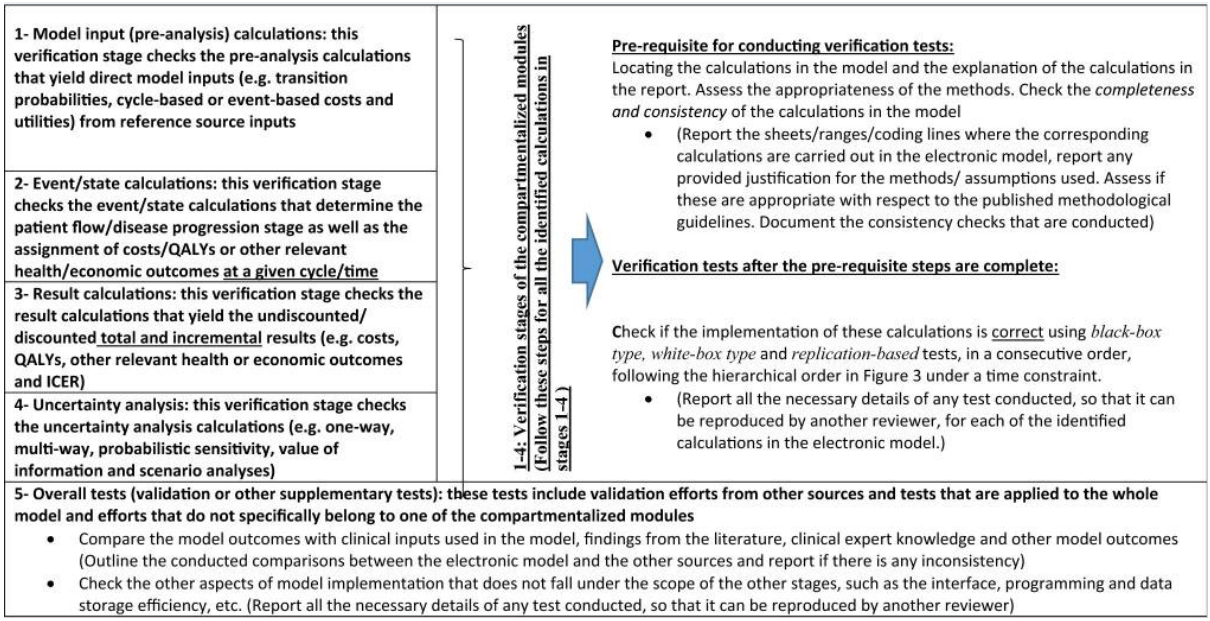
Table 61. Difference in ICER: Scenario results: Adverse event rates doubled

Model	ICER versus baseline (£/QALY)
Updated base case: ICER	48,660
Scenario: AE rates doubled	48,875
Difference	£+215 (0.44%)

Abbreviations: ICER = incremental cost-effectiveness ratio, AE = adverse events

Validation and transparency

B25. Priority question: In addition to the checks already performed, please also complete the TECH-VER checklist (Büyükkaramikli et al, 2019, https://pubmed.ncbi.nlm.nih.gov/31705406/).



Extensive internal validation was performed to check the model's performance based on the five-domain TECH-VER checklist tool. For each of the stages, black box, white box and replication-based tests were conducted as advised by the TECH-VER checklist tool. Details of some of the main checks are provided in Table 62. Black box tests are similar to those already conducted as outlined in Table B.3.38 of the company submission.

Table 62. TECH-VER checklist

Verification stage 1 : Model input/pre-analysis calculation		
Component	**Check**	**Result**
Survival analysis : Fitted survival curves in statistical software R correspond to curves estimated in the model	Set the general population hazard to zero on the "Overall Survival" sheet cells K20:K542. Survival curves generated in the model are provided on the "Efficacy" sheet	Survival curves produced in R (graphically presented in section B.3.3 of the company submission) match the survival curves used in the model

Costs are reflective of cycle length	Drug cost per cycle are calculated correctly in in sheet “Drug costs” cell D8, Treatment admin costs per cycle, sheet “Treatment administration” cell D8, Disease management costs per cycle, sheet “Disease management” cell D8:F10.	All costs have been adjusted correctly to calculate per cycle costs
Verification stage 2 : Event/State calculations		
Component	**Check**	**Result**
Patient distribution	Logic check implemented in sheet “CC-486 calculations” cells L14:L536 to prevent RFS exceeding OS. Logic check implemented in sheet “CC-486 Calculations” cells M14:M536 to prevent RFS: on treatment exceeding Total relapse. The same logic tests are implemented in sheet “Rydapt calculation” and “No Active Ther.Calculations” and “CC-486 FLT3 Calculations” for the comparators.	Patient distribution has been implemented correctly using the area under the curve approach and logic tests have been implemented to prevent RFS: on treatment exceeding total relapse and RFS exceeding OS
Assignment of costs/utilities to health states	Calculation of costs per health state disaggregated by their individual components are calculated by multiplying the patient distribution by their respective costs and provided in sheet “CC-486 Calculations” cells AE14:AP536. The same method is implemented in the sheets “Rydapt calculation” and “No Active Ther.Calculations” and “CC-486 FLT3 Calculations” for the comparators.	No issues identified in assigning costs/utilities for health state
Verification stage 3 : Result calculations		
Component	**Check**	**Result**
Summation of accumulated costs, QALYS and life years	Total cost, QALYs and Life years provided in sheet “Deterministic results” cells O12:P14 are calculated correctly	Totals cost, QALYs and life years have been summed up correctly in the model
Interpretation of results	ICER on sheet “Deterministic Results” cell P23 calculated as costs over QALYs if results in the north-	The ICER has been calculated correctly and dominance and extended dominance has been

	east quadrant of the CE plane	interpretated correctly on sheet "Deterministic Results" cell P23.
Discounting	Discount rate on sheet "CC-486 Calculations" calculated correctly in cells BZ12:CE536 and applied correctly. Same for the sheets "Rydapt calculation" and "No Active Ther.Calculations" and "CC-486 FLT3 Calculations" for the comparators.	Discounting has been applied and implemented correctly
Half-cycle correction	Half-cycle correction applied correctly	Half cycle correction has been implemented correctly to the patient distribution
Disaggregation of total costs/ QALYs	Disaggregated cost and QALYS calculated correctly in sheet "Deterministic results" cells N50:U126	Disaggregated costs/QALYs have been calculated correctly and sum to that of the accumulated total costs/QALYs
Verification stage 4 : Uncertainty analysis calculations		
Component	**Check**	**Result**
One-way sensitivity analysis	Sheet "One-Way inputs" consists of all the inputs that should be varied, and high/low values have been calculated correctly	All high/low values have been calculated correctly
Probabilistic sensitivity analysis	Check correct distributions and standard errors have been used and probabilistic value generated is reasonable given the standard error. Implementation of Cholesky decomposition matrix to make regression-based inputs probabilistic	Correct distributions and standard errors have been implemented. Inputs relating to overall survival relapse free survival have been made probabilistic based on the Cholesky decomposition matrix correctly. All probabilistic values are reasonable given the base value and standard error.
Verification stage 5 : Overall validation		
Navigation buttons	Working as intended	All navigation buttons have been tested and working as intended

B26. Please provide cross validations, i.e. comparisons with other relevant NICE TAs focused on similar, potentially relevant, indications (e.g. TA 454) as well as a cross validation with the study by Bewersdorf et al

(https://pubmed.ncbi.nlm.nih.gov/34525174/) and for each comparison elaborate on the identified differences regarding:

a. Model structure and assumptions, input parameters related to clinical effectiveness, health state utility values, resource use and costs

The following tables provide a comparison of the NICE TAs identified, Table 63 outlines the model structure and assumptions, Table 64 details the clinical effectiveness inputs, followed by health state utility values in Table 65. The scope of NICE TAs presented was kept broad. These were restricted to AML only as opposed to phase of treatment e.g., induction, consolidation and maintenance. The NICE TAs are further supplemented by the Bewersdorf et al. (2021)[60] US based oral azacitidine model. The company submission utilises the PLD from QUAZAR to model the outcomes which is expected to be more accurate than using the summary data that was available to the authors of this study. Importantly, the company's submission utilised a more recent data cut to inform the overall survival endpoint. Therefore, the reliability and accuracy of the company submission is expected to outweigh that of this article.

Previous submissions, for the most part, used partitioned survival modelling. There was a large variation in the health state structure which can be expected given the range of indications. A complex semi markov model with many health states was deemed complex and challenging to review by the ERG.

Regarding clinical effectiveness, the models reviewed included a wide variety of approaches including parametric survival models, cure models, flexible models (splines and mixture cure) as well as modelling the KM data directly. Again, the broad array of methods can, at least in part, be attributed to the varying populations across each economic evaluation. The views of the ERG were specific to each of these scenarios however there was broad agreement supporting the usage of standard parametric survival models given sufficient data was available.

A large array of health state utility values were identified since the health states differed across the models. For similar health states, variation of utility values was identified however these follow the logical ordering from higher to lower utility: functionally cured, relapse free, relapsed.

Two resource use parameters featured highly in the deterministic sensitivity analysis of incremental costs. These were Relative Dose Intensity (RDI) of oral azacitidine and the proportion of patients receiving SCT. In the case of this submission, the best source of data for these parameters was the QUAZAR AML-001 trial. Comparing against the Bewersdorf et al. (2021)[60] model, which also sourced data from QUAZAR AML-001, it was unclear how RDI was modelled if at all. In this model patients were similarly modelled to receive HSCT and the source of this parameter appears to align with the source used in this NICE submission; 15 of 238 oral azacitidine patients and 32 of 234 placebo patients received a stem cell transplant.[22]
Of the main cost outcomes, the costs for chemotherapy admin per oral administration, nurse visit and haematologist visit had the greatest impact on incremental costs. Comparison was restricted to the most recent NICE submission since i) unit costs are expected to change year on year and ii) the cost perspective of the Bewersdorf et al. (2021)[60] model was that of the US Health care system. NICE TA642[61] had a cost year of 2018 and included both nurse and haematologist visits. Haematologist visits were costed at £108 in NICE TA642 using PSSRU 2018. This was £58 lower than the value used in this submission based on the NHS reference costs 2019-2020. Similar Nurse visits were costed with the PSSRU at £37 in TA642. This was £62.30 lower than the value used in this submission. It is unclear from NICE TA642 exactly which admin cost was used for the oral Gilteritinib regimen since this information has been redacted. Further admin costs for subcutaneous and intravenous injections were far higher as can be expected.

Table 63. Model structure and assumptions

TA identifier	Model Type	Health States	ERG Critique
TA399 (Population not eligible for HSCT)[57]	Semi Markov	Remission, stable disease, relapse/post-progression and death	The main limitation was the assumption that no subsequent active treatment was given after the initial azacitidine or CCR treatment.
TA523[25]	Partitioned survival model	AML diagnosis/induction, complete response/remission, relapse, stem cell transplant and death	Allowing patients to move from relapse to CR leads to inconsistencies as CR following relapse is unlikely to occur without further (non-trial) treatment. Model does not accommodate response to subsequent treatment.

TA545[56]	Semi-Markov cohort state-transition model	Induction, complete remission (on and off treatment), refractory states (salvage therapy and non-curative therapy), HSCT, relapse states (salvage therapy and non-curative therapy), post HSCT with and without graft vs host disease, functional cure and death	The proposed model structure is complex and challenging to critique given the difficulties in determining the flow of patients. The company was requested to provide a clearer description of the assumptions and to explain the advantage of the state-transition model compared to a simpler and more conventional partitioned survival analysis models.
TA552[62]	Decision tree and partitioned survival	Newly diagnosed disease, remission, disease progression, death	Patients may not progress through a linear pathway: they could receive transplant before progression or after progression, and progression could occur before or after transplant
TA642[61]	Decision tree and partitioned survival	Alive event free, alive post event, death. Sub-models (With HSCT and No HSCT)	The sub-models (With HSCT and No HSCT) and the health states (event-free and post-event) were questioned in terms of appropriateness. Also, the approach to estimate health state occupancy over time raised concerns.
Bewersdorf et al. (2021)[60]	Partitioned survival analysis	Remission, post progression, death	N/A

Table 64. Clinical effectiveness inputs

TA identifier	Efficacy modelling approach	ERG Critique
TA399[57]	OS, RFS and PFS curves were constructed by fitting parametric survival models to data from the trial. The treatment effect was modelled using proportional hazards for all survival curves.	Usage of KM nonparametric curves as observed in the clinical trial provide the best source data with which to populate PFS and RFS model parameters, while minimizing the structural uncertainty of the cost-effectiveness results.
TA523[25]	A cure model (assuming the rate of death from the general population after the end of the trial) was used in the base case. Parametric models are explored in scenario analysis for transparency. A piecewise approach was used for EFS, where the KM curve is used prior to the trial cut-off, followed by a parametric tail after the cut-off.	The ERG considers that the approach taken by the company was the most appropriate, given the available data, because it avoids the need to make any assumptions about the data, e.g., proportional hazards, and it reflects the actual treatment effect observed in the trial.
TA545[56]	Standard parametric models (Gompertz) are used in the base case for OS for patients in the refractory state. Flexible survival analysis methods are used to capture the visible	The OA Gompertz curve selected by the company for its base-case had the best fit according to AIC/BIC, and the company also considered that it had the best visual fit,

	plateau in KM data and the more complex instantaneous risk of events.	stating that the spline-based models resulted in late-occurring plateaus.
TA552[62]	Parametric survival curves fitted to the patient-level data to extrapolate over the model time horizon.	Significant concerns related to survival analyses and extrapolation beyond the trial period because the available data was too immature to robustly estimate the survival benefit for post-transplant patients.
TA642[61]	EFS and OS fitted to the KM data following a parametric survival modelling approach. Cure point aligned with flattening of KM curves from a range of publications. Cure assumption: The model assumes that all patients who remain alive after 3 years are "cured".	Imposes two inappropriate structural constraints: (i) the cure assumption is applied to all surviving patients, irrespective of their relapse/progression status, and (ii) time to HSCT is assumed to be fixed.
Bewersdorf et al. (2021)[60]	Parametric survival curves fitted to the patient-level data to extrapolate over the model time horizon. In all cases individual log-logistic regression distributions were chosen based on fit, visual inspection and pragmatic modelling considerations.	N/A

Table 65. Health state Utility Values

TA identifier	Health state utility approach	Health state utility values	ERG Critique
TA399[57]	Utilities were mapped from trial-based disease specific EORTC QLQC30 data to EQ-5D utility values using published algorithms.	Post-progression/relapse: 0.623 Remission (CR/Cri): 0.771 Remission (PR.SD): 0.716	Utility values suitably mapped from HRQoL measurements from trial
TA523[25]	Data from the literature used in the base case and results from a TTO study were used in scenario analysis.	Induction: 0.648 Consolidation: 0.710 Monotherapy:810 Complete remission: 0.830 Relapse :0.655 (0.53-0.78) SCT Treatment: 0.613 SCT Recovery:0.810 Post-SCT Recovery: 0.826	For several health states, there were multiple values published in the literature, and the company did not clearly justify how these values were selected from the multiple sources. Over time long term survivors would have greater utility than general population.
TA545[56]	TTO and VAS and literature based approaches	Relapse/ Refractory: 0.568 Chemotherapy: 0.6574 Consolidation: 0.6574 HSCT procedure:0.6574 GVHD (post HSCT): 0.67 CR or CRp: 0.7400 Functionally cured: 0.820	In the absence of direct HRQoL data, the ERG considered the approach used to be reasonable and appropriately justified. One exception was utility value for functionally cured being too high.

TA552[62]	Utility values based on a vignette time-trade-off study conducted in members of the UK general population	Induction: 0.550 Remission (post-induction/consolidation): 0.656	The ERG was concerned about the generalisability of the utility values used in company model
TA642[61]	Trial based EQ-5D-5L utility scores and literature based scenario	Health state utilities confidential	Programming error post 3 years cure point.
Bewersdorf et al. (2021)[60]	Literature based	Relapsed AML: 0.53 Early remission: 0.66 Prolonged Remission: 0.82	N/A

b. And how these differences affect estimated outcomes per comparator / interventions (life years, QALYs, costs)

On the basis of the findings of B26.a the following scenario was performed.

Alternative cost assumptions for nurse and haematologist visits – increase in the unit cost by 40% for haematologist visit from £166.00 to £232.40 and nurse visit from £99.30 to £139.02. Scenario results for the ITT population are provided in Table 66 and a comparison with the base case ICER presented in Table 67. Results for the FLT3 subgroup are provided in Table 68 and a comparison between the base case is presented in Table 69.

Table 66. Scenario results : Increase in unit cost for nurse and haematologist visit

Technology	Total costs (£)	Total LYG	Total QALYs	Incremental costs (£)	Incremental LYG	Incremental QALYs	ICER versus baseline (£/QALY)
Watch and wait +BSC	[redacted]	2.799	[redacted]	-	-	-	-
Oral azacitidine	[redacted]	3.864	[redacted]	[redacted]	1.06	[redacted]	51,704

Abbreviations: ICER = incremental cost-effectiveness ratio; LYG = life years gained; QALY = quality-adjusted life year; BSC = best supportive care

Table 67. Difference in ICER Scenario results : Increase in unit cost for nurse and haematologist visit

Model	ICER versus baseline (£/QALY)
Base case	48,660
Scenario: 40% increase in costs	51,704
Difference	£+3,044 (6.23%)

Table 68. Scenario results : Increase in unit cost for nurse and haematologist visit – FLT3 subgroup

Technology	Total costs (£)	Total LYG	Total QALYs	Incremental costs (£)	Incremental LYG	Incremental QALYs	ICER versus baseline (£/QALY)	Pairwise ICER vs oral azacitidine
No active therapy	[redacted]	2.731	[redacted]	-	-	-	-	26,846
Oral azacitidine	[redacted]	4.828	[redacted]	[redacted]	2.10	[redacted]	26,846	-
Midostaurin	[redacted]	3.600	[redacted]	[redacted]	0.87	[redacted]	295,460	Oral azacitidine is dominant

Abbreviations: FLT3 = fms-like tyrosine kinase 3; ICER = incremental cost-effectiveness ratio; LYG = life years gained; QALY = quality-adjusted life year

Table 69. Difference in ICER: Scenario results: Increase in unit cost for nurse and haematologist visit - FLT3 subgroup

Model	ICER (£/QALY)
Base case	Midostaurin is strictly dominated by oral azacitidine
Scenario: 40% increase in costs	Midostaurin is strictly dominated by oral azacitidine
Difference in ICER	NA

Abbreviations: ICER = incremental cost-effectiveness ratio, RFS = relapse free survival

References

1. Hui, D., et al., *Concepts and definitions for "supportive care," "best supportive care," "palliative care," and "hospice care" in the published literature, dictionaries, and textbooks.* Supportive care in cancer : official journal of the Multinational Association of Supportive Care in Cancer, 2013. **21**(3): p. 659-685.
2. Dohner, H., et al., *Diagnosis and management of AML in adults: 2017 ELN recommendations from an international expert panel.* Blood, 2017. **129**(4): p. 424-447.
3. BMS, *Data on file: CC-486-AML-001 (QUAZAR) Clinical Study Report.* 2020.
4. NCCN, *NCCN Clinical Practice Guidelines in Oncology (NCCN Guidelines). Acute Myeloid Leukemia. Version 4. 2020. NCCN.org.* 2020.
5. EMA. *Rydapt: EPAR - Product information [online]*. [cited 2021 29 November 2021]; Available from: https://www.ema.europa.eu/en/documents/product-information/rydapt-epar-product-information_en.pdf.
6. Kindler, T., D.B. Lipka, and T. Fischer, *FLT3 as a therapeutic target in AML: still challenging after all these years.* Blood, 2010. **116**(24): p. 5089-102.
7. Daver, N., et al., *Targeting FLT3 mutations in AML: review of current knowledge and evidence.* Leukemia, 2019. **33**(2): p. 299-312.
8. Reid, J. and S. Bhurke. *QUAZAR AML-001: Update from EHA2021 on survival outcomes according to NPM1 and FLT3 gene mutational status at diagnosis*. 2021 [cited 2021 29 November 2021]; Available from: https://aml-hub.com/medical-information/quazar-aml-001-update-from-eha2021-on-survival-outcomes-according-to-npm1-and-flt3-gene-mutational-status-at-diagnosis.
9. Dohner, H., A. Wei, and G.J. Roboz. *Survival outcomes from the QUAZAR AML-001 trial with oral azacitidine for patients with acute myeloid leukemia in remission by NPM1 and FLT3 gene mutation status at diagnosis [online]*. 2021 [cited 2021 29 November 2021]; Available from: https://library.ehaweb.org/eha/2021/eha2021-virtual-congress/324539/hartmut.dhner.survival.outcomes.from.the.quazar.aml-001.trial.with.oral.html?f=menu%3D6%2Abrowseby%3D8%2Asortby%3D2%2Amedia%3D3%2Ace_id%3D2035%2Alabel%3D21988%2Aot_id%3D25551%2Amarker%3D1285.
10. Roboz, G.J., et al., *Design of the randomized, Phase III, QUAZAR AML Maintenance trial of CC-486 (oral azacitidine) maintenance therapy in acute myeloid leukemia.* Future Oncol, 2016. **12**(3): p. 293-302.
11. Wei, A.H., *Maintenance therapy for AML: are we there yet?* Blood, 2019. **133**(13): p. 1390-1392.
12. McMahon, C.M. and S.M. Luger, *Maintenance therapy in acute myeloid leukemia: What is the future?* Semin Hematol, 2019. **56**(2): p. 102-109.
13. *Vidaza. Summary of product characteristics. Available at: https://www.medicines.org.uk/emc/product/6468/smpc#gref. Last access: January 20, 2022.*
14. *Cytarabine. Summary of product characteristics. Available at: https://www.medicines.org.uk/emc/product/3802/smpc#gref (last access: January 20, 2022).*

15. NHS. *Pan-London Haemato-Oncology Clinical Guidelines Acute Leukaemias and Myeloid Neoplasms Part 2: Acute Myeloid Leukaemia*. 2020 [cited 2021 29 November 2021]; Available from: https://rmpartners.nhs.uk/wp-content/uploads/2020/01/Pan-London-AML-Guidelines-Jan-2020.pdf.
16. Garcia-Manero, G., et al., *Phase I study of oral azacitidine in myelodysplastic syndromes, chronic myelomonocytic leukemia, and acute myeloid leukemia.* J Clin Oncol, 2011. **29**(18): p. 2521-7.
17. Garcia-Manero, G., et al., *Efficacy and safety of extended dosing schedules of CC-486 (oral azacitidine) in patients with lower-risk myelodysplastic syndromes.* Leukemia, 2016. **30**(4): p. 889-96.
18. Laille, E., et al., *Pharmacokinetics and Pharmacodynamics with Extended Dosing of CC-486 in Patients with Hematologic Malignancies.* PLoS One, 2015. **10**(8): p. e0135520.
19. Ravandi, F., et al., *Impact of subsequent allogeneic hematopoietic stem cell transplant (HSCT) on overall survival (OS) outcomes in the Quazar AML-001 trial of Oral Azacitidine (CC-486) maintenance therapy for patients with acute myeloid leukemia (AML) in first remission who were not Eligible for HSCT at study entry.* Transplantation and Cellular Therapy, 2021. **27**: p. S131-S132.
20. Loke, J., et al., *The role of allogeneic stem cell transplantation in the management of acute myeloid leukaemia: a triumph of hope and experience.* Br J Haematol, 2020. **188**(1): p. 129-146.
21. Wei, A.H., et al., *Supplementary Appendix to: Wei AH et al. Oral Azacitidine Maintenance Therapy for Acute Myeloid Leukemia in First Remission. N Engl J Med. 2020; 383(26): 2526-2537.*
22. Wei, A.H., et al., *Oral Azacitidine Maintenance Therapy for Acute Myeloid Leukemia in First Remission.* N Engl J Med, 2020. **383**(26): p. 2526-2537.
23. EMA. *Onureg: EPAR- Product information [online]*. [cited 2021 29 November 2021]; Available from: https://www.ema.europa.eu/en/documents/product-information/onureg-epar-product-information_en.pdf.
24. Stone, R.M., et al., *Midostaurin plus Chemotherapy for Acute Myeloid Leukemia with a FLT3 Mutation.* N Engl J Med, 2017. **377**(5): p. 454-464.
25. NICE. *Midostaurin for untreated acute myeloid leukaemia [TA523] [online]*. [cited 2021 29 November 2021]; Available from: https://www.nice.org.uk/guidance/ta523.
26. Wei, A., et al. *1036 CC-486 Improves Overall Survival (OS) and Relapse-Free Survival (RFS) for Patients with Acute Myeloid Leukemia (AML) in First Remission after Intensive Chemotherapy (IC), Regardless of Amount of Consolidation Received: Results from the Phase III QUAZAR AML-001 Maintenance Trial [online]*. 2020 [cited 2021 29 November 2021]; Available from: https://ash.confex.com/ash/2020/webprogram/Paper138498.html.
27. Mayer, R.J., et al., *Intensive postremission chemotherapy in adults with acute myeloid leukemia.* New England Journal of Medicine, 1994. **331**(14): p. 896-903.
28. Schlenk, R.F., *Post-remission therapy for acute myeloid leukemia.* Haematologica, 2014. **99**(11): p. 1663.
29. ClinicalTrials.gov. *Efficacy of Oral Azacitidine Plus Best Supportive Care as Maintenance Therapy in Subjects With Acute Myeloid Leukemia (AML) in Complete Remission (QUAZAR AML-001) [online]*. [cited 2021 29 November 2021]; Available from:

https://clinicaltrials.gov/ct2/show/results/NCT01757535?term=QUAZAR+AML&draw=2&rank=1&view=results.
30. FDA. *NDA Multidisciplinary Review and Evaluation*. [cited 2021 29 November 2021]; Available from: https://www.accessdata.fda.gov/drugsatfda_docs/nda/2020/214120Orig1s000MultidisciplineR.pdf.
31. EMA. *CHMP Assessment Report: Oral Azacitidine*. [cited 2021 29 November 2021]; Available from: https://www.ema.europa.eu/en/documents/assessment-report/onureg-epar-public-assessment-report_en.pdf.
32. Milligan, D.W., et al., *Guidelines on the management of acute myeloid leukaemia in adults.* Br J Haematol, 2006. **135**(4): p. 450-74.
33. Heuser, M., et al., *Acute myeloid leukaemia in adult patients: ESMO Clinical Practice Guidelines for diagnosis, treatment and follow-up.* Ann Oncol, 2020. **31**(6): p. 697-712.
34. Rashidi, A., et al., *Maintenance therapy in acute myeloid leukemia: an evidence-based review of randomized trials.* Blood, 2016. **128**(6): p. 763-73.
35. Larson, R.A., et al., *Midostaurin reduces relapse in FLT3-mutant acute myeloid leukemia: the Alliance CALGB 10603/RATIFY trial.* Leukemia, 2021. **35**(9): p. 2539-2551.
36. Visser, O., et al., *Incidence, survival and prevalence of myeloid malignancies in Europe.* Eur J Cancer, 2012. **48**(17): p. 3257-66.
37. Kayser, S., et al., *The impact of therapy-related acute myeloid leukemia (AML) on outcome in 2853 adult patients with newly diagnosed AML.* Blood, 2011. **117**(7): p. 2137-45.
38. Kottaridis, P.D., et al., *The presence of a FLT3 internal tandem duplication in patients with acute myeloid leukemia (AML) adds important prognostic information to cytogenetic risk group and response to the first cycle of chemotherapy: analysis of 854 patients from the United Kingdom Medical Research Council AML 10 and 12 trials.* Blood, 2001. **98**(6): p. 1752-9.
39. Mead, A.J., et al., *FLT3 tyrosine kinase domain mutations are biologically distinct from and have a significantly more favorable prognosis than FLT3 internal tandem duplications in patients with acute myeloid leukemia.* Blood, 2007. **110**(4): p. 1262-70.
40. Whitman, S.P., et al., *FLT3 D835/I836 mutations are associated with poor disease-free survival and a distinct gene-expression signature among younger adults with de novo cytogenetically normal acute myeloid leukemia lacking FLT3 internal tandem duplications.* Blood, 2008. **111**(3): p. 1552-9.
41. Phillippo, D., et al. *NICE DSU Technical Support Document 18: Methods for Population-Adjusted Indirecr Comparisons in Submission to NICE. Appendix D: Worked example of MAIC and STC [online]*. [cited 2021 29 November 2021]; Available from: http://nicedsu.org.uk/wp-content/uploads/2017/05/Population-adjustment-TSD-FINAL.pdf.
42. Phillippo, D.M. *Calibration of Treatment Effects in Network Meta-Analysis using Individual Patient Data [online]*. 2019 [cited 2022 21 January 2022]; Available from: https://research-information.bris.ac.uk/ws/portalfiles/portal/218211125/David_Phillippo_PhD_Thesis.pdf.

43. Coyle, D. and J.A.V. Pierre, *Economic Evaluation of Azacitidine in Elderly Patients with Acute Myeloid Leukemia with High Blast Counts.* PharamcoEconomics, 2019.
44. Tikhonova, I.A., et al., *Azacitidine for Treating Acute Myeloid Leukaemia with More Than 30 % Bone Marrow Blasts: An Evidence Review Group Perspective of a National Institute for Health and Care Excellence Single Technology Appraisal.* Pharmacoeconomics, 2017. **35**(3): p. 363-373.
45. Stein, E., et al., *Cost Effectiveness of Midostaurin in the Treatment of Newly Diagnosed FLT3-Mutated Acute Myeloid Leukemia in the United States.* Pharmacoeconomics, 2019. **37**(2): p. 239-253.
46. Lambert, J., et al., *Gemtuzumab ozogamicin for de novo acute myeloid leukemia: final efficacy and safety updates from the open-label, phase III ALFA-0701 trial.* Haematologica, 2019. **104**(1): p. 113-119.
47. Wood, B., et al. *NICE DSU Technical Support Document 19: Partitioned Survival Analysis for Decision Modelling in Health Care: A Critical Review [online]*. 2017 [cited 2022 21 January 2022]; Available from: http://nicedsu.org.uk/wp-content/uploads/2017/06/Partitioned-Survival-Analysis-final-report.pdf.
48. Huls, G., et al., *Azacitidine maintenance after intensive chemotherapy improves DFS in older AML patients.* Blood, 2019. **133**(13): p. 1457-1464.
49. Oliva, E.N., et al., *Azacitidine for Post-Remission Therapy in Elderly Patients with Acute Myeloid Leukemia : Final Results of the Qoless AZA-Amle Randomized Trial.* Blood, 2019. **134**(Supplement_1): p. 117-117.
50. Robles, C., et al., *Low-dose cytarabine maintenance therapy vs observation after remission induction in advanced acute myeloid leukemia: an Eastern Cooperative Oncology Group Trial (E5483).* Leukemia, 2000. **14**(8): p. 1349-53.
51. Latimer, N. *NICE DSU Technical Support Document 14: Survival analysis for economic evaluations alongside clinical trials -Extrapolation with patient-level data [online]*. [cited 2021 29 November 2021]; Available from: http://nicedsu.org.uk/wp-content/uploads/2016/03/NICE-DSU-TSD-Survival-analysis.updated-March-2013.v2.pdf.
52. Hess, K. and G. R. *Muhaz: Hazard Function Estimation in Survival Analysis [online]*. 2021 [cited 2022 21 January 2022]; Available from: https://CRAN.R-project.org/package=muhaz.
53. Therneau, T. *A Package for Survival Analysis in r [online]*. 2021 [cited 2022 21 January 2022]; Available from: https://CRAN.R-project.org/package=survival.
54. Tremblay, G. and P.B. Haines, A., *A Criterion-based Approach for the Systematic and Transparent Extrapolation of Clinical Trial Survival Data.* 2015. **2**(2): p. 147-60.
55. Royston, P. and M.K. Parmar, *Flexible parametric proportional-hazards and proportional-odds models for censored survival data, with application to prognostic modelling and estimation of treatment effects.* Stat Med, 2002. **21**(15): p. 2175-97.
56. NICE. *Gemtuzumab ozogamicin for untreated acute myeloid leukaemia [TA545] [online]*. [cited 2021 29 November 2021]; Available from: https://www.nice.org.uk/guidance/ta545.
57. NICE. *Azacitidine for treating acute myeloid leukaemia with more than 30% bone marrow blasts [TA399] [online]*. [cited 2021 29 November 2021]; Available from: https://www.nice.org.uk/guidance/ta399/resources/azacitidine-for-treating-

acute-myeloid-leukaemia-with-more-than-30-bone-marrow-blasts-pdf-82602918570949.
58. Coens, C., et al., *International standards for the analysis of quality-of-life and patient-reported outcome endpoints in cancer randomised controlled trials: recommendations of the SISAQOL Consortium.* Lancet Oncol, 2020. **21**(2): p. e83-e96.
59. Yang, X. and S. Shoptaw, *Assessing missing data assumptions in longitudinal studies: an example using a smoking cessation trial.* Drug Alcohol Depend, 2005. **77**(3): p. 213-25.
60. Bewersdorf, J.P., et al., *Cost-effectiveness analysis of oral azacitidine maintenance therapy in acute myeloid leukemia.* Blood Advances, 2021. **5**(22): p. 4686-4690.
61. NICE. *Gilteritinib for treating relapsed or refractory acute myeloid leukaemia [TA642] [online]*. [cited 2021 29 November 2021]; Available from: https://www.nice.org.uk/guidance/ta642/evidence/final-appraisal-determination-committee-papers-pdf-8830117118.
62. NICE. *Liposomal cytarabine–daunorubicin for untreated acute myeloid leukaemia [TA552] [online]*. 2018 [cited 2022 21 January 2022]; Available from: nice.org.uk/guidance/ta552/resources/liposomal-cytarabinedaunorubicin-for-untreated-acute-myeloid-leukaemia-pdf-82607018513605.

Appendices

Appendix B.1

Parametric survival models were fit to the individual patient-level data from the QUAZAR AML-001 trial for each outcome (overall survival [OS] and relapse-free survival [RFS]). Survival analyses and assessments conducted and presented in the subsequent sections follow the structure outlined by Tremblay et al. 2016 and were supported by the metrics and criteria described in the National Institute for Health and Care Excellence (NICE) Technical Support Document (TSD) 14.

Parametric models were fit to extrapolate the probability of survival from event beyond the follow-up time of the trial. The exponential, Weibull, log-logistic, log-normal, generalized gamma, and Gompertz distributions were used. Exponential, Weibull, and Gompertz distributions were parameterized as proportional hazard (PH) models while log-logistic, log-normal, and generalized gamma distributions were parameterized as accelerated failure time (AFT) models. Suitability of survival plots were assessed as per the NICE TSD 14. Specifically, plots of the estimated parametric survival curves were overlayed with KM curves to visually assess their fit to the trial data and beyond. A time horizon of 40 years, or 480 months, was used when considering how well the parametric distributions extrapolated beyond the support of the trial data. Parametric models were fit to each individual treatment arm separately (hereafter referred to as “individual” models) as well as the pooled set patients with a treatment covariate (hereafter referred to as “joint” models; sometimes referred to as proportional treatment models). Log-cumulative hazard plots for estimated parametric models overlayed with KM curves were then used to assess the suitability of each parametric model. Particularly, a lack of parallel lines between treatment arms in the observed period would indicate a violation of the PH assumption and mean that distributions reliant on this assumption (ie, PH models) may not be the optimal choice. As described by Tremblay et al. 2016, joint AFT models and individual models may still be considered in the presence of evidence of PH violation. Model fit statistics, including the Akaike’s information criterion (AIC) and Bayesian information criterion (BIC), were used to assess and compare the fit of both individual and joint models (lower values

indicate better model fit). Finally, the clinical validity of the extrapolated OS and RFS curves was assessed by clinical experts.

Overall Survival

The probability of survival over time by treatment arm as estimated by the KM method is shown with KM curves in
Figure *18*. The median survival time for oral AZA and placebo was [redacted] months, respectively. This is aligned with the median survival time for oral AZA and placebo from the ITT population (oral AZA: 24.7 [95% CI: 18.7, 30.5]; placebo: 14.8 [95% CI: 11.7, 17.6]).

Figure 18. Kaplan–Meier curves by treatment arm – OS, ITT population censored for HSCT

The unstratified Cox PH model estimated oral AZA to provide a reduced rate of mortality compared to placebo ([redacted]). This is comparable to the HR estimated using the ITT population (HR: 0.73; 95% CI: 0.59, 0.90). The log-cumulative hazard plot and Schoenfeld residual plots showed violation of the PH assumption. A visual inspection of the log-cumulative hazard plot suggested that the

two lines were not parallel (Figure 19). Similarly, the Schoenfeld residual plot displayed a non-horizontal line and the Grambsch-Therneau global Schoenfeld residual test value was significant (p-value [redacted]; Figure 20).

Figure 19. Log-cumulative hazard plot from unstratified Cox PH model – OS, ITT population censored for HSCT

Figure 20. Schoenfeld residuals plot from unstratified Cox PH model – OS, ITT population censored for HSCT

The stratified Cox PH model estimated oral AZA to be more beneficial compared to placebo ( ). This is aligned with the HR estimated using the ITT population (HR: 0.69; 95% CI: 0.56, 0.86). According to the Schoenfeld residual plot and Grambsch-Therneau global Schoenfeld residual test, the PH assumption was violated since the line on the plot was not horizontal and the p-value was statistically significant (p-value ; Figure 21).

Figure 21.Schoenfeld residuals plot from stratified Cox PH model – OS, ITT population censored for HSCT

Parametric curves fit using joint models with a treatment covariate are shown in Figure 22 to Figure 27 while individual models are shown in Figure 28 to Figure 33. Note in these figures, KM curves are drawn with a solid line; parametric curves are drawn with a dashed line. Model fit statistics are presented in Table 70.

Based on the AIC and BIC, joint generalized gamma provided the best statistical fit for the ITT population censored for HSCT among all distributions (Table 70). Visually the oral AZA arm remained apart from the no active treatment arm which corresponds to clinical expectations and the extrapolated tails were clinically plausible. Overall, this aligns with the assessment for the ITT population where joint generalized gamma was also determined to have the optimal fit based on AIC/BIC, visual inspection, and clinical plausibility (Section A.10, Document A of the company submission [CS]).

Figure 22. Parametric curves fit to the OS outcome in ITT population censored for HSCT – Exponential distribution, joint model

Figure 23. Parametric curves fit to the OS outcome in ITT population censored for HSCT – Weibull distribution, joint model

Figure 24. Parametric curves fit to the OS outcome in ITT population censored for HSCT – Log-logistic distribution, joint model

Figure 25. Parametric curves fit to the OS outcome in ITT population censored for HSCT – Log-normal distribution, joint model

Figure 26. Parametric curves fit to the OS outcome in ITT population censored for HSCT – Generalized gamma distribution, joint model

Figure 27.Parametric curves fit to the OS outcome in ITT population censored for HSCT – Gompertz distribution, joint model

Figure 28. Parametric curves fit to the OS outcome in ITT population censored for HSCT – Exponential distribution, individual model

Figure 29.Parametric curves fit to the OS outcome in ITT population censored for HSCT – Weibull distribution, individual model

Figure 30.Parametric curves fit to the OS outcome in ITT population censored for HSCT – Log-logistic distribution, individual model

Figure 31.Parametric curves fit to the OS outcome in ITT population censored for HSCT – Log-normal distribution, individual model

Figure 32.Parametric curves fit to the OS outcome in ITT population censored for HSCT – Generalized gamma distribution, individual model

Figure 33.Parametric curves fit to the OS outcome in ITT population censored for HSCT – Gompertz distribution, individual model

Table 70. Model fit statistics (AIC and BIC) for parametric models of the OS outcome, ITT population censored for HSCT

Parametric Model	AIC	Ranks based on AIC	BIC	Ranks based on BIC
Joint models				
Exponential	[redacted]	[redacted]	[redacted]	[redacted]
Weibull	[redacted]	[redacted]	[redacted]	[redacted]
Log-logistic	[redacted]	[redacted]	[redacted]	[redacted]
Log-normal	[redacted]	[redacted]	[redacted]	[redacted]
Generalized Gamma	[redacted]	[redacted]	[redacted]	[redacted]
Gompertz	[redacted]	[redacted]	[redacted]	[redacted]
Individual models				
Exponential	[redacted]	[redacted]	[redacted]	[redacted]
Weibull	[redacted]	[redacted]	[redacted]	[redacted]
Log-logistic	[redacted]	[redacted]	[redacted]	[redacted]
Log-normal	[redacted]	[redacted]	[redacted]	[redacted]
Generalized Gamma	[redacted]	[redacted]	[redacted]	[redacted]
Gompertz	[redacted]	[redacted]	[redacted]	[redacted]

Abbreviations: AIC, Akaike's information criteria; BIC, Bayesian information criterion; HSCT, hematopoietic stem cell transplant; ITT, intention-to-treat; OS, overall survival.

Relapse-free survival

The probability of RFS over time by treatment arm as estimated by the KM method is shown with KM curves in Figure 34. The median survival time for oral AZA and placebo was [redacted], respectively. This is aligned with the median survival time for oral AZA and placebo from the ITT population (oral AZA: [redacted]).

Figure 34.Kaplan–Meier curves by treatment arm – RFS, ITT population censored for HSCT

The unstratified Cox PH model estimated oral AZA to provide increased benefit compared to placebo ([redacted]). is aligned with the HR estimated using the ITT population ([redacted]). The log-cumulative hazard plot and Schoenfeld residual plots showed violation of the PH assumption. A visual inspection of the log-cumulative hazard plot suggested that the two lines were not parallel (Figure 35). Similarly, the Schoenfeld residual plot displayed a non-horizontal line and the Grambsch-Therneau global Schoenfeld residual test value was significant (p-value <[redacted]; Figure 36).

Figure 35. Log-cumulative hazard plot from unstratified Cox PH model – RFS, ITT population censored for HSCT

Figure 36. Schoenfeld residuals plot from stratified Cox PH model – RFS, ITT population censored for HSCT

The stratified Cox PH model estimated oral AZA to be more beneficial compared to placebo ([redacted]). This is aligned with the HR estimated using the ITT population (HR: 0.65; 95% CI: 0.52, 0.81). According to the Schoenfeld residual plot and Grambsch-Therneau global Schoenfeld residual test, the PH assumption was violated since the line on the plot was not horizontal and the p-value was statistically significant (p-value = 0.001;Figure 37).

Figure 37. Schoenfeld residuals plot from stratified Cox PH model – RFS, ITT population censored for HSCT

Parametric curves fit using joint models with a treatment covariate are shown in Figure 38 to Figure 43 while individual models are shown in Figure 44 to Figure 49. Note in these figures, KM curves are drawn with a solid line; parametric curves are drawn with a dashed line. Model fit statistics are presented in Table 71.

The optimal RFS model for the ITT population censored for HSCT appears to be the joint log-logistic model. This model exhibits no cross-over of the treatment arms (Figure 40), has good visual fit, and has higher precision than the individual models, due to the higher statistical power of fitting a single model to both treatment arms. This rests on the assumption that the relative treatment effect can be modeled by an AFT

factor, which the (reasonably straight) lines in the log-cumulative hazard plot (Figure 35) supports. From a statistical fit perspective, the log-logistic distribution is the best fitting joint model in terms of AIC and BIC (Table 71). Of note, the joint log-logistic model was also determined to have the optimal fit for the ITT population (Section A.10, Document A of the CS).

Figure 38.Parametric curves fit to the RFS outcome in ITT population censored for HSCT – Exponential distribution, joint model

Figure 39.Parametric curves fit to the RFS outcome in ITT population censored for HSCT – Weibull distribution, joint model

Figure 40.Parametric curves fit to the RFS outcome in ITT population censored for HSCT – Log-logistic distribution, joint model

Figure 41.Parametric curves fit to the RFS outcome in ITT population censored for HSCT – Log-normal distribution, joint model

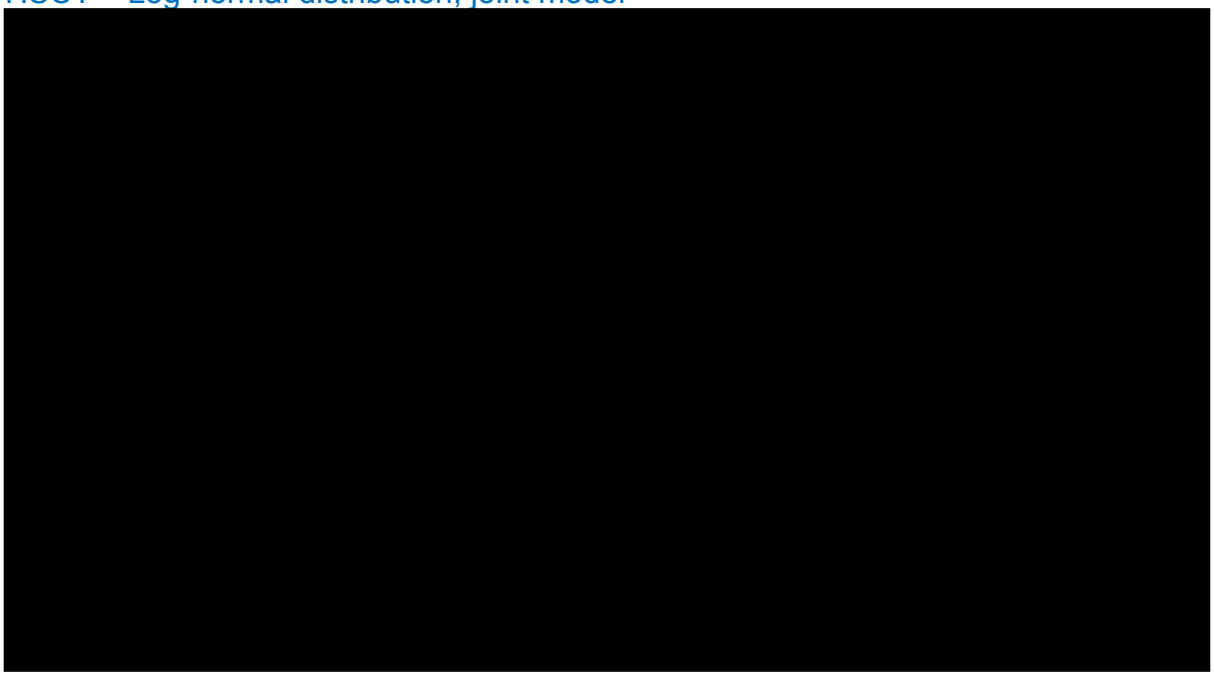


Figure 42. Parametric curves fit to the RFS outcome in ITT population censored for HSCT – Generalized gamma distribution, joint model

Figure 43. Parametric curves fit to the RFS outcome in ITT population censored for HSCT – Gompertz distribution, joint model

Figure 44. Parametric curves fit to the RFS outcome in ITT population censored for HSCT – Exponential distribution, individual model

Figure 45. Parametric curves fit to the RFS in ITT population censored for HSCT – Weibull distribution, individual model

Figure 46. Parametric curves fit to the RFS outcome in ITT population censored for HSCT – Log-logistic distribution, individual model

Figure 47. Parametric curves fit to the RFS outcome in ITT population censored for HSCT – Log-normal distribution, individual model

Figure 48. Parametric curves fit to the RFS outcome in ITT population censored for HSCT – Generalized gamma distribution, individual model

Figure 49. Parametric curves fit to the RFS outcome in ITT population censored for HSCT – Gompertz distribution, individual model

Table 71. Model fit statistics (AIC and BIC) for parametric models of the RFS outcome, ITT population censored for HSCT

Parametric Model	AIC	Ranks based on AIC	BIC	Ranks based on BIC
Joint models				
Exponential	[redacted]	[redacted]	[redacted]	[redacted]
Weibull	[redacted]	[redacted]	[redacted]	[redacted]
Log-logistic	[redacted]	[redacted]	[redacted]	[redacted]
Log-normal	[redacted]	[redacted]	[redacted]	[redacted]
Generalized Gamma	[redacted]	[redacted]	[redacted]	[redacted]
Gompertz	[redacted]	[redacted]	[redacted]	[redacted]
Individual models				
Exponential	[redacted]	[redacted]	[redacted]	[redacted]
Weibull	[redacted]	[redacted]	[redacted]	[redacted]
Log-logistic	[redacted]	[redacted]	[redacted]	[redacted]
Log-normal	[redacted]	[redacted]	[redacted]	[redacted]
Generalized Gamma	[redacted]	[redacted]	[redacted]	[redacted]
Gompertz	[redacted]	[redacted]	[redacted]	[redacted]

Abbreviations: AIC, Akaike's information criteria; BIC, Bayesian information criterion; HSCT, hematopoietic stem cell transplant; ITT, intention-to-treat; RFS, relapse free survival.

Appendix B.8

Parametric survival models were fit to the individual patient-level data from the QUAZAR AML-001 trial for each outcome (overall survival [OS] and relapse-free survival [RFS]). Survival analyses and assessments conducted and presented in the subsequent sections follow the structure outlined by Tremblay et al. 2016 and were supported by the metrics and criteria described in the National Institute for Health and Care Excellence (NICE) Technical Support Document (TSD) 14.

Parametric models were fit to extrapolate the probability of survival from event beyond the follow-up time of the trial. The exponential, Weibull, log-logistic, log-normal, generalized gamma, and Gompertz distributions were used. Exponential, Weibull, and Gompertz distributions were parameterized as proportional hazard (PH) models while log-logistic, log-normal, and generalized gamma distributions were parameterized as accelerated failure time (AFT) models. Suitability of survival plots were assessed as per the NICE TSD 14. Specifically, plots of the estimated parametric survival curves were overlayed with KM curves to visually assess their fit to the trial data and beyond. A time horizon of 40 years, or 480 months, was used when considering how well the parametric distributions extrapolated beyond the support of the trial data. Parametric models were fit to each individual treatment arm separately (hereafter referred to as “individual” models) as well as the pooled set patients with a treatment covariate (hereafter referred to as “joint” models; sometimes referred to as proportional treatment models). Log-cumulative hazard plots for estimated parametric models overlayed with KM curves were then used to assess the suitability of each parametric model. Particularly, a lack of parallel lines between treatment arms in the observed period would indicate a violation of the PH assumption and mean that distributions reliant on this assumption (ie, PH models) may not be the optimal choice. As described by Tremblay et al. 2016, joint AFT models and individual models may still be considered in the presence of evidence of PH violation. Model fit statistics, including the Akaike’s information criterion (AIC) and Bayesian information criterion (BIC), were used to assess and compare the fit of both individual and joint models (lower values

indicate better model fit). Finally, the clinical validity of the extrapolated OS and RFS curves was assessed by clinical experts.

Overall Survival

The probability of survival over time by treatment arm as estimated by the KM method is shown with KM curves in Figure 50. The median survival time for oral AZA and placebo was [redacted] months, respectively.

Figure 50. Kaplan–Meier curves by treatment arm – OS, ITT population restricted to EU-only patients

The unstratified Cox PH model estimated oral AZA to result in a reduced rate of mortality compared to placebo [redacted]. The log-cumulative hazard plot and Schoenfeld residual plot showed violation of the PH assumption. A visual inspection of the log-cumulative hazard plot suggested that the two lines were not parallel (Figure 51). Similarly, the Schoenfeld residual plot displayed a non-horizontal line and the Grambsch-Therneau global Schoenfeld residual test value was statistically significant (p-value [redacted]; Figure 52).

Similarly, the stratified Cox PH model estimated oral AZA to result in a reduced rate of mortality compared to placebo [redacted]. According to the Schoenfeld residual plot and Grambsch-Therneau global Schoenfeld residual test, the PH assumption was violated since the line on the plot was not horizontal and the p-value was statistically significant (p-value [redacted];Figure 53). Given the shape of the KM-estimated hazard functions and suspected violations of the PH assumption, individual model fits and joint AFT models (log-normal, log-logistic, generalized gamma) may be preferred over joint PH models (exponential, Weibull, Gompertz) because they do not assume hazards between treatment arms to be proportional.

Figure 51. Log-cumulative hazard plot from unstratified Cox PH model – OS, ITT population restricted to EU-only patients

Figure 52. Schoenfeld residuals plot from unstratified Cox PH model – OS, ITT population restricted to EU-only patients

Figure 53. Schoenfeld residuals plot from stratified Cox PH model – OS, ITT population restricted to EU-only patients

Parametric curves from joint models are shown in Figure 54 to Figure 59, while parametric curves from individual models are shown in Figure 60 to Figure 65. Note in these figures, KM curves are drawn with a solid line; parametric curves are drawn with

a dashed line. Model fit statistics (AIC, BIC) for all parametric distributions are presented in Table 72.

Based on the AIC and BIC, joint generalized gamma provided the best statistical fit for the ITT population restricted to the EU subgroup among all distributions (Figure 58). Visual inspection of the joint generalized gamma survival function supports this conclusion, in that the curve closely fits the data and provides sensible extrapolations with the probability of survival approaching zero by 40 years. Overall, this aligns with the assessment for the ITT population where joint generalized gamma was also determined to have the optimal fit based on AIC/BIC, visual inspection, and clinical plausibility (Section A.10, Document A of the company submission [CS]).

Figure 54. Parametric curves fit to the OS outcome in the ITT population restricted to EU-only patients – Exponential distribution, joint model

Figure 55. Parametric curves fit to the OS outcome in the ITT population restricted to EU-only patients – Weibull distribution, joint model

Figure 56.Parametric curves fit to the OS outcome in the ITT population restricted to EU-only patients – Log-logistic distribution, joint model

Figure 57.Parametric curves fit to the OS outcome in the ITT population restricted to EU-only patients – Log-normal distribution, joint model

Figure 58.Parametric curves fit to the OS outcome in the ITT population restricted to EU-only patients – Generalized gamma distribution, joint model

Figure 59. Parametric curves fit to the OS outcome in the ITT population restricted to EU-only patients – Gompertz distribution, joint model

Figure 60.Parametric curves fit to the OS outcome in the ITT population restricted to EU-only patients – Exponential distribution, individual model

Figure 61. Parametric curves fit to the OS outcome in the ITT population restricted to EU-only patients – Weibull distribution, individual model

Figure 62. Parametric curves fit to the OS outcome in the ITT population restricted to EU-only patients – Log-logistic distribution, individual model

Figure 63. Parametric curves fit to the OS outcome in the ITT population restricted to EU-only patients – Log-normal distribution, individual model

Figure 64. Parametric curves fit to the OS outcome in the ITT population restricted to EU-only patients – Generalized gamma distribution, individual model

Figure 65. Parametric curves fit to the OS outcome in the ITT population restricted to EU-only patients – Gompertz distribution, individual model

Table 72. Model fit statistics (AIC and BIC) for parametric models of the OS outcome in the ITT population restricted to EU-only patients

Parametric Model	AIC	Ranks based on AIC	BIC	Ranks based on BIC
Joint models				
Exponential	[redacted]	[redacted]	[redacted]	[redacted]
Weibull	[redacted]	[redacted]	[redacted]	[redacted]
Log-logistic	[redacted]	[redacted]	[redacted]	[redacted]
Log-normal	[redacted]	[redacted]	[redacted]	[redacted]
Generalized Gamma	[redacted]	[redacted]	[redacted]	[redacted]
Gompertz	[redacted]	[redacted]	[redacted]	[redacted]
Individual models – Oral AZA arm				
Exponential	[redacted]	[redacted]	[redacted]	[redacted]
Weibull	[redacted]	[redacted]	[redacted]	[redacted]
Log-logistic	[redacted]	[redacted]	[redacted]	[redacted]
Log-normal	[redacted]	[redacted]	[redacted]	[redacted]
Generalized Gamma	[redacted]	[redacted]	[redacted]	[redacted]
Gompertz	[redacted]	[redacted]	[redacted]	[redacted]
Individual models – Placebo arm				
Exponential	[redacted]	[redacted]	[redacted]	[redacted]
Weibull	[redacted]	[redacted]	[redacted]	[redacted]
Log-logistic	[redacted]	[redacted]	[redacted]	[redacted]
Log-normal	[redacted]	[redacted]	[redacted]	[redacted]
Generalized Gamma	[redacted]	[redacted]	[redacted]	[redacted]
Gompertz	[redacted]	[redacted]	[redacted]	[redacted]
Individual models – Sum of two arms				
Exponential	[redacted]	[redacted]	[redacted]	[redacted]
Weibull	[redacted]	[redacted]	[redacted]	[redacted]
Log-logistic	[redacted]	[redacted]	[redacted]	[redacted]
Log-normal	[redacted]	[redacted]	[redacted]	[redacted]
Generalized Gamma	[redacted]	[redacted]	[redacted]	[redacted]
Gompertz	[redacted]	[redacted]	[redacted]	[redacted]

Abbreviations: AIC, Akaike's information criteria; BIC, Bayesian information criterion; ITT, intention-to-treat; OS, overall survival.

The difference in mean time to event between treatment arms for OS was estimated via both the KM method over the duration of trial follow-up ([redacted] months; the minimum of the last observations across treatment arms) and via parametric models restricted to 40 years (Table 73). The 95% bootstrapped CIs for parametric curves are presented to assist with the inspection of uncertainty. The KM estimated difference in mean time to mortality between oral AZA and placebo was [redacted] months. Most parametric models estimated a larger increase in mean time to mortality for oral AZA compared

to placebo than did the KM estimator. All joint models estimate a significant increase in time to mortality for oral AZA compared to placebo.

Table 73. Difference in mean time to event for OS between oral AZA and placebo arms in the ITT population restricted to EU-only patients

Model	Difference in mean OS, months (oral AZA – placebo)	Difference in mean OS, months 95% CI, lower bound	Difference in mean OS, months 95% CI, upper bound
KM	[redacted]	NA	NA
Joint models			
Exponential	[redacted]	[redacted]	[redacted]
Weibull	[redacted]	[redacted]	[redacted]
Log-Logistic	[redacted]	[redacted]	[redacted]
Log-Normal	[redacted]	[redacted]	[redacted]
Generalized Gamma	[redacted]	[redacted]	[redacted]
Gompertz	[redacted]	[redacted]	[redacted]
Individual models			
Exponential	[redacted]	[redacted]	[redacted]
Weibull	[redacted]	[redacted]	[redacted]
Log-Logistic	[redacted]	[redacted]	[redacted]
Log-Normal	[redacted]	[redacted]	[redacted]
Generalized Gamma	[redacted]	[redacted]	[redacted]
Gompertz	[redacted]	[redacted]	[redacted]

Abbreviations: AZA, azacitidine; CI, confidence interval; ITT, intention-to-treat; KM, Kaplan–Meier; NA, not applicable; OS, overall survival.

Log-cumulative hazard plots for joint models and individual models are presented in Figure 66 and

Figure *67*, respectively. According to a visual assessment of the log-cumulative hazard plots, generalized gamma appears to be the best fit followed by log-normal. It should be noted, events early in time have created the stretching effect seen in the graphs but they represent a small number of events as the x-axis is on a log scale. As suggested in Tremblay et al. 2016, when log-cumulative hazard plots are not parallel, but relatively straight, AFT models (ie, log-normal, log-logistic and generalized gamma) with a treatment covariate and individual parametric models without a treatment covariate are preferred over parametric models with a treatment covariate that assume PH (ie, Weibull, exponential and Gompertz). These findings are consistent with the evidence presented above regarding model fit, AIC and BIC.

Figure 66. Log-cumulative hazard versus log time plots for the OS outcome in the ITT population restricted to EU-only patients – parametric model fits (dashed line) compared to KM fits (solid line) by treatment arm; joint models

Abbreviations: trt = treatment; pbo = placebo.

Figure 67. Log-cumulative hazard versus log time plots for the OS outcome in the ITT population restricted to EU-only patients – parametric model fits (dashed line) compared to KM fits (solid line) by treatment arm; individual models

Abbreviations: trt, treatment; pbo, placebo.

The marginal survival gain both pre- and post-extrapolation for each model is presented in Table 74. The cut-point to distinguish pre- and post-extrapolation time periods for the OS outcome was [redacted] months (the minimum of the last observations across treatment arms). According to the results, all models satisfied Criterion 5 in terms of having rate of survival gain in the extrapolated tail being lower than the rate of gain observed in the KM curve. In addition, for all models, the extrapolated tail rate of gain was lower compared to the pre-extrapolation rate of gain.

Table 74. Evaluation of Criterion 5 – estimated rate of OS gain per month by receiving oral AZA instead of placebo in the ITT population restricted to EU-only patients, before and after the trial cutoff

Models	Pre-extrapolation	Extrapolated tail
KM	[redacted]	-
Joint models		
Exponential	[redacted]	[redacted]
Weibull	[redacted]	[redacted]

Log-Logistic		[redacted]			[redacted]	
Log-Normal						
Generalized Gamma						
Gompertz						
Individual models						
Exponential						
Weibull						
Log-Logistic						
Log-Normal						
Generalized Gamma						
Gompertz						

Notes: The rate of survival gain in the pre-extrapolation period is defined as the difference in survival between oral AZA and placebo at [redacted] months divided by the number of months in the pre-extrapolation period (ie [redacted] months). The rate of survival gain in the post-extrapolation period is defined as the marginal relative difference in the extrapolated period (post cut-off) divided by the number of months post-cut-off. Negative values represent the rate of survival loss for oral AZA (ie, gain for placebo), which in the case of most fitted models indicate a crossing of curves. Abbreviations: AZA, azacitidine; ITT, intention-to-treat; KM, Kaplan–Meier; OS, overall survival.

Relapse-free Survival

The probability of RFS over time by treatment arm as estimated by the KM method is shown with KM curves in Figure 68. The median survival time for oral AZA and placebo was [redacted] months, respectively.

The unstratified Cox PH model estimated oral AZA to result in a reduced rate of relapse or mortality compared to placebo (HR: 0.56; 95% CI: 0.43, 0.73). The log-cumulative hazard plot and Schoenfeld residual plots showed violation of the PH assumption. A visual inspection of the log-cumulative hazard plot suggested that the two lines were not parallel (Figure 69). Similarly, the Schoenfeld residual plot displayed a non-horizontal line and the Grambsch-Therneau global Schoenfeld residual test value was statistically significant (p-value [redacted]; Figure 70**).**

Figure 68. Kaplan–Meier curves by treatment arm – RFS, ITT population restricted to EU-only patients

Figure 69. Log-cumulative hazard plot from unstratified Cox PH model– RFS, ITT population restricted to EU-only patients

Figure 70. Schoenfeld residuals plot from unstratified Cox PH model – RFS, ITT population restricted to EU-only patients

The stratified Cox PH model estimated oral AZA to be more beneficial compared to placebo [redacted]. According to the Schoenfeld residual plot and Grambsch-Therneau global Schoenfeld residual test, the PH assumption was violated since the line on the plot was not horizontal and the p-value was statistically significant (p-value [redacted]; Figure 71). Given the shape of the KM-estimated hazard functions and suspected violations of the PH assumption, individual model fits and joint AFT models (log-normal, log-logistic, generalized gamma) may be preferred over joint PH models (exponential, Weibull, Gompertz) because they do not assume hazards between treatment arms to be proportional.

Figure 71. Schoenfeld residuals plot from stratified Cox PH model – RFS, ITT population restricted to EU-only patients

Parametric curves from joint models are shown in Figure 72 to Figure 77, while parametric curves from individual models are shown in Figure 78 to Figure 83**.** Note in these figures, KM curves are drawn with a solid line; parametric curves are drawn with a dashed line. Model fit statistics (AIC, BIC) for all parametric distributions are presented in Table B.8.4.

The optimal RFS model for the ITT population restricted to the EU subgroup appears to be the joint log-logistic model (Figure 74). This model exhibits no cross-over of the treatment arms, has good visual fit to the data, and has a higher precision than the individual models, due to the higher statistical power of fitting a single model to both treatment arms. This rests on the assumption that the relative treatment effect can be modeled by an AFT factor, which the (reasonably straight) lines in the log cumulative hazard plot (Figure 69) supports. From a statistical standpoint, the log-logistic distribution is the best fitting joint model in terms of AIC and BIC (Table 75). Overall, this aligns with the assessment for the ITT population where the joint log-logistic model was also determined to have the optimal fit (**Section A.10**, **Document A** of the CS).

Figure 72. Parametric curves fit to the RFS outcome in the ITT population restricted to EU-only patients – Exponential distribution, joint model

Figure 73. Parametric curves fit to the RFS outcome in the ITT population restricted to EU-only patients – Weibull distribution, joint model

Figure 74. Parametric curves fit to the RFS outcome in the ITT population restricted to EU-only patients – Log-logistic distribution, joint model

Figure 75. Parametric curves fit to the RFS outcome in the ITT population restricted to EU-only patients – Log-normal distribution, joint model

Figure 76. Parametric curves fit to the RFS outcome in the ITT population restricted to EU-only patients – Generalized gamma distribution, joint model

Figure 77. Parametric curves fit to the RFS outcome in the ITT population restricted to EU-only patients – Gompertz distribution, joint model

Figure 78. Parametric curves fit to the RFS outcome in the ITT population restricted to EU-only patients – Exponential distribution, individual model

Figure 79. Parametric curves fit to the RFS outcome in the ITT population restricted to EU-only patients – Weibull distribution, individual model

Figure 80. Parametric curves fit to the RFS outcome in the ITT population restricted to EU-only patients – Log-logistic distribution, individual model

Figure 81. Parametric curves fit to the RFS outcome in the ITT population restricted to EU-only patients – Log-normal distribution, individual model

Figure 82. Parametric curves fit to the RFS outcome in the ITT population restricted to EU-only patients – Generalized gamma distribution, individual model

Figure 83. Parametric curves fit to the RFS outcome in the ITT population restricted to EU-only patients – Gompertz distribution, individual model

Table 75. Model fit statistics (AIC and BIC) for parametric models of the RFS outcome in the ITT population restricted to EU-only patients

Parametric Model	AIC	Ranks based on AIC	BIC	Ranks based on BIC
Joint models				
Exponential	[redacted]	[redacted]	[redacted]	[redacted]
Weibull	[redacted]	[redacted]	[redacted]	[redacted]
Log-logistic	[redacted]	[redacted]	[redacted]	[redacted]
Log-normal	[redacted]	[redacted]	[redacted]	[redacted]
Generalized Gamma	[redacted]	[redacted]	[redacted]	[redacted]
Gompertz	[redacted]	[redacted]	[redacted]	[redacted]
Individual models – Oral AZA arm				
Exponential	[redacted]	[redacted]	[redacted]	[redacted]
Weibull	[redacted]	[redacted]	[redacted]	[redacted]
Log-logistic	[redacted]	[redacted]	[redacted]	[redacted]
Log-normal	[redacted]	[redacted]	[redacted]	[redacted]
Generalized Gamma	[redacted]	[redacted]	[redacted]	[redacted]
Gompertz	[redacted]	[redacted]	[redacted]	[redacted]
Individual models – Placebo arm				
Exponential	[redacted]	[redacted]	[redacted]	[redacted]
Weibull	[redacted]	[redacted]	[redacted]	[redacted]
Log-logistic	[redacted]	[redacted]	[redacted]	[redacted]
Log-normal	[redacted]	[redacted]	[redacted]	[redacted]
Generalized Gamma	[redacted]	[redacted]	[redacted]	[redacted]
Gompertz	[redacted]	[redacted]	[redacted]	[redacted]
Individual models – Sum of two arms				
Exponential	[redacted]	[redacted]	[redacted]	[redacted]
Weibull	[redacted]	[redacted]	[redacted]	[redacted]
Log-logistic	[redacted]	[redacted]	[redacted]	[redacted]
Log-normal	[redacted]	[redacted]	[redacted]	[redacted]
Generalized Gamma	[redacted]	[redacted]	[redacted]	[redacted]
Gompertz	[redacted]	[redacted]	[redacted]	[redacted]

Abbreviations: AIC, Akaike's information criteria; BIC, Bayesian information criterion; ITT, intention-to-treat, RFS, relapse free survival.

The estimated difference in mean time to event between treatment arms for RFS was estimated via both the KM method over the duration of trial follow-up ([redacted] months; the minimum of the last observations across treatment arms) and via parametric models restricted to 40 years (Table 76). The 95% bootstrapped CIs for parametric curves are presented to assist with the inspection of uncertainty. The KM estimated difference in mean time to relapse or mortality between oral AZA and placebo was [redacted] months. Except for individual Gompertz, all parametric models estimated a larger

increase in mean time to relapse or mortality for oral AZA compared to placebo than did the KM estimator. All joint models estimate a significant increase in time to relapse or mortality for oral AZA compared to placebo.

Table 76. Difference in mean time to event for RFS between oral AZA and placebo arms in the ITT population restricted to EU-only patients

Model	Difference in mean RFS, months (oral AZA – placebo)	Difference in mean RFS, months 95% CI, lower bound	Difference in mean RFS, months 95% CI, upper bound
KM	[redacted]	NA	NA
Joint models			
Exponential	[redacted]	[redacted]	[redacted]
Weibull	[redacted]	[redacted]	[redacted]
Log-Logistic	[redacted]	[redacted]	[redacted]
Log-Normal	[redacted]	[redacted]	[redacted]
Generalized Gamma	[redacted]	[redacted]	[redacted]
Gompertz	[redacted]	[redacted]	[redacted]
Individual models			
Exponential	[redacted]	[redacted]	[redacted]
Weibull	[redacted]	[redacted]	[redacted]
Log-Logistic	[redacted]	[redacted]	[redacted]
Log-Normal	[redacted]	[redacted]	[redacted]
Generalized Gamma	[redacted]	[redacted]	[redacted]
Gompertz	[redacted]	[redacted]	[redacted]

Abbreviations: AZA, azacitidine; CI, confidence interval; ITT, intention-to-treat; KM, Kaplan–Meier; NA, not applicable; RFS, relapse free survival

Log-cumulative hazard plots for joint models and individual models for RFS are presented in Figure 84 and

Figure *85*, respectively. According to a visual assessment of the log-cumulative hazard plots, log-logistic appears to be the best fit. In comparison, the model fit for Gompertz is less optimal. These findings are consistent with the evidence presented above regarding model fit, AIC, BIC, and clinical plausibility.

Figure 84. Log-cumulative hazard versus log time plots for the RFS outcome in the ITT population restricted to EU-only patients – parametric model fits (dashed line) compared to KM fits (solid line) by treatment arm; joint models

Figure 85. Log-cumulative hazard versus log time plots for the RFS outcome in the ITT population restricted to EU-only patients – parametric model fits (dashed line) compared to KM fits (solid line) by treatment arm; individual models

The marginal survival gain both pre- and post-extrapolation for each model is presented in Table 77**.** The cut-point to distinguish pre- and post-extrapolation time periods for the RFS outcome was [redacted] months (the minimum of the last observations across treatment arms). According to the results, all the models, except the joint Gompertz model, satisfied Criterion 5 in terms of having rate of gain in the extrapolated tail being lower than the rate of gain observed in the KM curve. In addition, for all models, the extrapolated tail rate of gain was lower compared to the pre-extrapolation rate of gain.

Table 77. Evaluation of Criterion 5 – estimated rate of RFS gain per month by receiving oral AZA instead of placebo in the ITT population restricted to EU-only patients, before and after the trial cutoff

Models	Pre-extrapolation	Extrapolated tail
KM	[redacted]	-

Joint models						
Exponential		[redacted]			[redacted]	
Weibull						
Log-Logistic						
Log-Normal						
Generalized Gamma						
Gompertz						
Individual models						
Exponential		[redacted]			[redacted]	
Weibull						
Log-Logistic						
Log-Normal						
Generalized Gamma						
Gompertz						

Notes: The rate of survival gain in the pre-extrapolation period is defined as the difference in relapse-free survival between oral AZA and placebo at [redacted] months divided by the number of months in the pre-extrapolation period (ie [redacted] months). The rate of survival gain in the post-extrapolation period is defined as the marginal relative difference in the extrapolated period (post cut-off) divided by the number of months post-cut-off. Negative values represent the rate of survival loss for oral AZA (ie, gain for placebo), which in the case of most fitted models indicate a crossing of curves.
Abbreviations: AZA, azacitidine; ITT, intention-to-treat, KM, Kaplan–Meier; RFS, relapse free survival.

Appendix B.15

Table 78. Number of patients, mean value and standard error of the EQ-5D data collected on day 1 of each cycle.

Visit	Statistics	CC-486 [redacted]	Placebo [redacted]	Overall [redacted]
C1D1	N	[redacted]	[redacted]	[redacted]
	Mean			
	SE			
C2D1	N			
	Mean			
	SE			
C3D1	N			
	Mean			
	SE			
C4D1	N			
	Mean			
	SE			
C5D1	N			
	Mean			
	SE			
C6D1	N			
	Mean			
	SE			
C7D1	N			
	Mean			
	SE			
C8D1	N			
	Mean			
	SE			
C9D1	N			
	Mean			
	SE			
C10D1	N			
	Mean			
	SE			
C11D1	N			
	Mean			
	SE			
C12D1	N			
	Mean			
	SE			
C13D1	N			
	Mean			
	SE			
C14D1	N			
	Mean			
	SE			
C15D1	N			
	Mean			
	SE			
C16D1	N			

	Mean	[redacted]			[redacted]			
	SE							
C17D1	N							
	Mean							
	SE							
C18D1	N							
	Mean							
	SE							
C19D1	N							
	Mean							
	SE							
C20D1	N							
	Mean							
	SE							
C21D1	N							
	Mean							
	SE							
C22D1	N							
	Mean							
	SE							
C23D1	N							
	Mean							
	SE							
C24D1	N							
	Mean							
	SE							
C25D1	N							
	Mean							
	SE							
C26D1	N							
	Mean							
	SE							
C27D1	N							
	Mean							
	SE							
C28D1	N							
	Mean							
	SE							
C29D1	N							
	Mean							
	SE							
C30D1	N							
	Mean							
	SE							
C31D1	N							
	Mean							
	SE							
C32D1	N							
	Mean							
	SE							
C33D1	N							
	Mean							
	SE							
C34D1	N							
	Mean							

	SE	[redacted]			[redacted]			
C35D1	N							
	Mean							
	SE							
C36D1	N							
	Mean							
	SE							
C37D1	N							
	Mean							
	SE							
C38D1	N							
	Mean							
	SE							
C39D1	N							
	Mean							
	SE							
C40D1	N							
	Mean							
	SE							
C41D1	N							
	Mean							
	SE							
C42D1	N							
	Mean							
	SE							
C43D1	N							
	Mean							
	SE							
C44D1	N							
	Mean							
	SE							
C45D1	N							
	Mean							
	SE							
C46D1	N							
	Mean							
	SE							
C47D1	N							
	Mean							
	SE							
C48D1	N							
	Mean							
	SE							
C49D1	N							
	Mean							
	SE							
C50D1	N							
	Mean							
	SE							
C51D1	N							
	Mean							
	SE							
C52D1	N							
	Mean							
	SE							

C53D1	N	[redacted]			[redacted]			
	Mean							
	SE							
C54D1	N							
	Mean							
	SE							
C55D1	N							
	Mean							
	SE							
C56D1	N							
	Mean							
	SE							
C57D1	N							
	Mean							
	SE							
C58D1	N							
	Mean							
	SE							
C59D1	N							
	Mean							
	SE							
C60D1	N							
	Mean							
	SE							
C61D1	N							
	Mean							
	SE							
C62D1	N							
	Mean							
	SE							
C63D1	N							
	Mean							
	SE							
C64D1	N							
	Mean							
	SE							
C65D1	N							
	Mean							
	SE							
C66D1	N							
	Mean							
	SE							
C67D1	N							
	Mean							
	SE							
C68D1	N							
	Mean							
	SE							
C69D1	N							
	Mean							
	SE							
C70D1	N							
	Mean							
	SE							
C71D1	N							

	Mean	[redacted]	[redacted]	[redacted]
	SE			
C72D1	N			
	Mean			
	SE			
C73D1	N			
	Mean			
	SE			
C74D1	N			
	Mean			
	SE			
C75D1	N			
	Mean			
	SE			
C76D1	N			
	Mean			
	SE			
C77D1	N			
	Mean			
	SE			
C78D1	N			
	Mean			
	SE			
C79D1	N			
	Mean			
	SE			
C80D1	N			
	Mean			
	SE			
EOT	N			
	Mean			
	SE			

Table 79. Number of patients, mean value and standard error of the EQ-5D data collected on day 1 of each cycle captured for both relapse and relapse free health states.

Visit	Statistics	CC-486 [redacted]		Placebo [redacted]	
		Relapse	Relapse Free	Relapse	Relapse Free
C1D1	N	[redacted]	[redacted]	[redacted]	[redacted]
	Mean	[redacted]	[redacted]	[redacted]	[redacted]
	SE	[redacted]	[redacted]	[redacted]	[redacted]

C2D1	N	[redacted]	[redacted]	[redacted]	[redacted]
	Mean	[redacted]	[redacted]	[redacted]	[redacted]
	SE	[redacted]	[redacted]	[redacted]	[redacted]
C3D1	N	[redacted]	[redacted]	[redacted]	[redacted]
	Mean	[redacted]	[redacted]	[redacted]	[redacted]
	SE	[redacted]	[redacted]	[redacted]	[redacted]
C4D1	N	[redacted]	[redacted]	[redacted]	[redacted]
	Mean	[redacted]	[redacted]	[redacted]	[redacted]
	SE	[redacted]	[redacted]	[redacted]	[redacted]
C5D1	N	[redacted]	[redacted]	[redacted]	[redacted]
	Mean	[redacted]	[redacted]	[redacted]	[redacted]
	SE	[redacted]	[redacted]	[redacted]	[redacted]
C6D1	N	[redacted]	[redacted]	[redacted]	[redacted]
	Mean	[redacted]	[redacted]	[redacted]	[redacted]
	SE	[redacted]	[redacted]	[redacted]	[redacted]
C7D1	N	[redacted]	[redacted]	[redacted]	[redacted]
	Mean	[redacted]	[redacted]	[redacted]	[redacted]
	SE	[redacted]	[redacted]	[redacted]	[redacted]
C8D1	N	[redacted]	[redacted]	[redacted]	[redacted]
	Mean	[redacted]	[redacted]	[redacted]	[redacted]
	SE	[redacted]	[redacted]	[redacted]	[redacted]
C9D1	N	[redacted]	[redacted]	[redacted]	[redacted]
	Mean	[redacted]	[redacted]	[redacted]	[redacted]
	SE	[redacted]	[redacted]	[redacted]	[redacted]
C10D1	N	[redacted]	[redacted]	[redacted]	[redacted]
	Mean	[redacted]	[redacted]	[redacted]	[redacted]
	SE	[redacted]	[redacted]	[redacted]	[redacted]
C11D1	N	[redacted]	[redacted]	[redacted]	[redacted]
	Mean	[redacted]	[redacted]	[redacted]	[redacted]
	SE	[redacted]	[redacted]	[redacted]	[redacted]
C12D1	N	[redacted]	[redacted]	[redacted]	[redacted]
	Mean	[redacted]	[redacted]	[redacted]	[redacted]
	SE	[redacted]	[redacted]	[redacted]	[redacted]
C13D1	N	[redacted]	[redacted]	[redacted]	[redacted]
	Mean	[redacted]	[redacted]	[redacted]	[redacted]
	SE	[redacted]	[redacted]	[redacted]	[redacted]
C14D1	N	[redacted]	[redacted]	[redacted]	[redacted]
	Mean	[redacted]	[redacted]	[redacted]	[redacted]
	SE	[redacted]	[redacted]	[redacted]	[redacted]
C15D1	N	[redacted]	[redacted]	[redacted]	[redacted]
	Mean	[redacted]	[redacted]	[redacted]	[redacted]
	SE	[redacted]	[redacted]	[redacted]	[redacted]
C16D1	N	[redacted]	[redacted]	[redacted]	[redacted]
	Mean	[redacted]	[redacted]	[redacted]	[redacted]

	SE	[redacted]	[redacted]	[redacted]	[redacted]
C17D1	N	[redacted]	[redacted]	[redacted]	[redacted]
	Mean	[redacted]	[redacted]	[redacted]	[redacted]
	SE	[redacted]	[redacted]	[redacted]	[redacted]
C18D1	N	[redacted]	[redacted]	[redacted]	[redacted]
	Mean	[redacted]	[redacted]	[redacted]	[redacted]
	SE	[redacted]	[redacted]	[redacted]	[redacted]
C19D1	N	[redacted]	[redacted]	[redacted]	[redacted]
	Mean	[redacted]	[redacted]	[redacted]	[redacted]
	SE	[redacted]	[redacted]	[redacted]	[redacted]
C20D1	N	[redacted]	[redacted]	[redacted]	[redacted]
	Mean	[redacted]	[redacted]	[redacted]	[redacted]
	SE	[redacted]	[redacted]	[redacted]	[redacted]
C21D1	N	[redacted]	[redacted]	[redacted]	[redacted]
	Mean	[redacted]	[redacted]	[redacted]	[redacted]
	SE	[redacted]	[redacted]	[redacted]	[redacted]
C22D1	N	[redacted]	[redacted]	[redacted]	[redacted]
	Mean	[redacted]	[redacted]	[redacted]	[redacted]
	SE	[redacted]	[redacted]	[redacted]	[redacted]
C23D1	N	[redacted]	[redacted]	[redacted]	[redacted]
	Mean	[redacted]	[redacted]	[redacted]	[redacted]
	SE	[redacted]	[redacted]	[redacted]	[redacted]
C24D1	N	[redacted]	[redacted]	[redacted]	[redacted]
	Mean	[redacted]	[redacted]	[redacted]	[redacted]
	SE	[redacted]	[redacted]	[redacted]	[redacted]
C25D1	N	[redacted]	[redacted]	[redacted]	[redacted]
	Mean	[redacted]	[redacted]	[redacted]	[redacted]
	SE	[redacted]	[redacted]	[redacted]	[redacted]
C26D1	N	[redacted]	[redacted]	[redacted]	[redacted]
	Mean	[redacted]	[redacted]	[redacted]	[redacted]
	SE	[redacted]	[redacted]	[redacted]	[redacted]
C27D1	N	[redacted]	[redacted]	[redacted]	[redacted]
	Mean	[redacted]	[redacted]	[redacted]	[redacted]
	SE	[redacted]	[redacted]	[redacted]	[redacted]
C28D1	N	[redacted]	[redacted]	[redacted]	[redacted]
	Mean	[redacted]	[redacted]	[redacted]	[redacted]
	SE	[redacted]	[redacted]	[redacted]	[redacted]
C29D1	N	[redacted]	[redacted]	[redacted]	[redacted]
	Mean	[redacted]	[redacted]	[redacted]	[redacted]
	SE	[redacted]	[redacted]	[redacted]	[redacted]
C30D1	N	[redacted]	[redacted]	[redacted]	[redacted]
	Mean	[redacted]	[redacted]	[redacted]	[redacted]
	SE	[redacted]	[redacted]	[redacted]	[redacted]
C31D1	N	[redacted]	[redacted]	[redacted]	[redacted]

	Mean	[redacted]	[redacted]	[redacted]	[redacted]
	SE	[redacted]	[redacted]	[redacted]	[redacted]
C32D1	N	[redacted]	[redacted]	[redacted]	[redacted]
	Mean	[redacted]	[redacted]	[redacted]	[redacted]
	SE	[redacted]	[redacted]	[redacted]	[redacted]
C33D1	N	[redacted]	[redacted]	[redacted]	[redacted]
	Mean	[redacted]	[redacted]	[redacted]	[redacted]
	SE	[redacted]	[redacted]	[redacted]	[redacted]
C34D1	N	[redacted]	[redacted]	[redacted]	[redacted]
	Mean	[redacted]	[redacted]	[redacted]	[redacted]
	SE	[redacted]	[redacted]	[redacted]	[redacted]
C35D1	N	[redacted]	[redacted]	[redacted]	[redacted]
	Mean	[redacted]	[redacted]	[redacted]	[redacted]
	SE	[redacted]	[redacted]	[redacted]	[redacted]
C36D1	N	[redacted]	[redacted]	[redacted]	[redacted]
	Mean	[redacted]	[redacted]	[redacted]	[redacted]
	SE	[redacted]	[redacted]	[redacted]	[redacted]
C37D1	N	[redacted]	[redacted]	[redacted]	[redacted]
	Mean	[redacted]	[redacted]	[redacted]	[redacted]
	SE	[redacted]	[redacted]	[redacted]	[redacted]
C38D1	N	[redacted]	[redacted]	[redacted]	[redacted]
	Mean	[redacted]	[redacted]	[redacted]	[redacted]
	SE	[redacted]	[redacted]	[redacted]	[redacted]
C39D1	N	[redacted]	[redacted]	[redacted]	[redacted]
	Mean	[redacted]	[redacted]	[redacted]	[redacted]
	SE	[redacted]	[redacted]	[redacted]	[redacted]
C40D1	N	[redacted]	[redacted]	[redacted]	[redacted]
	Mean	[redacted]	[redacted]	[redacted]	[redacted]
	SE	[redacted]	[redacted]	[redacted]	[redacted]
C41D1	N	[redacted]	[redacted]	[redacted]	[redacted]
	Mean	[redacted]	[redacted]	[redacted]	[redacted]
	SE	[redacted]	[redacted]	[redacted]	[redacted]
C42D1	N	[redacted]	[redacted]	[redacted]	[redacted]
	Mean	[redacted]	[redacted]	[redacted]	[redacted]
	SE	[redacted]	[redacted]	[redacted]	[redacted]
C43D1	N	[redacted]	[redacted]	[redacted]	[redacted]
	Mean	[redacted]	[redacted]	[redacted]	[redacted]
	SE	[redacted]	[redacted]	[redacted]	[redacted]
C44D1	N	[redacted]	[redacted]	[redacted]	[redacted]
	Mean	[redacted]	[redacted]	[redacted]	[redacted]
	SE	[redacted]	[redacted]	[redacted]	[redacted]
C45D1	N	[redacted]	[redacted]	[redacted]	[redacted]
	Mean	[redacted]	[redacted]	[redacted]	[redacted]
	SE	[redacted]	[redacted]	[redacted]	[redacted]

C46D1	N	[redacted]	[redacted]	[redacted]	[redacted]
	Mean	[redacted]	[redacted]	[redacted]	[redacted]
	SE	[redacted]	[redacted]	[redacted]	[redacted]
C47D1	N	[redacted]	[redacted]	[redacted]	[redacted]
	Mean	[redacted]	[redacted]	[redacted]	[redacted]
	SE	[redacted]	[redacted]	[redacted]	[redacted]
C48D1	N	[redacted]	[redacted]	[redacted]	[redacted]
	Mean	[redacted]	[redacted]	[redacted]	[redacted]
	SE	[redacted]	[redacted]	[redacted]	[redacted]
C49D1	N	[redacted]	[redacted]	[redacted]	[redacted]
	Mean	[redacted]	[redacted]	[redacted]	[redacted]
	SE	[redacted]	[redacted]	[redacted]	[redacted]
C50D1	N	[redacted]	[redacted]	[redacted]	[redacted]
	Mean	[redacted]	[redacted]	[redacted]	[redacted]
	SE	[redacted]	[redacted]	[redacted]	[redacted]
C51D1	N	[redacted]	[redacted]	[redacted]	[redacted]
	Mean	[redacted]	[redacted]	[redacted]	[redacted]
	SE	[redacted]	[redacted]	[redacted]	[redacted]
C52D1	N	[redacted]	[redacted]	[redacted]	[redacted]
	Mean	[redacted]	[redacted]	[redacted]	[redacted]
	SE	[redacted]	[redacted]	[redacted]	[redacted]
C53D1	N	[redacted]	[redacted]	[redacted]	[redacted]
	Mean	[redacted]	[redacted]	[redacted]	[redacted]
	SE	[redacted]	[redacted]	[redacted]	[redacted]
C54D1	N	[redacted]	[redacted]	[redacted]	[redacted]
	Mean	[redacted]	[redacted]	[redacted]	[redacted]
	SE	[redacted]	[redacted]	[redacted]	[redacted]
C55D1	N	[redacted]	[redacted]	[redacted]	[redacted]
	Mean	[redacted]	[redacted]	[redacted]	[redacted]
	SE	[redacted]	[redacted]	[redacted]	[redacted]
C56D1	N	[redacted]	[redacted]	[redacted]	[redacted]
	Mean	[redacted]	[redacted]	[redacted]	[redacted]
	SE	[redacted]	[redacted]	[redacted]	[redacted]
C57D1	N	[redacted]	[redacted]	[redacted]	[redacted]
	Mean	[redacted]	[redacted]	[redacted]	[redacted]
	SE	[redacted]	[redacted]	[redacted]	[redacted]
C58D1	N	[redacted]	[redacted]	[redacted]	[redacted]
	Mean	[redacted]	[redacted]	[redacted]	[redacted]
	SE	[redacted]	[redacted]	[redacted]	[redacted]
C59D1	N	[redacted]	[redacted]	[redacted]	[redacted]
	Mean	[redacted]	[redacted]	[redacted]	[redacted]
	SE	[redacted]	[redacted]	[redacted]	[redacted]
C60D1	N	[redacted]	[redacted]	[redacted]	[redacted]
	Mean	[redacted]	[redacted]	[redacted]	[redacted]

	SE	[redacted]	[redacted]	[redacted]	[redacted]
C61D1	N	[redacted]	[redacted]	[redacted]	[redacted]
	Mean	[redacted]	[redacted]	[redacted]	[redacted]
	SE	[redacted]	[redacted]	[redacted]	[redacted]
C62D1	N	[redacted]	[redacted]	[redacted]	[redacted]
	Mean	[redacted]	[redacted]	[redacted]	[redacted]
	SE	[redacted]	[redacted]	[redacted]	[redacted]
C63D1	N	[redacted]	[redacted]	[redacted]	[redacted]
	Mean	[redacted]	[redacted]	[redacted]	[redacted]
	SE	[redacted]	[redacted]	[redacted]	[redacted]
C64D1	N	[redacted]	[redacted]	[redacted]	[redacted]
	Mean	[redacted]	[redacted]	[redacted]	[redacted]
	SE	[redacted]	[redacted]	[redacted]	[redacted]
C65D1	N	[redacted]	[redacted]	[redacted]	[redacted]
	Mean	[redacted]	[redacted]	[redacted]	[redacted]
	SE	[redacted]	[redacted]	[redacted]	[redacted]
C66D1	N	[redacted]	[redacted]	[redacted]	[redacted]
	Mean	[redacted]	[redacted]	[redacted]	[redacted]
	SE	[redacted]	[redacted]	[redacted]	[redacted]
C67D1	N	[redacted]	[redacted]	[redacted]	[redacted]
	Mean	[redacted]	[redacted]	[redacted]	[redacted]
	SE	[redacted]	[redacted]	[redacted]	[redacted]
C68D1	N	[redacted]	[redacted]	[redacted]	[redacted]
	Mean	[redacted]	[redacted]	[redacted]	[redacted]
	SE	[redacted]	[redacted]	[redacted]	[redacted]
C69D1	N	[redacted]	[redacted]	[redacted]	[redacted]
	Mean	[redacted]	[redacted]	[redacted]	[redacted]
	SE	[redacted]	[redacted]	[redacted]	[redacted]
C70D1	N	[redacted]	[redacted]	[redacted]	[redacted]
	Mean	[redacted]	[redacted]	[redacted]	[redacted]
	SE	[redacted]	[redacted]	[redacted]	[redacted]
C71D1	N	[redacted]	[redacted]	[redacted]	[redacted]
	Mean	[redacted]	[redacted]	[redacted]	[redacted]
	SE	[redacted]	[redacted]	[redacted]	[redacted]
C72D1	N	[redacted]	[redacted]	[redacted]	[redacted]
	Mean	[redacted]	[redacted]	[redacted]	[redacted]
	SE	[redacted]	[redacted]	[redacted]	[redacted]
C73D1	N	[redacted]	[redacted]	[redacted]	[redacted]
	Mean	[redacted]	[redacted]	[redacted]	[redacted]
	SE	[redacted]	[redacted]	[redacted]	[redacted]
C74D1	N	[redacted]	[redacted]	[redacted]	[redacted]
	Mean	[redacted]	[redacted]	[redacted]	[redacted]
	SE	[redacted]	[redacted]	[redacted]	[redacted]
C75D1	N	[redacted]	[redacted]	[redacted]	[redacted]

	Mean	[redacted]	[redacted]	[redacted]	[redacted]
	SE	[redacted]	[redacted]	[redacted]	[redacted]
C76D1	N	[redacted]	[redacted]	[redacted]	[redacted]
	Mean	[redacted]	[redacted]	[redacted]	[redacted]
	SE	[redacted]	[redacted]	[redacted]	[redacted]
C77D1	N	[redacted]	[redacted]	[redacted]	[redacted]
	Mean	[redacted]	[redacted]	[redacted]	[redacted]
	SE	[redacted]	[redacted]	[redacted]	[redacted]
C78D1	N	[redacted]	[redacted]	[redacted]	[redacted]
	Mean	[redacted]	[redacted]	[redacted]	[redacted]
	SE	[redacted]	[redacted]	[redacted]	[redacted]
C79D1	N	[redacted]	[redacted]	[redacted]	[redacted]
	Mean	[redacted]	[redacted]	[redacted]	[redacted]
	SE	[redacted]	[redacted]	[redacted]	[redacted]
C80D1	N	[redacted]	[redacted]	[redacted]	[redacted]
	Mean	[redacted]	[redacted]	[redacted]	[redacted]
	SE	[redacted]	[redacted]	[redacted]	[redacted]
EOT	N	[redacted]	[redacted]	[redacted]	[redacted]
	Mean	[redacted]	[redacted]	[redacted]	[redacted]
	SE	[redacted]	[redacted]	[redacted]	[redacted]

Table 80. presenting the completion rate of the EQ-5D questionnaire for both trial arms

Visit	Completion rate CC-486 [redacted]	Completion rate Placebo [redacted]	Completion rate Overall [redacted]
C1D1	[redacted]	[redacted]	[redacted]
C2D1	[redacted]	[redacted]	[redacted]
C3D1	[redacted]	[redacted]	[redacted]
C4D1	[redacted]	[redacted]	[redacted]
C5D1	[redacted]	[redacted]	[redacted]
C6D1	[redacted]	[redacted]	[redacted]
C7D1	[redacted]	[redacted]	[redacted]
C8D1	[redacted]	[redacted]	[redacted]
C9D1	[redacted]	[redacted]	[redacted]
C10D1	[redacted]	[redacted]	[redacted]
C11D1	[redacted]	[redacted]	[redacted]
C12D1	[redacted]	[redacted]	[redacted]
C13D1	[redacted]	[redacted]	[redacted]
C14D1	[redacted]	[redacted]	[redacted]
C15D1	[redacted]	[redacted]	[redacted]
C16D1	[redacted]	[redacted]	[redacted]
C17D1	[redacted]	[redacted]	[redacted]

C18D1	[redacted]	[redacted]	[redacted]
C19D1	[redacted]	[redacted]	[redacted]
C20D1	[redacted]	[redacted]	[redacted]
C21D1	[redacted]	[redacted]	[redacted]
C22D1	[redacted]	[redacted]	[redacted]
C23D1	[redacted]	[redacted]	[redacted]
C24D1	[redacted]	[redacted]	[redacted]
C25D1	[redacted]	[redacted]	[redacted]
C26D1	[redacted]	[redacted]	[redacted]
C27D1	[redacted]	[redacted]	[redacted]
C28D1	[redacted]	[redacted]	[redacted]
C29D1	[redacted]	[redacted]	[redacted]
C30D1	[redacted]	[redacted]	[redacted]
C31D1	[redacted]	[redacted]	[redacted]
C32D1	[redacted]	[redacted]	[redacted]
C33D1	[redacted]	[redacted]	[redacted]
C34D1	[redacted]	[redacted]	[redacted]
C35D1	[redacted]	[redacted]	[redacted]
C36D1	[redacted]	[redacted]	[redacted]
C37D1	[redacted]	[redacted]	[redacted]
C38D1	[redacted]	[redacted]	[redacted]
C39D1	[redacted]	[redacted]	[redacted]
C40D1	[redacted]	[redacted]	[redacted]
C41D1	[redacted]	[redacted]	[redacted]
C42D1	[redacted]	[redacted]	[redacted]
C43D1	[redacted]	[redacted]	[redacted]
C44D1	[redacted]	[redacted]	[redacted]
C45D1	[redacted]	[redacted]	[redacted]
C46D1	[redacted]	[redacted]	[redacted]
C47D1	[redacted]	[redacted]	[redacted]
C48D1	[redacted]	[redacted]	[redacted]
C49D1	[redacted]	[redacted]	[redacted]
C50D1	[redacted]	[redacted]	[redacted]
C51D1	[redacted]	[redacted]	[redacted]
C52D1	[redacted]	[redacted]	[redacted]
C53D1	[redacted]	[redacted]	[redacted]
C54D1	[redacted]	[redacted]	[redacted]
C55D1	[redacted]	[redacted]	[redacted]
C56D1	[redacted]	[redacted]	[redacted]
C57D1	[redacted]	[redacted]	[redacted]
C58D1	[redacted]	[redacted]	[redacted]
C59D1	[redacted]	[redacted]	[redacted]
C60D1	[redacted]	[redacted]	[redacted]
C61D1	[redacted]	[redacted]	[redacted]

C62D1	[redacted]	[redacted]	[redacted]
C63D1	[redacted]	[redacted]	[redacted]
C64D1	[redacted]	[redacted]	[redacted]
C65D1	[redacted]	[redacted]	[redacted]
C66D1	[redacted]	[redacted]	[redacted]
C67D1	[redacted]	[redacted]	[redacted]
C68D1	[redacted]	[redacted]	[redacted]
C69D1	[redacted]	[redacted]	[redacted]
C70D1	[redacted]	[redacted]	[redacted]
C71D1	[redacted]	[redacted]	[redacted]
C72D1	[redacted]	[redacted]	[redacted]
C73D1	[redacted]	[redacted]	[redacted]
C74D1	[redacted]	[redacted]	[redacted]
C75D1	[redacted]	[redacted]	[redacted]
C76D1	[redacted]	[redacted]	[redacted]
C77D1	[redacted]	[redacted]	[redacted]
C78D1	[redacted]	[redacted]	[redacted]
C79D1	[redacted]	[redacted]	[redacted]
C80D1	[redacted]	[redacted]	[redacted]
EOT	[redacted]	[redacted]	[redacted]
Overall	[redacted]	[redacted]	[redacted]

Table 81. Demographic and Disease Characteristics and HRQoL Scores at Baseline (HRQoL Evaluable Population)

Characteristic	Level	CC-486	Placebo	Overall
		[redacted]	[redacted]	[redacted]
Age (years)		[redacted]	[redacted]	[redacted]
	N	[redacted]	[redacted]	[redacted]
	Mean (SD)	[redacted]	[redacted]	[redacted]
	Median	[redacted]	[redacted]	[redacted]
	Q1, Q3	[redacted]	[redacted]	[redacted]
	Min, Max	[redacted]	[redacted]	[redacted]
Age, n(%)		[redacted]	[redacted]	[redacted]
	< 65	[redacted]	[redacted]	[redacted]
	65 - 74	[redacted]	[redacted]	[redacted]
	≥ 75	[redacted]	[redacted]	[redacted]
	Missing	[redacted]	[redacted]	[redacted]

Gender, n(%)		[redacted]	[redacted]	[redacted]
	Male	[redacted]	[redacted]	[redacted]
	Female	[redacted]	[redacted]	[redacted]
	Missing	[redacted]	[redacted]	[redacted]
Race, n(%)		[redacted]	[redacted]	[redacted]
	White	[redacted]	[redacted]	[redacted]
	Black	[redacted]	[redacted]	[redacted]
	Asian	[redacted]	[redacted]	[redacted]
	Other	[redacted]	[redacted]	[redacted]
	Missing	[redacted]	[redacted]	[redacted]
Geographic region, n(%)		[redacted]	[redacted]	[redacted]
	North America	[redacted]	[redacted]	[redacted]
	Europe	[redacted]	[redacted]	[redacted]
	Asia	[redacted]	[redacted]	[redacted]
	Australia	[redacted]	[redacted]	[redacted]
	South America	[redacted]	[redacted]	[redacted]
	Missing	[redacted]	[redacted]	[redacted]
WHO AML classification, n(%)		[redacted]	[redacted]	[redacted]
	AML with recurrent genetic abnormalities	[redacted]	[redacted]	[redacted]
	AML with myelodysplasia-related changes	[redacted]	[redacted]	[redacted]
	Therapy-related myeloid neoplasma	[redacted]	[redacted]	[redacted]
	AML not otherwise specified	[redacted]	[redacted]	[redacted]
	Missing	[redacted]	[redacted]	[redacted]

Type of AML, n(%)		[redacted]	[redacted]	[redacted]
	Primary	[redacted]	[redacted]	[redacted]
	Secondary	[redacted]	[redacted]	[redacted]
	Missing	[redacted]	[redacted]	[redacted]
Response status (CR/CRi) after induction therapy (with or without consolidation therapy), n(%)		[redacted]	[redacted]	[redacted]
	CR	[redacted]	[redacted]	[redacted]
	CRi	[redacted]	[redacted]	[redacted]
	Missing	[redacted]	[redacted]	[redacted]
Prior history of MDS or CMML, n(%)		[redacted]	[redacted]	[redacted]
	Yes	[redacted]	[redacted]	[redacted]
	Primary	[redacted]	[redacted]	[redacted]
	Secondary	[redacted]	[redacted]	[redacted]
	No	[redacted]	[redacted]	[redacted]
	Missing	[redacted]	[redacted]	[redacted]
Cytogenetic risk category at time of induction therapy, n(%)		[redacted]	[redacted]	[redacted]
	Intermediate	[redacted]	[redacted]	[redacted]
	Poor	[redacted]	[redacted]	[redacted]
	Missing	[redacted]	[redacted]	[redacted]

Consolidation therapy following induction, n(%)		[redacted]	[redacted]	[redacted]
	Yes	[redacted]	[redacted]	[redacted]
	1 Cycle	[redacted]	[redacted]	[redacted]
	2 Cycles	[redacted]	[redacted]	[redacted]
	3 Cycles	[redacted]	[redacted]	[redacted]
	4 Cycles	[redacted]	[redacted]	[redacted]
	Missing	[redacted]	[redacted]	[redacted]
	No	[redacted]	[redacted]	[redacted]
	Missing	[redacted]	[redacted]	[redacted]
ECOG performance status, n(%)		[redacted]	[redacted]	[redacted]
	0	[redacted]	[redacted]	[redacted]
	1	[redacted]	[redacted]	[redacted]
	2	[redacted]	[redacted]	[redacted]
	3	[redacted]	[redacted]	[redacted]
	Missing	[redacted]	[redacted]	[redacted]
Minimal residual disease status from central pathology report, n(%)		[redacted]	[redacted]	[redacted]
	Positive	[redacted]	[redacted]	[redacted]
	Negative	[redacted]	[redacted]	[redacted]
	Missing	[redacted]	[redacted]	[redacted]
Bone marrow blast (%)				
	N	[redacted]	[redacted]	[redacted]
	Mean (SD)	[redacted]	[redacted]	[redacted]
	Median	[redacted]	[redacted]	[redacted]

	Q1, Q3	[redacted]	[redacted]	[redacted]
	Min, Max	[redacted]	[redacted]	[redacted]
Time from initial AML diagnosis to randomization (months)		[redacted]	[redacted]	[redacted]
	N	[redacted]	[redacted]	[redacted]
	Mean (SD)	[redacted]	[redacted]	[redacted]
	Median	[redacted]	[redacted]	[redacted]
	Q1, Q3	[redacted]	[redacted]	[redacted]
	Min, Max	[redacted]	[redacted]	[redacted]
Time from start of induction therapy to randomization (months)		[redacted]	[redacted]	[redacted]
	N	[redacted]	[redacted]	[redacted]
	Mean (SD)	[redacted]	[redacted]	[redacted]
	Median	[redacted]	[redacted]	[redacted]
	Q1, Q3	[redacted]	[redacted]	[redacted]
	Min, Max	[redacted]	[redacted]	[redacted]

Characteristic	Level	CC-486	Placebo	Overall
		[redacted]	[redacted]	[redacted]
Age (years)		[redacted]	[redacted]	[redacted]
	N	[redacted]	[redacted]	[redacted]
	Mean (SD)	[redacted]	[redacted]	[redacted]
	Median	[redacted]	[redacted]	[redacted]
	Q1, Q3	[redacted]	[redacted]	[redacted]
	Min, Max	[redacted]	[redacted]	[redacted]
Age, n(%)		[redacted]	[redacted]	[redacted]
	< 65	[redacted]	[redacted]	[redacted]
	65 - 74	[redacted]	[redacted]	[redacted]
	≥ 75	[redacted]	[redacted]	[redacted]

	Missing	[redacted]	[redacted]	[redacted]
Gender, n(%)				
	Male	[redacted]	[redacted]	[redacted]
	Female	[redacted]	[redacted]	[redacted]
	Missing	[redacted]	[redacted]	[redacted]
Race, n(%)				
	White	[redacted]	[redacted]	[redacted]
	Black	[redacted]	[redacted]	[redacted]
	Asian	[redacted]	[redacted]	[redacted]
	Other	[redacted]	[redacted]	[redacted]
	Missing	[redacted]	[redacted]	[redacted]
Geographic region, n(%)				
	North America	[redacted]	[redacted]	[redacted]
	Europe	[redacted]	[redacted]	[redacted]
	Asia	[redacted]	[redacted]	[redacted]
	Australia	[redacted]	[redacted]	[redacted]
	South America	[redacted]	[redacted]	[redacted]
	Missing	[redacted]	[redacted]	[redacted]
WHO AML classification, n(%)				
	AML with recurrent genetic abnormalities	[redacted]	[redacted]	[redacted]
	AML with myelodysplasia-related changes	[redacted]	[redacted]	[redacted]
	Therapy-related myeloid neoplasma	[redacted]	[redacted]	[redacted]
	AML not otherwise specified	[redacted]	[redacted]	[redacted]
	Missing	[redacted]	[redacted]	[redacted]

Type of AML, n(%)				
	Primary	[redacted]	[redacted]	[redacted]
	Secondary	[redacted]	[redacted]	[redacted]
	Missing	[redacted]	[redacted]	[redacted]
Response status (CR/CRi) after induction therapy (with or without consolidation therapy), n(%)				
	CR	[redacted]	[redacted]	[redacted]
	CRi	[redacted]	[redacted]	[redacted]
	Missing	[redacted]	[redacted]	[redacted]
Prior history of MDS or CMML, n(%)				
	Yes	[redacted]	[redacted]	[redacted]
	Primary	[redacted]	[redacted]	[redacted]
	No	[redacted]	[redacted]	[redacted]
	Missing	[redacted]	[redacted]	[redacted]
Cytogenetic risk category at time of induction therapy, n(%)				
	Intermediate	[redacted]	[redacted]	[redacted]
	Poor	[redacted]	[redacted]	[redacted]
	Missing	[redacted]	[redacted]	[redacted]

Consolidation therapy following induction, n(%)				
	Yes	[redacted]	[redacted]	[redacted]
	1 Cycle	[redacted]	[redacted]	[redacted]
	2 Cycles	[redacted]	[redacted]	[redacted]
	3 Cycles	[redacted]	[redacted]	[redacted]
	4 Cycles	[redacted]	[redacted]	[redacted]
	Missing	[redacted]	[redacted]	[redacted]
	No	[redacted]	[redacted]	[redacted]
	Missing	[redacted]	[redacted]	[redacted]
ECOG performance status, n(%)				
	0	[redacted]	[redacted]	[redacted]
	1	[redacted]	[redacted]	[redacted]
	2	[redacted]	[redacted]	[redacted]
	3	[redacted]	[redacted]	[redacted]
	Missing	[redacted]	[redacted]	[redacted]
Minimal residual disease status from central pathology report, n(%)				
	Positive	[redacted]	[redacted]	[redacted]
	Negative	[redacted]	[redacted]	[redacted]
	Missing	[redacted]	[redacted]	[redacted]
Bone marrow blast (%)				
	N	[redacted]	[redacted]	[redacted]
	Mean (SD)	[redacted]	[redacted]	[redacted]

	Median	[redacted]	[redacted]	[redacted]
	Q1, Q3	[redacted]	[redacted]	[redacted]
	Min, Max	[redacted]	[redacted]	[redacted]
Time from initial AML diagnosis to randomization (months)				
	N	[redacted]	[redacted]	[redacted]
	Mean (SD)	[redacted]	[redacted]	[redacted]
	Median	[redacted]	[redacted]	[redacted]
	Q1, Q3	[redacted]	[redacted]	[redacted]
	Min, Max	[redacted]	[redacted]	[redacted]
Time from start of induction therapy to randomization (months)		[redacted]	[redacted]	[redacted]
	N	[redacted]	[redacted]	[redacted]
	Mean (SD)	[redacted]	[redacted]	[redacted]
	Median	[redacted]	[redacted]	[redacted]
	Q1, Q3	[redacted]	[redacted]	[redacted]
	Min, Max	[redacted]	[redacted]	[redacted]

Table 82. Demographic and Disease Characteristics and HRQoL Scores at Baseline (Intent-to-Treat Population with Any Missing Utility Measurement)

Characteristic	Level	CC-486	Placebo	Overall
		[redacted]	[redacted]	[redacted]
Age (years)		[redacted]	[redacted]	[redacted]
	N	[redacted]	[redacted]	[redacted]
	Mean (SD)	[redacted]	[redacted]	[redacted]
	Median	[redacted]	[redacted]	[redacted]
	Q1, Q3	[redacted]	[redacted]	[redacted]
	Min, Max	[redacted]	[redacted]	[redacted]

Age, n(%)				
	< 65	[redacted]	[redacted]	[redacted]
	65 - 74	[redacted]	[redacted]	[redacted]
	≥ 75	[redacted]	[redacted]	[redacted]
	Missing	[redacted]	[redacted]	[redacted]
Gender, n(%)		[redacted]	[redacted]	[redacted]
	Male	[redacted]	[redacted]	[redacted]
	Female	[redacted]	[redacted]	[redacted]
	Missing	[redacted]	[redacted]	[redacted]
Race, n(%)				
	White	[redacted]	[redacted]	[redacted]
	Black	[redacted]	[redacted]	[redacted]
	Asian	[redacted]	[redacted]	[redacted]
	Other	[redacted]	[redacted]	[redacted]
	Missing	[redacted]	[redacted]	[redacted]
Geographic region, n(%)		[redacted]	[redacted]	[redacted]
	North America	[redacted]	[redacted]	[redacted]
	Europe	[redacted]	[redacted]	[redacted]
	Asia	[redacted]	[redacted]	[redacted]
	Australia	[redacted]	[redacted]	[redacted]
	South America	[redacted]	[redacted]	[redacted]
	Missing	[redacted]	[redacted]	[redacted]
WHO AML classification, n(%)		[redacted]	[redacted]	[redacted]
	AML with recurrent genetic abnormalities	[redacted]	[redacted]	[redacted]
	AML with myelodysplasia-related changes	[redacted]	[redacted]	[redacted]

	Therapy-related myeloid neoplasma	[redacted]	[redacted]	[redacted]
	AML not otherwise specified	[redacted]	[redacted]	[redacted]
	Missing	[redacted]	[redacted]	[redacted]
Type of AML, n(%)		[redacted]	[redacted]	[redacted]
	Primary	[redacted]	[redacted]	[redacted]
	Secondary	[redacted]	[redacted]	[redacted]
	Missing	[redacted]	[redacted]	[redacted]
Response status (CR/CRi) after induction therapy (with or without consolidation therapy), n(%)				
	CR	[redacted]	[redacted]	[redacted]
	CRi	[redacted]	[redacted]	[redacted]
	Missing	[redacted]	[redacted]	[redacted]
Prior history of MDS or CMML, n(%)		[redacted]	[redacted]	[redacted]
	Yes	[redacted]	[redacted]	[redacted]
	Primary	[redacted]	[redacted]	[redacted]
	No	[redacted]	[redacted]	[redacted]
	Missing	[redacted]	[redacted]	[redacted]
Cytogenetic risk category at time of induction therapy, n(%)				
	Intermediate	[redacted]	[redacted]	[redacted]

	Poor	[redacted]	[redacted]	[redacted]
	Missing	[redacted]	[redacted]	[redacted]
Consolidation therapy following induction, n(%)				
	Yes	[redacted]	[redacted]	[redacted]
	1 Cycle	[redacted]	[redacted]	[redacted]
	2 Cycles	[redacted]	[redacted]	[redacted]
	3 Cycles	[redacted]	[redacted]	[redacted]
	4 Cycles	[redacted]	[redacted]	[redacted]
	Missing	[redacted]	[redacted]	[redacted]
	No	[redacted]	[redacted]	[redacted]
	Missing	[redacted]	[redacted]	[redacted]
ECOG performance status, n(%)		[redacted]	[redacted]	[redacted]
	0	[redacted]	[redacted]	[redacted]
	1	[redacted]	[redacted]	[redacted]
	2	[redacted]	[redacted]	[redacted]
	3	[redacted]	[redacted]	[redacted]
	Missing	[redacted]	[redacted]	[redacted]
Minimal residual disease status from central pathology report, n(%)				
	Positive	[redacted]	[redacted]	[redacted]
	Negative	[redacted]	[redacted]	[redacted]
	Missing	[redacted]	[redacted]	[redacted]

Bone marrow blast (%)				
	N	[redacted]	[redacted]	[redacted]
	Mean (SD)	[redacted]	[redacted]	[redacted]
	Median	[redacted]	[redacted]	[redacted]
	Q1, Q3	[redacted]	[redacted]	[redacted]
	Min, Max	[redacted]	[redacted]	[redacted]
Time from initial AML diagnosis to randomization (months)				
	N	[redacted]	[redacted]	[redacted]
	Mean (SD)	[redacted]	[redacted]	[redacted]
	Median	[redacted]	[redacted]	[redacted]
	Q1, Q3	[redacted]	[redacted]	[redacted]
	Min, Max	[redacted]	[redacted]	[redacted]
Time from start of induction therapy to randomization (months)				
	N	[redacted]	[redacted]	[redacted]
	Mean (SD)	[redacted]	[redacted]	[redacted]
	Median	[redacted]	[redacted]	[redacted]
	Q1, Q3	[redacted]	[redacted]	[redacted]
	Min, Max	[redacted]	[redacted]	[redacted]

Table 83. Demographic and Disease Characteristics and HRQoL Scores at Baseline (Intent-to-Treat Population with Complete Utility Measurements)

Characteristic	Level	CC-486	Placebo	Overall

		[redacted]	[redacted]	[redacted]
Age (years)				
	N	[redacted]	[redacted]	[redacted]
	Mean (SD)	[redacted]	[redacted]	[redacted]
	Median	[redacted]	[redacted]	[redacted]
	Q1, Q3	[redacted]	[redacted]	[redacted]
	Min, Max	[redacted]	[redacted]	[redacted]
Age, n(%)		[redacted]	[redacted]	[redacted]
	< 65	[redacted]	[redacted]	[redacted]
	65 - 74	[redacted]	[redacted]	[redacted]
	≥ 75	[redacted]	[redacted]	[redacted]
	Missing	[redacted]	[redacted]	[redacted]
Gender, n(%)		[redacted]	[redacted]	[redacted]
	Male	[redacted]	[redacted]	[redacted]
	Female	[redacted]	[redacted]	[redacted]
	Missing	[redacted]	[redacted]	[redacted]
Race, n(%)		[redacted]	[redacted]	[redacted]
	White	[redacted]	[redacted]	[redacted]
	Black	[redacted]	[redacted]	[redacted]
	Asian	[redacted]	[redacted]	[redacted]
	Other	[redacted]	[redacted]	[redacted]
	Missing	[redacted]	[redacted]	[redacted]
Geographic region, n(%)				
	North America	[redacted]	[redacted]	[redacted]
	Europe	[redacted]	[redacted]	[redacted]
	Asia	[redacted]	[redacted]	[redacted]
	Australia	[redacted]	[redacted]	[redacted]
	South America	[redacted]	[redacted]	[redacted]
	Missing	[redacted]	[redacted]	[redacted]

WHO AML classification, n(%)		[redacted]	[redacted]	[redacted]
	AML with recurrent genetic abnormalities	[redacted]	[redacted]	[redacted]
	AML with myelodysplasia-related changes	[redacted]	[redacted]	[redacted]
	AML not otherwise specified	[redacted]	[redacted]	[redacted]
Type of AML, n(%)		[redacted]	[redacted]	[redacted]
	Primary	[redacted]	[redacted]	[redacted]
	Secondary	[redacted]	[redacted]	[redacted]
	Missing	[redacted]	[redacted]	[redacted]
Response status (CR/CRi) after induction therapy (with or without consolidation therapy), n(%)				
	CR	[redacted]	[redacted]	[redacted]
	CRi	[redacted]	[redacted]	[redacted]
	Missing	[redacted]	[redacted]	[redacted]
Prior history of MDS or CMML, n(%)				
	Yes	[redacted]	[redacted]	[redacted]
	Primary	[redacted]	[redacted]	[redacted]
	Secondary	[redacted]	[redacted]	[redacted]
	No	[redacted]	[redacted]	[redacted]
	Missing	[redacted]	[redacted]	[redacted]

Cytogenetic risk category at time of induction therapy, n(%)				
	Intermediate	[redacted]	[redacted]	[redacted]
	Poor	[redacted]	[redacted]	[redacted]
	Missing	[redacted]	[redacted]	[redacted]
Consolidation therapy following induction, n(%)				
	Yes	[redacted]	[redacted]	[redacted]
	1 Cycle	[redacted]	[redacted]	[redacted]
	2 Cycles	[redacted]	[redacted]	[redacted]
	3 Cycles	[redacted]	[redacted]	[redacted]
	4 Cycles	[redacted]	[redacted]	[redacted]
	Missing	[redacted]	[redacted]	[redacted]
	No	[redacted]	[redacted]	[redacted]
	Missing	[redacted]	[redacted]	[redacted]
ECOG performance status, n(%)				
	0	[redacted]	[redacted]	[redacted]
	1	[redacted]	[redacted]	[redacted]
	2	[redacted]	[redacted]	[redacted]
	Missing	[redacted]	[redacted]	[redacted]
Minimal residual disease status from central pathology report, n(%)				

	Positive	[redacted]	[redacted]	[redacted]
	Negative	[redacted]	[redacted]	[redacted]
	Missing	[redacted]	[redacted]	[redacted]
Bone marrow blast (%)				
	N	[redacted]	[redacted]	[redacted]
	Mean (SD)	[redacted]	[redacted]	[redacted]
	Median	[redacted]	[redacted]	[redacted]
	Q1, Q3	[redacted]	[redacted]	[redacted]
	Min, Max	[redacted]	[redacted]	[redacted]
Time from initial AML diagnosis to randomization (months)				
	N	[redacted]	[redacted]	[redacted]
	Mean (SD)	[redacted]	[redacted]	[redacted]
	Median	[redacted]	[redacted]	[redacted]
	Q1, Q3	[redacted]	[redacted]	[redacted]
	Min, Max	[redacted]	[redacted]	[redacted]
Time from start of induction therapy to randomization (months)				
	N	[redacted]	[redacted]	[redacted]
	Mean (SD)	[redacted]	[redacted]	[redacted]
	Median	[redacted]	[redacted]	[redacted]
	Q1, Q3	[redacted]	[redacted]	[redacted]
	Min, Max	[redacted]	[redacted]	[redacted]

NICE National Institute for Health and Care Excellence

Patient organisation submission

Oral azacitidine for maintenance treatment of acute myeloid leukaemia after induction therapy [ID3892]

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. [Please note that declarations of interests relevant to this topic are compulsory].

Information on completing this submission

- Please do not embed documents (such as a PDF) in a submission because this may lead to the information being mislaid or make the submission unreadable
- We are committed to meeting the requirements of copyright legislation. If you intend to include **journal articles** in your submission you must have copyright clearance for these articles. We can accept journal articles in NICE Docs.
- Your response should not be longer than 10 pages.

About you

1.Your name	[redacted]
2. Name of organisation	Leukaemia Care
3. Job title or position	[redacted]
4a. Brief description of the organisation (including who funds it). How many members does it have?	**Leukaemia Care** is a national blood cancer charity, founded in 1969. We are dedicated to ensuring that anyone affected by blood cancer receives the right information, advice and support. Approximately 85-90% of our income comes from fundraising activities – such as legacies, community events, marathons etc. Leukaemia Care also receives funding from a wide range of pharmaceutical companies, but in total those funds are less than 15% of our annual income. Leukaemia Care has undertaken a voluntary commitment to adhere to specific policies that regulate our involvement with the pharmaceutical industry set out in our code of practice here: https://media.leukaemiacare.org.uk/wp-content/uploads/Leukaemia-CARE-Code-of-Practice-pdf.pdf.
4b. Has the organisation received any funding from the manufacturer(s) of the technology and/or comparator products in the last 12 months? [Relevant	Novartis £1,887.95 (£292.95 ASH video and £1,595 honorarium) Pfizer £10,000 support services

manufacturers are listed in the appraisal matrix.] If so, please state the name of manufacturer, amount, and purpose of funding.	
4c. Do you have any direct or indirect links with, or funding from, the tobacco industry?	No
5. How did you gather information about the experiences of patients and carers to include in your submission?	Information was gathered through Leukaemia Care's patient survey 'Living with Leukaemia (2017), which included responses from 443 AML patients. Data and quotes were also gathered from a new survey (2021), conducted for the purpose of this submission, on patients' opinions on treatment options in AML. Some statistics were taken from an ALAN (Acute Leukaemia Advocates Network) report. Additional information (e.g., quotes) was gathered through a one-to-one conversation with a patient who had previously received azacitidine.
Living with the condition	
6. What is it like to live with the condition? What do carers	Acute myeloid leukaemia (AML) is a rapidly progressing form of leukaemia. As of 2018, there are 3089 new cases in the UK a year, and 2,628 deaths. Generally, only around 20% of people diagnosed with AML will survive for 5 years or more after their diagnosis.

experience when caring for someone with the condition?	The rapidly progressing nature of this condition means that 53% of AML patients are diagnosed via emergency presentation (NCIN/NCRAS routes to diagnosis report). This compares to a cancer average of 21%. Additionally, 79% of patients start treatment within a week of their diagnosis. Being diagnosed with AML can also have a huge emotional impact, prompting patients (and their families) to experience feelings of disbelief, denial, anger, fear, blame, guilt, isolation and depression. In our survey, 42% of AML patients reported that they have felt depressed or anxious more often since their diagnosis and 5% said they feel constantly depressed or anxious since diagnosis. The emotional impact does not only affect the patient in isolation and is often also felt by carers and family members. This can place huge emotional strain on families and friends, many of whom may be affected by the diagnosis. As such, improvements in a patients' treatment and prognosis will also have a wider impact on the lives of their family and friends. Relapse rates are high in AML with about 50% of all patients who achieved remission after their initial treatment relapsing. Evidence indicates that having relapsed from initial treatment worsens a patient's quality of life further. Relapsed patients are more likely to feel isolated all of the time, they are also the most likely group to experience anxiety (74%). Additionally, relapsed patients will have to experience the physical and emotional effects of sometimes gruelling treatment again. The negative financial impact of having AML is felt by the majority of patients; 56% of patients reported increased costs and/or reduced income which is higher than the average for other leukaemia types (43%). Due to the nature of AML, 79% had to stop working or their time in education altogether (compared with 45% across other leukaemia types). This negative impact is increased if carers such as family/household members have to reduce hours or stop working in order to care for their loved one with AML. This undoubtedly adds additional stress and worry for patients and their families and reduces their quality of life after diagnosis further.

Current treatment of the condition in the NHS	
7. What do patients or carers think of current treatments and care available on the NHS?	In our recent survey for AML patients for the purpose of this submission, when asked if they thought existing treatment options for AML on the NHS were sufficient 77.8% of respondents said either no or not sure. Some of the backbone therapies, e.g., stem cell transplant and chemotherapy, often have high levels of toxicity and severe/long-term side effects. An AML patient told us "*The treatment is quite cruel and doesn't take into account the patient*". Another major reason for adults with AML to claim that current treatments available on the NHS are insufficient is that there is currently no potential cure in this setting. Other therapies and comparators available in the relapsed setting include salvage chemotherapy, which is used if a patient has not responded to prior chemotherapy treatments. However, salvage chemotherapy only extends patient lives by a matter of months.
8. Is there an unmet need for patients with this condition?	Yes. As AML has poor prognosis, patients think that more treatment options are needed in this setting. When asked what was not being addressed by existing treatment options one patient commented *"I would strive for a much higher 'cure rate' and a less gruelling treatment regime. More treatments that offer hope of a full cure, not just remission."* Moreover, as AML has high relapse rates patients also want access to more drugs which can prevent relapse. One AML patient told us *"I relapsed after 1st diagnosis, so improved treatment may have prevented the relapse".* The negative psychological, financial, and quality of life impact relapse has means that preventing it before happening is in patients' best interests. The need for drugs which prevent relapse is also highlighted by patient's desire not to have a second round of treatment that goes back to the backbone therapies which patients have described as *"gruelling"* and sometimes intolerable.

	In some cases, when patients run out of treatment options, best supportive care before death is the only option. There is therefore an unmet need that more treatment options need to be made available in this setting.
Advantages of the technology	
9. What do patients or carers think are the advantages of the technology?	Azacitidne as maintenance therapy could prevent relapse before it happens. Relapsing leads to lower chances of overall survival and has a negative impact on patients' quality of life, e.g., their mental health. Hence having treatments which could prevent relapse before it happens, such as azacitidine, could improve patient experience and save lives. Our recent survey showed that 77.8% of AML patients would be willing to have additional treatments if it could potentially prevent relapse Two patients we spoke to who have had azacitidine previously said that they experienced *"no major side effects"* and that the *"side effects were minimal".* After relapsing following a stem cell transplant, one patient we spoke to with AML said he was given azacitidine to "*try to kickstart my stem cell transplant into working and then keep me on it to prevent my AML from returning".* He commented that *"when I had been handed over to [the consultant's] care nobody had much hope for my survival".* This patient is now in full remission and attributes being alive today partly thanks to azacitidine. This patient had intravenous azacitidine and comments that the *"side effects were virtually non-existent for me. I found myself feeling tired, but not exhausted, for a few days after each course of treatment. I did not experience any physical effects like nausea."* *"The main side effect on my life from intravenous azacitidine was the time it took. Initially I was on seven days of treatment once a month which took a fair bit of time and impacted on my working life. The first days of treatment meant being at the hospital for a full day, by the time I'd had a blood test, waited for the results and then waited for the drug to be made up by the pharmacy. Subsequent visits were shorter, but it still involved travelling to the hospital, including weekends when there were often public transport*

	problems. I was later moved to five days of treatment every six weeks, but it was still a significant time commitment, trying to work holidays around it, for example." The use of oral azacitidine in tablet form in this setting could therefore have a major improvement on patient's quality of life and treatment experience. The patient who had intravenous azacitidine said hypothetically *"having the drug orally would have made a major difference as it would have freed up a significant part of my time and enabled me to lead a much more 'normal' life."*
Disadvantages of the technology	
10. What do patients or carers think are the disadvantages of the technology?	An increased number of treatments is not often desirable, as patients will have to endure more side effects without the full guarantee such treatments will be effective in preventing relapse. There were more adverse events reported in the QUAZAR AML-001 clinical trial in those who took azacitidine as maintenance therapy vs. those who took a placebo. However, those who stopped treatment with azacitidine due to adverse events only accounted for 12.3% of patients, which shows that majority of patients could tolerate the drug. Furthermore, in this trial azacitidine was shown to improve overall survival (OS). Average OS for those who took oral azacitidine was 24.7 months, and by comparison those who took the placebo only had an average OS of 14.8 months. As previously mentioned, our survey showed that 77.8% of AML patients would be willing to have additional treatments if it could potentially prevent relapse, 16.7% were not sure, and only 5.6% (1 person) said no. In general patients would rather take additional treatments to try to prevent relapse, even if the outcome is not guaranteed. As this patient outlines below, taking more treatments and enduring more potential side effects is in most cases preferable to relapse: *"I would have been happy to receive oral azacitidine to prevent relapse as the chemotherapy conditioning me for my stem cell transplant had been fairly extreme. After going through all that, I would have been happy to take a milder form of chemotherapy to remain in remission. It would seem to me to be a small*

	price to pay to remain free of AML and prolong my life." The potential benefits of oral azacitidne as maintenance therapy therefore outweigh the disadvantages of the therapy e.g., side effects.
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

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

- AML is rapidly progressing with poor prognosis. As such the psychological, physical and financial impact of an AML diagnosis on a patient and their loved ones is significant. Relapse rates are high in AML affecting a patient's quality of life further.
- The majority of patients surveyed (77.8%) said they would be willing to have additional treatments to prevent relapse. A drug, such as azacitidne, which has been shown to improve overall survival in the clinical trial, would therefore be welcome by relapsed patients and their families.
- Patients we spoke to who had taken azacitidine reported the side effects as being *"minimal*" and "*virtually non-existent*".
- Patients favour an oral therapy as it reduces travel time, financial burden and allows patients to spend more time with friends and family, thus enabling them to lead a more 'normal' life.

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

Your privacy

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

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

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

……

NICE National Institute for Health and Care Excellence

Professional organisation submission

Oral azacitidine for maintenance treatment of acute myeloid leukaemia after induction therapy [ID3892]

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	University Hospitals Birmingham (UHB) NHS Foundation Trust and University of Birmingham (UoB), Royal College of Pathologists, British Society for Haematology

3. Job title or position	[redacted]
4. Are you (please tick all that apply):	☑ an employee or representative of a healthcare professional organisation that represents clinicians? ☑ a specialist in the treatment of people with this condition? ☑ a specialist in the clinical evidence base for this condition or technology? ☐ other (please specify):
5a. Brief description of the organisation (including who funds it).	Both RCPath and BSH are charities, representing clinicians and scientists involved in diagnostics and treatment of haematological disease.
5b. Has the organisation received any funding from the manufacturer(s) of the technology and/or comparator	No

products in the last 12 months? [Relevant manufacturers are listed in the appraisal matrix.] If so, please state the name of manufacturer, amount, and purpose of funding.	
5c. 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.)	To improve the survival of patients with acute myeloid leukaemia and to prevent relapse following remission after induction chemotherapy.
7. What do you consider a clinically significant treatment response? (For example, a reduction in tumour size by x cm, or a	For oral azacitidine, a clinically significant response will be maintaining remission from the leukaemia.

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?	The risk of relapse for patients is substantial. For the small number of patients with good risk genetic abnormalities in their leukaemia, the risk of relapse is up to 30% at 2 years. For those with intermediate and high risk genetics the risk is much higher. Intermediate risk patients have between 40 and 70% and high risk it is in excess of 70% and usually inevitable. The number of patients in the higher risk groups increases with age. While allogeneic transplant can be offered to younger patients (if fit enough) to reduce the risk of relapse, the majority of patients above the age of 60 cannot undergo such an intensive procedure. The median age of incidence of AML is between 70-75 years. The options for treatment at relapse are limited in the older age group; the vast majority will only have low dose palliative chemotherapy or conservative management with transfusions and antibiotic treatment alone. There is therefore an unmet clinical need for patients to reduce the risk of relapse for unfit patients and the older, frailer age group.
What is the expected place of the technology in current practice?	
9. How is the condition currently treated in the NHS?	In summary, the treatment of acute myeloid leukaemia can be either intensive or non-intensive depending on the fitness of the patient. Intensive treatment is aimed to be curative, with induction chemotherapy followed by consolidation with chemotherapy and/or an allogeneic stem cell transplant depending on the risk stratification of the patient. Non-intensive chemotherapy protocols are palliative aiming to increase the overall survival of the patient and improve their quality of life.
• Are any clinical guidelines used in the treatment of	Broadly, guidelines are based on the European Leukaemia Net guidelines from 2017 (https://doi.org/10.1182/blood-2016-08-733196) . More specific guidelines have been written by regional cancer alliances in England e.g. West Midlands Cancer Alliance (https://wmcanceralliance.nhs.uk/images/Documents/Haematology/Final_Guidance_for_Acute_Myeloid_Leukaemia_Treatment_in_Adults_in_the_West_Midlands_v18_clean.pdf) and Pan London Cancer Alliance (https://rmpartners.nhs.uk/wp-content/uploads/2020/01/Pan-London-AML-Guidelines-Jan-2020.pdf) The guidance elaborates on the pathways outlined in the NICE guidance for AML (https://pathways.nice.org.uk/pathways/blood-and-

the condition, and if so, which?	bone-marrow-cancers/leukaemia#path=view%3A/pathways/blood-and-bone-marrow-cancers/myeloid-leukaemia.xml&content=view-node%3Anodes-acute-myeloid-leukaemia)
• Is the pathway of care well defined? Does it vary or are there differences of opinion between professionals across the NHS? (Please state if your experience is from outside England.)	Yes, outside of clinical trials chemotherapy treatment is uniform across England, following the pathways described in the guidelines. The one major change due to COVID-19 is that for some patients where it is deemed that the risks from COVID are high, may have treatment with venetoclax and azacitidine or low dose cytarabine (http://www.cureleukaemia.co.uk/page/news/523/aml-working-party-covid-19-recommendations). This is to reduce the inpatient stay for patients, the degree of cytopenias and infection and their risk of contracting COVID19. There is variation in the implementation of this guidance depending on the perceived risks in each region and the ability of hospitals to go ahead with intensive chemotherapy.
• What impact would the technology have on the	The technology would be used following intensive chemotherapy. In the phase 3 trial looking at maintenance oral azacitidine (http://doi.org/10.1056/NEJMoa2004444 , QUAZAR AML-001) 80% of patients started oral azacitidine in remission after at least one consolidation course of chemotherapy, but all had an improved survival compared with placebo. For patients suitable for intensive chemotherapy but unsuitable for allogeneic stem cell transplant this gives them an increased survival and reduces their risk of relapse. Many patients received maintenance after only a single

current pathway of care?	consolidation course, so it may improve their quality of life; they may not need prolonged admissions to hospital for second or third consolidation courses of chemotherapy (each of which is typically between 4-6 weeks).
10. Will the technology be used (or is it already used) in the same way as current care in NHS clinical practice?	Currently maintenance chemotherapy is not routinely prescribed for patients with AML except those with FLT3 mutations who receive midostaurin (a tyrosine kinase inhibitor) for up to 12 months following completion of consolidation chemotherapy if they do not proceed to allogeneic stem cell transplant.
• How does healthcare resource use differ between the technology and current care?	Patients are followed up every month to 6 weeks (at the discretion of the clinician) following completion of consolidation chemotherapy to monitor their recovery and for relapse. Patients with a genetic marker for minimal residual disease (MRD) may have a marrow aspirate to measure their MRD every 3 months for the first 2 years. As time from completion of chemotherapy increases, the time between follow up appointments lengthens at the discretion of the clinician. With the use of oral azacitidine maintenance, patients will need to be reviewed every 28 days, prior to each 14 day course of treatment. However, similarly to many patients receiving oral chemotherapy agents since the advent of the COVID-19 pandemic, this may be done remotely as long as patients remain well.
• In what clinical setting should the	The drug treatment will be used exclusively in secondary care by practising haematologists.

technology be used? (For example, primary or secondary care, specialist clinics.)	
• What investment is needed to introduce the technology ? (For example, for facilities, equipment, or training.)	Haematology clinics are well versed in the use of oral chemotherapy agents in the outpatient settings and would be staffed to implement this. However, there may be an increased resource in terms of the number of patients returning to clinic 4 weekly having chemotherapy.
11. Do you expect the technology to provide clinically meaningful	Yes, I expect an increased number of patients for whom we cannot offer a consolidative allogeneic stem cell transplant to have a longer survival without leukaemia and therefore an improved quality of life.

benefits compared with current care?	
• Do you expect the technology to increase length of life more than current care?	Yes, the phase 3 data suggests that patients will live longer.
• Do you expect the technology to increase health-related quality of life more than current care?	Living without leukaemia will imply that these patients are less likely to require transfusions and get life threatening infections. Compared with placebo, there were increased numbers of patients with gastro-intestinal toxicity (nausea, vomiting, diarrhoea and constipation) and fatigue (https://doi.org/10.1186/s13045-021-01142-x). Although most are grade 1 or 2, given the drug needs to be taken chronically this could affect the patient's quality of life. The paper also mentions neutropenia and thrombocytopenia but in general haematologists can manage these problems. The fatigue and quality of life scores were not significantly different between patients on treatment and placebo. (https://www.nejm.org/doi/suppl/10.1056/NEJMoa2004444/suppl_file/nejmoa2004444_appendix.pdf)
12. Are there any groups of people for whom the	Patients for whom it is not possible to find a suitable stem cell donor may benefit from treatment with oral azacitidine. There are reduced numbers of donors of Asian and African origin on the international stem cell donor panels. In the absence of a related donor, it can be very difficult to find a suitable donor for these patients. Oral azacitidine offers an opportunity for them to improve their survival.

technology would be more or less effective (or appropriate) than the general population?	
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	Haematology clinics are well versed in the use of oral chemotherapy agents in the outpatient settings and would be staffed to implement this. However, there may be an increased resource in terms of the number of patients returning to clinic 4 weekly having chemotherapy.

concomitant treatments needed, additional clinical requirements, factors affecting patient acceptability or ease of use or additional tests or monitoring needed.)	
14. Will any rules (informal or formal) be used to start or stop treatment with the technology? Do these include	Treatment will be stopped if the leukaemia recurs. There may be occasions where patients cannot tolerate the treatment. Neither of these situations would require additional testing.

any additional testing?	
15. Do you consider that the use of the technology will result in any substantial health-related benefits that are unlikely to be included in the quality-adjusted life year (QALY) calculation?	No, I would expect the increase in overall survival should increase the QALY calculation.
16. Do you consider the technology to be innovative in its potential to make	Up until now, most post-consolidation maintenance chemotherapy regimens have not shown an improvement in overall survival, although some have shown an improvement in relapse free survival. Additionally delivering this form of chemotherapy is practical because it is an oral tablet and seems reasonably well tolerated, giving a reasonable quality of life.

a significant and substantial impact on health-related benefits and how might it improve the way that current need is met?	
• Is the technology a 'step-change' in the management of the condition?	It should significantly improve the survival of patients following chemotherapy who cannot have a transplant.
• Does the use of the technology address any particular unmet need of the	It should significantly improve the survival of patients following chemotherapy who cannot have a transplant. They may not be able to have a transplant because of their co-morbidities, frailty (especially following intensive chemotherapy) or because do not have a suitable donor.

patient population?	
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?	Compared with placebo, there were increased numbers of patients with gastro-intestinal toxicity (nausea, vomiting, diarrhoea and constipation) and fatigue (https://doi.org/10.1186/s13045-021-01142-x). Although most are grade 1 or 2, given the drug needs to be taken chronically this could affect the patient's quality of life. The paper also mentions neutropenia and thrombocytopenia but in general haematologists can manage these problems reasonably well. The fatigue and quality of life scores were not significantly different between patients on treatment and placebo (https://www.nejm.org/doi/suppl/10.1056/NEJMoa2004444/suppl_file/nejmoa2004444_appendix.pdf)
Sources of evidence	
18. Do the clinical trials on the technology reflect current UK clinical practice?	Yes they do.
• If not, how could the results be extrapolate	

d to the UK setting?	
• What, in your view, are the most important outcomes, and were they measured in the trials?	Overall survival (OS) is most important followed by relapse free survival. This is the first trial showing improved OS with an oral maintenance regime after consolidation chemotherapy. In addition it appears to be tolerable, important if it is to be taken long term during maintenance.
• If surrogate outcome measures were used, do they adequately predict long-term clinical outcomes?	N/A
• Are there any adverse effects that	Not routinely used yet, so I cannot comment.

were not apparent in clinical trials but have come to light subsequently?	
19. Are you aware of any relevant evidence that might not be found by a systematic review of the trial evidence?	Only data from a single phase 3 trial is available.
20. Are you aware of any new evidence for the comparator treatment(s) since the	The comparator was placebo.

publication of NICE technology appraisal guidance TA523?	
21. How do data on real-world experience compare with the trial data?	None yet available.
Equality	
22a. Are there any potential equality issues that should be taken into account when considering this treatment?	Need to make sure it is available to all those who not eligible for transplant including ethnic minorities.

22b. Consider whether these issues are different from issues with current care and why.	At the moment there is no further treatment available for patients without an allogeneic stem cell donor.
Topic-specific questions	
23. How are subcutaneous azacitidine and low dose cytarabine used in clinical practice after induction therapy for adults with acute myeloid leukaemia who have complete	Neither are used routinely after induction chemotherapy and consolidation. The evidence for maintenance with subcutaneous (s/c) azacitidine comes from the HOVON97 trial (http://doi.org/10.1182/BLOOD-2018-10-879866) which showed an improvement in relapse free survival (RFS) but no improvement in OS at 1 year. The NCRI AML16 study similarly did not show an improved OS at 5 years (https://library.ehaweb.org/eha/2015/20th/103225/alan.burnett.a.comparison.of.limited.consolidation.chemotherapy.therapy.or.not.html?f=m1). Giving s/c azacitidine is more difficult than oral and requires more day unit visits unless it can be delivered at home. Although there was evidence that MRD negative patients had a significant improved survival from this study, it has not become routinely used. S/c cytarabine is mostly used palliative for patients to control blood counts if they are unfit for other forms of treatment usually at presentation and sometimes relapse.

disease remission, or complete remission with incomplete blood count recovery, who are not eligible for, including those who choose not to proceed to, haematopoietic stem cell transplantation?	
Key messages	

24. In up to 5 bullet points, please summarise the key messages of your submission.

- Improves overall survival of patients who cannot have allogeneic stem cell transplant after consolidation chemotherapy
- Is reasonably well tolerated in comparison to placebo
- Is deliverable in outpatient setting or even as a remote clinic because it is oral
- May help patients from ethnic minorities for who may not have have an unrelated donor available
- May reduce the intensity of consolidation required as most patients received after one to two consolidation courses.

Thank you for your time.

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

……

Your privacy

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

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

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

……

in collaboration with:

Erasmus School of Health Policy & Management

Maastricht University

Oral azacitidine for maintenance treatment of acute myeloid leukaemia after induction therapy [ID3892]

Produced by Kleijnen Systematic Reviews (KSR) Ltd, in collaboration with Erasmus University Rotterdam (EUR) and Maastricht University Medical Centre (UMC)

Authors

Jeremy Howick, Reviews Manager, KSR Ltd, United Kingdom (UK)
Willem Witlox, Health Economist, Maastricht UMC, The Netherlands
Charlotte Ahmadu, Health Economist, KSR Ltd, UK
Sabine Grimm, Health Economist, Maastricht UMC, The Netherlands
Nigel Armstrong, Health Economics Manager, KSR Ltd, UK
Kevin McDermott, Systematic Reviewer, KSR Ltd, UK
Thomas Otten, Health Economist, Maastricht UMC, The Netherlands
Caro Noake, Information Specialist, KSR Ltd, UK
Robert Wolff, Managing Director, KSR Ltd, UK
Manuela Joore, Health Economist, Maastricht UMC
Jos Kleijnen, Founder and Owner, KSR Ltd, UK

Correspondence to Jeremy Howick, Kleijnen Systematic Reviews Ltd
Unit 6, Escrick Business Park
Riccall Road, Escrick
York, YO19 6FD
United Kingdom

Date completed 18/02/2021

Source of funding: This report was commissioned by the National Institute for Health Research (NIHR) Evidence Synthesis Programme as project number NIHR135436.

Declared competing interests of the authors 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.

Confidential comparator prices are highlighted in green throughout the report.

Any de-personalised data are highlighted in pink 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 Evidence Synthesis Programme. Any errors are the responsibility of the authors.

This report should be referenced as follows:

Howick J, Witlox W, Ahmadu C, Grimm S, Armstrong N, McDermott K, Otten T, Noake C, Wolff R Joore M, Kleijnen J. Oral azacitidine for maintenance treatment of acute myeloid leukaemia after induction therapy [ID3892]: a Single Technology Assessment. York: Kleijnen Systematic Reviews Ltd, 2022.

Contributions of authors

Jeremy Howick 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. Willem Witlox acted as health economic project lead, critiqued the company's economic evaluation and contributed to the writing of the report. Manuela Joore, Sabine Grimm, Charlotte Ahmadu, and Nigel Armstrong acted as health economists on this assessment, critiqued the company's economic evaluation and contributed to the writing of the report. Kevin McDermott and Charlotte Ahmadu acted as systematic reviewers, critiqued the clinical effectiveness methods and evidence and contributed to the writing of the report. Caro Noake critiqued the search methods in the submission and contributed to the writing of the report. Robert Wolff and 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

AE	Adverse events
AESI	Adverse events of special interest
AIC	Akaike Information Criterion
AML	Acute myeloid leukaemia
BCS	Best case scenario
BIC	Bayesian information criterion
BNF	British National Formulary
BSC	Best supportive care
BSCH	British Committee for Standards in Haematology
CENTRAL	Cochrane Central Register of Controlled Trials
CI	Confidence interval
CMML	Chronic myelomonocytic leukaemia
CR	Complete remission
CRD	Centre for Reviews and Dissemination
CRi	Complete remission with incomplete blood count recovery
CS	Company Submission
CSR	Clinical study report
DARE	Database of Abstracts of Reviews of Effects
DSA	Deterministic sensitivity analyses
DSU	Decision Support Unit
ECOG	Eastern Cooperative Oncology Group
EMA	European Medicines Agency
eMIT	Electronic market information tool
EP	Extension phase
EQ-5D	European Quality of Life-5 Dimensions
ERG	Evidence Review Group
ESMO	European Society for Medical Oncology
EU	European Union
FLT3	FMS-like tyrosine kinase 3
FDA	Food and Drug Administration
GI	Gastrointestinal
HCRU	Healthcare resource utilisation
HRQoL	Health-related quality of life
HSCT	Hematopoietic stem cell transplantation
ICER	Incremental cost effectiveness ratio
HR	Hazard ratio
IC	Intensive chemotherapy
ITC	Indirect treatment comparison
ITD	Internal tandem duplications
ITT	Intention to treat
KM	Kaplan-Meier
KSR	Kleijnen Systematic Reviews
LYs	Life years
MAIC	Matched adjusted indirect comparison
MDS	Myelodysplastic syndrome
MeSH	Medical subject headings
MHRA	Medicines and Healthcare Products Regulatory Agency
MID	Minimally important difference
MRD	Measurable residual disease
n	Number of patients in the category
N	Number of patients evaluable
NCCN	National Comprehensive Cancer Network

NHS	National Health Service
NICE	National Institute for Health and Care Excellence
OS	Overall survival
PAS	Patient access scheme
PFS	Progression-free survival
PSA	Probabilistic sensitivity analysis
PSM	Partition support model
PSS	Personal Social Services
QALY	Quality adjusted life year
QD	Per day
RBC	Red Blood Count
RCT	Randomised controlled trial
RFS	Relapse-free survival
RoB	Risk of bias
RTK	Receptor tyrosine kinases
SAE	Serious adverse event
SLR	Systematic literature review
SmPC	Summary of product characteristics
SoC	Standard of care
SOC	System organ class
STA	Single Technology Appraisal
STC	Stem cell transplant
STM	State transition model
TEAE	Treatment emergent adverse events
TKD	Tyrosine kinase domain
TSD	Technical support documents
UK	United Kingdom
UMC	University Medical Centre
WHO	World Health Organization

Table of Contents

Table of Tables

Table of Figures

1. EXECUTIVE SUMMARY

This summary provides a brief overview of the key issues identified by the Evidence Review Group (ERG) as being potentially important for decision making. If possible, it also includes the ERG's preferred assumptions and the resulting incremental cost effectiveness ratios (ICERs).

Section 1.1 provides an overview of the key issues. Section 1.2 presents the key model outcomes. Section 1.3 discusses the decision problem, Section 1.4 issues relate to the clinical effectiveness, and Section 1.5 issues related to the cost effectiveness. Other key issues are discussed in Section 1.6 while a summary in presented in Section 1.7.

Background information on the condition, technology and evidence and information on key as well as non-key issues are in the main ERG report, see Sections 2 (background), 3 (decision problem), 4 (clinical effectiveness) and 5 (cost effectiveness) for more details.

All issues identified represent the ERG's view, not the opinion of the National Institute for Health and Care Excellence (NICE).

1.1 Overview of the ERG's key issues

Table 1.1: Summary of key issues

ID3892	Summary of issue	Report sections
1	Low dose cytarabine and subcutaneous azacitidine are part of standard therapy according to NICE guidance yet were not viewed by the company to be part of BSC.	2, 3.2, 4.2
2	Most patients in the QUAZAR trial received one dose or no doses of consolidation therapy, resulting in a selection bias that could have exaggerated the benefits of oral azacitidine.	2, 3.2, 4.2
3	Few patients in the QUAZAR trial were recruited from UK sites, and there were relevant differences between the UK and analysed populations; this limits the generalisability to UK clinical practice.	3.2.3
4	HRQoL and fatigue were measured on day 1 of each 28-day cycle, when adverse events were less likely to arise.	3.2
5	Randomisation of patients in RATIFY trial occurred at induction and not maintenance phase, potentially introducing a high risk of bias in any analysis at the maintenance phase.	3.3
6	The SLR eligibility criteria would not have identified the RATIFY trial; other midostaurin studies may also have been missed.	3.3
7	HSCT was not included as a separate health state but was implicitly included in the modelling through the survival analysis, increasing the likelihood of bias.	4.2.2
8	Some patients in QUAZARAML-001 trial received fewer cycles of consolidation therapy than is standard practice in the UK. This limits the applicability of the results to a UK setting.	4.2.3
9	Patient baseline characteristics in model are not subgroup-specific (for example in the FLT3 subgroup, consolidation subgroup or Europe subgroup); patient baseline characteristics may not align with the subgroups being analysed.	4.2.3

ID3892	Summary of issue	Report sections
10	Survival analyses of the FLT3 subgroup are likely to be biased due to limitations associated with the indirect comparison.	4.2.6
11	In the company's base-case analysis, only grade 3 and 4 AEs are applied with a maximum frequency of one and a duration of 1 week, which may underestimate the real impact of AEs.	4.2.7
12	The current source of utility values may not accurately reflect the relapse utility.	4.2.8
13	Some resource use estimates appear inconsistent with expert opinion and require further justification.	4.2.9
14	Treatment effectiveness in the FLT3 subgroup was analysed for the different comparisons separately; preventing comparison of oral azacitidine, midostaurin, watch and wait plus BSC.	5.1
AE = adverse event; BSC = best supportive care; ERG = Evidence Review Group; NICE = National Institute for Health and Care Excellence; HRQoL = health-related quality of life; HSCT = hematopoietic stem cell transplantation; SLR = systematic literature review; UK = United Kingdom		

1.2 Overview of key model outcomes

NICE technology appraisals compare how much a new technology improves length (overall survival) and quality of life in a quality adjusted life year (QALY). An ICER is the ratio of the extra cost per QALY gained.

Overall, the technology is modelled to affect QALYs by:

- Increased relapse-free survival (RFS), with an incremental of 0.853 years (80% of total incremental life years (LYs)) in the oral azacitidine arm (2.088 years) compared with watch and wait with best supportive care (BSC) arm (1.235 years).
- Increased post-relapse survival, with an incremental of 0.211 years (20% of total incremental LYs) in the oral azacitidine arm (1.779 years) compared with watch and wait with BSC arm (1.568 years).

Overall, the technology is modelled to affect costs by:

- The higher drug costs (additional cost of [redacted], [redacted] of total incremental costs) and disease management costs (additional cost of [redacted]) in RFS on-treatment compared with watch and wait plus BSC.
- The lower disease management costs (reduced cost of [redacted]) in RFS off-treatment compared with watch and wait plus BSC.

The company performed and presented the results of probabilistic sensitivity analyses (PSA), deterministic sensitivity analyses (DSA) as well as scenario analyses. The parameters that had the greatest effect on the ICER based on the company's deterministic sensitivity analyses were:

- Health state utility – RFS on treatment
- Health state utility – RFS off treatment
- Oral azacitidine relative dose intensity

Company Submission (CS) scenarios that have the greatest impact on the ICER (not including scenarios related to discount rates and time horizon) were:

- Using the QUAZAR AML-001 Europe only population (decreased ICER to [redacted])
- Cure modelling with a 5-year cure point (decreased ICER to [redacted])
- Utility values based on Joshi 2019 for all health states (decreased ICER to [redacted])

1.3 The decision problem: summary of the ERG's key issues

The ERG identified one issue related to the comparators used in the CS (see Table 1.2), one issue related to the population (see Table 1.3), and one issue related to the outcomes (see Table 1.5).

Table 1.2: Key issue 1 Appropriateness of excluding low dose cytabarine and subcutaneous azacitidine as part of best supportive care

Report Section	**2, 3.2, 4.2**
Description of issue and why the ERG has identified it as important	Low dose cytarabine and subcutaneous azacitidine were not viewed by the company to be part of BSC. Yet, they are part of standard therapy according to NICE guidance. Failure to include these treatments may have overestimated the benefits and safety of azacitidine.
What alternative approach has the ERG suggested?	None.
What is the expected effect on the cost effectiveness estimates?	Unknown.
What additional evidence or analyses might help to resolve this key issue?	Additional evidence about the use of low dose cytabarine and subcutaneous azacitidine as BSC in this population, for example from independent clinical experts.
BSC = best supportive care; ERG = Evidence Review Group; NICE = National Institute for Health and Care Excellence	

Table 1.3: Key issue 2. Patients in the main (QUAZAR) trial may not have received sufficient consolidation therapy

Report section	**2, 3.2, 4.2**
Description of issue and why the ERG has identified it as important	Most patients in the QUAZAR trial received one dose or no doses of consolidation therapy, whereas at least one dose is recommended by NICE. This generated a non-representative sample for the trial that may have exaggerated the apparent benefits of oral azacitidine.
What alternative approach has the ERG suggested?	The ERG base-case included patients who had received at least one cycle of consolidation therapy.
What is the expected effect on the cost effectiveness estimates?	Excluding patients with no cycles of consolidation therapy from the analysis increased the ICER through an improvement in LYs gained for patients in the comparator arm (decrease in oral azacitidine arm).
What additional evidence or analyses might help to resolve this key issue?	Further evidence on proportions of patients in UK clinical practice receiving no/1/2+ cycles of consolidation.
ERG = Evidence review group; ICER = Incremental cost effectiveness ratio; LYs = life years; NICE = National Institute for Health and Care Excellence; UK = United Kingdom	

1.4 The clinical effectiveness evidence: summary of the ERG's key issues

Key issues 1 and 2, detailed in Section 1.3 apply to this Section as well. Key issue 3 concerns the generalisability to the UK setting, and key issue 4 concerns the way in which certain AEs were measured.

Table 1.4: Key issue 3. Few UK patients and questionable generalisability to UK NHS setting

Report section	**3.2.3**
Description of issue and why the ERG has identified it as important	Only 35 (out of 472) patients in the QUAZAR trial were recruited from UK sites, and there are notable differences between the UK population and the populations analysed. This limits the generalisability to the UK setting.
What alternative approach has the ERG suggested?	None.
What is the expected effect on the cost effectiveness estimates?	Unclear.
What additional evidence or analyses might help to resolve this key issue?	The ERG has no further suggestions.
ERG = Evidence Review Group; UK = United Kingdom	

Table 1.5: Key issue 4. The way Health related quality of life (HRQoL) and fatigue were measured could have exaggerated the benefits and safety of oral azacitidine

Report section	**3.2**
Description of issue and why the ERG has identified it as important	HRQoL and fatigue were measured on day 1 of each 28-day cycle. This may have missed AEs, given that patients would have been off oral azacitidine for 14 days prior.
What alternative approach has the ERG suggested?	None.
What is the expected effect on the cost effectiveness estimates?	The ICER is likely to increase, as AEs are expected to occur more frequently in the oral azacitidine arm.
What additional evidence or analyses might help to resolve this key issue?	The ERG has no further suggestions.
AEs = adverse events; ERG = Evidence Review Group; HRQoL = health-related quality of life; ICER = incremental cost effectiveness ratio	

Table 1.6: Key issue 5. Randomisation of patients in ITC RATIFY trial

Report section	**3.3**
Description of issue and why the ERG has identified it as important	Randomisation of patients in the RATIFY trial (which was not prospectively designed to determine the independent effect of midostaurin as maintenance therapy) occurred at induction and not maintenance phase, potentially introduces bias in any analysis at the maintenance phase.

Report section	3.3
What alternative approach has the ERG suggested?	If possible, conducting an analysis using an RCT of midostaurin where patients were randomised maintenance phase.
What is the expected effect on the cost effectiveness estimates?	Unclear
What additional evidence or analyses might help to resolve this key issue?	The ERG has no further suggestions.
ERG = Evidence Review Group; RCT = randomised controlled trial	

Table 1.7: Key issue 6. ITC SLR eligibility criteria may have missed relevant studies

Report section	3.3
Description of issue and why the ERG has identified it as important	The SLR eligibility criteria lists 'oral azacitidine' as the only intervention which implies that it would not have identified studies of relevant treatments which were not compared to oral azacitidine, such as the RATIFY trial. Other midostaurin studies may also have been missed.
What alternative approach has the ERG suggested?	It is unlikely that there are other studies of midostaurin in the FLT3 population.
What is the expected effect on the cost effectiveness estimates?	Unclear.
What additional evidence or analyses might help to resolve this key issue?	An updated ITC SLR eligibility criteria with midostaurin and placebo as intervention/comparator.
ERG = Evidence Review Group; ITC = indirect treatment comparison; SLR = systematic literature review	

1.5 The cost effectiveness evidence : summary of the ERG's key issues

A full summary of the cost effectiveness evidence review conclusions can be found in Section 6.4 of this report. The company's cost effectiveness results are presented in Section 5, the ERG's summary and detailed critique in Section 4, and the ERG's amendments to the company's model and results are presented in Section 6. The key issues in the cost effectiveness evidence are discussed in the key issue Tables below.

Table 1.8: Key issue 7: HSCT not appropriately reflected in the modelling

Report section	4.2.2
Description of issue and why the ERG has identified it as important	HSCT was not included as a separate health state but was implicitly included in the modelling through the survival analysis. This may induce bias. The potential impact of HSCT on HRQoL is also not captured, apart from a disutility associated with HSCT. Therefore, the impact on HSCT is not accurately reflected in the HRQoL measures.
What alternative approach has the ERG suggested?	Explore a change in the modelling approach by which HSCT is included as a health state in the model (with utility and survival benefit) and survival of patients without HSCT is based on ITT with HSCT censored population. Alternatively, consider adding a utility

	benefit for patients with HSCT and reflecting increased proportion of HSCT in comparator arm by choosing appropriate survival distributions (that allow for a longer tail). Do not use a HSCT disutility decrement if the utility benefit of having HSCT is not captured.
What is the expected effect on the cost effectiveness estimates?	The company's approach likely induces bias in favour of oral azacitidine, as higher proportions of patients in the placebo arm receive HSCT. Currently the patients who receive placebo incur only costs and a disutility as opposed to a benefit in HRQoL, and potentially the benefit on survival is under-estimated.
What additional evidence or analyses might help to resolve this key issue?	Overlay the KM curves of ITT versus ITT with HSCT censored in one plot. Provide AIC/BIC fit for the individual distributions per treatment arm for the HSCT censored analysis. Show all distributions in one plot. Enable a scenario in the economic model where individual distributions can be chosen and modelled together with assumptions about survival for patients with HSCT and their HRQoL. Provide evidence on utility benefit for patients post-HSCT in this or similar population.
AIC = Akaike Information Criterion; BIC = Bayesian information criterion; ERG = Evidence Review Group; HRQoL = health-related quality of life; HSCT = hematopoietic stem cell transplantation; ITT = intention to treat; KM = Kaplan-Meier	

Table 1.9: Key issue 8: QUAZAR trial not representative in terms of consolidation therapy

Report section	**4.2.3**
Description of issue and why the ERG has identified it as important	QUAZAR AML-001 trial likely not representative of UK clinical practice population, amongst others due to differences in consolidation therapy use. Most patients in QUAZARAML-001 trial did not receive at least 2 cycles of consolidation therapy. The company provided a subgroup of patients that received at least 1 cycle of consolidation therapy, but their response lacked a detailed description of the assessment of the NICE TSD DSU 14 criteria.
What alternative approach has the ERG suggested?	Exclude patients with fewer than 2 cycles of consolidation therapy from the analysis and provide a detailed description of the assessment of the NICE TSD DSU 14 criteria in the consolidation subgroup.
What is the expected effect on the cost effectiveness estimates?	Excluding patients with no cycles of consolidation therapy from the analysis increased the ICER through an improvement in life years gained for patients in the comparator arm (decrease in oral azacitidine arm). The impact of selecting other curves on the ICER was not explored due to the lack of a detailed description of the assessment of the NICE TSD DSU 14 criteria.
What additional evidence or analyses might help to resolve this key issue?	Explore excluding patients with fewer than 2 cycles of consolidation therapy. Provide evidence (for example from independent expert opinion) of proportions of patients in UK clinical practice receiving no / 1 / 2+ cycles of consolidation therapy and compare these with the

Report section	4.2.3
	proportions observed in the QUAZAR AML-001 trial. Provide a detailed description of the assessment of the NICE TSD DSU 14 criteria in the consolidation subgroup.
DSU = Decision Support Unit; ERG = Evidence Review Group; ICER = incremental cost effectiveness ratio; NICE = National Institute for Health and Care Excellence; TSD = TECHNICAL SUPPORT DOCUMENT = technical support documents; UK = United Kingdom	

Table 1.10: Key issue 9: Subgroup specific patient baseline characteristics

Report section	4.2.3
Description of issue and why the ERG has identified it as important	Patient baseline characteristics in model are not subgroup-specific (i.e., for FLT3 subgroup, consolidation subgroup or Europe subgroup), and patient baseline characteristics do therefore not align with the subgroups being analysed.
What alternative approach has the ERG suggested?	For subgroup analysis, use subgroup-specific patient baseline characteristics.
What is the expected effect on the cost effectiveness estimates?	Unknown – probably minor.
What additional evidence or analyses might help to resolve this key issue?	An updated model with updated patient baseline characteristics per subgroup.
ERG = Evidence Review Group	

Table 1.11: Key issue 10: Bias and lack of detail in survival analyses for the FLT3 subgroup

Report section	4.2.6
Description of issue and why the ERG has identified it as important	Survival analyses of the FLT3 subgroup are likely to be extremely biased due to limitations associated with the indirect comparison (see Sections 3.4 and 4.2.6 for a more detailed critique). In addition, details for the survival analyses of OS and RFS in the FLT3 subgroup were lacking, including 1) log-cumulative hazard plots, 2) AIC/BIC statistics for individual models, 3) plots showing all joint models in one plot and 4) evaluation of criterion 5 (OS/RFS gain pre and post extrapolation).
What alternative approach has the ERG suggested?	Bias due to limitations associated with the ITC may not be resolvable. In addition, a detailed description of the assessment of the NICE TSD DSU 14 criteria in the FLT3 subgroup. An analysis excluding patients without consolidation therapy for the FLT3 subgroup
What is the expected effect on the cost effectiveness estimates?	Unknown.

What additional evidence or analyses might help to resolve this key issue?	A detailed description of the assessment of the NICE TSD DSU 14 criteria in the FLT3 subgroup, including but not limited to 1) log-cumulative hazard plots, 2) AIC/BIC statistics for individual models, 3) plots showing all joint models in one plot and 4) evaluation of criterion 5 (OS/RFS gain pre and post extrapolation). An analysis excluding patients without consolidation therapy for the FLT3 subgroup would be potentially useful.
AIC = Akaike Information Criterion; BIC = Bayesian information criterion; DSU = Decision Support Unit; ERG = Evidence Review Group; ICER = Incremental cost effectiveness ratio; NICE = National Institute for Health and Care Excellence; OS = overall survival; RFS = relapse-free survival; TSD = Technical support documents	

Table 1.12: Key issue 11: Underestimation of adverse events

Report section	**4.2.7**
Description of issue and why the ERG has identified it as important	In the company's base-case analysis, only grade 3 and 4 AEs are applied with a maximum frequency of one and a duration of 1 week. This may underestimate the real impact of AEs.
What alternative approach has the ERG suggested?	Further research is required to ascertain that AEs are not underestimated.
What is the expected effect on the cost effectiveness estimates?	As AEs have a higher prevalence in the treatment than in the comparator arm, underestimating the impact of AEs will benefit the treatment arm.
What additional evidence or analyses might help to resolve this key issue?	Further information on the AE duration and reoccurrence would help to resolve this issue.
AE = adverse event; ERG = Evidence Review Group	

Table 1.13: Key issue 12: Uncertainty in the choice of HRQoL upon relapse

Report section	**4.2.8**
Description of issue and why the ERG has identified it as important	There is uncertainty around the choice of utility value which was applied upon relapse, and it is unclear why the data from the QUAZAR AML-001 trial was not used. The current source (Joshi et al.) may not accurately reflect the relapse utility.
What alternative approach has the ERG suggested?	Explore alternative sources for the relapse utility value. Analyses should be conducted to investigate the impact of the relapse utility data as per the QUAZAR AML-001 trial.
What is the expected effect on the cost effectiveness estimates?	The size of the overall impact is unclear, and its effect may depend on modelling choices made for treatment effectiveness. Using an alternative utility value sourced from Tremblay decreased the ICER.
What additional evidence or analyses	Further explore the utility value for relapse as calculated based on the QUAZAR AML-001 trial data – and provide a scenario analysis using this/provide justification for why this is likely not adequate.

Report section	4.2.8
might help to resolve this key issue?	
ERG = Evidence Review Group; ICER – incremental cost effectiveness ratio	

Table 1.14: Key issue 13: Lack of clarity about some resource use items

Report section	4.2.9
Description of issue and why the ERG has identified it as important	Some resource use estimates appear inconsistent with expert opinion and require further justification.
What alternative approach has the ERG suggested?	Provide further justification, potentially updated analysis.
What is the expected effect on the cost effectiveness estimates?	Unknown.
What additional evidence or analyses might help to resolve this key issue?	Further justification, and updated analysis.
ERG = Evidence Review Group	

Table 1.15: Key issue 14: Lack of a fully incremental analysis for all comparators in the FLT3 subgroup

Report section	Section 5.1
Description of issue and why the ERG has identified it as important	Treatment effectiveness in the FLT3 subgroup was analysed for the different comparisons separately so a fully incremental analysis was not performed. The use of different analyses is problematic as it does not allow for comparison of oral azacitidine, midostaurin and watch and wait plus BSC.
What alternative approach has the ERG suggested?	Perform a fully incremental analysis for all comparators in the FLT3 subgroup.
What is the expected effect on the cost effectiveness estimates?	Unknown.
What additional evidence or analyses might help to resolve this key issue?	A fully incremental analysis for all comparators in the FLT3 subgroup.
BSC = best supportive care; ERG = Evidence Review Group	

1.6 Other key issues: summary of the ERG's view

The company claimed that the end-of-life criteria were met. However, the ERG's analysis did not find that the first criterion (less than 24 months survival) was met, and that there was considerable uncertainty regarding the survival benefit of oral azacitidine (see Sections 2, and 7).

1.7 Summary of the ERG's view

The updated CS base-case probabilistic and deterministic ICERs were £48,332 and £48,660 per QALY gained, respectively. For the FLT3 subgroup, midostaurin was dominated by oral azacitidine, and the probabilistic ICER for oral azacitidine versus watch and wait plus BSC in this subgroup was £25,403 per QALY gained. The estimated ERG base-case ICER (probabilistic), based on the ERG preferred assumptions highlighted in Section 6.1, was £52,731 per QALY gained. The most influential adjustment was using the consolidation subgroup instead of the intention to treat (ITT) population. The ICER increased most in the scenario analysis assuming a post-hematopoietic stem cell transplantation (HSCT) utility increment for the proportion of patients treated with HSCT. For the FLT3 subgroup, midostaurin was dominated by oral azacitidine in the ERG base-case and the probabilistic ICER for oral azacitidine versus watch and wait plus BSC was £25,275 per QALY gained.

There is large remaining uncertainty about the effectiveness and cost effectiveness of oral azacitidine, which can be partly resolved by the company by conducting further analyses. The appropriate number of cycles of consolidation therapy in UK clinical practice and the most appropriate curves for the modelling of OS and RFS in the consolidation subgroup are unknown. In addition, the current approaches (both in the CS and ERG base-case) to reflect HSCT in the modelling and to incorporate HRQoL are likely biased. Results of the FLT3 subgroup are likely biased and updated baseline patient characteristics reflective of this subgroup are required, as well as a detailed description of survival analyses. Therefore, the ERG believes that neither the CS nor the ERG report contains an unbiased ICER of oral azacitidine compared with relevant comparators.

Table 1.16: Summary of ERG's preferred assumptions and ICER

Technologies	Total costs (£)	Total QALYs	Incremental costs (£)	Incremental QALYs	ICER (£/QALY)
CS deterministic base-case					
Oral azacitidine	[redacted]	[redacted]			
w&w+BSC	[redacted]	[redacted]	[redacted]	[redacted]	48,660
Matter of judgement (1-consolidation subgroup)					
Oral azacitidine	[redacted]	[redacted]			
w&w+BSC	[redacted]	[redacted]	[redacted]	[redacted]	53,574
Matter of judgement (2-Relapse utility based on Tremblay)					
Oral azacitidine	[redacted]	[redacted]			
w&w+BSC	[redacted]	[redacted]	[redacted]	[redacted]	47,478
Matter of judgement (3-no temporary disutility for HSCT)					
Oral azacitidine	[redacted]	[redacted]			
w&w+BSC	[redacted]	[redacted]	[redacted]	[redacted]	48,729

Technologies	Total costs (£)	Total QALYs	Incremental costs (£)	Incremental QALYs	ICER (£/QALY)
Deterministic ERG base-case					
Oral azacitidine	[redacted]	[redacted]			
w&w+BSC	[redacted]	[redacted]	[redacted]	[redacted]	53,291
Probabilistic ERG base-case					
Oral azacitidine	[redacted]	[redacted]			
w&w+BSC	[redacted]	[redacted]	[redacted]	[redacted]	52,731

BSC = best supportive care; CS = Company Submission; ERG = Evidence Review Group; HSCT = hematopoietic stem cell transplantation; ICER = incremental cost effectiveness ratio; QALYs = quality adjusted life years

2. CRITIQUE OF COMPANY'S DEFINITION OF DECISION PROBLEM

Table 2.1: Statement of the decision problem (as presented by the company)

	Final scope issued by NICE	Decision problem addressed in the CS	Rationale if different from the final NICE scope	ERG Comment
Population	Adults with AML who have complete disease remission, or complete remission with incomplete blood count recovery, following induction therapy with or without consolidation treatment who are not eligible for, including those who choose not to proceed to HSCT	As per final scope	N/A	The population is in line with the NICE scope
Intervention	Oral azacitidine as maintenance treatment	As per final scope	N/A	The intervention is in line with the NICE scope
Comparator(s)	Midostaurin Established clinical management without oral azacitidine (which may include a watch and wait strategy with BSC, low dose cytarabine or subcutaneous azacitidine)	Midostaurin Established clinical management without oral azacitidine (which may include a "watch and wait" strategy with BSC)	Low dose cytarabine and subcutaneous azacitidine are not used in clinical practice as maintenance treatments for AML in the population eligible for maintenance treatment with oral azacitidine (as confirmed by two UK AML treating clinicians) and are therefore not considered as comparators to oral azacitidine (further detail provided in Section B.1.1)	• The ERG does not understand why the company did not use the comparators described in the final NICE scope. • Low dose cytarabine and subcutaneous azacitidine may be legitimate active comparators and they are mentioned in the final NICE scope. • The company does not provide evidence upon

	Final scope issued by NICE	Decision problem addressed in the CS	Rationale if different from the final NICE scope	ERG Comment
				which the expert opinion was based. • Participants with FLT3 would have received midostaurin in routine practice
Outcomes	The outcome measures to be considered include: • overall survival • relapse free survival • adverse effects of treatment • health-related quality of life	As per final scope	N/A	The population is in line with the NICE scope
Economic analysis	The reference case stipulates that the cost effectiveness of treatments should be expressed in terms of incremental cost per QALY. The reference case stipulates that the time horizon for estimating clinical and cost effectiveness should be sufficiently long to reflect any differences in costs or outcomes between the technologies being compared.	As per final scope	As per final scope	Economic analysis is in line with the reference case

	Final scope issued by NICE	Decision problem addressed in the CS	Rationale if different from the final NICE scope	ERG Comment
	Costs will be considered from an NHS and PSS perspective. The availability of any commercial arrangements for the intervention, comparator and subsequent treatment technologies will be taken into account. The availability of any managed access arrangement for the intervention will be taken into account.			
Subgroups to be considered	None	N/A	N/A	N/A
Special considerations including issues related to equity or equality	The availability and cost of biosimilar and generic products should be taken into account. Guidance will only be issued in accordance with the marketing authorisation. Where the wording of the therapeutic indication does not include specific treatment combinations, guidance will be issued only in the context of the evidence that has underpinned the marketing authorisation granted by the regulator.	None identified	N/A	No issues related to equity or equality were raised

	Final scope issued by NICE	Decision problem addressed in the CS	Rationale if different from the final NICE scope	ERG Comment
Based on Table 1 and pages 10 to 12 of the CS[1] AML = acute myeloid leukaemia; BSC = best supportive care; CS = Company Submission; DCIS = ductal carcinoma in situ; ERG = Evidence Review Group; HSCT = haematopoietic stem cell transplantation; N/A = not applicable; NHS = National Health Service; NICE = National Institute of Health and Care Excellence; PSS = Personal Social Services; QALY = quality adjusted life year; UK = United Kingdom				

2.1 Population

The population defined in the scope is: Adults with acute myeloid leukaemia who have complete disease remission, or complete remission with incomplete blood count recovery, following induction therapy with or without consolidation treatment who are not eligible for, including those who choose not to proceed to, haematopoietic stem cell transplantation.[2] The population in the Company Submission (CS) is the same as the population defined in the final National Institute for Health and Care Excellence (NICE) scope.[1]

The population considered in the CS matches the population in the main clinical trial for azacitidine in this indication, the QUAZAR trial.[3, 4]

In 2020, the United States Food and Drug Administration (FDA) approved oral azacitidine as maintenance therapy for adults with acute myeloid leukaemia (AML) in first remission.[5] In 2021 the European Medicines Agency (EMA) approved oral azacitidine for adult patients with AML who achieved complete remission (CR) or CR with incomplete blood count recovery (CRi) following induction therapy with or without consolidation treatment and who are not candidates for, including those who choose not to proceed to, hematopoietic stem cell transplantation (HSCT).[6, 7] Additionally, NICE recommends a consolidation phase lasting "several months."[8]

2.2 Intervention

The intervention (oral azacitidine as maintenance treatment) is in line with the scope.

Azacitidine is a DNA methyltransferase inhibitor and epigenetic modifier. Oral azacitidine is administered according to the following standard dose: 300 mg azacitidine orally per day. Each repeated cycle consists of a treatment period of 14 days followed by a treatment free period of 14 days (28-day treatment cycle). Patients are to be treated with an anti-emetic 30 minutes prior to each dose of oral azacitidine for the first 2 treatment cycles and may be omitted after 2 cycles if there is no nausea and vomiting (see summary of product characteristics (SmPC) for further details). Treatment should be discontinued if more than 15% blasts are observed in peripheral blood or bone marrow, or if unacceptable toxicity arises. The company notes that a dose schedule modification for AML disease relapse, with 5% to 15% blasts in peripheral blood or bone marrow, in conjunction with a clinical assessment, an extension of the dosing schedule from 14 to 21 days of repeated 28-day cycles should be considered. Dosing should not exceed 21 days during any 28-day period. Oral azacitidine should be discontinued if more than 15% blasts are observed in either the peripheral blood or bone marrow or at the physician's discretion. Dose interruption and/or dose reduction (to 200 mg) for haematologic and non-haematologic adverse reactions are recommended based on clinical and laboratory findings (see SmPC Section 4.2 for further detail).

2.3 Comparators

The description of the comparators in the NICE scope is as follows: "*Midostaurin, Established clinical management without oral azacitidine (which may include a watch and wait strategy with best supportive care, low dose cytarabine or subcutaneous azacitidine)*".[2] The company used the following comparators: "*Midostaurin, established clinical management without oral azacitidine (which may include a "watch and wait" strategy with best supportive care)*."[1] That is, the company did not include low dose cytarabine or subcutaneous azacitidine as comparators. The company's rationale for using different comparators from those described in the final NICE scope was:

"*Low-dose cytarabine and subcutaneous azacitidine are not used in clinical practice as maintenance treatments for AML in the population eligible for maintenance treatment with oral azacitidine (as confirmed by two UK AML treating clinicians) and are therefore not considered as comparators to oral azacitidine*"[1]

With respect to low-dose cytarabine and subcutaneous azacitidine as comparators listed in the final NICE scope yet that were not considered by the company, the company noted the following in their response to request for clarification: "*BMS do not consider them to be "best supportive care" but rather examples of active treatments that target the underlying leukaemia*."[9]

The company also consulted with two United Kingdom (UK) AML clinical experts who, according to the company, "*unequivocally confirmed that these treatments [midostaurin, low dose subcutaneous azacitidine] are not used in UK clinical practice for AML maintenance. The clinical experts could only provide very limited examples where these treatments could be used in situations resembling maintenance treatment, such as those patients whose disease was in partial remission, or patients who showed signs of early relapse. We believe that these situations might be miscategorised as maintenance treatment.*"[1]

In addition, in the final NICE scope, a number of technology appraisals have been published for treating people with untreated disease, and name treatments including cytarabine (NICE Technology Appraisal Guidance 552,[10] NICE Technology Appraisal Guidance 545),[11] (option for adults who are not eligible for HSCT and have AML with 20-30% blasts and multilineage dysplasia according to the World Health Organization (WHO) classification) subcutaneous azacitidine (NICE Technology Appraisal Guidance 218).[12] At least some of these guidelines note that the treatments (including cytarabine) are recommended for consolidation therapy. For example, NICE Guidance 552 state: "*For consolidation (5 to 8 weeks after the start of the last induction): daunorubicin 29 mg/m^2 and cytarabine 65 mg/m^2 on days 1 and 3. A subsequent course of consolidation may be given when there is no disease progression or unacceptable toxicity.*")[10]

Moreover, the NICE Pathways document for myeloid leukaemia makes the following recommendation for AML: "*Midostaurin is recommended, within its marketing authorisation, as an option in adults for treating newly diagnosed acute FLT3-mutation-positive myeloid leukaemia with standard daunorubicin and cytarabine as induction therapy, with high-dose cytarabine as consolidation therapy, and alone after complete response as maintenance therapy. It is recommended only if the company provides midostaurin with the discount agreed in the patient access scheme*."[13]

ERG comments:

- The Evidence Review Group (ERG) does not understand why the company did not use the comparators described in the final NICE scope.
- The claim that low dose cytarabine and subcutaneous azacitidine may be active comparators does not impinge on the fact that they are potentially legitimate comparators and that they are mentioned in the final NICE scope.
- While NICE guidance implies that consolidation therapy is standard practice, the number of cycles of consolidation therapy is not clear.
- The company does not provide evidence upon which the expert opinion was based.

2.3.1 Placebo comparators

The company used a placebo comparator whereas placebo was not listed as a comparator in the final NICE scope.[2] In response to the ERGs request to justify this, the company noted the following: "*At the time of study design of the QUAZAR AML-001 trial (2011), no therapies were approved in multiple regions for use in the AML maintenance setting. Furthermore, there is currently no standard of care for maintenance therapy in AML, and in routine clinical practice, many patients are unlikely to receive further active treatment after achieving remission. Therefore, placebo was determined to be the appropriate comparator for oral azacitidine in the QUAZAR AML 001 trial and its selection was agreed upon with regulatory agencies (eg, the US FDA; this was an FDA special protocol assessment trial)*."[9] In other places, the company seems to suggest that there is standard of care:

- The company also states, in response to the same request for clarification, that "*BMS considers the "watch and wait" comparator to be represented by the placebo arm within the QUAZAR study, with best supportive care common to both randomised treatment groups*"[9]
- In their response to another question, the company notes that "*Both experts confirmed that consolidation therapy is standard of care in UK clinical practice*" This is a more sensible approach to determining standard of care (SoC).[9]

ERG comments:

- The ERG notes that the final NICE scope specifies that watch and wait is an acceptable strategy with best supportive care (BSC).
- The ERG notes an ambiguity in the company's view on whether BSC was established or not.
- The ERG does not understand why the company used placebo as a comparator when it was not included in the final NICE scope.
- The ERG acknowledges that the placebo comparator can be considered as a "watch and wait" strategy, however the final NICE scope states that watch and wait must be accompanied by "BSC, low dose cytarabine or subcutaneous azacitidine."[2]

2.3.2 Comparators for the FLT3 subgroup

On page 15 of the CS, the company states that NICE recommends midostaurin for patients with fms-like tyrosine kinase 3 (FLT3).[1] A recent clinical trial also demonstrated a benefit of midostaurin as maintenance therapy in AML with FLT3.[14] The company also notes, on page 16 of the CS that "*[a]pproximately 25% of AML patients are FLT3-mutation-positive and approximately 30-40% of will achieve first remission. A majority of these patients go on to receive HSCT leaving approximately 10% of these patients who are likely to have midostaurin maintenance in the UK (as confirmed by UK clinicians)*."[1]

On page 37 of their submission, the company noted that "*FLT3-ITD/TKD mutations appeared to confer a negative prognosis in the placebo arm, but this was not apparent in the oral azacitidine arm (data cut-off date, 15 July 2019)*."[1]

In section B.2.7. of their submission, the company reports doing a multivariate analysis which confirmed the independent prognostic impact of FLT3. They report that oral azacitidine also significantly improved overall survival (OS) independent of FLT3 mutation status (hazard ratio (HR) 0.72; p=0.003) and that FLT3-ITD/TKD mutations at diagnosis appeared to have a negative prognostic influence in the placebo arm.

ERG comment:

- It appears that patients with FLT3 were included in the QUAZAR trial yet did not receive midostaurin. For patients with FLT3, this may have led to unfavourable outcomes for this subgroup, compared with what would have happened in routine practice. For the same reason, the benefits of oral azacitidine would have been exaggerated compared to what would have been the case if the control group received routine care.

While the company reports the OS and RFS independent of FLT3 status, this is not done for other efficacy or safety outcomes 2.4 Outcomes

The NICE final scope lists the following outcome measures:

- overall survival (OS)
- relapse-free survival (RFS)
- adverse effects of treatment
- health-related quality of life (HRQoL)

These were all assessed in the QUAZAR trial.

2.5 Other relevant factors

According to the company, oral azacitidine is innovative because most patients with AML experience disease relapse after induction chemotherapy, so effective maintenance treatment for patients who attain remission may play a role in preventing disease relapse and prolonging OS. Also, according to the company, oral azacitidine addresses a substantial unmet need for a well-tolerated and easily administered AML maintenance treatment that significantly prolongs survival among patients with AML who are in remission after IC, without compromising HRQoL.[1]

According to the company, this appraisal fulfils the end-of-life criteria as specified by NICE because the median survival of patients in the placebo **plus BSC** group in the QUAZAR trial was 14.8 months, which is lower than 24 months, and because oral azacitidine plus BSC prolonged life by 9.9 months compared with placebo plus BSC. However, the ERG's analysis did not find that oral azacitidine meets the end-of-life criteria (see Section 7).

The company also states that no equality issues related to the use of azacitidine for the maintenance treatment of adults with AML have been identified or are foreseen.[1]

ERG comment:

- The ERG has raised several problems with the evidence of effectiveness (see Key Issues). There is therefore uncertainty regarding the estimates upon which the claims that the end-of-life criteria.
- The ERG does not consider there to be any equality issues related to the use of oral azacitidine for the maintenance treatment of adults with AML.

3. CLINICAL EFFECTIVENESS

3.1 Critique of the methods of review(s)

The company performed a systematic literature review (SLR) to identify and summarise the available randomised controlled trial (RCT) evidence for maintenance treatment options for adult patients (≥18 years) with AML who have achieved CR/CRi and are not candidates for stem cell transplant (SCT).

3.1.1 Searches

Appendix D of the CS provided details of the systematic literature searches used to identify efficacy and safety data for maintenance treatments for patients with AML. Searches were originally conducted in January 2020 and updated twice, first in February 2021 and again in June 2021. The ERG has presented only the major limitations of each search strategy in the report. A summary of the resources searched are provided in Table 3.1.

Table 3.1: Resources searched for clinical efficacy and safety. Jan 2020, Feb 2021 & June 2021.

Search strategy element	Resource	Host/Source	Date range	Date searched
Databases	Embase	Ovid	RCTs 2005-date of search SRs 2015-date of search	Original: 18.1.20 Updated: 19.2.21 & 11.6.21
	MEDLINE & MEDLINE In-Process			
	CENTRAL			
	CDSR			
	DARE			18.1.20*
	HTA			
Trials registries	ClinicalTrials.gov	Internet		Original: 18.1.20 Updated: 19.2.21 & 11.6.21
Conference Proceedings	ASCO	Internet	2020-2021	Original: 18.1.20 Updated: 19.2.21 & 11.6.21
	ASH		2020-2021	Original: 18.1.20 Updated: 19.2.21 & 11.6.21
	EBMT		2020-2021	Original: 18.1.20 Updated: 19.2.21 & 11.6.21
	EHA		2020-2021	Original: 18.1.20 Updated: 19.2.21 & 11.6.21

	SOHO		2020-2021	Original: 18.1.20 Updated: 19.2.21 & 11.6.21
Additional searches	FDA Database	Internet		Original: 18.1.20 Updated: 19.2.21 & 11.6.21
	Bibliographies of relevant systematic review articles checked for additional relevant references			
CENTRAL = Cochrane Central Register of Controlled Trials; CDSR = Cochrane Database of Systematic Reviews; DARE = Database of Abstracts of Reviews of Effects; HTA = Health Technology Assessment Database; ASCO = American Society of Clinical Oncology; ASH = American Society of Hematology; EBMT = European Society for Blood and Marrow Transplantation; EHA = European Hematology Association; SOHO = Society of Hematologic Oncology; SRs = systematic reviews; RCTs = randomised controlled trials; FDA = Food and Drug Administration *No updates required as no new records have been added to DARE/HTA since the original searches were run				

ERG comment:

- Searches were reported for a good range of resources, including one trials registry, five conference proceedings and the Food and Drug Administration (FDA) website. Initially only the latest iteration of each search was provided, but after a request at clarification all strategies for both the original and updated Ovid searches were provided and these searches were clearly structured and well documented.
- The strategy for the January 2020 combined Ovid search provided at clarification, also contained a search of the two Centre for Reviews and Dissemination (CRD) databases, Database of Abstracts of Reviews of Effects (DARE) and Health Technology Assessment Database (HTA). These resources were not previously mentioned in the CS. Given their archival nature these were appropriately removed from the later updates
- A single search strategy was provided for the three Ovid searches: MEDLINE, Embase, Cochrane Library (Cochrane Database of Systematic Reviews (CDSR) and Cochrane Central Register of Controlled Trials (CENTRAL)) and for the original search the two CRD databases (DARE and HTA). The strategy was split by resource and combined both free text and the appropriate subject headings and field tags for each database. Whilst correctly combined and reproducible, given its length a more reader friendly approach may have been to split the search into four separate searches.
- In addition to the searches listed above, the bibliographies of relevant systematic review articles checked for additional relevant references.
- Searches were structured to combine terms for the condition AML and where appropriate, a trials filter including a date restriction of 2005-present, and a systematic reviews filter restricted to results published between 2015-present. The date limits were introduced based on clinician experience and the literature, the CS stated that "…*maintenance therapies for the population of interest were deemed to be a more recent introduction to the AML space with any studies predating 2005 to be highly unlikely or not reflective of current clinical practice*".[1]
- The ERG queried whether any separate adverse event (AE) searches were performed. The company responded that no additional searches had been run to identify safety data associated with maintenance treatments in the population of interest, other than those stated in the clinical SLR (Appendix D). Guidance by CRD[15] recommends that if searches have been limited by a study design filter, additional searches should be undertaken to ensure that AEs that are long-term, rare or unanticipated are not missed. Whilst relevant observational studies may have been identified by the

SRs retrieved as part of the search reported in Appendix D, as they are listed in the exclusion criteria (Appendix D, Table B.5.3.) it is unclear what impact this may have had on the review.

- The ERG noted that the search reported for ClinicalTrials.gov, did not reflect a proper search syntax (including any limits used) for this resource (Appendix D, Table B.5.2.). It also appears to report the number of included studies arising from the search rather than the number of records retrieved. However, the number retrieved (n=1,222) is provided in the PRISMA flowchart (Figure B.5.1.). Whilst the ERG is unable to fully critique the complete search strategy due to the lack of transparency and reproducibility, the numbers reported in the flow chart combined with the keywords provided in Table B.5.2. and the searches reported in Ovid, suggest that it is unlikely that any key studies would have been missed.
- The reporting of the number of included studies, rather than hits recalled in Table B.5.2. is also continued for the other grey literature searches, as with ClinicalTrials.gov full numbers retrieved are provided in the PRISMA flow chart (FDA (n=27) and conference proceedings (n=2,881)).

3.1.2 Inclusion criteria

As stated above, a SLR was conducted to identify and summarise the relevant evidence. Study eligibility criteria are presented in table 3.2.

Table 3.2: Eligibility criteria used in search strategy for RCT and non-RCT evidence

	Inclusion criteria	Exclusion criteria

Population	• Male and female adults (≥18 years)
	• Histologically confirmed de novo AML or AML secondary to prior myelodysplastic disease
	  o Receiving maintenance treatment after first CR/CRi (following induction with intensive chemotherapy, with or without consolidation, in 1L)
	  o SCT ineligible at CR/CRi
	Intermediate/poor cytogenetic risk or favourable-risk cytogenetics
		• Patients <18 years
		• Relapsed or refractory AML
		• Prior bone marrow or STC
		• Ineligible for intensive induction chemotherapy at 1L
		• Achieved CR/CRi following therapy with hypomethylating agents or prior therapy with hypomethylating agents for MDS within 4 months of developing AML

Interventions	• Oral azacitidine (ONUREG®)	

	Inclusion criteria	Exclusion criteria
Comparators	• BSC (e.g., hydroxyurea) • Azacitidine (IV, SC, oral) • Decitabine • LDAC (cytarabine) • Idarubicin • Daunorubicin • Gemtuzumab ozogamicin • Venetoclax+decitabine/azacitidine/cytarabine • Glasdegib • Enasidenib (IDH2) • Ivosidenib (IDH1) • Sorafenib (FLT3) • midostaurin (FLT3) • Immunotherapies (BCG vaccination, IFN-a, IL-2) • 6-Mercaptopurine • Ceplene + IL-2 • Tipifarnib • Desatinib • Lenalidomide • Quizartinib • rhIL-11 • Lomustine • Methotrexate • Norethandrolone • OCV-501 (WT1 peptide vaccine) • Lirilumab • ATRA • Histamine dihydrochloride and IL-2 • Thioguanine, cyclophosphamide • FLT3L • Nivolumab	• SCT • High Intensity Therapies: • "7+3" • 7+3+midostaurin • 7+3+gemtuzumab ozogamicin • 7+3+cladribine • 7+3+mitoxantrone • HiDAC (cytarabine) • HiDAC+midostaurin • Vyxeos (daunorubicin+cytarabine) • Therapies for R/R AML: • FLAG-IDA, MEC, Gilteritinib
Outcomes	• Effectiveness: • OS • RFS/event-free survival/disease-free survival/PFS • Time to relapse from CR/CRi • Time to discontinuation from treatment • Safety/tolerability: • Any AEs (e.g., neutropenia, infections) • Treatment-related adverse events • SAEs • Withdrawals due to AEs • Patient-reported outcomes (FACIT-Fatigue, EQ-5D)	• Studies that do not report any relevant outcomes

	Inclusion criteria	Exclusion criteria
Study design	• RCTs in any country (phases II, III & II/III) Systematic reviews and meta-analyses of RCTs (included at the title and abstract stage only)	• Non-randomised, single-arm, or observational studies • Open-label extension phases of RCTs • Pre-clinical studies, case reports, expert opinion articles, letters, narrative (non-systematic) reviews • Phase I Pilot studies
Language restrictions	Articles in English	All non-English articles

Based on Table B**.5.3** appendix D, CS.[1]
AEs = adverse events; CS = Company Submission; 1L = first line; 7 + 3 = cytarabine + daunorubicin; AML = acute myeloid leukaemia; ATRA = all-trans retinoic acid; BCG = Bacille Calmette-Guérin; BSC = best supportive care; CR(i) = complete remission (with incomplete platelet recovery); EQ-5D = EuroQol-5 Dimension; FACIT = Functional Assessment of Chronic Illness Therapy; FLAG-IDA = fludarabine-cytarabine-filgrastim-idarubicin; FLT3(L) = FMS-like tyrosine kinase 3 (ligand); HiDAC = High Dose Ara-C; IDH 1(2) = isocitrate dehydrogenase 1(2); IFN-a = interferon alpha; IL-2(11) = interleukin-2(11); IV = intravenous; MEC = mitoxantrone, etoposide, and cytarabine; MDS = myelodysplastic syndrome; LDAC = Low Dose Ara-C; OS = overall survival; PFS = progression-free survival; RCT = randomised controlled trial; RFS = relapse-free survival; R/R = relapsed/refractory; SAEs = serious adverse events; SC = subcutaneous; SCT = stem cell transplant; WT1 = Wilms tumor gene 1.
Note: BSC is shaded grey as it is not an active comparator.

3.1.3 Critique of data extraction

The company states that *"All studies meeting the inclusion criteria were extracted into a pre-specified data extraction template by one reviewer. The extractions were validated by a second reviewer against the original publication."*[16] Best practice would be to have two independent reviewers extract data in duplicate. The company do not provide sufficient information in the CS or the associated appendices to determine that the two reviewers completed these tasks independently. Furthermore, the company failed to provide any detail on how disagreements were resolved. By failing to conduct independent data extraction and to resolve disagreements either by consensus or by a third independent adjudicating reviewer, data is more prone to error and bias.

3.1.4 Quality assessment

According to the CS, quality assessment was conducted using the Cochrane risk of bias assessment tool for randomised controlled trials. It is stated that two reviewers conducted the assessments independently and then after comparing evaluations, consensus was reached. This represents an accepted process for conducting quality assessments, although the use of a third independent reviewer to adjudicate disagreements would have represented the optimal process. Appendix D of the CS contains quality assessments for both the QUAZAR AML-001 and RATIFY trials.[16]

3.1.5 Evidence synthesis

The company conducted a SLR to identify and summarise the available RCT evidence for maintenance treatment options for adult patients (≥18 years) with AML who have achieved CR/CRi and are not candidates for SCT. The literature search was originally performed on 18th January 2020 and updated

twice, on 19th February 2021 and on 11th June 2021.[16] Because the company only used one trial for their analysis, they did not conduct a meta-analysis.[1]

The QUAZAR AML-001 provided data for outcomes, namely OS, RFS, adverse effects of treatment and HRQoL. Two separate data cut-off points were considered in the reporting of results. The CS clarifies that the primary analysis for OS and analyses for all other outcomes was based on the primary data cut-off on 15th July 2019 however a further data cut-off on 8th September 2020 provided longer follow-up and more mature survival data. Data is presented from the 15th July 2019 cut-off except for OS data which are presented for both the July 2019 and September 2020 analyses. As all efficacy data supporting the use of oral azacitidine for the treatment of AML are provided solely by the QUAZAR AML-001 trial, no meta-analysis was conducted.

3.2 Critique of trials of the technology of interest, their analysis and interpretation (and any standard meta-analyses of these)

This Section of the report details the sources of evidence in the CS for the clinical effectiveness of azacitidine.

Section B.2.2 of the CS[1] identified one ongoing, Phase III, RCT, QUAZAR AML-001 (NCT01757535), which is the sole source of data on the clinical effectiveness of azacitidine in patients with AML in CR/Cri after intensive chemotherapy (IC) with or without consolidation chemotherapy, and were not candidates for HSCT, as per the NICE final scope.

3.2.1 Details of the included trial: QUAZAR AML-001 trial

The QUAZAR AML-001 trial (n=472) is an ongoing, international, multicentre, Phase 3, randomised, two-arm, double-blind, placebo-controlled, parallel group study of oral azacitidine plus BSC versus placebo plus BSC as maintenance treatment, consisting of four phases: pre-randomisation phase (screening phase within 28 days prior to randomisation), randomisation and double-blind treatment phase, follow-up phase and extension Phase (EP) (See Figure 3.1 for the study design's overview).

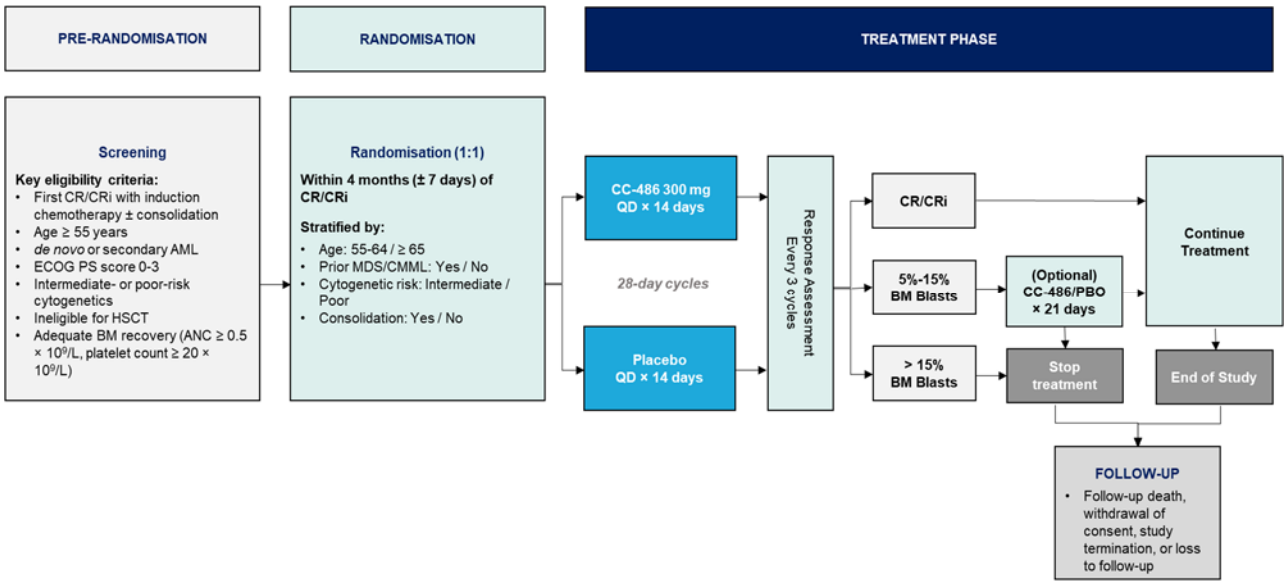
All screening procedures to confirm AML diagnosis, verify CR/CRi status, etc were conducted 28 days before randomisation. Eligible patients were then randomised 1:1 to receive 300 mg oral azacitidine daily or matching placebo for the first 14 days of each 28-day cycle. The study stipulated that randomisation must have occurred within 4 months (± 7 days) of achieving the first CR/CRi status to allow time for eligible patients to receive (and recover from) up to four cycles of consolidation chemotherapy, and to avoid enriching the study population with patients with better prognoses.[1] Following randomisation, crossover between treatment arms was permitted. Study assessments were conducted during the treatment phase and AML relapse was not considered an AE for the purposes of the safety analysis.[1] Patients with subsequent evidence of AML relapse (≥5% and ≤15% blasts in the peripheral blood or bone marrow) had the option to continue treatment with an extended dose schedule to 300 mg per day (QD) for 21 days, provided it was in the best interest of the patient to do so as judged by the investigator.[1] Table 3.3 further describes the QUAZAR AML-001 study methodology.

In the follow-up phase, the AEs of patients were collected up to 28 days after the last dose of study treatment, and patients were followed up for survival every month for the first year, and then every three months until death.[4] The EP was due to protocol amendment 2.[4] Subjects who demonstrated clinical benefit from receiving oral azacitidine and gave consent, were permitted to continue to receive the study drug after unblinding, and were followed for at least another 12 months, until death, withdrawal of consent, study closure, or were lost to follow-up.[4] For patients randomised into the

placebo group, the clinical study report (CSR) reports that "*upon consent, any subject who was discontinued from the Treatment Phase (irrespective of randomisation group), could enter the EP and be followed for survival (without receiving azacitidine).*"[4]

Table B.2.4. of the CS details the key inclusion and exclusion criteria for patients included in the QUAZAR AML-001 study. They included adults aged ≥55 years who were newly diagnosed, histologically confirmed de novo acute AML or AML secondary to myelodysplastic syndromes (MDS) or chronic myelomonocytic leukaemia (CMML), had undergone induction with IC with or without consolidation therapy, and had either achieved first CR or CRi status within 4 months (± 7 days) prior to randomisation.[1]

Figure 3.1: Study design, QUAZAR AML-001 (NCT01757535) study



Source: Figure B.2.1 of the CS[1]
AML = acute myeloid leukaemia; ANC = absolute neutrophil count; BM = bone marrow; CC-486 = azacitidine; CMML = chronic myelomonocytic leukaemia; CR = complete remission; CRi = complete remission with incomplete blood count recovery; ECOG PS = Eastern Cooperative Oncology Group performance status; HSCT = haematopoietic stem cell transplant; IWG = International Working Group; MDS = myelodysplastic syndromes; QD = once daily.

Table 3.3: Summary of study methodology, QUAZAR AML-001

Study design	International, multicentre, Phase 3, randomised, two-arm, double-blind, placebo-controlled, parallel group
Study objective	Primary objective • Evaluate whether maintenance treatment with oral azacitidine improved OS compared with placebo Secondary objectives: • Determine the effect of oral azacitidine on RFS, safety and tolerability, HRQoL and HCRU (hospitalisations, medications, clinic visits, medical/diagnostic procedures, and treatment for AEs)
Locations	Conducted at 148 sites in 23 countries (Australia, Austria, Belgium, Brazil, Canada, Czech Republic, Finland, France, Germany, Ireland, Israel, Italy, Korea, Lithuania, Mexico, Poland, Portugal, Russia, Spain, Taiwan, Turkey, United Kingdom [N=8], and the United States)

Study status	• Ongoing First patient, first visit: 10th May 2013 • First data cut-off date: 15th July 2019 (all outcomes) • Second data cut-off date: 8th September 2020 (OS only)
Study treatments	Treatment was assigned by a central randomisation procedure using an IVRS. Patients were randomly assigned in a 1:1 ratio to receive: • oral azacitidine tablets 300 mg QD (N=238) or • matching placebo (N=234)
Blinding	Patients, investigators, site staff and clinical and medical personnel were unaware of treatment assignments until study closure and database lock for the primary analysis (data cut-off 15th July 2019). The EP until the most recent data cut (8th September 2020) was unblinded.
Concomitant medication(s)	Permitted: • BSC (including, but not limited to RBC and platelet transfusions, use of an ESA, antibiotic, antiviral, and antifungal therapy, nutritional support, and G-CSFs for patients experiencing neutropenic infections, pre-treatment or post-treatment with a serotonin (5-HT3) receptor antagonist (or other anti-emetic medication) Disallowed: • Cytotoxic chemotherapeutic agents or experimental agents, romiplostim and other TSAs (e.g., interleukin-11), hydroxyurea, lenalidomide, pomalidomide, thalidomide, arsenic trioxide, interferon and retinoids
Study outcome(s)	Primary outcome: • OS Secondary outcomes: • RFS • Time to relapse from CR/CRi • Time to discontinuation from treatment • HRQoL assessment (FACIT-Fatigue Scale, EQ-5D-3L) • HCRU assessment
Pre-planned subgroups	Additional pre-planned exploratory subgroup analyses were performed where an adequate number of patients were available to allow meaningful interpretation. Analyses were performed within the following subgroups for the OS and RFS outcomes: • Age (<65, ≥65, ≥75 years) • Gender • Race • CR/CRi status at randomisation and at first achieving response • CR/CRi status at randomisation and use of consolidation • Prior history of MDS or CMML • Cytogenic risk category at induction therapy • MRD status at screening (prior to randomisation) • CR/CRi status at randomisation and MRD status at screening • Consolidation therapy following induction • Geographic region • ECOG PS • WHO AML classification • Type of first line subsequent therapy
Based on Table B.2.3. of the CS[1]	

AEs = adverse events; AML = acute myeloid leukaemia; ANC = absolute neutrophil count; BSC = best supportive care; CMML = chronic myelomonocytic leukaemia; CR = complete remission; CRi = complete remission with incomplete blood count recovery; CS = Company Submission; ECOG PS = Eastern Cooperative Oncology Group performance status; EP = extended phase; HRQoL – health related quality of life; IC = intensive chemotherapy; MDS = myelodysplastic syndromes; OS = overall survival; RFS = relapse-free survival; SCT = stem cell transplantation; WHO = World Health Organization

ERG comment:

- As the CS stated that, *"patients with subsequent evidence of AML relapse (≥5% and ≤15% blasts in the peripheral blood or bone marrow) had the option to continue treatment with an extended dose schedule to 300 mg QD for 21 days, provided it was in the best interest of the patient to do so as judged by the INV,"*[1] the ERG asked the company to justify the choice of 21 days and provide details about how the investigators made their judgement about what was in the best interest of the patient. In their response to clarification, concerning the choice of 21 days, the company stated that, *"the proposed mechanism of action of oral azacitidine is to expose potential leukemic cells extended drug exposure over the treatment cycle. Both 14 and 21 days of treatment with oral azacitidine were studied in initial Phase I/II studies in patients with AML/MDS/CMML,"* whilst in response to investigator judgement, the company stated that, *"in part this would depend on the overall condition of the patient and their performance status, the ability to and ease of obtaining other therapy options for the patient, and the preferred option of the patient."*[9]
- The CS reports that the QUAZAR AML-001 trial was designed to allow for up to 4 cycles of consolidation chemotherapy, yet no patient received 4 cycles of consolidation and only 4% of patients received 3 cycles of consolidation therapy following induction therapy, with most who did receive consolidation, receiving only 1 cycle of treatment (See Table 3.4). The company also stated that they raised an enquiry on the number of cycles of consolidation chemotherapy to be given after achievement of remission with two UK clinical experts, one of which was an investigator for the trial and both experts estimated that [redacted] of the UK transplant ineligible AML population would receive 4 to 5 cycles of consolidation whilst [redacted] would be expected to receive 3 cycles of treatment (see Table 3.5).[9] The ERG would like to further query why then the trial was designed to account for the possibility of 4 cycles of consolidation treatment if it is expected that this will not be seen in clinical practice. Further discussions on the possible effect of consolidation treatments on survival will be discussed in Section 3.2.5.7 of this report. The company's clinical experts also estimated that between [redacted] to [redacted] of patients in this population would not be expected to receive any consolidation treatment following induction.[9] The ERG retains its misgivings concerning these estimations and recommends that independent clinical expert advice may be necessary to determine what the role of consolidation treatment following induction before maintenance treatment with oral azacitidine would be in UK clinical practice.
- In its clarification letter, the ERG also asked the company to clarify if the recovery time following consolidation, before maintenance therapy with oral azacitidine in the trial, would be representative of UK clinical practice. In their response to clarification, the company stated that the decision to start maintenance therapy with oral azacitidine within 120 days or 4 months of achieving CR/CRi with intensive induction chemotherapy was made following feedback from a UK clinician.[9] They stated that, *"each consolidation cycle was estimated to take around 6 weeks, and the expert explained that patients would start maintenance therapy as soon as blood count recovery was achieved which would approximately be around 30-35 days after the last consolidation cycle. Both experts stated that the 4-months timeframe of the QUAZAR AML-001 trial would typically allow*

patients to receive 2 cycles of consolidation, including sufficient recovery time before maintenance treatment with oral azacitidine started."[9]

- The NICE final scope in Table 2.1 clearly defines the population as interest as those, *"who are not eligible for, including those who choose not to proceed to, haematopoietic stem cell transplantation."*[4] The ERG noted that 10% (n=47) of patients in the trial received subsequent HSCT, with 13.7% (n=32) of patients on the placebo arm and 6.3% (n=15) of patients on the oral azacitidine arm, thus the ERG asked the company to clarify what the company meant by 'subsequent'.[1] In its response to clarification, the company stated that, *"all 47 patients from the QUAZAR trial who received a subsequent HSCT, did so after treatment discontinuation. The majority of them (41) had relapsed on study drug (32 on placebo and 9 on oral azacitidine) and received HSCT as salvage therapy. The remaining 6 patients, all of which were in the oral azacitidine arm, were transplanted while still in CR1."*[9] The ERG had serious concerns about why six patients still in first remission received HSCT as a curative therapy, especially as the company in its response to clarification states that, *"multiple factors play a role to determine the HSCT eligibility of an AML patient in first CR. It is recommended to take a decision individually for each patient based on an assessment of their risk of relapse if treated with chemotherapy alone, in the context of the mortality risk associated with transplant… it is not uncommon for transplant eligibility of a patient to change over time,"*[9] and if they continued in the study or discontinued due to now being 'transplant eligible'. The company in their response to clarification stated that, *"in 6 patients, representing 2.5% of patients randomised to receive oral azacitidine, a decision was made after randomisation by the treating clinician to proceed with transplant in CR1. The sponsor considers this to be reflective of real-world practice, and to restrict this potentially curative option from the patient pathway to be an unethical scenario. The protocol required that treatment with oral azacitidine or placebo ceased in the event of subsequent AML therapy, including transplant, however the patients were followed up for survival."*[9] The ERG were thus sceptical about if there would be a survival benefit due to HSCT (and has been explored in Section 3.2.5.1 of this report) and if it has been captured appropriately in the model (discussed in Section 4.2.2 of this report). In conclusion, although transplant ineligibility was a trial inclusion criterion and 10% of patients in the trial received subsequent HSCT, the extent to which this would reflect in clinical practice is what necessitates further discussion and perhaps the input of an independent UK clinical expert.

Table 3.4: Baseline characteristics - consolidation therapies received, QUAZAR AML-001 trial

Parameter	Total (N=472)
Yes	378 (80)
1 Cycle	212 (45)
2 Cycles	147 (31)
3 Cycles	19 (4)
4 Cycles	0
No	94 (20)

Based on Table 9 of the CL response[9]
CL = clarification letter; N = number

Table 3.5: Estimated proportions of patients receiving consolidation therapy in UK clinical practice

Number of consolidation cycles	Proportion of patients receiving number (N) of consolidation cycles in UK clinical practice	
	Advisor 1	Advisor 2
Patients receiving consolidation	[redacted]	[redacted]
N = 1	[redacted]	[redacted]
N = 2	[redacted]	[redacted]
N = 3	[redacted]	[redacted]
N = 4	[redacted]	[redacted]
N = 5	[redacted]	[redacted]
N = 0	[redacted]	[redacted]
Based on Table 8 of the CL response[9] N = number; UK = United Kingdom		

3.2.2 Statistical analyses of the QUAZAR AML-001 trial

Statistical analyses in the QUAZAR AML-001 trial have been summarised in Table 3.6.

Table 3.6: Summary of statistical analyses, QUAZAR AML-001 trial

Hypothesis objective	The null hypothesis for testing the primary efficacy outcome is that the OS distributions for oral azacitidine and placebo are equivalent
Statistical analysis for key outcomes[a]	Primary outcome (OS) • KM methods were used to estimate the survival distribution functions for each treatment group • Survival distributions were compared using a stratified log-rank test, stratifying by age at time of induction therapy, prior history of MDS, cytogenic risk category, received consolidation therapy following induction therapy • A stratified Cox proportional hazards model was used to estimate the HR ratio with interaction terms of treatment and time and with a p-value of 0.006 • CIs for survival estimates at 6 months, 1-year and 2-years were calculated with Greenwood's variance formula • A sequential gate-keeping approach was used to control the overall type 1 error in order to perform hypothesis testing on multiple outcomes, OS was tested first at the two-side 0.05 significance level • Other than the pre-specified sequential testing of OS and RFS, no additional alpha adjustments for multiplicity were made Key secondary outcome • RFS was analysed using the same methods as those for OS • To preserve the overall alpha level at 0.05 across the OS and RFS outcomes, formal statistical inference for the RFS analyses can only be made if superiority of oral azacitidine is demonstrated for OS, at the two-sided 0.05 significance level

Sample size, power calculation	• The equality of the OS curves were compared between the oral azacitidine and placebo treatment groups using a stratified log-rank test • Assuming a median OS of 16 months in the placebo group, a median OS of 22.9 months in the oral azacitidine group (43% improvement), and a study duration of 60 months with a drop-out rate of 5% from both treatment groups, over the duration of the study, the design requires 330 deaths and approximately 460 patients (230 per treatment group) to be randomised in order achieve at least 90% power to detect a constant HR of 0.70 and demonstrate a statistically significant difference in OS • It was assumed that the OS distribution was exponential with a constant failure (hazard) rate and that accrual was non-uniform during an accrual period of 36 months with 25% of the patients accrued during each of the first 2 years of enrolment (50% accrued at 24 months) and the remaining 50% accrued during the last year of enrolment • Sample size calculations were based on a one-sided alpha of 0.025 with one interim analysis for futility after 30% of the events have occurred
Data management, patient withdrawals	**Missing data** • Missing individual data were generally treated as missing and no values were imputed. **Discontinuations** • Patients who discontinued study treatment for any reason were to undergo end-of-treatment procedures • Additionally, all discontinued patients were followed for 28 days following the last dose of study treatment or until the date of the last study visit (whichever was longer) for AEs • After the follow-up visit, patients were followed for survival by telephone, every month for the first year and then every 3 months until death, withdrawal of consent for further follow-up, study end, or until the patient was lost to follow-up • Discontinued patients were not replaced

Based on Table B.2.9. of the CS[1]
AE = adverse event; ANCOVA = analysis of covariance; CI = confidence interval; CMH = Cochran-Mantel-Haenszel; EQ-5D-3L = European Quality of Life 5 dimensions, 3 levels; FACIT = Functional Assessment of Chronic Illness Therapy; HCRU = healthcare resource utilisation; HR = hazard ratio; HRQoL = health-related quality of life; ITT = intention-to-treat; KM = Kaplan-Meier; MDS = myelodysplastic syndromes; MID = minimally important difference; MMRM = mixed model repeated measures; OS = overall survival; RFS = relapse-free survival
[a] Analysis performed using the ITT population

ERG comment: The statistical analysis of the key outcomes appeared appropriate. Several protocol amendments were made (Table 3 of the CSR[4]) and the reduction of number of clinical visits in a cycle from 2 (days 1 and 15) to 1, with the day 15 visit being optional and at the discretion of the investigator, beginning from cycle 25, might not have been appropriate in capturing the HRQoL outcome and AEs as mentioned in Key Issue 4.

3.2.3 Baseline characteristics of the QUAZAR AML-001 trial

Table 3.7 summarises the key baseline demographic and disease characteristics of patients in the QUAZAR AML-001 study. The majority of patients in the study were in the age range ≥65 to <75 years (61%), male (52%), White (87.5%), from study centres in Europe (67%), had a de novo AML diagnosis (91%), had an Eastern Cooperative Oncology Group (ECOG) performance of either Grade 0 (48%) or Grade 1 (44%), were categorised as being of intermediate cytogenetic risk (86%), had a negative measurable residual disease (MRD) status at randomisation (52%), were HSCT ineligible due to their age (65%), and received at least one subsequent AML therapy (65%).

Table 3.7: Key patient baseline characteristics, QUAZAR AML-001 trial (ITT population)

Baseline characteristic	Oral azacitidine (N=238)	Placebo (N=234)
Age (years)		
Median (range)	68 (55–86)	68 (55–82)
Age category, n (%)		
≥55 to <65 years	66 (28)	68 (29)
≥65 to <75 years	144 (61)	142 (61)
≥75 years	27 (11)	24 (10)
≥85 years	1 (0)	0
Sex, n (%)		
Male	118 (50)	127 (54)
Female	120 (50)	107 (46)
Type of AML, n (%)		
Primary (de novo)	213 (89)	216 (92)
Secondary	25 (11)	18 (8)
Time since original AML diagnosis (months) to randomisation		
Median (range)	4.2 (1.5–9.2)	4.2 (1.4–10.9)
Time from CR/CRi to randomization, days		
Median (range)	84 (7–154)	86 (7–263)
ECOG performance status, n (%)		
Grade 0	116 (49)	111 (47)
Grade 1	101 (42)	106 (45)
Grade 2–3	21 (9)	17 (7)
Cytogenetic risk category defined by NCCN at diagnosis, n (%)		
Intermediate	203 (85)	203 (87)
Poor	35 (15)	31 (13)
Reason ineligible for transplant[a], n (%)		
Age	154 (65)	152 (65)
Comorbidities	52 (22)	50 (21)
Performance Status	14 (6)	9 (4)

Baseline characteristic	Oral azacitidine (N=238)	Placebo (N=234)
Not acceptable or available donor	37 (15.5)	35 (15.0)
Patient decision	19 (8)	32 (14)
Unfavourable cytogenetics	6 (3)	10 (4)
Other	28 (12)	21 (9)
Received subsequent HSCT	15 (6.3)	32 (13.7)
Received consolidation therapy following induction therapy, n (%)		
Yes	186 (78)	192 (82)
1 Cycle	110 (46)	102 (44)
2 Cycles	70 (29)	77 (33)
3 Cycles	6 (2.5)	13 (6)
4 Cycles	0	0
No	52 (22)	42 (18)
CR/CRi status at randomisation[b], n (%)		
CR	183 (77)	177 (76)
CRi	50 (21)	44 (19)
Not in CR/CRi	5 (2)	11 (5)
Missing	0	2 (1.0)
Time from start of induction therapy to randomisation, months		
Median (range)	4.0 (1.4–8.8)	4.0 (1.3–15.1)
ATC Dictionary Level Preferred named[c], n (%)		
Subjects with at least one subsequent AML therapy	137 (57.6)	170 (72.6)
Intensive chemotherapy	69 (29)	88 (38)
Low intensity therapy	[redacted]	[redacted]
Other	[redacted]	[redacted]
Missing	[redacted]	[redacted]
Subsequent AML therapies reported for ≥ 10% of subjects in either treatment group, n (%)		
Antineoplastic and immunomodulating agents	[redacted]	[redacted]
Cytarabine	83 (34.9)	92 (39.3)
Fludarabine	32 (13.4)	48 (20.5)
Azacitidine	31 (13.0)	47 (20.1)
Hydroxycarbamide	28 (11.8)	34 (14.5)
Idarubicin	20 (8.4)	33 (14.1)

Based on Tables B.2.6. and B.2.7. of the CS[1]
AML = acute myeloid leukaemia; ANC = absolute neutrophil count; ATC= Anatomical Therapeutic Chemical; BM = bone marrow; CMML = chronic myelomonocytic leukaemia; CR = complete remission; CRi = complete remission with incomplete blood count recovery; CS = Company Submission; ECOG = Eastern Cooperative Oncology Group; HSCT = haematopoietic stem cell transplant; IWG = International Working Group; MDS =

Baseline characteristic	Oral azacitidine (N=238)	Placebo (N=234)
myelodysplastic syndrome; MRD = measurable residual disease; max = maximum; min = minimum; n = number of patients in the category; N = number of patients evaluable; NCCN = National Comprehensive Cancer Network; SD = standard deviation; WHO = World Health Organisation. [a]A patient may have had more than one reason [b]CR/CRi at randomisation was programmatically derived based on IWG for AML response criteria using BM data collected during screening, and ANC and platelets closest to randomisation date. For a patient with BM blasts <5%, and both ANC <1.0 x 109/L and platelet count <100 x 109/L, the patient was considered not in CR/CRi [c]Coded using WHO Drug Dictionary version March 2019. A subject with multiple occurrences of a drug class or drug preferred name is counted only once in the specific ATC classification or preferred name, respectively **Note**: time interval in days was calculated as the difference between the randomisation date and the date of interest (e.g., date of original AML diagnosis) plus one day. Time interval presented in months is transformed from days to months by using the conversion formula: months = days/30.4375.		

ERG comment:

- The ERG does not believe that the baseline characteristics of the QUAZAR AML-001 population sufficiently describes a patient population unfit to receive HSCT. In the trial, 89% of patients were <75 years and 92% had an ECOG performance status of grade 0-1. The ERG queried the company on the parameters that disqualified patients in the trial from being eligible for HSCT, and in their response to clarification, the company stated that, *"the population under investigation within QUAZAR-AML-001 represents a cohort that were deemed fit enough to undergo intensive induction chemotherapy, and that achieved CR/CRi. This may be reflected in the lower proportion of patients recruited with ECOG 2-3 performance status compared to ECOG 0-1."*[9] A table was provided to demonstrate the reasons for transplant ineligibilities which has already been summarised in Table 3.7, but as further information was not given, the ERG's concerns on the eligibility criteria persist.
- Due to the high percentage of patients receiving subsequent AML therapies, 72.6% on the placebo arm and 52.6% on the oral azacitidine arm, the ERG asked for clarification on if these therapies were administered during the maintenance treatment period and what those therapies were, to which in response the company stated, *"subsequent AML therapies did not occur with concomitant AML maintenance treatment in the QUAZAR-AML-001 study,"* and provided a table of those therapies. The ERG infers that these subsequent therapies would have been administered post study discontinuation. On the placebo arm [redacted] and [redacted] of patients received subsequent intensive chemotherapy and low-intensity chemotherapy, respectively.[9] While [redacted] and [redacted] of patients on the oral azacitidine arm received subsequent intensive chemotherapy and low-intensity chemotherapy, respectively.[9] The ERG also queried the company for evidence that low-dose cytarabine and subcutaneous azacitidine are not part of BSC, the company in their response stated that, *"Low dose cytarabine, and subcutaneous azacitidine… have licensed indications to treat acute myeloid leukaemia… BMS do not consider them to be "best supportive care" but rather examples of active treatments that target the underlying leukaemia… The Pan-London AML guidelines refer to low dose cytarabine or azacitidine as potential treatment options for patients who are not fit for intensive chemotherapy. Similarly, the ELN 2017 guidelines refers to both treatments as 'selected conventional care regimens' for patients not considered to be candidates for intensive chemotherapy. This guidance also clearly distinguishes them from 'best supportive care',*

describing BSC as follows: "…for patients who cannot tolerate any antileukemic therapy, or who do not wish any therapy. In the QUAZAR phase 3 study, subsequent therapies received by patients included both low dose cytarabine and subcutaneous azacitidine (along with other chemotherapeutic regimens and HSCT)."[9] In the table of subsequent AML therapies provided by the company in the response to clarification, [redacted] of patients received cytarabine whilst [redacted] of patients received azacitidine.[9] It is unclear if these were low-dose and subcutaneous therapies, respectively. Section 3.2.1 of this report has addressed the issue of the potential survival benefit of subsequent therapies.

3.2.3.1. Representativeness of UK clinical practice

In its clarification letter, the ERG asked the company to discuss the generalisability of the study's population to what is expected to be the UK's AML transplant ineligible population. The company stated that, *"BMS considers that the baseline disease demographics of the QUAZAR-AML-001 trial, which includes a majority of patients from Europe (65%), align to the UK's AML (ineligible for HSCT following induction) population with some caveats."*[9] Caveats of which were age (the trial limited its patient population to patients of age to ≥ 55 years), and cytogenetic risk (the study included patients with intermediate and poor cytogenetics whereas patients with favourable risk cytogenetics are less likely to proceed to HSCT in first CR).[9] The company thus provided tables of demographic and disease characteristics across the trial's ITT population, (displayed in Table 3.8) EU-subgroup, UK-subgroup, and UK clinical practice (as estimated by the company's two UK clinical experts) which has been displayed in Tables 3.8 and 3.9.

The ERG earlier expressed its concerns about 92% of the trial's ITT population having an ECOG performance status of grade 0-1, which is similarly duplicated across the trial's EU-subgroup ([redacted]), and UK-subgroup ([redacted]). However, the company's two UK clinical experts estimated that of the percentage of AML patients in UK clinical practice who would be transplant ineligible, [redacted] would also have an ECOG grade of 0-1. The ERG has some concerns about whether these estimates have accurately captured what is expected to be seen in clinical practice for the patient population who would receive oral azacitidine as maintenance therapy. Also notable is that [redacted] of patients in the UK-subgroup were transplant ineligible because there was no appropriate/acceptable donor, this is quite startling in comparison with the ITT population ([redacted]), EU-subgroup, ([redacted]) and what has been estimated would be UK clinical practice ([redacted]). Additionally, [redacted] patient in the UK-subgroup received 3-4 cycles of consolidation therapy with most in this subgroup receiving just 1 cycle ([redacted]) whereas the company's UK clinical experts have estimated that [redacted] of transplant ineligible AML patients in UK clinical practice would receive 2 cycles of consolidation treatment. Granted that the trial's UK-subgroup has a small number of patients, the ERG still has serious concerns about if the UK-subgroup of patients in the QUAZAR AML-001 trial can be expected to be representative of the UK clinical practice. The ERG however felt that the EU-subgroup would be more in line with what is expected to be seen in UK clinical practice.

Table 3.8: Demographic characteristics across QUAZAR AML-001 trial ITT, EU-subgroup, UK-subgroup, and UK clinical practice

Parameter	ITT (N=472)	EU (N=[redacted])	UK ([redacted])	UK clinical practice[a]
Age (years)				
Mean (SD)	67.9 (5.66)	[redacted]	[redacted]	[redacted]
Median (min, max)	68.0 (55, 86)	[redacted]	[redacted]	[redacted]

Age Category – n (%)				
≥ 55 to < 65 years	134 (28.4)	[redacted]	[redacted]	[redacted]
≥ 65 to < 75 years	286 (60.6)	[redacted]	[redacted]	[redacted]
≥ 75 years	52 (11.0)	[redacted]	[redacted]	[redacted]
≥ 85 years	1 (0.2)	[redacted]	[redacted]	[redacted]
Sex – n (%)				
Male	245 (51.9)	[redacted]	[redacted]	[redacted]
Female	227 (48.1)	[redacted]	[redacted]	[redacted]
Race – n (%)				
White	413 (87.5)	[redacted]	[redacted]	[redacted]

Based on Table 18 of CL response[9]

AML = acute myeloid leukemia; CL = clarification letter; EU = European Union; ITT = intention to treat; UK = United Kingdom

[a]Data presented was provided by one of the two UK clinical experts in AML. The other expert stated that the range of percentages across the ITT, EU subgroup and UK subgroup seem representative for the UK.

Table 3.9: Disease baseline characteristics across QUAZAR AML-001 trial ITT, EU-subgroup, UK-subgroup, and UK clinical practice

Parameter	ITT (N=472)	EU ([redacted])	UKf ([redacted])	UK clinical practice[a]
Initial AML classification, n (%)				
AML with recurrent genetic abnormalities	85 (18)	[redacted]	[redacted]	[redacted]
ML with myelodysplasia - related changes	91 (19)	[redacted]	[redacted]	[redacted]
Therapy-related myeloid neoplasms	2 (0.4)	[redacted]	[redacted]	[redacted]
AML not otherwise specified	293 (62)	[redacted]	[redacted]	[redacted]
Missing	1 (0.2)	[redacted]	[redacted]	[redacted]
Type of AML, n (%)				
Primary (de novo)	429 (91)	[redacted]	[redacted]	[redacted]
Secondary	43 (9)	[redacted]	[redacted]	[redacted]
Time since original AML diagnosis (months) to randomisation				
Median (range)	4.2 (1.4–10.9)	[redacted]	[redacted]	[redacted]
Prior history of MDS/CMML, n (%)				
Primary	37 (8)	[redacted]	[redacted]	[redacted]
Secondary	0	[redacted]	[redacted]	[redacted]
Missing	2 (0.4)	[redacted]	[redacted]	[redacted]
ECOG performance status, n (%)				
Grade 0	227 (48)	[redacted]	[redacted]	[redacted]
Grade 1	207 (44)	[redacted]	[redacted]	[redacted]
Grade 2–3	38 (8)	[redacted]	[redacted]	[redacted]
Cytogenetic risk category defined by NCCN at diagnosis, n (%)				

Parameter	ITT (N=472)	EU ([redacted])	UKf ([redacted])	UK clinical practice[a]
Intermediate	406 (86)	[redacted]	[redacted]	[redacted]
Poor	66 (14)	[redacted]	[redacted]	[redacted]
Reason ineligible for transplant[b], n (%)				
Age	306 (65)	[redacted]	[redacted]	[redacted]
Comorbidities	102 (22)	[redacted]	[redacted]	[redacted]
Performance Status	23 (5)	[redacted]	[redacted]	[redacted]
Not acceptable or available donor	72 (15)	[redacted]	[redacted]	[redacted]
Patient decision	51 (11)	[redacted]	[redacted]	[redacted]
Unfavourable cytogenetics	16 (3)	[redacted]	[redacted]	[redacted]
Other	49 (10)	[redacted]	[redacted]	[redacted]
Received consolidation therapy following induction therapy, n (%)				
Yes	378 (80)	[redacted]	[redacted]	[redacted]
1 Cycle	212 (45)	[redacted]	[redacted]	[redacted]
2 Cycles	147 (31)	[redacted]	[redacted]	[redacted]
3 Cycles	19 (4)	[redacted]	[redacted]	[redacted]
4 Cycles	0	[redacted]	[redacted]	[redacted]
No	94 (20)	[redacted]	[redacted]	[redacted]
MRD status at randomisation[a], n (%)				
Negative	244 (52)	[redacted]	[redacted]	[redacted]
Positive	219 (46)	[redacted]	[redacted]	[redacted]
Missing	9 (2)	[redacted]	[redacted]	[redacted]
Response achieved after induction therapy (with or without consolidation therapy), n (%)				
CR	384 (81)	[redacted]	[redacted]	[redacted]
CRi	88 (19)	[redacted]	[redacted]	[redacted]
CR/CRi status at randomization, n (%)				
CR	360 (76)	[redacted]	[redacted]	[redacted]
CRi	94 (20)	[redacted]	[redacted]	[redacted]
Not in CR/CRi	16 (3)	[redacted]	[redacted]	[redacted]
Missing	2 (0.4)	[redacted]	[redacted]	[redacted]
Time from start of induction therapy to randomisation, months				
Median (range)	4.0 (1.3–15.1)	[redacted]	[redacted]	[redacted]
Time from induction therapy to first achieving CR/CRi, days				
Median (range)	35.0 (13.0–455.0)	[redacted]	[redacted]	[redacted]
Time since first achieving CR/CRi to randomisation, days				
Median (range)	85.0 (7.0–263.0)	[redacted]	[redacted]	[redacted]
Bone marrow blasts, %				
Median (range)	2.0 (0.0–6.5)	[redacted]	[redacted]	[redacted]
Peripheral blood blasts, %				
Median (range)	0.0 (0.0–2.0)	[redacted]	[redacted]	[redacted]
Based on Table 17 of the CL response[9]				

Parameter	ITT (N=472)	EU ([redacted])	UKf ([redacted])	UK clinical practice[a]
AML = acute myeloid leukemia; CL = clarification letter; CMML = chronic myelomonocytic leukaemia; CR = complete remission; CRi = complete remission with incomplete blood count recovery; ECOG PS = Eastern Cooperative Oncology Group performance status; EU = European Union; ITT = intention to treat; MDS = myelodysplastic syndromes; NCCN = National Comprehensive Cancer Network; UK = United Kingdom [a]The presented percentages were provided by two UK clinical experts in AML. When the experts provided different numbers, these were summarised as a range. E.g., one expert said 50% and the other stated 40%, then this is reported as 40-50%. [b]Note: Only one expert provided feedback on the percentages.				

3.2.4 Risk of bias assessment of the QUAZAR AML-001 trial

Table 3.10 details the risk of bias assessment (RoB) of the QUAZAR AML-001 study conducted by the company using the seven-item NICE quality assessment checklist.[17]As stated in Section 3.1.4 of this report, the company clarified that two reviewers independently conducted quality assessments and after comparing assessments and resolved any disagreements by consensus.

ERG comment: The ERG undertook a RoB assessment of the QUAZAR AML-001 trial using the same criteria, based on Wei et al., 2020[18] and mostly agreed with the company's assessment except that it was unclear if the concealment of treatment allocation continued until unblinding in the EP as the paper only states that, *"patients, investigators, study site staff and Celgene clinical and medical personnel were unaware of treatment assignments."*

Table 3.10: Quality assessment of the QUAZAR AML-001 trial, NICE checklist

Question	CS	ERG
Was randomisation carried out appropriately?	Yes	Yes.
Was the concealment of treatment allocation adequate?	Yes	Unclear. Insufficient details on processes or methods provided
Were the groups similar at the outset of the study in terms of prognostic factors?	Yes, baseline characteristics were balanced between treatment arms	Yes
Were the care providers, participants and outcome assessors blind to treatment allocation?	Yes	Yes
Were there any unexpected imbalances in dropouts between groups?	No; based on patient disposition	No
Is there any evidence to suggest that the authors measured more outcomes than they reported?	No; all primary and secondary trial endpoints were reported.	No
Did the analysis include an intention-to-treat analysis? If so, was this appropriate and were appropriate methods used to account for missing data?	Yes	Yes
Based on Table B.5.20. of Appendix D[16] CS = Company Submission; ERG = Evidence Review Group		

3.2.5 Efficacy results from the QUAZAR AML-001 trial

The company submitted efficacy analyses for primary and secondary endpoints based on the primary data base lock of 15th July 2019 cut-off point when 329 events occurred to allow for a fully powered OS analysis. A further data cut-off point of 8th September 2020 providing longer follow-up and more mature survival data, has also been used to present OS data.

3.2.5.1 Primary outcome: overall survival (OS)

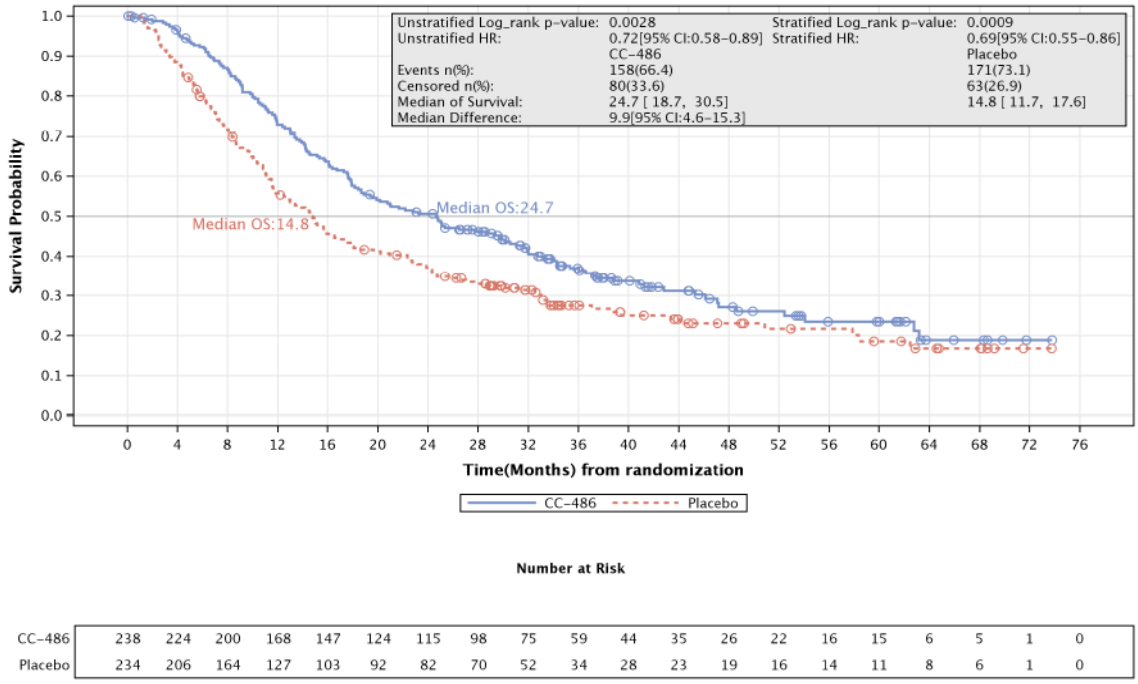
The company defined OS as, *"the number of days from the date of randomisation until the date of death from any cause, calculated as (date of death – date of randomisation + 1). Patients surviving at the end of the follow-up period or who were lost to follow-up were censored at the date last known to be alive. For patients who withdrew consent, the last date known alive was considered the date of consent withdrawal from the study."*[1] In the QUAZAR AML-001 trial ITT population, at the primary database lock (15th July 2019, 41.2 months), oral azacitidine is reported to have a significantly longer OS when compared to placebo, and as demonstrated in Figure 3.2, at a median OS difference of 9.9 months (median OS: 24.7 months for oral azacitidine versus 14.8 months for placebo; stratified HR: 0.69 (95% confidence interval (CI) 0.55-0.86), p<0.001). At the more recent 8th September 2020 data cut-off point (51.7 months), oral azacitidine continued to be associated with a significantly longer OS when compared to placebo, maintaining the median OS difference of 9.9 months (See Figure 3.3). One year after randomisation, survival rates in the oral azacitidine arm were reportedly higher when compared to the placebo arm (72.8% versus 55.8%; difference 17.0 percentage points (95% CI: 8.4-25.6)). A summary of OS findings in the ITT population has been tabulated in Tables 3.11 and 3.12.

ERG comment:

- The OS data collected during the EP at the 8th September 2020 cut-off was used to inform the cost effectiveness model. During the EP, the two study arms were unblinded and in its response to clarification, the company stated that *"any patient who provided additional consent could enter the EP and was followed-up for survival for at least another 12 months until death, withdrawal of consent, study closure or loss to follow-up… crossover from the placebo group to the oral azacitidine group was not permitted in the EP."*[9] The ERG agrees with the company's statements on this cut-off point being able to provide more mature and robust data when compared to the primary database lock. Additionally, as patients were censored more frequently for the OS analysis using the primary database lock, ("because they were alive without documented relapse at time of data-cut-off."[9]) the EP cut-off point is more liable to confirm the sustained benefit of oral azacitidine over placebo.
- The ERG had some concerns about if OS is variable to prior consolidation therapy and has explored this possibility in Section 3.2.5.7 of this report. As time goes by, the proportion of patients surviving in both arms comes close to each other as observed in Figure 3.2 at around 48 months and overlaps at 64 months and onwards, this can also be seen in the 3-year survival time point in Table 3.11 (72.8% in the azacitidine arm compared to 55.8% in the placebo arm at 1-year, 50.6% versus 37.1% at 2-years, and [redacted] versus [redacted] survival estimate by year 3). From the Kaplan-Meier (KM) curve in Figure 3.2, one could conclude that these observations allude to the survival benefit of azacitidine rapidly diminishing after 3 years, when compared to placebo.
- As stated in Section 3.2.1, the ERG was concerned about if the 10% of patients in the trial who had received subsequent salvage and curative HSCT would experience improved survival benefit. The OS results from the sensitivity analyses censoring for patients who received any subsequent AML therapy showed that the OS HR is consistent with the ITT population (see Table 3.11). Thus, if the

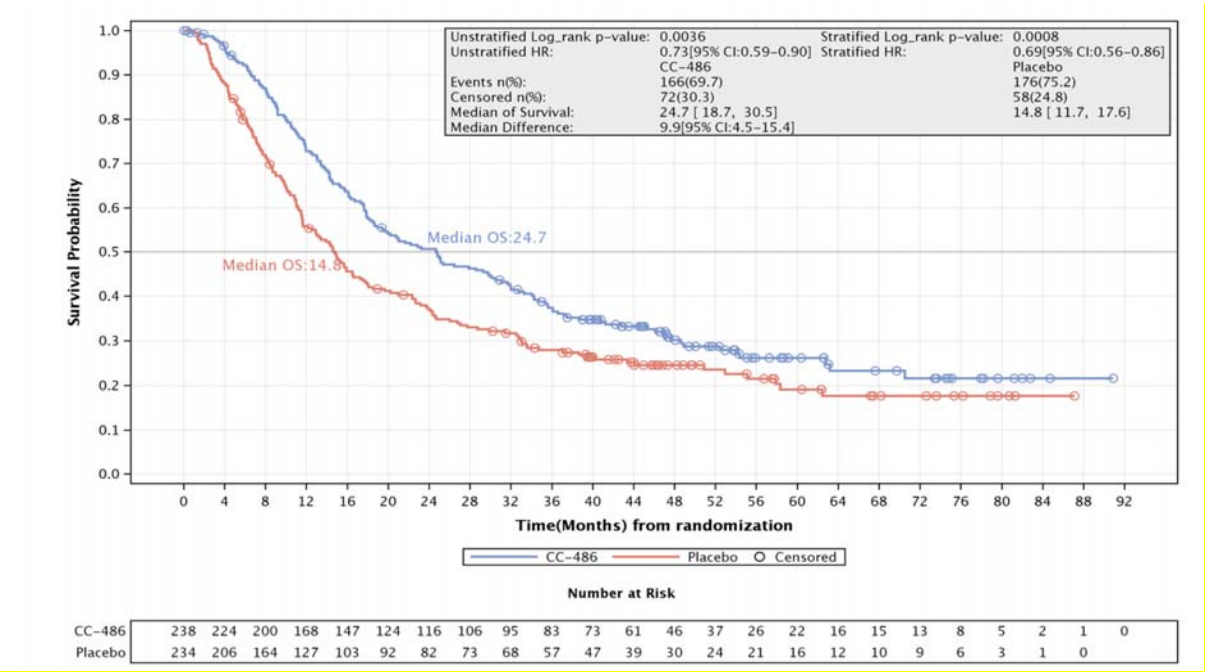
10% of patients had not received subsequent HSCT, their prospects of OS would be much lower. As stated in Section 3.2.1, if receiving HSCT as salvage or curative therapies following oral azacitidine use would be expected to be seen in clinical practice for this patient population, then the survival curve for the ITT population would be expected to be indicative of the UK clinical practice population.

Figure 3.2: KM analysis of OS data, QUAZAR AML-001 trial (ITT population) (data cut-off point, 15 July 2019)



Source: Figure B.2.3. of the CS[1]
CC-486 = oral azacitidine; CI = confidence interval; CS = Company Submission; HR = hazard ratio; ITT = intention-to-treat; KM = Kaplan-Meier; OS = overall survival

Figure 3.3: KM analysis of OS data, QUAZAR AML-001 trial (ITT population) (data cut-off point, 8 September 2020)



Source: Figure B.2.4. of the CS[1]
CC-486 = oral azacitidine; CI = confidence interval; CS = Company Submission; HR = hazard ratio; ITT = intention-to-treat; KM = Kaplan-Meier; OS = overall survival.

Table 3.11: Summary of OS, QUAZAR AML-001 trial (ITT population)

Endpoint	Oral azacitidine (N=238)	Placebo (N=234)	Difference (95% CI)
15th July 2019			
Patients with event (death), n (%)	158 (66.4)	171 (73.1)	-
Patients censored, n (%)	80 (33.6)	63 (26.9)	-
Median OS, months (95% CI)[a]	24.7 (18.7-30.5)	14.8 (11.7-17.6)	9.9 (4.6-15.3)
HR (95% CI)[b]	0.69 (0.55, 0.86)[e]		-
p-value[c]	0.0009		-
1-year survival estimate (95% CI)[d]	0.728 [redacted]	0.558 [redacted]	0.170 (0.084-0.256)
2-year survival estimate (95% CI)[d]	0.506 [redacted]	0.371 [redacted]	0.135 (0.045-0.225)
8th September 2020			
Patients with event (death), n (%)	[redacted]	[redacted]	[redacted]
Patients censored, n (%)	[redacted]	[redacted]	[redacted]
Median OS, months (95% CI)[a]	24.7 (18.7-30.5)	14.8 (11.7-17.6)	9.9 (4.5-15.4)
HR (95% CI)[b]	0.69 (0.56-0.86)	-	
p-value[c]	0.0008	-	

Endpoint	Oral azacitidine (N=238)	Placebo (N=234)	Difference (95% CI)
1-year survival estimate (95% CI)[d]	[redacted]	[redacted]	[redacted]
2-year survival estimate (95% CI)[d]	[redacted]	[redacted]	[redacted]
3-year survival estimate (95% CI)[d]	[redacted]	[redacted]	[redacted]

Based on Table B.2.11. of the CS[1]
CI = confidence interval; CS = Company Submission; HR = hazard ratio; ITT = intention-to-treat; KM = Kaplan–Meier; n = number of patients in the category; N = number of patients evaluable; OS = overall survival.
[a]Median estimate of OS was derived using the KM method. Difference was calculated as oral minus placebo. The CI for the difference was derived using Kosorok's method.
[b]The HR is from a Cox proportional hazards model stratified by age, cytogenetic risk category, and receipt of consolidation therapy or not.
[c]The p-value is two-sided from a log-rank test stratified by age, cytogenetic risk category, and receipt of consolidation therapy or not.
[d]KM methods were used to estimate the one-year, two-year and three-year survival probabilities. The CIs for the difference in the one-year and two-year survival probabilities were derived using Greenwood's variance estimate.
[e]HRs were not provided in the primary publication, as the proportional hazards assumption appeared to be violated, as indicated by the significant treatment-by-time interaction.

Note: percentages are based on the number of patients in each treatment group, unless otherwise specified.

Table 3.12: Pre-specified sensitivity analyses of overall survival for subsequent therapies, QUAZAR AML-001 trial

Sensitivity analysis	HR[a] (oral azacitidine versus placebo)	95% CI for HR	P value[a]
Censored for all subsequent therapy (including post treatment transplant)	[redacted]	[redacted]	[redacted]
Censored for post-treatment transplant	[redacted]	[redacted]	[redacted]

Based on Table 17 of the CSR[4]
CI = confidence interval; CSR = clinical study report; HR = hazard ratio
[a]The HR is from a Cox proportional hazards model stratified by age, cytogenetic risk category and received consolidation therapy or not. The nominal p-value is 2-sided from a log-rank test, stratified as described for the HR.

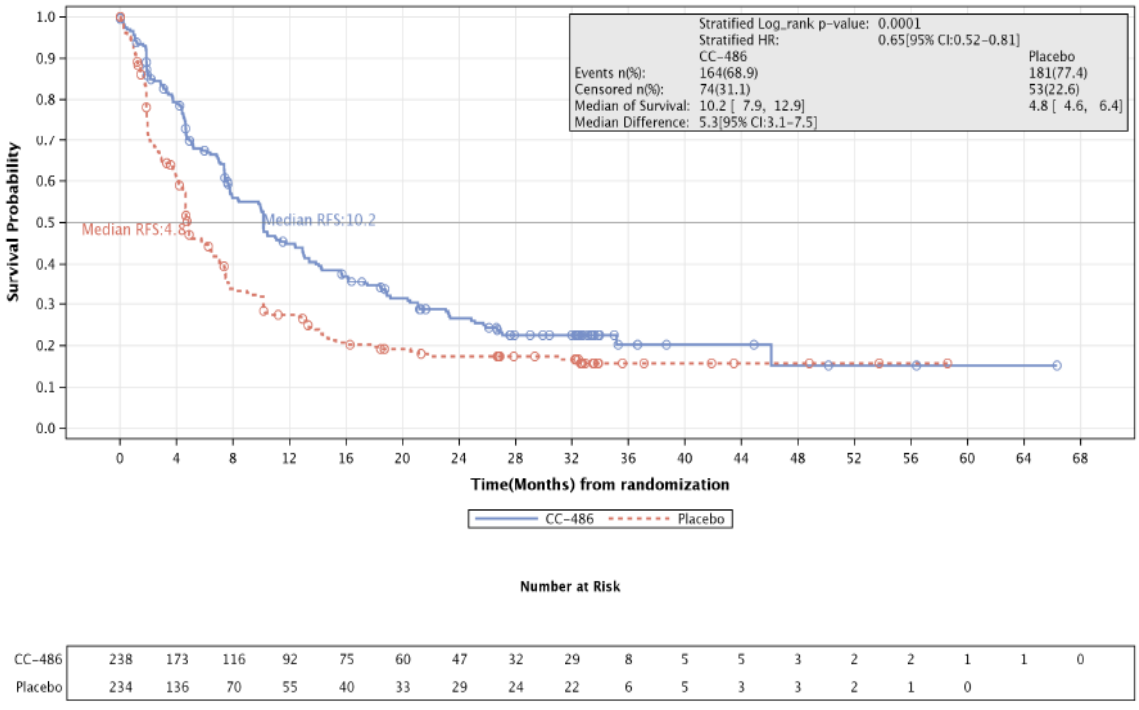
3.2.5.2 Secondary outcome: relapse-free survival (RFS)

A summary of RFS findings in the ITT population has been tabulated in Table 3.13. The company defined RFS as, *"the time from the date of randomisation to the date of documented relapse or death, whichever occurred first. Patients who were still alive without documented relapse, or who were lost to follow-up or withdrew consent without documented relapse, were censored at the date of their last response assessment."*[1] The CSR states that *"documented relapse was defined as the earliest of ≥ 5% bone marrow blasts from the central pathology report, appearance of blasts in the peripheral blood with confirmation of bone marrow blasts ≥ 5% within 100 days (i.e., approximately 3 cycles), or at least 2 peripheral blood blasts ≥ 5% within 30 days."*[4] As demonstrated in Figure 3.4, at the primary database lock, RFS was reported to be significantly longer with oral azacitidine when compared to the placebo

arm and a clinically meaningful difference in median difference RFS of 5.3 months (median RFS: 10.2 months in azacitidine arm versus 4.8 months in placebo; (HR 0.65 (95% CI: 0.52-0.81), p<0.0001).[1] Higher RFS rates were also observed in the oral azacitidine group when compared to the placebo group at six months (67.4% versus 45.2%), one year (44.9% versus 27.4%), and two years (26.6% versus 17.4%).[1]

ERG comment: The ERG queried why the 2020 data cut-off point where patients were followed up in the EP was not used to inform the results of this outcome, and in its response to clarification, the company stated that, *"the objective of the EP was to follow-up on survival of trial participant. Since it was not required in the EP to collect bone marrow and peripheral blood. While there were some isolated bone marrow or peripheral blood samples recorded after the July 2019 database lock (and unblinding of the study),* [redacted] *. Since bone marrow and peripheral blood samples were not collected routinely and also not according to the methodology as required in the treatment phase of the trial, we consider that reliable RFS data can only be drawn from the July 2019 data cut-off point."*[9] The probabilities of RFS was consistently higher for the azacitidine group when compared to the placebo group across different time points as demonstrated in Figure 3.4. The ERG also has concerns about what the effect of prior chemotherapy (and the number of cycles of,) may have on RFS. As the number of cycles of consolidation therapy was not balanced on both arms, the ERG queries the meaningfulness of this outcome.

Figure 3.4: KM analysis of RFS, QUAZAR AML-001 trial (ITT population) (data cut-off point, 15 July 2019)



Source: Figure B.2.5. of the CS[1]
CC-486 = oral azacitidine; CI = confidence interval; CS = Company Submission; HR = hazard ratio; ITT = intention-to-treat; KM = Kaplan-Meier; RFS = relapse-free survival

Table 3.13: Summary of RFS, QUAZAR AML-001 trial (ITT population) (data cut-off point, 15 July 2019)

Endpoint	Oral azacitidine (N=238)	Placebo (N=234)	Difference (95% CI)
Patients with event (relapse or death), n (%)	164 (68.9)	181 (77.4)	-
Patients censored, n (%)	74 (31.1)	53 (22.6)	-
Median RFS, months (95% CI)[a]	10.2 (7.9-12.9)	4.8 (4.6-6.4)	5.3 (3.1-7.5)
HR (95% CI)[b]	0.65 (0.52-0.81)		-
p-value[c]	0.0001		-
6-month RFS estimate (95% CI)[d]	[redacted]	[redacted]	[redacted]
1-year RFS estimate (95% CI)[d]	[redacted]	[redacted]	[redacted]
2-year RFS estimate (95% CI)[d]	[redacted]	[redacted]	[redacted]

Table B.2.12. of the CS[1]
CI = confidence interval; CS = Company Submission; HR = hazard ratio; ITT = intention-to-treat; KM = Kaplan-Meier; n = number of patients in the category; N = number of patients evaluable; RFS = relapse-free survival
[a]Median estimate of RFS was derived using the KM method. Difference was calculated as oral azacitidine minus placebo. The CI for the difference was derived using Kosorok's method.
[b]The HR is from a Cox proportional hazards model stratified by age, cytogenetic risk category, and receipt of consolidation therapy or not.
[c]The p-value is two-sided from a log-rank test stratified by age, cytogenetic risk category, and receipt of consolidation therapy or not.
[d]KM methods were used to estimate the six-month, one-year, and two-year RFS probabilities. The CIs for the difference in these RFS probabilities were derived using Greenwood's variance estimate.
Note: percentages are based on the number of patients in each treatment group, unless otherwise specified.

3.2.5.3 Secondary outcome: time to relapse

The company defined time to relapse as, *"the time from the date of randomisation to the date of documented relapse."*[1] As demonstrated in Table 3.14, 154 (64.7%) patients in the oral azacitidine group and 179 (76.5%) patients in the placebo group experienced a programmatically derived documented relapse. The median time-to-relapse was 10.2 months in the oral azacitidine group and 4.9 months in the placebo group.

ERG comment: It is unclear if this *"programmatically-derived documented relapse"* follows the study's definition of a relapse being *"the earliest of ≥ 5% bone marrow blasts from the central pathology report, appearance of blasts in the peripheral blood with confirmation of bone marrow blasts ≥ 5% within 100 days (i.e., approximately 3 cycles), or at least 2 peripheral blood blasts ≥ 5% within 30 days."*[4]

Table 3.14: Summary of time to relapse, QUAZAR AML-001 trial (ITT population)

Parameter	Oral azacitidine (N=238)	Placebo (N=234)
Patients relapsed, n (%)	154 (65)	179 (76)
Patients died without relapse, n (%)	[redacted]	[redacted]
Patients censored, n (%)	[redacted]	[redacted]
Median time to relapse, months (95% CI)[a]	10.2 (8.3-13.4)	4.9 (4.6-6.4)
6-month relapse rate estimate (95% CI)[b]	0.31 (0.25-0.37)	0.54 (0.48-0.61)
1-year relapse rate estimate (95% CI)[b]	0.53 (0.46-0.59)	0.72 (0.65-0.77)
2-year relapse rate estimate (95% CI)[b]	0.69 (0.62-0.75)	0.82 (0.76-0.86)

Based on Table B.2.13. of the CS[1]
CI = confidence interval; CR = complete remission; CRi = complete remission with incomplete blood count recovery; CS = Company Submission; KM = Kaplan-Meier; n = number of patients in the category; N = number of patients evaluable
[a]Unstratified KM analysis.
[b]Estimates of relapse rates are based on the cumulative incidence function from a competing risk analysis with death as a competing risk of relapse from CR/Cri.
Note: percentages are based on the number of patients in each treatment group, unless otherwise specified.

3.2.5.4 Secondary outcome: time to discontinuation from treatment

The company defined time to discontinuation from treatment as, *"the time from the date of randomisation to the date of discontinuation from investigational product."*[1] At the 15th July 2019 primary database lock, the majority of patients on both treatment arms had discontinued treatment. Table 3.15 shows that slightly more patients on the oral azacitidine arm continued with treatment as compared to the placebo arm and the median time to treatment discontinuation for any reason was 11.4 months for patients on oral azacitidine compared to 6.1 months for patients on the placebo study arm.

The company in its response to clarification states that, *"per protocol, treatment was discontinued when patients stopped benefitting from the study treatment. In the majority of the cases it was disease relapse that led to treatment discontinuation, and as the difference in RFS demonstrates, this occurred later in the oral azacitidine arm compared to the placebo arm."*[9]

ERG comment: The six patients (2.5%) still in first remission who received subsequent HSCT as curative therapy as mentioned in Section 3.2.1 of this report were all on the oral azacitidine treatment arm.

Table 3.15: Summary of time to discontinuation from treatment, QUAZAR AML-001 trial (ITT population)

Parameter	Oral azacitidine (N=238)	Placebo (N=234)	Difference (95% CI)
Patients with treatment discontinuation, n (%)	[redacted]	[redacted]	[redacted]
Patients with treatment discontinuation due to relapse, n (%)	[redacted]	[redacted]	[redacted]
Patients with treatment discontinuation due to adverse events, n (%)	[redacted]	[redacted]	[redacted]

Patients with treatment discontinuation due to eligibility for bone marrow or stem cell transplant, n (%)	6 (2.5)	0	-
Patients censored, n (%)	[redacted]	[redacted]	[redacted]
Median time to treatment discontinuation, months (95% CI)[a]	11.4 (9.8-13.6)	6.1 (5.1-7.4)	[redacted]
6-month treatment discontinuation rate estimate (95% CI)[b]	[redacted]	[redacted]	[redacted]
1-year treatment discontinuation rate estimate (95% CI)[b]	[redacted]	[redacted]	[redacted]
2-year treatment discontinuation rate estimate (95% CI)[b]	[redacted]	[redacted]	[redacted]

Based on Table B.2.14. of the CS[1] and Table 5 of CL response[9]
CI = confidence interval; CL = clarification letter; CS = Company Submission; ITT = intention-to-treat; KM = Kaplan–Meier; n = number of patients in the category; N = number of patients evaluable
[a]Median estimate of time to discontinuation is from an unstratified KM analysis. Differences were calculated as oral azacitidine minus placebo. The CIs for the differences were derived using Kosorok's method.
[b]KM methods were used to estimate the treatment discontinuation rate. Differences were calculated as oral azacitidine minus placebo. The CIs for the difference were derived using Greenwood's variance estimate.
Note: percentages are based on the number of patients in each treatment group, unless otherwise specified.

3.2.5.5 Secondary outcome: health-related quality of life (HRQoL)

Two scales were used to analyse HRQoL in the QUAZAR AML-001 study- FACIT-Fatigue scale and EQ-5D-3L health utility index. These scales were both analysed as change from baseline and the proportion of patients with clinically meaningful improvement based on a prespecified minimally important difference (MID).[1] The company indicated that the FACIT-Fatigue Scale and the EQ-5D-3L questionnaires were ideally completed prior to dosing and prior to interaction with study personnel on day 1 of every cycle, beginning on day 1 of cycle 1 and the treatment discontinuation visit.[1]

There were 225 (94.5%) patients in the oral azacitidine group and 219 (93.6%) patients in the placebo group who had a valid quality of life assessment at baseline and at least one valid baseline assessment, and were included in the HRQoL-evaluable population for the FACIT-Fatigue scale.[1] At baseline, mean scores on both the FACIT-Fatigue scale and the EQ-5D-3L health utility index were similar across the oral azacitidine and placebo groups as shown in Table 3.16.[1]

ERG comment:

- The CSR states that, *"the completion rates, based on the number of subjects in the ITT population, declined over time for both groups… with the oral azacitidine group having a significantly higher proportion of completion than the placebo group at the Cycle Day 1 Visit and thereafter."*[4] As the HRQoL evaluable population appears to be randomised patients who had a valid QoL assessment on day 1 cycle 1 and at least one valid post baseline assessment, the ERG thus requested for the additional results tables not included in the CSR, to be sent. However, upon closer inspection of the HRQoL results tables (Tables 14.3.6.2.1-14.3.6.2.4) were all missing. The ERG could not assess how significant the difference in proportion of completion between study arms was, and therefore, cannot comment on the interpretability of this comparison between oral azacitidine and placebo.

- The ERG also had some concerns about how HRQoL and fatigue were measured. In its clarification letter, the ERG asked the company to justify the impact and direction of potential bias by measuring HRQoL and fatigue on day 1 of each 28-day cycle given that study treatments were administered in the first 14 days of each cycle. The company in their response to clarification stated that, *"in theory, assessing HRQoL at the start of each treatment cycle is less likely to capture the effect of treatment-related symptomatic AEs [adverse events] on HRQoL, especially if AEs are short-lived or when treatment cycles are long. Therefore, detrimental effects on HRQoL caused by AEs may be more likely to be underestimated for oral azacitidine (vs. placebo/SOC). Despite this, it is believed that the impact would be marginal. The negative impact of AEs is not anticipated to have a long-lasting effect in most cases, as dose would likely be modified to address the issue. Those AEs with longer-lasting effects would be captured by the HRQoL instrument on day 1 of each 28-day cycle."*[9] The ERG has highlighted this as a key concern in Section 1.3 of this report.

Table 3.16: Mean baseline HRQoL scores, QUAZAR AML-001 trial (HRQoL-evaluable population)

HRQoL Domain	Oral azacitidine (N=225)	Placebo (N=219)	Overall (N=444)
FACIT-Fatigue scale[c], mean (SD)	[redacted]	[redacted]	[redacted]
EQ-5D-3L health utility index[c], mean (SD)	[redacted]	[redacted]	[redacted]

Based on Table B.2.15. of the CS[1]
CS = Company Submission; EQ-5D-3L = European Quality of Life – 5 dimensions, 3 levels; FACIT = Functional Assessment of Chronic Illness Therapy; HRQoL = health-related quality of life; N = number of patients evaluable; SD = standard deviation
[c]A higher score indicates a lower level of fatigue for FACIT-Fatigue, better health state for the EQ-5D-3L.

3.2.5.6 Secondary outcome: healthcare resource utilisation (HCRU)

Healthcare resource utilisation (HCRU) for hospitalisations was analysed as the total number of hospitalisations, total number of days hospitalised, rate of hospitalisations, days of hospitalisation per person-year of exposure, and associated relative risk of hospitalisation (with 95% CI) (see Table 3.17).[1] After adjustment for duration of study drug exposure, the results showed that oral azacitidine was associated with significantly fewer hospitalisation events per person-year (0.48 versus 0.64; p=0.0068) and a lower number of days hospitalised per person-year (7.89 versus 13.36; p<0.0001) than placebo.[1] These results suggest that maintenance treatment with oral azacitidine will lead to a reduction in HCRU associated with hospitalisations.[1]

Table 3.17: Summary of hospitalisation data, QUAZAR AML-001 trial (safety population)

Hospitalisation parameter	Oral azacitidine (N=236)	Placebo (N=233)
Total person-years exposure, years	363.8	234.9
Number of patients hospitalised, n (%)	108 (45.8)	118 (50.6)
Number of hospital events	173	151
Rate/person-year (2-sided 95% CI)[a]	0.48 [redacted]	0.64 [redacted]
Relative risk (2-sided 95% CI)[b]	0.740 (0.595-0.920)	
Two-sided p-value[b]	0.0068	

Number of days hospitalised	2872	3139
Rate/person-year (2-sided 95% CI)[a]	7.89 [redacted]	13.36 [redacted]
Relative risk (2-sided 95% CI)[b]	0.591 (0.562-0.621)	
Two-sided p-value[b]	<0.0001	

Based on Table B.2.16. of the CS[1]
CI = confidence interval; CS = Company Submission; n = number of patients in the category; N = number of patients evaluable.
[a]The 95% CI for the rate per person-year of exposure is based on the Exact method.
[b]The 95% CI for the relative risk estimate and associated nominal p-value testing that the relative risk is equal to one are based on asymptotic methods.

3.2.5.7 Subgroup analysis

The company conducted several subgroup analyses to determine whether the OS and RFS findings in the full study population would be consistent across patient demographic and disease-related subgroups. These subgroups included age at induction therapy, sex, race, geographic region, CR/CRi status at randomisation, cytogenetic risk category, receipt of consolidation therapy after induction, ECOG performance status score, prior MDS or CMML, and MRD status at screening.[1] These results have been demonstrated in Figures 3.5, 3.6, 3.7 and 3.8. They show that the OS and RFS benefit seen in the study population is consistent across key demographic and disease-related subgroups.

Among patients from Europe ([redacted]), the median OS was [redacted] for the oral azacitidine arm and [redacted] for the placebo arm ([redacted]) and median RFS was [redacted] in the oral azacitidine arm and [redacted] in the placebo arm ([redacted]), demonstrating a comparative effectiveness of oral azacitidine over placebo that is greater in the EU-subgroup than in the ITT population.[1] The company have stated that, *"the results for the European subgroup may be more reflective of UK clinical practice,"* and the ERG has already explored the differences/similarities between the ITT population and EU-subgroup in Section 3.2.3 of this report.[1]

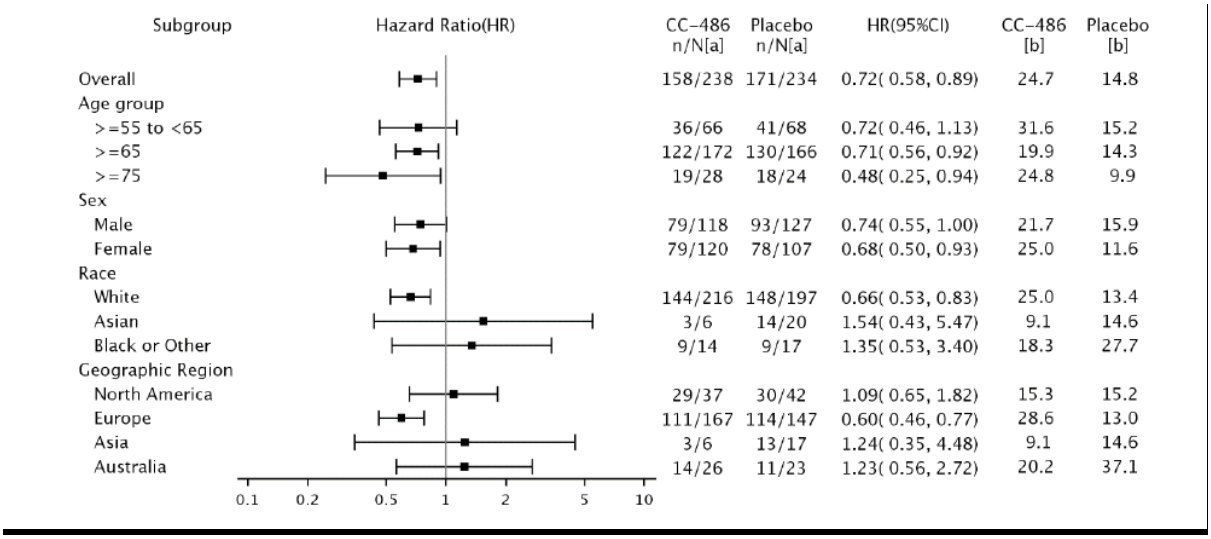
ERG comment:

- The improvement in OS is particularly notable in patients older than 65 years (HR = 0.71, 95% CI = 0.56-0.92) and those older than 75 years (HR = 0.48, 95% CI = 0.25-0.94) as demonstrated in Figure 3.5. As most AML patients in these subgroups would be considered transplant ineligible, these findings can be considered to be clinically meaningful. Also notable is the substantial difference in OS HR between patients in Europe (HR = 0.60, 95% CI = 0.46-0.77) and other geographical regions ([redacted]). Patients older than 75 years on oral azacitidine also had better improvement in RFS (HR = 0.40, 95% CI = 0.20-0.79) as demonstrated in Figure 3.7.
- MRD status, which is the detection of disease after treatment, is known to be a prognostic factor for survival in AML and 52% of patients in the trial had a negative MRD status at randomisation while 46% had a positive MRD status (see Table B.2.7. of CS[1]). Patients in the oral azacitidine arm

with a positive MRD status appear to have more improved OS and RFS when compared to the placebo arm, than patients with a negative MRD status (see Figure 3.6 and 3.8).

- Another subject of note is the effect of prior consolidation therapy following induction, and the number of cycles of consolidation therapy. Standard consolidation therapy for AML patients as stated by the FDA[19], would nominally consist of 3 to 4 cycles of high dose cytarabine, however in the QUAZAR AML-001 trial, 20% of patients received no consolidation treatment, 4% of patients received 3 cycles, 31% received 2 cycles and the majority, 45% received only 1 cycle of treatment (see Table B.2.7. of CS[1]). However, as discussed in Section 3.2.3, it is unclear how many cycles of consolidation therapy would be required for transplant ineligible AML patients following induction, in UK clinical practice. In its clarification letter, the ERG asked the company to comment on the potential implications of treating patients with one cycle of consolidation therapy and not giving consolidation therapy to 20% of patients in the QUAZAR AML-001 trial. In their response to clarification, the company provided a KM plot for subgroups of patients who received 0, 1 and ≥ 2 cycles of consolidation as displayed in Figure 3.9, and stated that, *"overall, administration of consolidation therapy was associated with treatment benefits in both the oral azacitidine group and the placebo group."*[9] From figures 3.6 and 3.8, it would seem that patients who received 1 to 2 cycles of consolidation would benefit from oral azacitidine maintenance therapy when compared to a placebo.
- In its clarification letter, the ERG requested for a subgroup analysis of patients with secondary AML at baseline, as only 9% of patients in the QUAZAR AML-001 trial had secondary AML, which is known to be associated with poor response to chemotherapy and poor outcomes. The company in their response provided a KM plot of OS in the ITT population with secondary AML, and stated that, *"the subgroup analysis for the secondary AML population demonstrates improved OS for the oral azacitidine group compared with the placebo group with a median OS of* [redacted] *months and* [redacted] *months, respectively (*[redacted]*)."*[9] The plot has been displayed in Figure 3.10, and the ERG does concur that the plot does demonstrates the comparative effectiveness of oral azacitidine to placebo in this small subgroup of patients. However, the KM plot demonstrated that survival benefit with oral azacitidine use improved with increased number of consolidation cycles. Thus, the role (and number of cycles) of consolidation treatment following induction in improving survival following cannot be underemphasised.

Figure 3.5: Forest plot of OS by demographic subgroup, QUAZAR AML-001 trial (ITT population)



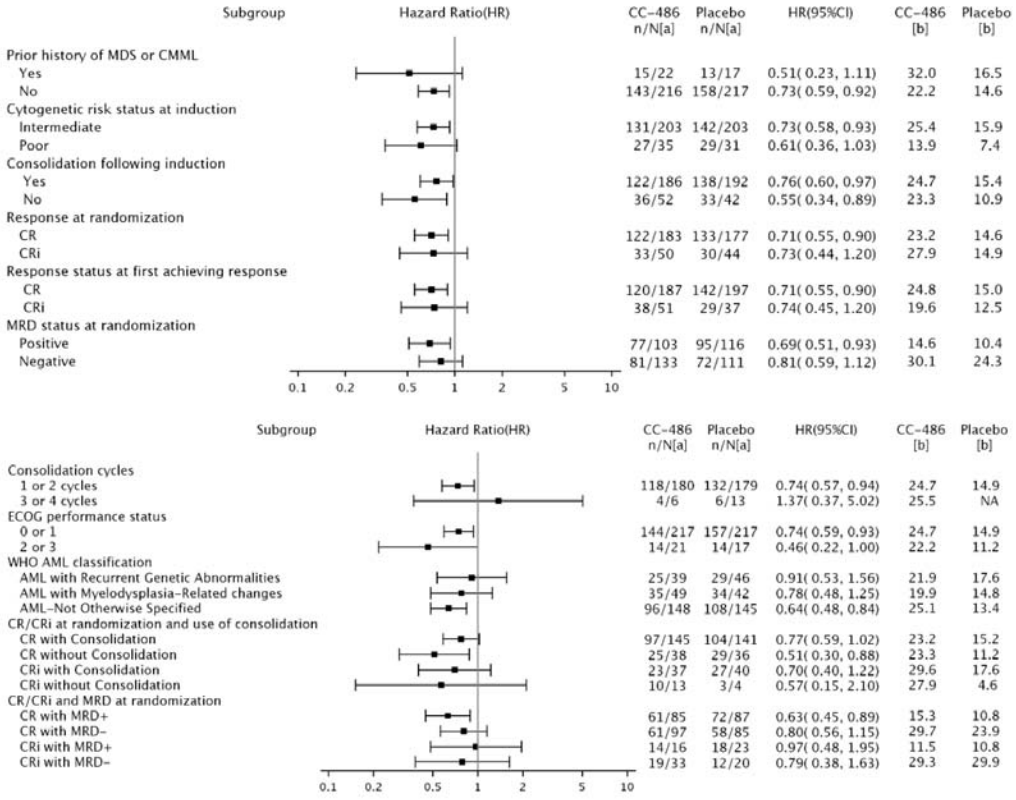
Source: Figure B.2.6. of the CS[1]

CC-486 = oral azacitidine; CI = confidence interval; CS = Company Submission; OS = overall survival.

[a]Number of events/number of patients.

[b]Median OS in months.

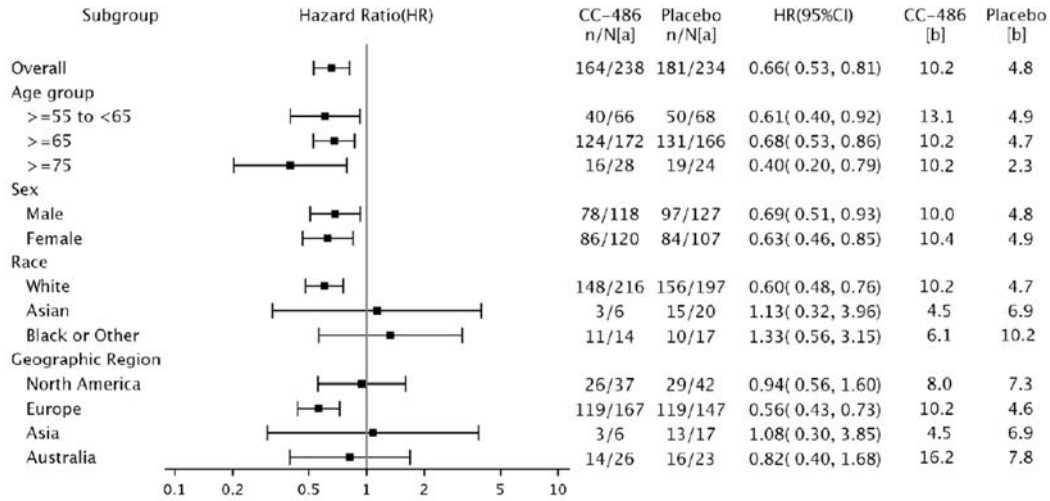
Figure 3.6: Forest plot of OS by disease-related subgroup, QUAZAR AML-001 trial (ITT population)



Source: Figure B.2.7. of the CS[1]

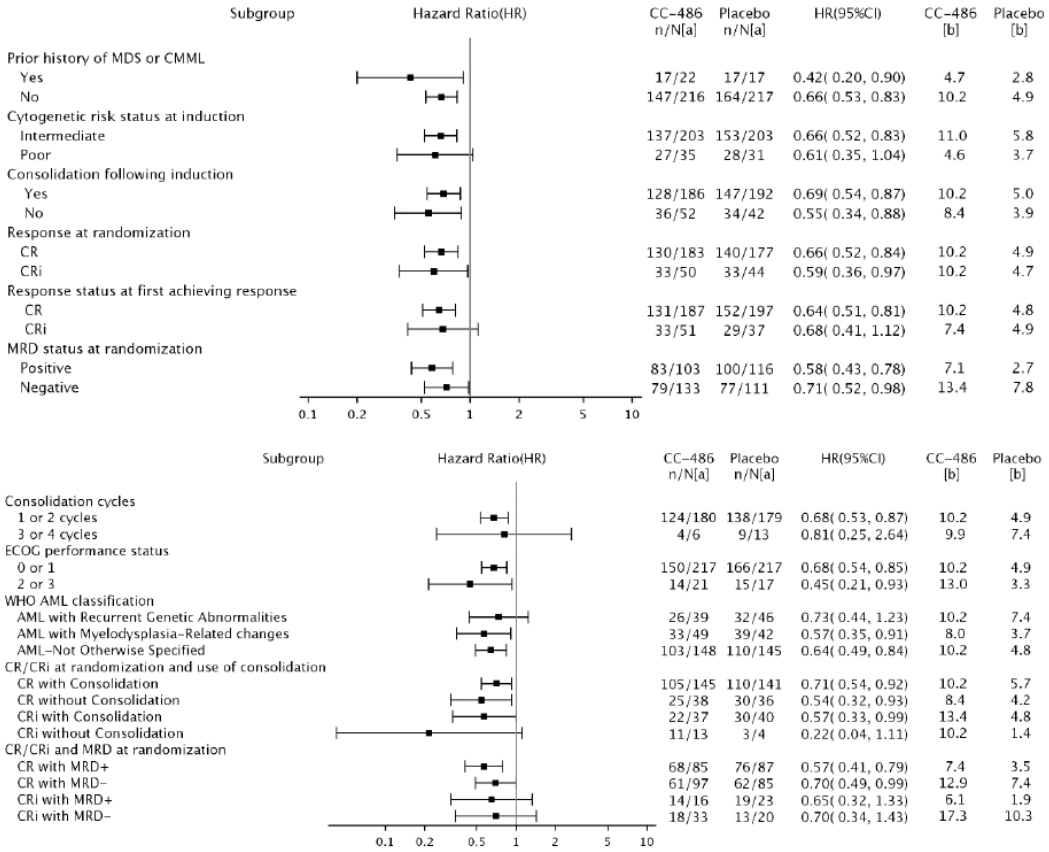
AML = acute myeloid leukaemia; CI = confidence interval; CMML = chronic myelomonocytic leukaemia; CR = complete remission; CRi = complete remission with incomplete blood count recovery; CS = Company Submission; ECOG = Eastern Cooperative Oncology Group; MDS = myelodysplastic syndromes; MRD = measurable residual disease; OS = overall survival; WHO = World Health Organization
[a]Number of events/number of patients.
[b]Median OS in months.

Figure 3.7: Forest plot of RFS by demographic subgroup, QUAZAR AML-001 trial (ITT population)



Source: Figure B.2.8. of the CS[1]
CC-486 = oral azacitidine; CI = confidence interval; CS = Company Submission; RFS = relapse-free survival
[a]Number of events/number of patients.
[b]Median OS in months.

Figure 3.8: Forest plot of RFS by disease-related subgroup, QUAZAR AML-001 trial (ITT population)



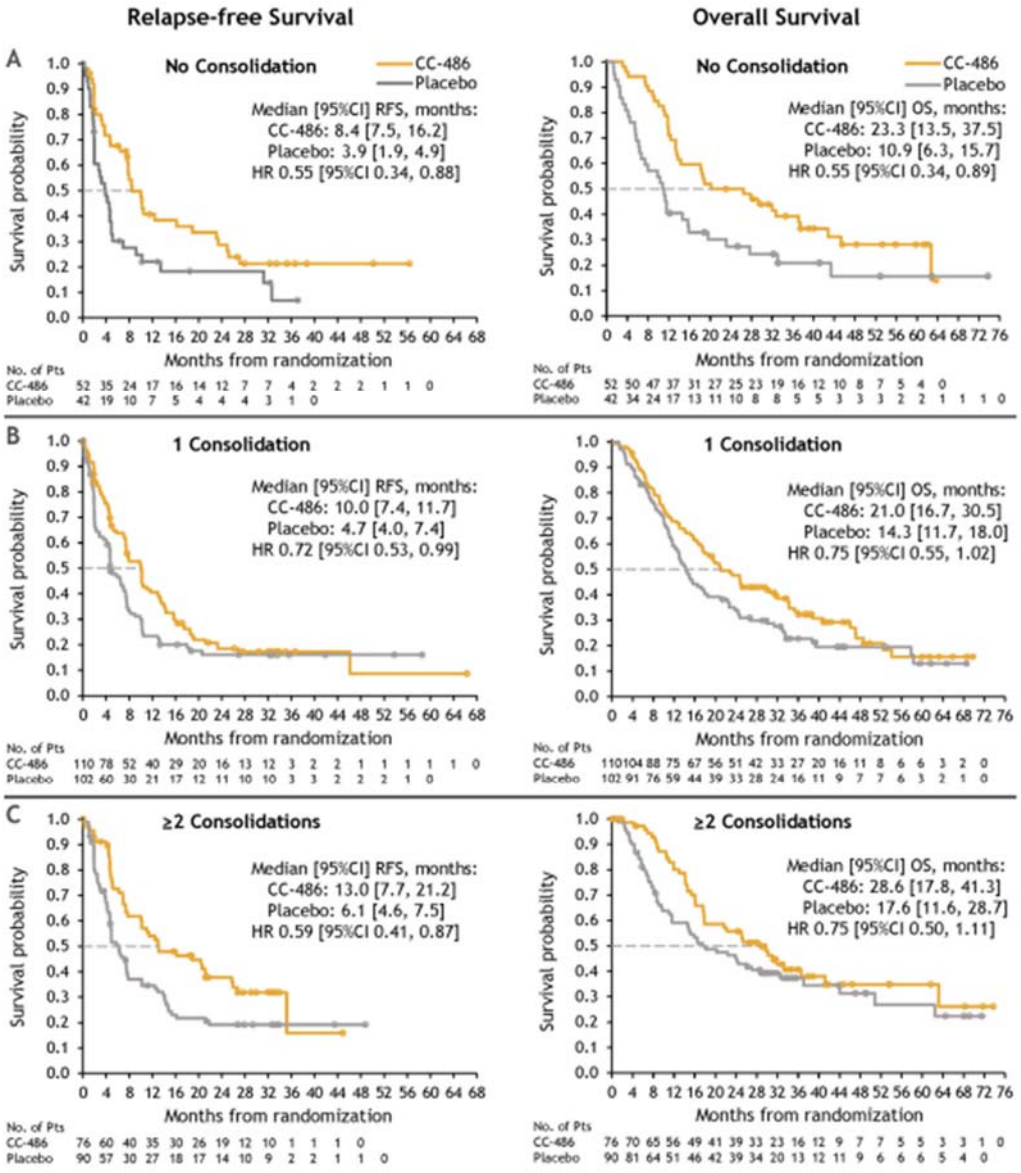
Source: Figure B.2.9. of the CS[1]

AML = acute myeloid leukaemia; CI = confidence interval; CMML = chronic myelomonocytic leukaemia; CR = complete remission; CRi = complete remission with incomplete blood count recovery; CS = Company Submission; ECOG = Eastern Cooperative Oncology Group; MDS = myelodysplastic syndromes; MRD = measurable residual disease; RFS = relapse-free survival; WHO = World Health Organisation.

[a]Number of events/number of patients.

[b]Median RFS in months.

Figure 3.9: KM plot of RFS and OS from time to randomisation in consolidation sub-group, QUAZAR AML-001 trial



Source: Figure 2 of CL response[9]

CI = confidence interval; CL = clarification letter; HR = hazard ratio; KM = Kaplan-Meier; No. = number; OS = overall survival; RFS = relapse-free survival

Figure 3.10: KM plot of overall survival- ITT population with secondary AML at baseline, QUAZAR AML-001 trial (September 2020 data cut-off)

Source: Figure 3 of CL response[9]
AML = acute myeloid leukaemia; ITT = intention to treat; KM = Kaplan-Meier; CI = confidence interval; CL = clarification letter

3.2.6 Safety results from the QUAZAR AML-001 trial

This section reports on the safety results discussed in section B.2.10 of the company submission.

ERG comments:

- The ERG requested for clarification on the scale used in defining the severity of AEs. The company in its response to clarification provided details of treatment emergent adverse event (TEAE) grading. They stated that, *"TEAEs were graded using NCI-CTCAE (National Cancer Institute - Common Terminology Criteria for Adverse Events) version 4.0."*[9]

- They also stated that AEs that were not defined in the Common Terminology Criteria for Adverse Events (CTCAE) were evaluated for severity according to the following scale:
  - Grade 1 = Mild – transient or mild discomfort; no limitation in activity; no medical intervention/therapy required
  - Grade 2 = Moderate – mild to moderate limitation in activity, some assistance may be needed; no or minimal medical intervention/therapy required

- Grade 3 = Severe – marked limitation in activity, some assistance usually required; medical intervention/therapy required, hospitalization is possible
- Grade 4 = Life threatening – extreme limitation in activity, significant assistance required; significant medical intervention/therapy required, hospitalization or hospice care probable
- Grade 5 = Death - the event results in death

- The ERG in its clarification letter also asked the company to provide the follow-up period for adverse events reporting. In their response to clarification, they stated that, *"TEAEs included adverse events that started between the first dose date and up to 28 days after the last dose date of study treatment or until the date of the last study visit (whichever was longer)."*[9] Additionally, no information was supplied to indicate the duration of reported adverse events. This was of particular concern as the ERG notes that in section B.3.4.4 of the CS, the company states that *"the duration of each AE was informed by clinical advisor opinion that each of the included events would last for approximately 1 week."*[1] Of further note was that the percentage of patients that experienced an AE at least once during the trial follow-up (as reported in Table B.3.8 of the CS[1]) was then used to model the frequency per AE. This implies that AEs were assumed to occur a maximum of once per patient for a duration of 1 week.

- The ERG sought clarification on this statement and requested that the company provide further information and justification of this position. In the response to clarification, the company stated that *"…additional clinical expert opinion was sought from the same two clinicians regarding this question. Feedback from the UK clinical experts indicated that the assumption of 1-week duration for all AE was not unreasonable as some AE may have lower duration whilst others may have higher. Explicit references were made to febrile, diarrhoea, vomiting and neutropenia as having a duration equal to or lower than 1 week and anaemia and thrombocytopenia typically having duration longer than 1 week. Furthermore, the clinicians noted that in clinical practice, the strategy is to predict and prevent AE recurrence hence modelling one event per patient may be a reasonable simplifying assumption."*[9] The ERG reviewed the CSR and associated documentation and could not identify any data that reliably demonstrates the duration of AEs. We express concerns about the selection of the '1 week' predicted AE duration that was recommended by the clinical experts.

3.2.6.1 Extent of exposure

The mean treatment exposure as defined in the safety population in the oral azacitidine group was [redacted] compared to [redacted] in the placebo group (see Table 3.18).[1] The median average daily dose of azacitidine was [redacted].[1] The mean number of treatment cycles in the oral azacitidine group was [redacted] and [redacted] in the placebo group, with the mean cycle length of greater than 28 days in both groups.[1] The mean RDI in the oral azacitidine group was [redacted] and [redacted] in the placebo group.[1] [redacted] of patients in the oral azacitidine group received >85% to ≤100% of planned dose intensity compared to [redacted] of patients in the placebo group.

Table 3.18: Treatment Exposure in QUAZAR AML-001 trial, safety population

Parameter	Oral azacitidine (N=236)	Placebo (N=233)
Treatment duration[a], months		
Mean (SD)	[redacted]	[redacted]

Parameter	Oral azacitidine (N=236)	Placebo (N=233)
Median (min, max)	11.6 (0.5, 74.3)	5.7 (0.7, 68.5)
Treatment duration[b], person-years	[redacted]	[redacted]
Average length of cycle[c], days		
Mean (SD)	[redacted]	[redacted]
Median (min, max)	[redacted]	[redacted]
Average number of days dosed per cycle[d]		
Mean (SD)	[redacted]	[redacted]
Median (min, max)	[redacted]	[redacted]
Number of cycles		
Mean (SD)	[redacted]	[redacted]
Median (min, max)	12.0 (1.0, 80.0)	6.0 (1.0, 73.0)
Number of treatment cycles initiated, n (%)		
1 or more	236 (100.0)	233 (100.0)
2 or more	[redacted]	[redacted]
3 or more	[redacted]	[redacted]
4 or more	[redacted]	[redacted]
5 or more	[redacted]	[redacted]
6 or more	[redacted]	[redacted]
12 or more	[redacted]	[redacted]
18 or more	[redacted]	[redacted]
24 or more	[redacted]	[redacted]
30 or more	[redacted]	[redacted]
Relative dose intensity (%)[e]		
Mean (SD)	[redacted]	[redacted]
≤ 75%, n (%)	[redacted]	[redacted]
> 75% to ≤ 85%, n (%)	[redacted]	[redacted]
> 85% to ≤ 100%, n (%)	[redacted]	[redacted]
> 100%, n (%)	[redacted]	[redacted]

Based on Table B.2.21 of the CS[1]
CS = Company Submission; max = maximum; min = minimum; n = number of patients in the category; N = number of patients evaluable; SD = standard deviation.
[a]Treatment duration in months is defined as (treatment end date — first dose date +1)/30.4375. Treatment end date is last dose date plus 14 days (the prescribed rest period of each cycle), or the death date, whichever is earlier.
[b]Total person-years of treatment duration is calculated as the sum of treatment duration(days)/365.25 across all patients.
[c]Average cycle length is defined as treatment duration in days/number of cycles.

Parameter	Oral azacitidine (N=236)	Placebo (N=233)
[d]Average number of days dosed per cycle is defined as total number of days dosed during the entire treatment period/number of cycles. [e]Relative dose intensity is defined as the ratio of dose intensity to the planned dose intensity (300 mg/day x 14 days/28 days = 150 mg/day for all subjects).		

ERG comment: It was noted by the ERG that the duration of exposure was increased in the oral azacitidine group compared to the placebo group and the company was asked to comment on the impact that this may have with respect to AEs. In the response to clarification the company stated that *"when examining the incidence of TEAEs, it is important to note that duration of exposure to study treatment in the oral azacitidine group (11.6 months) was approximately twice as long as exposure in the placebo group (5.7 months). Since subjects on oral azacitidine have a longer exposure, the chances of having a TEAE would be greater."*[9] The company did not provide any further insight into how this was considered in the analysis and in comparisons between the groups and therefore this is highlighted.

3.2.6.2 Treatment Emergent Adverse Events

In the QUAZAR AML-001 trial, 97.9% of patients in the oral azacitidine group and 96.6% of those in the placebo group experienced at least one TEAE (see Table 3.19).[1] When TEAEs were considered by the investigators to be related to study treatment, 89.8% of patients in the oral azacitidine group were affected compared to 51.5% of patients in the placebo group.[1] Furthermore, the frequency of serious TEAEs, TEAEs of grade 3/4 and TEAEs leading to death were all elevated in the oral azacitidine group compared to the placebo group (serious TEAEs: oral azacitidine 33.5%, placebo 25.3%; grade 3/4 TEAEs: oral azacitidine 71.6%, placebo 63.1%; TEAEs leading to death: oral azacitidine 3.8%; placebo 1.7%).[1]

Table 3.19: Summary of ≥1 TEAEs, QUAZAR AML-001 study (safety population)

Category	Oral azacitidine (N=236)	Placebo (N=233)
TEAEs, n (%)	231 (97.9)	225 (96.6)
TEAEs related to study treatment, n (%)	212 (89.8)	120 (51.5)
Serious TEAEs, n (%)	79 (33.5)	59 (25.3)
Treatment-related serious TEAEs, n (%)	22 (9.3)	5 (2.1)
Grade 3/4 TEAEs[a], n (%)	169 (71.6)	147 (63.1)
Treatment-related Grade 3/4 TEAEs[a], n (%)	113 (47.9)	54 (23.2)
TEAEs leading to death, n (%)	9 (3.8)	4 (1.7)
TEAEs leading to dose reduction, n (%)	37 (15.7)	6 (2.6)
TEAEs leading to dose interruption, n (%)	102 (43.2)	40 (17.2)
TEAEs leading to dose reduction and interruption, n (%)	24 (10.2)	3 (1.3)

Category	Oral azacitidine (N=236)	Placebo (N=233)
TEAEs leading to study treatment discontinuation, n (%)	31 (13.1)	10 (4.3)

Source: Table B.2.22 of the CS[1]
CS = Company Submission; n = number of patients in the category; N = number of patients evaluable; TEAE = treatment-emergent adverse event.
[a]Graded using Common Terminology Criteria for Adverse Events version 4.0.
Notes: AML relapse as defined by MedDRA high-level group term leukaemia's is excluded. AEs were evaluated from the first dose date through 28 days after the last dose of study treatment.

In the trial, the most commonly reported TEAEs were gastrointestinal (GI) events, which occurred more frequently in the oral azacitidine group (91.1%) than in the placebo group (61.8%).[1] No clear data was tabulated in the CS alongside this text to illustrate this. The CS provided tabulated data on TEAEs reported in >10% of patients in the safety population and which is summarised in Table 3.20 below. TEAEs of any grade reported by 10% of patients of either arm of the safety population was broadly similar. The most frequently reported TEAEs of any grade in the oral azacitidine arm (versus placebo) were GI events including nausea (65% versus 24%), vomiting (60% versus 10%) and diarrhoea (50% versus 21%) although these were generally mild to moderate severity (grade 1/2).[1] The CS reports that grade 3/4 GI TEAEs only occurred in 14.4% of patients in the oral azacitidine group and [redacted] of patients in the placebo group, and included diarrhoea (5.1% versus 1.3%), vomiting (3.0% versus 0%), nausea (2.5% versus 0.4%), and constipation (1.3% versus 0%).[1] Although GI events were the most common TEAEs observed during maintenance treatment with oral azacitidine, the CS reports that a relatively small percentage of patients who experienced these events required dose reduction ([redacted] for oral azacitidine versus [redacted] for placebo), dose interruption ([redacted] or treatment discontinuation ([redacted].[1]

The most reported haematologic TEAEs of any grade (versus placebo) were neutropenia (44% versus 26%), thrombocytopenia (33% versus 27%), and anaemia (20% versus 18%).[1] The occurrence of more serious haematologic TEAEs at grade 3/4 followed a similar pattern with neutropenia 41% versus 24 %), thrombocytopenia (22% versus 21%), and anaemia (14% versus 13%).[1]

Table 3.20: TEAEs reported in >10% of patients in QUAZAR AML-001 trial, safety population

Event	Oral azacitidine (N=236)		Placebo (N=233)	
	Any grade	Grade 3/4	Any grade	Grade 3/4
TEAEs, n (%)	231 (98)	169 (72)	225 (97)	147 (63)
Nausea	153 (65)	6 (3)	55 (24)	1 (<1)
Vomiting	141 (60)	7 (3)	23 (10)	0 (0)
Diarrhoea	119 (50)	12 (5)	50 (21)	3 (1)
Neutropenia	105 (44)	97 (41)	61 (26)	55 (24)
Constipation	91 (39)	3 (1)	56 (24)	0 (0)
Thrombocytopenia	79 (33)	53 (22)	63 (27)	50 (21)
Fatigue	70 (30)	7 (3)	45 (19)	2 (1)

Event	Oral azacitidine (N=236)		Placebo (N=233)	
	Any grade	Grade 3/4	Any grade	Grade 3/4
Anaemia	48 (20)	33 (14)	42 (18)	30 (13)
Asthenia	44 (19)	2 (1)	13 (6)	1 (<1)
Pyrexia	36 (15)	4 (2)	44 (19)	1 (<1)
Arthralgia	32 (14)	2 (1)	24 (10)	1 (<1)
Abdominal pain	31 (13)	2 (1)	16 (7)	0 (0)
Upper respiratory tract infection	31 (13)	1 (<1)	32 (14)	0 (0)
Decreased appetite	30 (13)	2 (1)	15 (6)	2 (1)
Cough	29 (12)	0 (0)	39 (17)	0 (0)
Febrile neutropenia	28 (12)	27 (11)	18 (8)	18 (8)
Back pain	28 (12)	3 (1)	23 (10)	2 (1)
Leukopenia	25 (11)	18 (8)	19 (8)	14 (6)
Pain in extremity	25 (11)	1 (<1)	12 (5)	0 (0)
Dizziness	25 (11)	0 (0)	21 (9)	0 (0)
Headache	23 (10)	0 (0)	26 (11)	1 (<1)
Peripheral oedema	21 (9)	0 (0)	24 (10)	1 (<1)

Based on Table B.2.23 of the CS[1]
CS = Company Submission; n = number of patients in the category; N = number of patients evaluable; TEAE = treatment-emergent adverse event
Notes: TEAEs were evaluated from the first dose date through 28 days after the last dose of study treatment. Events were coded according to preferred terms from the Medical Dictionary of Regulatory Activities, version 22 and were graded with the use of the National Cancer Institute Common Terminology Criteria for Adverse Events, version 4.0.

ERG comment: In its request for clarification, the ERG asked the company to provide TEAEs by severity and system class. In its response to clarification, the company provided what has been included in this report as Table 3.21. The table illustrates that grade 3/4 TEAEs for all System Organ Classes (SOC) reported for ≥ 2% of subjects were generally higher in the oral azacitidine group (versus placebo) with particular emphasis on SOC's blood and lymphatic system disorders ([redacted]), and gastrointestinal disorders ([redacted]). The most common grade 3/4 TEAEs reported with oral azacitidine were neutropenia (41.1% versus 23.6%), thrombocytopenia (22.5% versus 21.5%), anaemia (14.0% versus 12.9%), and febrile neutropenia (11.4% versus 7.7%). These figures are broadly consistent with TEAEs reported in >10% of patients that are detailed Table 3.20.

As the CS states that, *"the percentage of patients with haematologic AEs within each treatment group were generally consistent over time up to Cycle 12,"* it is unclear if these AEs which were higher on the oral azacitidine arm when compared to placebo, are expected to remain for a lifetime in patients who have experienced them.[1]

Table 3.21: TEAEs with a severity of Grade 3 or 4 by System Organ Class and Preferred Term Reported for ≥ 2% of Subjects in the CC-486 group Excluding AML Relapse (Safety Population)

System Organ Class Preferred Term[a]	Oral azacitidine (N=236)			Placebo (N=233)		
	Grade 3 n (%)	Grade 4 n (%)	Grade 3 or 4 n (%)	Grade 3 n (%)	Grade 4 n (%)	Grade 3 or 4 n (%)
Subjects with at least one Grade 3 or 4 TEAE[b]	[redacted]	[redacted]	169 (71.6)	[redacted]	[redacted]	147 (63.1)
Blood and lymphatic system disorders	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
Neutropenia	[redacted]	[redacted]	97 (41.1)	[redacted]	[redacted]	55 (23.6)
Thrombocytopenia	[redacted]	[redacted]	53 (22.5)	[redacted]	[redacted]	50 (21.5)
Anaemia	[redacted]	[redacted]	33 (14.0)	[redacted]	[redacted]	30 (12.9)
Febrile neutropenia	[redacted]	[redacted]	27 (11.4)	[redacted]	[redacted]	18 (7.7)
Leukopenia	[redacted]	[redacted]	18 (7.6)	[redacted]	[redacted]	14 (6.0)
Infections and infestations	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
Pneumonia	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
Gastrointestinal disorders	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
Diarrhoea	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
Vomiting	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
Nausea	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
Metabolism and nutrition disorders	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
Hypokalaemia	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
General disorders and administration site conditions	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
Fatigue	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
Investigations	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]

System Organ Class Preferred Term[a]	Oral azacitidine (N=236)			Placebo (N=233)		
	Grade 3 n (%)	Grade 4 n (%)	Grade 3 or 4 n (%)	Grade 3 n (%)	Grade 4 n (%)	Grade 3 or 4 n (%)
Blood uric acid increased	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
Vascular disorders	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
Hypertension	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
Eye disorders	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
Cataract	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
Nervous system disorders	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
Syncope	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]

Based on Table 12 of CL response[9]

CL = clarification letter; MedDRA = Medical Dictionary for Regulatory Activities; N = number of patients evaluable; TEAE = treatment-emergent adverse event

[a]Coded using MedDRA version 22.0. A subject with multiple TEAEs within a preferred term/system organ class is counted once for that preferred term/system organ class in each severity grade and once in the combined severity grade grouping.

[b]Graded using Common Terminology Criteria for Adverse Events (CTCAE) version 4.0.

Notes: Acute myeloid leukaemia relapse as defined by MedDRA high-level group term leukaemia's are excluded. Treatment-emergent adverse events include adverse events that started between first dose date and the date 28 days after the last dose date of study treatment.

3.2.6.3 Serious Treatment Emergent Adverse events

Serious TEAEs were reported for 33% of the patients in the oral azacitidine group and 25% in the placebo group.[1] The CS states that the most common serious TEAEs were infections with 17% of patients in the oral azacitidine group and 8% of patients in the placebo group and cites the Wei et al. 2019[18] as the source for this but does not explicitly present the data in the CS. Table 3.22 below represents serious TEAEs that were reported in ≥1% of patients in either treatment arm.[1] Febrile neutropenia occurred in 7% of patients in the oral azacitidine arm compared to 4% in the placebo arm, while pneumonia occurred in 4% of those in the oral azacitidine arm compared to 3% in the placebo arm.[1] In general, the rates of serious of TEAEs between both arms were similar.

Table 3.22: Serious TEAEs reported in ≥1% of patients in either treatment arm, QUAZAR AML-001 trial (safety population)

Event	Oral azacitidine (N=236)	Placebo (N=233)
Serious TEAEs, n (%)	79 (33)	59 (25)
Febrile neutropenia	16 (7)	9 (4)
Pneumonia	9 (4)	7 (3)
Pyrexia	5 (2)	1 (0.4)
Cellulitis	4 (2)	1 (0.4)
Sepsis	4 (2)	5 (2)
Influenza	3 (1)	0 (0)
Diarrhoea	3 (1)	0 (0)
Back pain	3 (1)	0 (0)
Atrial fibrillation	3 (1)	0 (0)
Cholecystitis	3 (1)	2 (1)
Anaemia	2 (1)	3 (1)
Thrombocytopenia	2 (1)	3 (1)

Source: Table B.2.24 of the CS[1]
CS = Company Submission; n = number of patients in the category; N = number of patients evaluable; TEAE = treatment-emergent adverse event.
Notes: Events were coded according to preferred terms from the Medical Dictionary of Regulatory Activities. A patient is counted only once for multiple events within preferred term/system organ class.

3.2.6.4 AEs leading to dose reduction, dose interruption, and/or discontinuation of treatment

The CS reports that AEs leading to dose reduction were reported for 16% of patients in the oral azacitidine group and 3% of patients in the placebo group.[1] No tabulated data was reported in the CS to represent this and so the trial CSR was reviewed to review the data. Page 171 of the CSR contains tabulated data with accompanying text stating *"incidence of TEAEs in the CC-486 group leading to dose reduction were reported for 15.7% of subjects and 2.6% of subjects in the placebo group. In the CC-486 group, TEAEs leading to dose reduction were primarily in the System Organ Classes of Blood and Lymphatic System Disorders (8.1%) and Gastrointestinal Disorders (5.5%). Treatment-emergent adverse events leading to dose reduction for more than 2 subjects in the CC-486 group were*

neutropenia (5.5% versus 0.4% in the placebo group), diarrhoea (3.4% versus 0%), thrombocytopenia (1.7% versus 1.3%), and nausea (1.7% versus 0%).”[4]

There were 43% of patients in the oral azacitidine group and 17% of patients in the placebo group who experienced TEAE’s that led to dose interruption.[1] These were primarily in the SOC of blood and lymphatic system disorders (26.7%), GI disorders (13.1%), and infections and infestations (12.7%) (referenced by CSR). The most frequent AEs leading to dose interruption (reported for ≥1% of patients in either treatment arm) were (versus placebo) neutropenia (20% versus 6%), thrombocytopenia (8% versus 2%), nausea (6% versus 0.4%), diarrhoea (4% versus 1%), vomiting (4% versus 0%), febrile neutropenia (2% versus 0.4%), and alanine aminotransferase increased (2% versus 1%).[1]

There were 13% of patients in the oral azacitidine group and 4% of patients in the placebo group who experienced at least one TEAE leading to study treatment discontinuation (excluding AML relapse).[1] In the oral azacitidine group, AEs leading to treatment discontinuation reported by >1 patient (versus placebo) included nausea (2% versus 0%), diarrhoea (2% versus 0%), vomiting (1% versus 0%), abdominal pain (1% versus 0%), fatigue (1% versus 0%), and thrombocytopenia (0.4% versus 1%).[1]

Discontinuation of study treatment because of AEs was reported for 13% of patients in the oral azacitidine group and 4% of patients in the placebo group.[1] In the oral azacitidine group, AEs leading to treatment discontinuation reported by >1 patient in either treatment arm included nausea (2% versus 0%), diarrhoea (2% versus 0%), vomiting (1% versus 0%), abdominal pain (1% versus 0%), fatigue (1% versus 0%), and thrombocytopenia (0.4% versus 1%) for oral azacitidine versus placebo, respectively.[1]

ERG comment: The ERG notes that there was a lack of tabulated data provided in the CS to support the statements and information included in the text. While the source was referenced, this lack of data presentation meant that the provided submission did not have optimal clarity in these sections. The data described in the CS and presented in the CSR emphasises that dose reduction, interruption and treatment discontinuation were all elevated on the oral azacitidine arm when compared to placebo.

3.2.6.5 Deaths

In general, few deaths were reported during the QUAZAR AML-001 trial. Most of these occurred after cycle 6: [redacted] patients in the oral azacitidine group and [redacted] patients in the placebo group.[1] AEs led to death in nine patients (4%) on the oral azacitidine arm (two dying from sepsis, two from cerebral haemorrhage, one from both sepsis and multiorgan failure, and one each from intracranial haemorrhage, cardiogenic shock, aspiration pneumonia, and suicide).[1] Whilst on the placebo arm, AEs led to death in four patients (2%) in the placebo group (two died from multiorgan failure, one from cerebral haemorrhage, and one from general health deterioration).[1] [redacted] leading to death were considered by the investigator in the trial to be treatment related.

3.2.6.6 Treatment emergent Adverse Events of Special Interest (AESI’s)

The CS reports that adverse events of special interest (AESI) were slightly increased in the oral azacitidine group compared to the placebo group (97% versus 91%) and states that ‘*events were largely manageable with dose modifications and standard therapeutic interventions and a few events were fatal (oral azacitidine vs. placebo:* [redacted] *or lead to discontinuation of study therapy (oral azacitidine vs. placebo:* [redacted]*’*[1]. Further scrutiny of the CSR reveals that the breakdown of reported ASEIs were higher in the oral azacitidine group compared to the placebo group for events deemed more severe, [redacted] versus [redacted] for serious; and [redacted] versus [redacted] for grade 3/4;

with [redacted] and [redacted] for events leading to death[4]. Table 3.23 below is based on B.5.25 in the CS and represents the comparisons of AESI's between the two groups.

Table 3.23: **Summary of treatment-related AESI (any grade), QUAZAR AML-001 study (safety population)**

AESI	Oral azacitidine (N=236)	Placebo (N=233)
Myelosuppression, n (%)	153 (65)	107 (46)
Haemorrhagic events, n (%)	51 (22%)	46 (20)
Infections, n (%)	147 (62)	123 (53)
Renal failure, n (%)	7 (3)	5 (2)
Hepatic failure, n (%)	3 (1)	3 (1)
Ischaemic colitis, n (%)	0 (0)	0 (0)
Cardiac events, n (%)	47 (20)	39 (17)
Psychiatric disorder, n (%)	11 (5)	8 (3)
Tumour lysis syndrome, n (%)	1 (<1)	1 (<1)
Interstitial lung disease, n (%)	2 (<1)	1 (<1)
Gastrointestinal events, n (%)	215 (91)	155 (67)
Anxiety, confusional state, insomnia, n (%)	34 (14)	30 (13)

Source: based on table B.5.25, CS [1]
AESI = adverse event of special interest; CS = Company Submission; n = number of patients in the category; N = number of patients evaluable.

3.3 Critique of trials identified and included in the indirect comparison and/or multiple treatment comparison

The company considered two comparators as relevant for this submission- a 'watch and wait' strategy with BSC, and midostaurin in patients with FLT3-mutation positive AML.[1] As there were no head-to-head studies comparing the efficacy of oral azacitidine as maintenance treatment to midostaurin as maintenance therapy in subjects with FLT3-ITD and/or FLT3-TKD (FLT3 mutation) positive AML, the company conducted an indirect treatment comparison (ITC) to assess this.

Section B.2.9.1 of the CS stated that two studies were identified from the SLR for the indirect comparison.[1] Appendix D is cited for further details of this SLR.[16] The company reported that RATIFY was the only study identified in the SLR that provided an analysis of midostaurin as maintenance treatment in AML, although subjects were not randomised at the maintenance phase, but for induction. The company also identified other sources of substantial heterogeneity between the QUAZAR AML-001 and RATIFY trials such as the difference in inclusion and exclusion criteria where the RATIFY trial included significantly younger patients and excluded patients based on FLT3 mutational status when compared to the QUAZAR AML-001 trial, baseline characteristics being substantially different, and inflated effect size estimation of survival outcomes.[9]

ERG comment: The SLR inclusion and exclusion criteria (Table B.5.3 in Appendix D) lists the only intervention as "*oral azacitidine*", which would also imply the exclusion of any studies of any treatment not compared to oral azacitidine.[16] Therefore, the ERG asked the company to clarify if this SLR reported in Appendix D is the one use to obtain studies for the indirect comparison, to which the company responded that it the eligibility criteria outlined in Appendix D were used to identify all trials assessing the efficacy and safety of maintenance therapies in AML.[9] They went on to state that studies included in the SLR were then assessed for the feasibility to be included in an ITC versus oral azacitidine. It therefore remains unclear as to how the RATIFY trial was included, given that it compared only midostaurin with placebo and not azacitidine. It also raises the question as to whether other midostaurin trials were missed, although the ERG considers that this is unlikely in the FLT3 population. Patients being randomised in RATIFY for induction and not maintenance introduces a high risk of bias in any analysis at the maintenance phase. However, it is even more unlikely that there has been an RCT of midostaurin: the ERG found one more study where midostaurin was administered for maintenance, but no separate analysis at the maintenance phase was reported. [14] In the response to clarification, the company also state that, *"the RATIFY trial was not prospectively designed to determine the independent effect of midostaurin as maintenance therapy."*[9]

3.4 Critique of the indirect comparison and/or multiple treatment comparison

The QUAZAR AML-001 trial initially consisted of 472 patients but after matching the inclusion/exclusion criteria of QUAZAR AML-001 to RATIFY i.e., removing the individual patient data for subjects without *FLT3* mutations and CRi, [redacted] AML patients (placebo, n=[redacted]; azacitidine, n=[redacted]) with a *FLT3*-mutation who had achieved CR remained within the primary analysis population.

Unmatched (n=472) and matched ([redacted]) results demonstrated that the HR is favourable for oral azacitidine when compared to midostaurin ([redacted]) and ([redacted]), respectively (see Table 3.24). Similarly, ITC results for RFS demonstrated that the HR is favourable for oral azacitidine when compared to midostaurin ([redacted]) and ([redacted]), respectively (see Table 3.25).

Table 3.24: ITC results for OS

Scenario	Oral azacitidine versus midostaurin HR (95% CI)
Unmatched	[redacted]
Primary Analysis	
Matched[a]	[redacted]

Source: Table B.2.19. of CS[1]
CI = confidence interval; CRi = complete remission with incomplete blood count recovery; CS = Company Submission; HR = hazard ratio; ITC = indirect treatment comparison; OS = overall survival
[a]Matching: 416 patients with CRi and no FLT3 mutation were removed from the unmatched population to align with inclusion/exclusion criteria in RATIFY.

Table 3.25: ITC results for RFS

Scenario	Oral azacitidine versus midostaurin HR (95% CI)
Unmatched	[redacted]
Primary Analysis	
Matched[a]	[redacted]

Source: Table B.2.20. of the CS[1]

CI = confidence interval; CRi = complete remission with incomplete blood count recovery; CS = Company Submission; HR = hazard ratio; ITC = indirect treatment comparison; RFS = relapse-free survival
[a]Matching 416 patients with CRi and no FLT3 mutation were removed from the unmatched population to align with inclusion/exclusion criteria in RATIFY.

ERG comment: The company were asked in the clarification letter to explain the likely effect on the bias of population adjustment of the indirect comparison with reference to TSD18.[20] Their response was to explain that a simulated treatment comparison would be preferred to a matched adjusted indirect comparison (MAIC) due to the lack of overlap between the populations in QUAZAR and RATIFY.[9] However, they considered that there was insufficient overlap even for an simulated treatment comparison given the estimated [redacted] overlap [redacted]) in age. The ERG would concur that population adjustment is unlikely to reduce any bias beyond the matching that had been performed and would not mitigate the main problem of selection bias in the non-randomised comparison of midostaurin versus placebo at the maintenance phase.

3.5 Additional work on clinical effectiveness undertaken by the ERG

The ERG conducted a quality assessment of the QUAZAR AML-001 study using similar criteria to the company's assessment, the results of which have been discussed in Section 3.2.4 of this report.

3.6 Conclusions of the clinical effectiveness section

The CS and response to clarification provided sufficient details for the ERG to appraise the literature searches. Despite some issues with transparency and reproducibility, searches were carried out on a good range of resources. Additional searches included conference proceedings, HTA organisations and the checking of reference lists in relevant SRs identified during the searches. The strategies provided contained a good use of free text terms, appropriate subject headings and study design filters.

The clinical effectiveness evidence for oral azacitidine in the CS is based on the QUAZAR AML-001 trial. The QUAZAR AML-001 is an ongoing, phase 3, randomised, double-blind, placebo-controlled trial comparing oral azacitidine (300 mg azacitidine orally once daily) plus BSC versus placebo plus BSC. This trial provides evidence for oral azacitidine in its expected position in the clinical pathway: as maintenance treatment for patients with AML who have achieved CR or CRi following induction therapy, with or without consolidation chemotherapy, and were not candidates for HSCT. The trial consisted of four phases: pre-randomisation phase, (screening phase within 28 days prior to randomisation) randomisation and double-blind treatment phase, (1:1 randomisation to study treatment until discontinuation following AML relapse) follow-up phase, (follow-up up to 28 days after last dose of study treatment for AEs, and then every month for the first year and then every 3 months until death) and EP (unblinding to receive azacitidine if subject did not meet study discontinuation criteria (or not receive azacitidine if in placebo group) and followed for survival for at least another 12 months until death, withdrawal of consent, study closure, or lost to follow-up). Data from the QUAZAR AML-001 trial were used as the main data for the economic modelling in this submission.

Detailed efficacy results are presented in Section 3.2.3 while detailed safety results are presented in Section 3.2.4:

- In the QUAZAR AML-001 trial, oral azacitidine/BSC significantly improved OS at both 15th July 2019 and 8th September 2020 data cut-off points when compared to placebo, meeting its primary endpoint. At a median follow-up of 41.2 months (primary database lock), oral azacitidine was associated with a significantly longer OS compared with placebo, with a clinically meaningful difference in median OS of 9.9 months (median OS: 24.7 months versus 14.8 months; HR 0.69 (95% CI: 0.55-0.86), p<0.001).

- Survival rates were higher in the oral azacitidine group than in the placebo group at one year after randomisation (72.8% versus 55.8%; difference 17.0 percentage points (95% CI: 8.4-25.6)). Higher RFS rates were observed in the oral azacitidine group than in the placebo group at six months (67.4% versus 45.2%), one year (44.9% versus 27.4%), and two years (26.6% versus 17.4%).
- The median time to relapse was 10.2 months in the oral azacitidine group and 4.9 months in the placebo group, 81.1% of patients on oral azacitidine had discontinued from the study compared to 88.9% of patients on the placebo arm by 15th July 2019, and oral azacitidine was associated with significantly fewer hospitalisation events per person-year (0.48 versus 0.64; p=0.0068) and a lower number of days hospitalised per person-year (7.89 versus 13.36; p<0.0001) than placebo. Overall, the results from the trial are favourable for oral azacitidine.
- The incidences of TEAEs were similar for the two treatment arms - 97.9% of patients in the oral azacitidine group and 96.6% of those in the placebo group experienced at least one TEAE during the study. The proportion of patients who experienced at least one TEAE considered by the study investigator to be related to study treatment was higher in the oral azacitidine group than in the placebo group (89.8% versus 51.5%). The rates of serious TEAEs (oral azacitidine: 33.5%; placebo: 25.3%), grade 3/4 TEAEs (oral azacitidine: 71.6%; placebo: 63.1%) and TEAEs leading to death (oral azacitidine: 3.8%; placebo: 1.7%) were notably higher in the oral azacitidine group when compared to the placebo group. The most common TEAEs were GI events, which occurred more frequently in the oral azacitidine group (91.1%) than in the placebo group (61.8%). The most common haematologic TEAEs were neutropenia, thrombocytopenia, and anaemia (which were among the most common grade 3/4 TEAEs reported with oral azacitidine).

No meta-analyses were carried out; however, the company conducted an ITC comparing the efficacy of oral azacitidine as maintenance treatment to midostaurin as maintenance therapy in subjects with FLT3-ITD and/or FLT3-TKD (FLT3 mutation) positive AML. The RATIFY trial was the only study identified in the SLR that provided an analysis of midostaurin as maintenance treatment in AML, although subjects were not randomised at the maintenance phase, but for induction. The QUAZAR AML-001 trial initially consisted of 472 patients but the individual patient data for subjects without FLT3 mutations were removed to match the inclusion/exclusion criteria of the QUAZAR AML-001 trial.

The company conducted a feasibility assessment of the RATIFY trial which identified significant heterogeneities in trial design (although the RATIFY trial included a maintenance therapy phase, the 205 patients who entered the maintenance phase were not re-randomised prior to the start of maintenance therapy), patient age (the inclusion criteria for QUAZAR AML-001 was ≥55 years compared with RATIFY which included patients aged 18-59 years), cytogenetic risk (favourable cytogenetic risk patients were included in RATIFY but not in QUAZAR AML-001), AML mutational status, HSCT eligibility (HSCT eligibility was not a formal exclusion criterion in the RATIFY trial and 57% of patients underwent HSCT while the QUAZAR AML-001 trial excluded patients who were eligible for HSCT at study screening and 6% of patients on the oral azacitidine arm underwent HSCT), history of consolidation therapy, and different time zero definitions of time-to-event outcomes.

QUAZAR-001 is an ongoing, phase 3, randomised, double-blind, placebo-controlled trial comparing oral azacitidine (300 mg azacitidine orally once daily) plus BSC versus placebo plus BSC as maintenance treatment. Administration of oral azacitidine or placebo continued until more than 15% blasts were present or unacceptable adverse effects occurred. The primary outcome for the trial was OS, and they also measured RFS, disease-free survival, PFS, time to relapse from CR/CRi, time to

discontinuation of treatment, safety, AEs (including TRAEs, and SAEs), withdrawals due to AEs, and patient reported outcomes (including EQ-5D).

4. COST EFFECTIVENESS

4.1 ERG comment on company's review of cost effectiveness evidence

Two sets of systematic literature searches were performed to identify available cost effectiveness and cost-utility studies (CS Appendices G and H).

4.1.1 Searches performed for cost effectiveness section

Appendix G of the CS reported literature searches used to identify published cost effectiveness and cost-utility studies.[16] Searches were conducted in February 2020 and updated in June 2021. A summary of the resources searched are provided in Table 4.1. The following paragraphs contain summaries and critiques of all searches related to cost effectiveness presented in the company submission.

Table 4.1: Resources searched for cost effectiveness and cost-utility studies. Feb 2020 and June 2021.

Search strategy element	Resource	Host/source	Date range of most recent search	Date searched
Databases	Embase	Ovid	2005-2021/06/11	12.2.20 Updated 11.6.21
	MEDLINE & MEDLINE In-Process		2005-2021/06/10	12.2.20 Updated 11.6.21
	NHS EED		All years	12.2.20*
HTAs	CADTH			12.2.20 Updated June 21
	NICE			12.2.20 Updated June 21
Conference proceedings	ASCO		2020-2021	12.2.20 Updated 11.6.21
	ASH			
	EHA			
	ISPOR EU			
	ISPOR US			
Additional searches	Handsearching of reference lists of relevant SRs identified during database searches.			

ASCO = American Society of Clinical Oncology; ASH = American Society of Hematology; CADTH = Canadian Agency for Drugs and Technologies in Health; EHA = European Hematology Association; EU = European Union; HTA = health technology assessment; ISPOR EU = International Society for Pharmacoeconomics and Outcomes Europe; ISPOR US = International Society for Pharmacoeconomics and Outcomes United States; NHS EED = NHS Economic Evaluation Database; NICE = National Institute for Health and Care Excellence; SRs = systematic reviews

*No updates required as no new records have been added to NHS EED since the original searches were run

ERG Comment:

- A good range of resources were searched for the economic SLR. Initially only the latest iteration of each search was provided, but after a request at clarification all strategies for both the original and updated Ovid searches were provided and these searches were clearly structured.
- The update search for MEDLINE and Embase provided in the CS was run as a single search split into facets by resource. The strategy combined both free text and the appropriate subject headings and field tags for each database. The strategy also utilised both the CADTH economics filter, and a second filter created by an information specialist based on terms identified in NICE economic analyses.
- It was unclear in the original Embase strategy, provided at clarification (Table G.2) at which point records were exported.[9] The ERG was unsure whether a line combining the two sets of results from lines #40 and #41 had been omitted in error from the reported strategy. Furthermore, the reported number of records identified for database searches within the economic evaluations search in the PRISMA flow diagram (CS Appendix G, Figure B.5.38[16]) appeared higher than the sum of the strategies reported, it is unclear where the error lies.
- The original MEDLINE and Embase strategies provided at clarification were run as separate searches, along with an additional search of NHS EED not reported in the original CS. Whilst a host was not named for the NHS EED search the syntax appears to be that of Ovid.
- As in the clinical effectiveness section, Table B.5.22. (Appendix G) reported the number of included studies, rather than total number of hits recalled for the grey literature searches. The full numbers retrieved were provided in the PRISMA flow chart (Figure B.5.38.) (HTA agencies (n=181) and Conference proceedings (n=2,607)).

4.1.2 Searches performed for health-related quality-of-life section

Appendix H of the CS reported literature searches used to identify health utility values for adults (≥18 years) with AML. Searches were conducted in February 2020 and updated in June 2021. A summary of the resources searched are provided in Table 4.2. The following paragraphs contain summaries and critiques of all searches related to HRQoL presented in the company submission.

Table 4.2: Resources searched for health-related quality-of-life studies. Feb 2020 and June 2021

Search strategy element	Resource	Host/Source	Date range of most recent search	Date searched
Databases	Embase	Ovid	1974-2021/06/10	12.2.20 Updated 11.6.21
	MEDLINE & MEDLINE In-Process		1946-2021/06/10	12.2.20 Updated 11.6.21
	CENTRAL		Up to 05/21	12.2.20 Updated 11.6.21
Additional searches	ScHARRHUD	Internet	All years	21.6.21
	Bibliographies of relevant SR articles were reviewed to obtain any additional, relevant references.			
CENTRAL = Cochrane Central Register of Controlled Trials; SR = systematic review				

ERG Comment:

- As previously reported in Section 3.1, only the strategy reporting the last update search was provided in the CS for the combined Ovid search. The ERG included a request for the original strategy along with the other missing searches at clarification. However, the update search appeared to have been resubmitted in error, rather than the original search strategy in the response, therefore the ERG was unable to fully critique these searches
- The update search for MEDLINE, Embase and CENTRAL provided in the CS was run as a single search split into facets by resource. The strategy combined a good range of free text terms and the appropriate subject headings and field tags for each database. The strategy also utilised both the CADTH health utilities/quality of life filter and a second filter containing cancer specific utility terms and other relevant quality of life terms not found in the CADTH filter. As with previous combined searches, the length of the strategy combined with some redundant lines affected the transparency of the searches, however this is unlikely to have adversely impacted on the recall of results.

4.1.3 Inclusion/exclusion criteria for cost effectiveness

A SLR of cost effectiveness studies was conducted to inform the economic model structure. The aim was to identify published economic evaluations of interventions which address the decision problem. The eligibility criteria for the study selection were included in table B.5.23 of appendix G of the CS and appear relevant for the task at hand. In- and exclusion criteria for the review on cost effectiveness studies are presented in Table 4.3. Any non-English studies were excluded during screening.

Table 4.3: Eligibility criteria for the systematic literature reviews

	Inclusion criteria	Exclusion criteria
Patient population	Male and female adults (≥18 years) with de novo AML or AML secondary to prior myelodysplastic disease receiving high intensity first-line (induction with or without consolidation), with or without maintenance treatment	Patients <18 years Relapsed or refractory AML
Intervention	Any non-transplant therapy	SCT
Comparator	Any non-transplant therapy	SCT
Outcomes(s) (Published economic evaluations)	Comparison of costs and consequences	Studies that do not report any relevant outcomes
Study design	Primary studies, systematic reviews, and meta-analyses that include comparative economic analyses (cost-utility, cost-benefit, cost effectiveness,	Publications focusing on economic burden Assessments from HTA agencies without full reviewer's reports

	Inclusion criteria	Exclusion criteria
	cost-minimisation, or cost-consequence studies	
Language	English	Any other language
Publication Types	All full-text articles from 2005-present Abstracts published over past 2 years	Full-text articles pre-2005 Abstracts published more than 2 years ago
Based on CS Appendix Table B.5.23[16] AML = acute myeloid leukaemia; HTA = Health Technology Assessments; ISPOR EU = International Society for Pharmacoeconomics and Outcomes Research Europe; ISPOR US = International Society for Pharmacoeconomics and Outcomes Research United States; SCT = stem cell transplant		

ERG comment: The ERG agrees that the eligibility criteria are suitable to fulfil the company's objective to identify cost effectiveness studies. The rationales for excluding cost effectiveness studies after full paper reviewing are considered appropriate given the defined in- and exclusion criteria.

4.1.4 Inclusion/ exclusion criteria for health-related quality of life searches

Regarding HRQoL studies, inclusion and exclusion criteria are presented in CS appendix Table B.5.29. These were broadly in line with the in- and exclusion criteria presented in Table 4.1 but differed regarding in- and exclusion criteria in the outcomes category. For HRQoL studies, inclusion criteria included direct utility values at baseline and utility increments or decrements by health state (using a number of different generic and disease specific HRQoL measures). No separate systematic review was performed on cost and resource use. Instead, the company reviewed the four included health technology assessments relevant to the decision problem (two NICE and two CADTH assessments of midostaurin and gemtuzumab ozogamicin, respectively).

ERG comment: The ERG agrees that the eligibility criteria are mainly suitable to fulfil the company's objective to identify cost effectiveness studies. The rationales for excluding cost effectiveness studies after full paper reviewing are considered appropriate given the defined in- and exclusion criteria. However, it might have been useful to not exclude studies on HSCT as these may have provided necessary detail to inform a HSCT health state in the model. Regarding cost and resource use studies, no justification was provided for only including identified HTAs rather than also looking into other included cost effectiveness studies or conducting an independent review.

4.1.5 Screening and data extraction

Appendix G of the CS emphasises that *"Two reviewers independently reviewed the study records, citation titles, and abstracts identified in the literature search to assess study eligibility."*[16] This appears to suggest that the results of the search were independently screened but does not indicate how disagreements were resolved. It is further emphasised *"that Citations of potentially eligible articles were independently screened by two reviewers in full-text form for formal inclusion in the final review. Disagreements between reviewers were resolved during a consensus meeting or by a third reviewer, as necessary."*."[16] This could be read to suggest that once the two reviewers had determined which articles were eligible, the citations within them were independently screened by two reviewers in full text form with any disagreements resolved by consensus or by a third reviewer. It is not clear the proportions of

disagreements resolved by each method, nor was it clear if the third reviewer independently reviewed or not.

Studies that were deemed eligible for inclusion were then subject to the data extraction process. Appendix G of the CS clarifies that "*Data extraction was performed for the studies meeting the outlined inclusion criteria. Information from the full-text articles, conference abstracts, and HTA reports was extracted into an Excel-based data extraction form by one reviewer and validated by a second reviewer.*"[16]

ERG comment: Best practice requires that two reviewers independently conduct extraction.[21]

4.1.6 Conclusions of the cost effectiveness review

A total of 21 records were identified in the search – these consisted of seven published economic evaluations, 10 conference abstracts and four HTA reports. Although some economic evaluations focused exclusively on therapies in the induction or consolidation phases (CPX-351 and high dose arabinoside/daunorubicin, respectively), no studies were identified that evaluated the cost effectiveness of maintenance treatment specifically. Rather, in studies where maintenance treatment was included, economic evaluations were limited to understanding the impact of a therapeutic regimen across all treatment phases (i.e., induction, consolidation, and maintenance). The economic evaluations of midostaurin were the only studies that included maintenance treatment as part of the treatment regimen, with the remaining studies focusing on the therapeutic regimens of the induction/consolidation treatment phases. The nine included HRQoL studies reporting health utility values by health state exhibited considerable variation in study designs and utility values reported by health states. Utility values reported in the four economic studies (out of the nine included studies) were sourced from the literature but not included in the current review, as they did not align with the inclusion criteria. No conclusions were formulated regarding the review for costs and resource use.

ERG comment: Eligibility criteria were largely suitable for the SLR performed. However, there is some doubt over whether the most appropriate sources for costs and resource use were identified with this review.

4.2 Summary and critique of company's submitted economic evaluation by the ERG

4.2.1 NICE reference case checklist

Table 4.4: NICE reference case checklist

Element of HTA	Reference case	ERG comment on CS
Perspective on outcomes	All direct health effects, whether for patients or, when relevant, carers	In line with reference case
Perspective on costs	NHS and PSS	In line with reference case
Type of economic evaluation	Cost utility analysis with fully incremental analysis	In line with reference case
Time horizon	Long enough to reflect all important differences in costs or outcomes between the technologies being compared	In line with reference case

Element of HTA	Reference case	ERG comment on CS
Synthesis of evidence on health effects	Based on systematic review	In line with 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.	In line with reference case
Source of data for measurement of health-related quality of life	Reported directly by patients and/or carers	In line with reference case
Source of preference data for valuation of changes in health-related quality of life	Representative sample of the UK population	In line with reference case
Equity considerations	An additional QALY has the same weight regardless of the other characteristics of the individuals receiving the health benefit	In line with 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	In line with reference case
Discounting	The same annual rate for both costs and health effects (currently 3.5%)	In line with reference case

CS = Company Submission; ERG = Evidence Review Group; EQ-5D = European Quality of Life-5 Dimensions; HTA = Health Technology Assessment; NHS = National Health Service; PSS = Personal social services; QALY = Quality adjusted life year; UK = United Kingdom

4.2.2 Model structure

As none of the studies identified in the SLR focused on oral azacitidine as maintenance treatment, a *de novo* model was required to assess the cost effectiveness of oral azacitidine compared with relevant comparators. The analysis was based on a three-health state partitioned survival model, using a cycle length of 28 days to align with treatment cycles for therapies considered in the model and other existing AML models[22, 23]. The partitioned survival analysis approach was deemed in line with other AML submissions and the key objectives of AML maintenance treatment: preventing progression and prolonging life. In addition, the OS and RFS data were considered relatively mature. The model was developed in Microsoft Excel.

The model consisted of the following health states: RFS; relapse (according to IWG 2003 response criteria in AML[3]) and death (Figure 4.1). In the RFS health state, patients could be either on- or off-treatment with oral azacitidine. In the comparator arm (watch and wait with BSC), all patients were considered off-treatment. Patients who relapse cannot achieve remission (i.e., move back from relapse to RFS), but the company considered that any remissions would be captured through OS.

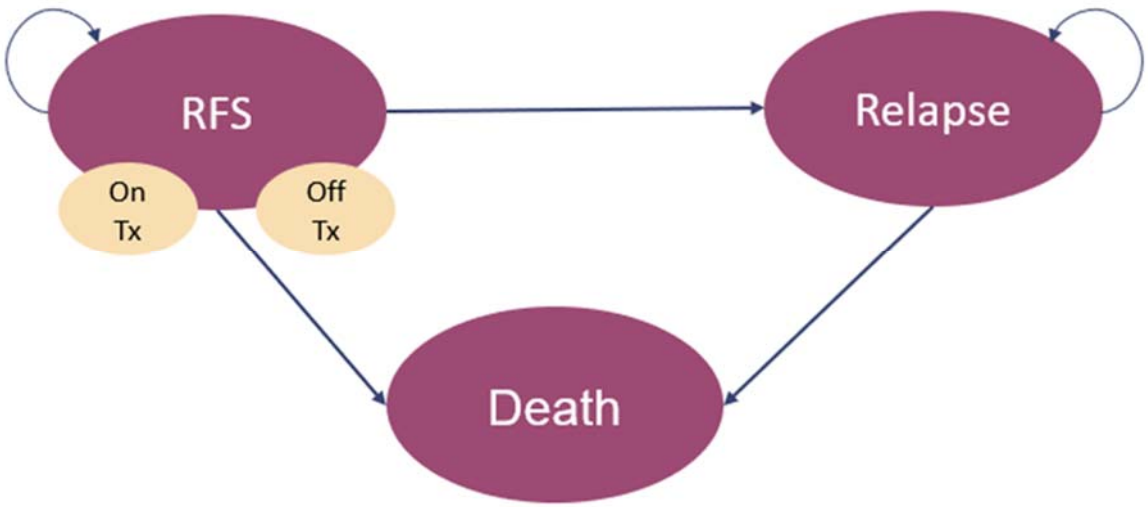
HSCT was modelled as part of subsequent treatments rather than explicitly as a separate heath state. The company considered that oral azacitidine is licenced for patients who are not suitable for transplant, and therefore it would be unlikely in clinical practice that patients will go on to receive HSCT after oral azacitidine unless they had relapsed. However, the company noted that 6.3% of the patients in QUAZAR AML-001 did receive HSCT in the oral azacitidine arm (mostly after relapse). In the placebo

arm, even more patients received HSCT (13.7%). These patients were not censored in the time-to-event analysis that informed the health state allocation. In addition, the company considered that including HSCT as a health state in the model would require inputs that were not captured in the QUAZAR AML-001 trial nor available from the literature for this population, such as the proportion of patients achieving a successful transplant and the outcomes following the transplant. The approach of not including HSCT as a health state aligned with other AML models, although not with the midostaurin submission TA 523[13]. AEs were modelled as events rather than health states.

The model base-case does not include a cure point, as this was not deemed appropriate in AML maintenance where the goal is to avoid disease progression and prolong life but not necessarily to cure patients (as would be the goal in induction treatments). A cure point of 5-years was explored as a scenario analysis, and patients who are assumed to be cured followed a standardised mortality ratio of 2.0 in line with the midostaurin submission.[13]

A lifetime horizon (i.e., 30 years) was applied to ensure all costs and QALYs were captured. This was considered appropriate given that the mean starting age of the cohort was 67.9 years. Therefore, by the end of the 30-year time horizon, the mean age was 97.9 years and <1% of patients in the model remained alive. Half-cycle correction was applied to the calculation of LYs and QALYs as transitions could occur continuously rather than at the start and end of a model cycle.

Figure 4.1: Model structure in CS



Based on CS Figure B.3.1
RFS = relapse-free survival, tx = treatment

ERG comment: The main concerns of the ERG relate to: a) implicitly modelling HSCT rather than including it as a separate health state; and b) the use of a partitioned survival model without exploring a state transition model approach alongside it.

a) HSCT was received by 6.3% of patients in the oral azacitidine arm and in 13.7% of patients in the placebo arm of the QUAZAR AML-001 ITT population. HSCT was not included as a separate health state but was implicitly included in the modelling through the survival analysis of the QUAZAR AML-001 ITT population (of which a proportion of patients received HSCT at some point). In addition, costs and a temporary disutility associated with undergoing HSCT were included

in the modelling. The ERG is concerned that this way of handling HSCT in the model may cause biases, one because survival analysis of OS and RFS may be biased, and two because no benefit in HRQoL post HSCT was captured in the model (instead HSCT was actually penalised with the short-term disutility). The company, in response to the clarification response[9], provided justification for their modelling decision:

- The company stated that insufficient data were collected to allow modelling of HSCT and that they were unaware of any published literature that reported HSCT data in patients who were initially in CR/CRi and ineligible for HSCT following induction therapy with or without consolidation. The ERG notes that the company's SLR excluded studies focusing on HSCT and that it therefore remains unknown whether this literature may have been available or not. The ERG acknowledges that there may be uncertainty (due to sparse evidence) about the impact of HSCT on survival and HRQoL. However, this uncertainty is currently not explored, rather it is assumed that HSCT has no positive impact on HRQoL and the chosen survival distributions may not appropriately capture the impact of HSCT as they do not fit well the end of the KM curves, particularly in the placebo arm.
- The company performed survival analysis on RFS and OS censored for HSCT (Appendix B.1 of clarification response[9]), which showed alignment with survival analysis in the ITT population. Regarding OS, the number of censored patients were not reported for the ITT population with data cut-off 8th September 2020, so the ERG could not assess how many more patients were in fact censored due to censoring for HSCT. For RFS, at 72 months of follow-up, compared with the ITT survival analysis, in the oral azacitidine arm [redacted] censored in this analysis, indicating that most patients receiving HSCT in this arm were already censored in the ITT population. In the placebo arm, [redacted] were censored at 72 months. The median survival is very similar between ITT and ITT censored for HSCT populations. However, the ERG notes that extrapolations may still differ, as HSCT will likely not affect median survival. It should also be noted that, based on visual inspection, all joint models and some individual models appeared to vastly under-estimate the latter parts of the KM curves in both arms, and this occurred more in the placebo arm than in the oral azacitidine arm. Since the proportional hazards assumption was violated and the log cumulative hazard plots did not indicate parallel lines (Figures 19 and 35 of the clarification response document), joint distributions may not be indicated.
- For OS, the [redacted] had the best statistical fit. However, the individual statistical fit per treatment arm was not presented and it is therefore impossible to assess which curves should be used individually. If the individual curves with the best joint statistical fit were to be used, this would indicate [redacted] even when HSCT was censored. The ERG considers that the company's new analysis is not supporting the statement that survival analysis is aligned between ITT and ITT censored for HSCT analyses. Furthermore, the ERG considers it concerning that this analysis indicates a waning of treatment effectiveness even when HSCT is censored.
- Regarding RFS, there appears to be no difference in KM curves between the ITT analysis and ITT censored for HSCT analysis, supporting the company's view that HSCT will likely not have an effect on RFS (as HSCT was primarily conducted as a salvage treatment following relapse).
- The company's approach aligns with other models in AML, although not with the midostaurin submission (TA523), and was supported by clinical opinion. The ERG considers that, whilst

this is useful to know, it does not address the uncertainty around the impact of differential HSCT in both treatment arms on the relative effectiveness of oral azacitidine versus watch and wait with BSC.

In conclusion, the impact of HSCT remains an area of uncertainty in this model that does not appear to be appropriately explored. The implications of a HRQoL effect of HSCT can be explored with this model, but to assess the impact on OS, additional detail needs to be provided on the survival analysis (including overlaying the KM curves in one plot, providing AIC/BIC fit for the individual distributions and showing all distributions in one plot, thus enabling differential selection of survival analysis, enabling this scenario in the economic model together with assumptions about survival for patients with HSCT and their HRQoL).

b) The NICE Decision Support Unit (DSU) TSD19 recommended the use of state transition models (STMs) alongside partitioned survival models (PSMs) to verify the plausibility of PSM extrapolations and to explore key clinical uncertainties in the extrapolation period. This was not done by the company, and the ERG was concerned that the chosen partitioned survival analysis model may not be fully validated. However, the ERG also considered that survival data were relatively mature and that a different modelling approach may therefore not significantly change model outcomes. The company, in their clarification response, did not provide an alternative state transition modelling approach for validation, but did elaborate on additional concerns around the use of state transition modelling in this case: there may be issues with estimating the "relapse to death" transition, which would be disproportionally driven by patients who relapsed earlier versus later (due to them generating more follow-up information) and which would be based on only those patients who had relapsed and died, thereby potentially being subject to selection bias. Furthermore, time-varying hazards would necessitate a semi-Markov or individual patient level simulation. Given the data maturity, any differences would only arise in the extrapolations. The ERG agrees that, given the likely difficulties with estimating the relapse to death transition, it is questionable whether an alternative modelling approach can provide better estimates of long-term survival. It is, however, important to explore this uncertainty with available alternative survival distributions.

4.2.3 Population

Consistent with the NICE scope, the marketing authorisation for oral azacitidine and the patient population in the QUAZAR AML-001 trial, the population considered in the CS (Section B.1.1) was adult patients with AML who achieved CR or CRi following induction therapy with or without consolidation treatment and who are not candidates for, including those who choose not to proceed to, HSCT. This was inclusive of both the categories of FLT3 mutations; tyrosine kinase domain (TKD) and internal tandem duplications (ITD).

Two patient groups were considered in the company's model:

- The QUAZAR AML-001 ITT population, compared with watch and wait with BSC, as informed by the placebo arm of the QUAZAR AML-001 ITT population (n=472)
- The QUAZAR AML-001 FLT3 subpopulation (FLT3-ITD and/or FLT3-TKD), compared with midostaurin, as informed by the indirect comparison ([redacted] in the MAIC)

The modelled baseline patient characteristics were presented in Table B.3.2 of the CS. These have been taken from the ITT population of QUAZAR AML-001 as they were considered to be representative of

the patient population in the UK that would be eligible for maintenance treatment with oral azacitidine. These patient characteristics were also used for the FLT3 subgroup analysis.

In an update post clarification response, the company also considered a scenario analysis using the EU population of the QUAZAR AML-001 study.

ERG comment: The main concerns of the ERG relate to: a) the generalisability of the ITT population to the UK setting and the relevance of the EU subgroup; b) results of the FLT3 subgroup being biased due to limitations with the indirect comparison; c) subgroup with at least 2 cycles of consolidation therapy.

a) The ERG questioned the generalisability of the QUAZAR AML-001 ITT population to the UK setting. The company maintained the ITT population in their base-case but argued that the EU subgroup may potentially be relevant because: there may be greater alignment in the diagnostic and treatment pathway between the UK and the rest of Europe; and one clinical expert highlighted broad differences in the healthcare environment in America versus Europe (in QUAZAR AML-001, 65% of patients were from Europe). In response to the clarification questions, the company provided a scenario analysis using the EU subgroup of QUAZAR AML-001 and stated that a further analysis of the FLT3 subgroup for EU patients would not have been possible due to sample size limitations. It should be noted that ideally, the company would also use the patient baseline characteristics of this subgroup in their scenario analysis. The ERG continues to use the ITT population in its base-case.
b) The company analysed the FLT3 subgroup using an indirect comparison of the QUAZAR AML-001 trial intervention arm with the RATIFY study to inform the comparator arm in which patients would be treated with midostaurin in the UK. The results of this subgroup analysis are likely to be extremely biased due to limitations associated with the indirect comparison (see Sections 3.4 and 4.2.6 for a more detailed critique). Patient baseline characteristics should be updated in this analysis to reflect those of the subgroup. An analysis excluding patients without consolidation therapy would be potentially useful.
c) The ERG doubted whether pre-treatment in the QUAZAR AML-001 trial was representative of clinical practice, given that the majority of patients in the QUAZAR AML-001 trial received only one cycle of consolidation therapy, and approximately 20% of the patients received no consolidation. In response to clarification question B5, the company argued that clinical experts confirmed that in UK clinical practice, *"there are patients that complete induction chemotherapy (and achieve remission) but may not receive consolidation chemotherapy"*. The ERG is still unclear whether QUAZAR AML-001 is indeed representative of UK clinical practice in terms of pre-treatment, especially given that a majority in the QUAZAR AML-001 trial did not receive at least 2 cycles of consolidation therapy. The company helpfully performed a scenario analysis using a subgroup with at least 1 cycle of consolidation therapy, which increased the ICER by approximately £5,000 per QALY gained. A scenario using a subgroup with at least 2 cycles of consolidation therapy was not provided, due to lack of data.

4.2.4 Interventions and comparators

The intervention is oral azacitidine with BSC. Oral azacitidine is available as 200 mg or 300 mg film-coated tablets to be taken orally with or without food. Consistent with the licence, the recommended starting dose of oral azacitidine is 300 mg once daily for the first 14 days of every 28-day treatment

cycle until disease progression or unacceptable toxicity. The SmPC of oral azacitidine recommends discontinuation upon blast counts >15% or unacceptable toxicities in the QUAZAR AML-001 trial.

The comparators considered were watch and wait with BSC and midostaurin. The NICE scope listed the following comparators: midostaurin; and established clinical management without oral azacitidine (which may include a "watch and wait" strategy with BSC, low dose cytarabine or subcutaneous azacitidine). The company justified the selection of the comparators with the fact that low dose cytarabine and subcutaneous azacitidine were not used in clinical practice as maintenance treatments for AML in the population eligible for maintenance treatment with oral azacitidine (as confirmed by two UK AML treating clinicians) and are therefore not considered as comparators to oral azacitidine. Watch and wait with BSC represents the SoC in current clinical practice because there are currently no approved or funded therapies indicated for this population for the independent maintenance treatment of AML in the UK. BSC included medications such as antibiotics, antifungals, and hydroxyurea (details in Table B.3.23 of the CS[1]).

For AML patients with mutations in FLT3, NICE recommended the use of midostaurin as an option for treating newly diagnosed acute FLT3 mutated AML patients. Midostaurin is an oral, type III, multi-target receptor tyrosine kinases (RTK) inhibitor that acts on FLT3 and multiple other RTKs. For patients in complete response, midostaurin is administered orally at 50mg twice daily as single agent maintenance treatment until relapse for up to 12 cycles of 28 days each.

ERG comment: The main concerns of the ERG relate to: a) exclusion of cytarabine and subcutaneous azacitidine and b) the appropriateness of the modelled subsequent treatments.

a) The ERG questioned whether the exclusion of cytarabine and subcutaneous azacitidine was indeed appropriate, given that these treatments were listed as comparators in the scope. In response to clarification question B3,[9] the company stated that these were neither recommended by NICE in the maintenance treatment population, nor was their use mentioned or endorsed as maintenance treatments in either the ELN (2017)[24] or British Committee for Standards in Haematology (BSCH) (2006) guidelines[25]. In addition, the company sought expert clinical advice from two UK AML clinicians, who confirmed that these treatments are not used in UK clinical practice for AML maintenance. The only uses of these treatments according to clinical experts were in very specific situations such as patients whose disease was in partial remission, or patients who showed signs of early relapse, which the company did not consider aligning with the definition of maintenance treatment.
b) The ERG questioned the appropriateness of the modelled subsequent treatments, including subcutaneous azacitidine and low dose cytarabine as well as salvage chemotherapy. The company, in response to clarification question B3b, stated that these were based on the QUAZAR AML-001 trial and validated by UK clinical experts. Both UK clinical expert opinions confirmed the use of subcutaneous azacitidine, low dose cytarabine and salvage chemotherapy as subsequent therapy in the treatment pathway in England. The ERG considers that the use of these subsequent treatments in the model is likely appropriate.

4.2.5 Perspective, time horizon and discounting

The analysis was constructed from the perspective of the NHS and the PSS in England and Wales. A discount rate of 3.5% per annum was applied for costs and benefits in line with the NICE reference case.

ERG comment: The approach is in concordance with the NICE reference case.

4.2.6 Treatment effectiveness and extrapolation

The main source of evidence on treatment effectiveness used for intervention and comparators is the QUAZAR AML-001 trial[26] (NCT01757535). This is an ongoing multicentre phase 3 trial of oral azacitidine plus BSC (n=238) versus placebo plus BSC (n=234) as maintenance treatment in patients with AML who have achieved CR or CRi after induction with or without consolidation chemotherapy and were not candidates for HSCT. The most recent data cut-off, from September 2020 (median follow-up time 51.7 months, data maturity of [redacted]), was used for the estimation of OS. Data from the July 2019 data cut-off (median follow-up time 41.2 months, data maturity of [redacted]) were used for the estimation of RFS, as the [redacted].

The nature of the hazard functions over time and the proportionality of the hazards between oral azacitidine and SoC as observed in the QUAZAR AML-001 trial data was examined using log-cumulative hazard plots and Schoenfeld residual plots. To estimate OS and RFS over the 30-year time horizon, parametric survival curves were fitted to QUAZAR AML-001 trial level data and used to extrapolate survival beyond the study time horizon. Six parametric models were considered (exponential, Weibull, Gompertz, log normal, log-logistic, and generalised gamma) and were assessed with regards to 1) visual inspection of model fit (using KM curves), 2) information criteria (AIC and BIC), 3) degree of agreement with log-cumulative hazard and Schoenfeld residual plots, 4) the marginal survival benefit in the observed and the extrapolated period, and 5) clinical considerations based on expert engagement, external literature and other relevant treatment and indication-specific domain knowledge. Details on each of these considerations are provided in Table 4.5.

Table 4.5: Selection of approach to estimate and extrapolate OS and RFS for ITT population

	OS (September 2020)	RFS (July 2019)
Fit to the observed data based on visual comparison with the Kaplan-Meier curves	Company stated that the individual and joint generalised gamma, and the individual and joint Gompertz had a good visual fit to the observed data. Company considered extrapolations of the individual generalised gamma and Gompertz, and the joint Gompertz implausible.	Company stated that all parametric models fitted do not fit well to the 'tail end' of the placebo RFS curve.
Fit to the observed data based on AIC and BIC	CS Table B.3.3. Joint models Generalised gamma had the lowest AIC and BIC, while the log-normal also provided a reasonable statistical fit. Individual models Placebo: generalised gamma had the lowest AIC and BIC, while the log-normal also	CS Table B.3.5. Joint models log-logistic had the lowest AIC and BIC, while the Gompertz also provided a reasonable statistical fit. Individual models Placebo: Gompertz had the lowest AIC and BIC, while

	OS (September 2020)	RFS (July 2019)
	provided a reasonable statistical fit. Oral azacitidine: AIC and BIC of generalised gamma, log-normal and log-logistic were within [redacted] respectively.	the log-logistic also provided a reasonable statistical fit. Oral azacitidine: AIC and BIC of Gompertz and log-logistic were within [redacted].
Agreement with log cumulative hazard and Schoenfeld residual plots	Company stated that proportional hazards assumption was violated, suggesting that log-logistic, log-normal, and generalized gamma were most appropriate.	Company stated that proportional hazards assumption was violated, suggesting that log-logistic, log-normal, and generalized gamma were most appropriate.
Marginal survival benefit in the observed and extrapolated period	Company stated that marginal survival benefit was broadly consistent between observed and extrapolated period for joint generalised gamma and other models.	Company stated that the log-logistic model exhibited much lower marginal survival in the extrapolation versus the observed period.
Clinical considerations	Company provided clinical reason (trial hazards decrease over time due to patient heterogeneity with respect to hazards/prognosis) why crossing of curves may be inevitable but spurious. Expert consultations suggested that crossing was not considered clinically likely	Company stated that the placebo RFS curve appeared to plateau sharply, and even cross the RFS curve of oral azacitidine. Expert consultations suggested such a cross-over was not clinically plausible, and it was more likely that this was due to statistical noise driven by small sample size
Base-case approach	Joint generalised gamma model	Joint log-logistic model
Scenario analyses	Cure model and hybrid model	Joint log-normal model
AIC = Akaike Information Criterion; BIC = Bayesian information criterion; CS = Company submission; OS = overall survival; RFS = Relapse free survival		

4.2.6.1 FLT-3 subgroup

The process of selecting the approach to estimate and extrapolate OS and RFS for oral azacitidine and midostaurin in the FLT-3 subgroup (based on an indirect treatment comparison between the QUAZAR AML-001 and RATIFY studies) is summarized in Table 4.6.

Table 4.6: Selection of approach to estimate and extrapolate OS and RFS for the FLT-3 subgroup.

	OS (September 2020)	RFS (July 2019)
Fit to the observed data based on visual comparison with the Kaplan-Meier curves	Not explicitly discussed.	Not explicitly discussed.
Fit to the observed data based on AIC and BIC	Appendix D Table B.5.16. Standard parametric models: Generalised gamma had the best statistical fit. Spline models: AIC of all spline models within [redacted]. 1 internal knot, normal linear predictor and 1 internal knot, odds linear predictor had the lowest BIC.	Appendix D Table B.5.16. Standard parametric models: Gompertz had the best statistical fit. Spline models: 1 internal knot, hazard linear predictor and 1 internal knot, odds linear predictor had the best statistical fit.
Agreement with log cumulative hazard and Schoenfeld residual plots	Company stated that proportional hazards assumption is violated, suggesting that proportional hazards models and AFT models were considered less appropriate. Individual models were fit to the QUAZAR AML-001 FLT3 IPD and digitized KM data from the RATIFY maintenance subgroup trial.	Company stated that proportional hazards assumption is violated, suggesting that proportional hazards models and AFT models were considered less appropriate. Individual models were fit to the QUAZAR AML-001 FLT3 IPD and digitized KM data from the RATIFY maintenance subgroup trial.
Marginal survival benefit in the observed and extrapolated period	Not explicitly discussed.	Not explicitly discussed.
Clinical considerations	Company stated that for the generalized gamma, the oral azacitidine arm remained apart from the no active treatment arm aligning with the clinical expectations. The Gompertz was not considered given the observed plateau. Company stated that all of the spline models led to crossing of curves, which	Company considered the generalized gamma to be a plausible option given there was no plateau. Although the Gompertz was best fitting, the plateau seen in the extrapolations was not considered to be plausible by the company. Company stated that the hazard linear predictor led to a divergence of oral

	OS (September 2020)	RFS (July 2019)
	was not expected based on clinical opinion	azacitidine and no active treatment curves after the point of crossover. Informed by clinical advisor opinion, the one knot odds linear predictor was deemed to be plausible. The one knot normal linear predictor may also be considered.
Base-case approach	Generalised gamma model	1 knot odds linear model
Scenario analyses	Log-normal model and 1 knot odds linear spline model	Generalised gamma model
AIC = Akaike Information Criterion; BIC = Bayesian information criterion; CS = Company submission; OS = Overall survival; RFS = Relapse free survival		

ERG comment: The main concerns of the ERG relate to: a) potential bias resulting from QUAZAR AML-001 trial limitations, b) use of the consolidation subgroup and lack of detail in the survival analyses c) treatment waning of oral azacitidine d) bias and lack of detail in survival analyses of the FLT3 subgroup, and e) survival analyses in the EU subgroup.

a) The company used the QUAZAR AML-001 trial[26] as its main source of evidence to inform the economic model. However, as critiqued in Sections 3.2 and 3.3, this trial has several limitations likely inducing biased results. Therefore, all cost effectiveness analyses are also subject to potential bias resulting from these trial limitations.
b) The ERG noted that, although consolidation therapy following induction therapy is recommended in NHS clinical practice,[8] approximately 20% of patients in the QUAZAR AML-001 trial did not receive any consolidation therapy. Upon request, the company provided survival analyses only including patients that received at least 1 cycle of consolidation therapy (comprising 78% of the oral azacitidine arm and 82% of the placebo arm). The company argues this analysis to be inappropriate because, based on consultation of experts in AML, there are patients that complete induction therapy but may not receive consolidation therapy. Although it is unclear to the ERG how many cycles of consolidation therapy are recommended in NHS clinical practice, it considers the subgroup including patients that received at least 1 cycle of consolidation therapy to be more appropriate for survival analyses than the QUAZAR AML-001 ITT population and adopted this subgroup in the ERG base-case. In response to clarification question B5, the company selected the joint generalized gamma and joint log-logistic for the modelling of OS and RFS respectively in the consolidation subgroup. The ERG agrees that based on the AIC and BIC statistics reported in Table 29 and Table 30 of the clarification response, these curves appear to have the best statistical fit to the trial data. However, although the company states that selection of curves was based on the criteria described in the NICE DSU TSD14 and that it overall aligns with the assessment for the ITT population, full details of these criteria were not provided in response to clarification question B5. Hence, the ERG was unable to determine the most appropriate curves for the modelling of OS and RFS (e.g., whether individual modelling would be more appropriate than joint modelling). Therefore, the ERG adopts the company's approach of modelling OS (joint generalized gamma) and RFS (join log-logistic) in its base-case, but stresses that full details of the NICE DSU TSD14 criteria are necessary for the ERG to perform a thorough assessment. Furthermore, more detail on

proportions of patients receiving zero, one and two+ consolidation cycles in UK clinical practice should be provided.

c) The company assumed no treatment waning of oral azacitidine in its base-case. In clarification question B9, the ERG requested that the company justify this assumption by providing HR plots with numbers of patients at risk over time. The company provided these plots for OS and RFS and stated that given the use of AFT models in their base-case (joint generalised gamma and log-logistic), the HRs varied over time and exhibited a natural waning effect. The company also stated that the treatment waning effect was further increased by incorporating general population mortality by deterministically setting the HR to 1 at 150 months. In addition, in response to clarification question B9c, the company explored treatment waning by selecting individual curves for the extrapolation of OS (individual log-normal) and RFS (individual log-logistic), increasing the ICER to £54,017. The ERG agrees that survival distributions can be chosen to reflect treatment waning, making additional treatment waning assumptions likely obsolete. However, a HR plot over time should be provided for the consolidation subgroup as well, comparing the modelled HR to the one observed in the trial.
d) Survival analyses of the FLT3 subgroup are likely to be extremely biased due to limitations associated with the indirect comparison (see Sections 3.4 and 4.2.6 for a more detailed critique). Although the ERG appreciates that the company used several criteria from NICE DSU TSD14 for selection of the most appropriate model, the ERG considers any approach or chosen model likely to be biased. In addition, details for the survival analyses of OS and RFS in the FLT3 subgroup were lacking, including 1) log-cumulative hazard plots, 2) AIC/BIC statistics for individual models, 3) plots showing all joint models in one plot and 4) evaluation of criterion 5 (OS/RFS gain pre and post extrapolation).
e) In response to clarification question B4, the company's results for the EU subgroup (choosing the joint generalised gamma to model OS and the joint log-logistic to model RFS) indicated [redacted], the latter being driven by [redacted] in the RFS: off treatment for watch and wait with BSC in the EU subgroup compared with the ITT population. The ERG considers that these results are contingent on the selected survival distributions and could be quite different when other choices are made. Whilst the company chose the models with the best global statistical fit, it was unclear whether the joint modelling was really appropriate. Especially for RFS, the log cumulative hazard plot did not really exhibit straight lines. Choosing individual distributions for RFS may have a significant impact on the ICER. For example, choosing the best fitting Gompertz to model RFS increases the ICER to £65,497 per QALY gained. It must be noted that the Gompertz likely lacks face validity as it predicts a very flat tail for the placebo arm, but this shows that results are highly dependent on model choice. The use of individual log-logistic distributions for RFS in both arms (best statistical fit in the oral azacitidine arm and second-best fit in the placebo arm) only increased the ICER marginally.

4.2.7 Adverse events

For the ITT population, the model included grade 3 and 4 AEs occurring in 5% or more of the patients according to the QUAZAR AML-001 trial. Leukopenia was excluded because, according to clinical advisors, the impact would already be captured by other AEs. For the FLT3 population, grade 3 and 4 AEs occurring in more than 10% of patients treated with midostaurin were included based on the ITT population of the maintenance phase in the RATIFY trial.[27] AE rates for patients treated with oral azacitidine were obtained from the FLT3 subgroup of the QUAZAR AML-001 trial. Following expert

opinion, a one-week duration was assumed for all AEs. All AE disutilities were applied in the first model cycle.

ERG comment: The main concerns of the ERG relate to: a) the modelled duration of AEs, b) the modelled frequency of AEs, c) the inclusion of grade 1 and 2 AEs and d) the application of different cut-off points for the inclusion of AEs in oral azacitidine and midostaurin.

a) The company applied a 1-week duration to all AEs based on expert opinion. The ERG questioned the validity of this duration in clarification questions B13 and B14 a) and b). The company responded by submitting the report of the expert interviews [#302], detailing that the 1-week duration was a simplifying assumption, justified by the duration of some AEs being longer and of some AE being shorter than 1 week. This simplifying assumption seems arbitrary to the ERG. The expert opinion report [#302] further details that neutropenia is rather an ongoing problem, which does not resolve on its own. This detail of the expert opinion, however, is not applied in the model as the duration of neutropenia in the model is also 1 week.
b) The company applied a maximum frequency of 1 per patient per AE. The ERG questioned the validity of assuming that each patient can experience each AE only once. Therefore, the ERG asked for clarification including a scenario analysis in clarification question B14. The company responded that this would be a plausible assumption as 1) it is not affected by discounting as all AEs are applied in the first cycle and therefore conservative, 2) this was a common assumption, and 3) the strategy in clinical practice would be to predict and prevent AE and thereby avoiding re-occurrence. The ERG disagrees with the reasoning given by the company, as it seems questionable that clinical practice would be 100% effective at preventing reoccurrence of AEs. The company did not comply with a request for a sensitivity analysis applying AEs in every cycle stating that this was an unreasonable assumption. While the ERG agrees that this would be an unreasonable assumption to implement in the base-case, the ERG would have nevertheless found it worthwhile to know how sensitive the model results would have been to such a change.
c) The company included only grade 3 and 4 events in their base-case. The ERG requested the inclusion of grade 1 and 2 AEs for two reasons: 1) due to a presumed shorter duration of grade 1 and 2 AEs, any effect they could have on utility measurement is less likely to be captured, especially because the quality-of-life measurement is on day 1 of the on-treatment period and after a period without treatment and 2) there was a higher prevalence of low-grade AEs in the treatment arm than in the comparator arm. The company complied with the request. The scenario analysis raised the ICER. The ERG did not consider grade 1 and 2 AEs in their analyses.
d) The company applied different cut-off points for oral azacitidine (≥5%) and midostaurin (≥10%) regarding the inclusion of AEs. The ERG asked for clarification regarding the reasons for applying different cut-off points (clarification question B12). The company responded that there was a lack of published evidence for AEs in midostaurin but that the higher cut-off point for midostaurin would generally increase the company's ICER and therefore be conservative.

4.2.8 Health-related quality of life

HRQoL in the ITT population of the QUAZAR AML-001 trial was measured using the EQ-5D-3L on each day one of the 28-day cycle. For patients receiving oral azacitidine this implied that the HRQoL assessment was on the first treatment day of every model cycle. In total 442 subjects were included in the analysis. Utilities were calculated using the UK value.[28]

Utility values for the RFS on- and off-treatment health states were derived by applying a linear mixed effects model with random intercepts. The optimal model was defined based on the level of significance, magnitude of coefficients and AIC and BIC statistics. The company used a model with only an intercept (Model 1) in their original model. However, in response to clarification question B.16, the company provided a corrected base-case using a model that included AEs as covariate (Model 3) but using only its intercept. This was done to reflect the utility excluding AEs that would be included separately through disutilities sourced from the literature. Two other models were considered, one including the treatment arm and the other including a treatment arm and AEs. The same utility value (Model 3 intercept [redacted], SE [redacted]) was applied to RFS regardless of whether the patient was on- or off-treatment, as validated by expert opinion.

In Section B.3.4.1 the company states that the QUAZAR AML-001 trial did not capture HRQoL post relapse. According to the CS, the SLR identified three relevant sets of utilities.[29] [22] [30] The study by Joshi [30] was selected out of these studies because it obtained utilities for AML using a composite time trade-off methodology from the UK general population.

The utility difference between RFS and relapse in Joshi[30] was subtracted from the RFS utility calculated with the linear mixed effects model, resulting in a utility upon relapse of [redacted].

A one-off 28-day utility decrement of 0.21[31] was applied to the proportion of patients receiving HSCT. No utility benefits resulting from subsequent treatments including HSCT were included in the modelling.

Health state utility values

A summary of all utility values used in the company's base-case cost effectiveness analysis and potential alternative values is provided in Table 4.7.

Table 4.7: Health state utility values

Health state	Utility value (base-case)	Reference	Utility values (Joshi)	Reference	Utility values (Tremblay)	Reference	Utility values (Stein)	Reference
RFS: on treatment	[redacted]	3	0.89 (0.15)	30	0.81 (0.2)	22	0.87 (0.2)	32
RFS: off treatment	[redacted]	3	0.89 (0.15)	30	0.83 (0.2)	22	0.87 (0.2)	32
Relapse	[redacted]	Based on 301	0.51 (0.46)	30	0.53 (0.2)	22	0.62 (0.2)	32

Based on CS Section B.3.4.3 and Model page 'Utilities'[1]
CS = Company Submission ; RFS = relapse-free survival
Note: Calculations made by applying the difference of relapse utility and relapse free utility from Joshi to the RFS utility from the QUAZAR AML-001 trial.

Disutility values

In the first model cycle and for the duration for only one model cycle disutilities were applied as a one-off for the following AEs: neutropenia, thrombocytopenia, anaemia, febrile neutropenia, diarrhoea, vomiting, nausea, and fatigue.

ERG comment: The main concerns of the ERG relate to: a) the calculation of the relapse health state b) the timing of the measurement of HRQoL, c) the application of a HRQoL benefit following HSCT, and d) age-adjusted utility values.

a) The ERG questions the choice of the source for the calculation of the utility of patients upon relapse. In Section B.3.4.1 the company states that their study has not assessed relapse utility, however, Appendix B.15 Table 78 of the clarification response[9] shows the EQ-5D measurement per cycle per treatment for a number of patients at a number of measurement points. The mean remains well above [redacted] for most measurement points. The ERG acknowledges that this may be biased as only few patients were observed in the relapse state and possibly not for long. Further justification for not using the relapse utility from the trial would be welcome.
   Instead of using their own data, the company calculated the relapse utility based on Joshi[30] arguing that this was the best choice, because the article used a time-trade off utility from a UK population. The ERG questioned why a composite time trade-off methodology was particularly desirable given that it was not referred to by the NICE reference case. Furthermore, the sample size in relapse in Joshi is small (n=23) resulting in a large standard error. The company considered Stein[32] and Tremblay[30] as alternative sources for relapse utilities. Although Tremblay was suggested to be used by experts, both are not ideal measurements either as they were elicited in US populations.
   While it is unclear what the most appropriate utility value would be, the relapse utility from Tremblay was implemented in the ERG base-case as utility measurements were used in TA523 and were mapped onto the EQ-5D. The company should further explore calculating the relapse utility based on the QUAZAR AML-001 trial data.
b) The QUAZAR AML-001 trial measured the HRQoL on every day 1 of a 28-day treatment cycle. Oral azacitidine was given on day 1 to 14 (or 1-7/1-21) of every treatment cycle, followed by a period of 14 days without treatment. Therefore, TRAEs are likely to occur during the first 14 days of every 28 days cycle and to diminish in the 14 days of rest thereafter. Therefore, the ERG believes that the utility estimates are likely biased. In response to clarification question B19 about this issue, the company responded that it believes that the impact is marginal as AE disutilities from the literature would capture the effect of the timing of the measurement.
c) According to the company, 6.3% of patients treated with oral azacitidine and 13.7% of the watch and wait with BSC patients from the QUAZAR AML-001 ITT population underwent HSCT. The company applied cost and disutilities for patients receiving the treatment, but no utility benefit after having gone through with HSCT. Patients are in that case penalized for receiving HSCT while in previous technical appraisals (TA523[13] and TA642[33]) a curative effect of HSCT was assumed. Table 46 of the CS related to TA523 shows that patients were assumed to return to baseline RFS utility after treatment with HSCT for a period. Following a full recovery from HSCT a utility increase of 0.016 was modelled. Table 29 of the CS related to TA642 suggests that utility was assumed to increase by 0.05 from RFS after HSCT. Upon request to apply a utility benefit for having undergone HSCT in clarification question B17, the company replied that due to lack of evidence adding an arbitrary utility benefit would increase uncertainty without increasing clarity. The ERG finds this argument unconvincing as the company's way of modelling is likely inducing

bias (i.e., assuming only a disutility and no utility benefit). The ERG considers that patients are likely to experience a net utility benefit after undergoing HSCT. The disutility applied was therefore removed in the ERG base-case analysis. Further, to model the effect of assumptions of previous STAs (TA523[34] and TA642[33]) a scenario analysis was conducted, in which the effect of applying a return to RFS utility after relapse was explored for the proportion of patients undergoing HSCT. The scenario analysis assumed the effect of the post HSCT utility to last 1.67 years, which was the average life expectancy after relapse.

d) The updated base-case RFS: on treatment and RFS: off treatment utility was [redacted]. Considering the median age of the QUAZAR AML-001 ITT population (68 years), this utility was higher than the age-adjusted general population norm (0.785)[35] in the UK. The ERG requested a justification and a scenario analysis capping the utility at general population levels in clarification question B18. The scenario analysis increased the ICER. The ERG maintains the company's trial utility values in their base-case (since they are directly estimated for the population in question) but wishes to highlight that these appear relatively high.

4.2.9 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 AEs.

Unit prices were based on the eMIT 2020,[36] NHS reference prices,[37] and British National Formulary (BNF).[38]

Resource use and costs data identified in the review

No SLR was performed specifically to inform cost and resource use. According to the CS, the SLR of cost effectiveness analyses identified seven studies reporting UK relevant resource use and cost information. Out of these, the company only considered the four identified HTAs (two by NICE and two by CADTH, on midostaurin and gemtuzumab ozogamicin) to identify cost and resource use in their submission. However, only one HTA (TA523[34]) was used to inform end-of-life costs.

Treatment costs

Oral azacitidine with Patient Access Scheme (PAS)

Costs for the acquisition of oral azacitidine were calculated based on treatment dose, number of administrations per cycle and the number of cycles in which it was applied. As per the SmPC the starting dose was 300mg repeated through days 1 to 14 of each 28-day treatment cycle. The model assumed a relative dose intensity of [redacted] based on the QUAZAR AML-001 trial. The price for 14 doses of 300 mg was set at [redacted]. Including [redacted] this amounted to a price per cycle of [redacted]. These costs were applied to all patients in the intervention population who were in the RFS on-treatment health state.

Midostaurin

For the comparator population of the FLT-3 subgroup, midostaurin was assumed to be administered twice daily as a single agent until relapse. The cost of midostaurin per unit was £100.18,[37] with four units administered daily to a cost-per cycle of £10,658.89. These costs were applied to the subgroup analysis of the FLT-3 comparator population in RFS.

Treatment administration costs

Treatment administration costs were applied to patients receiving oral azacitidine, midostaurin and subsequent treatments. Costs for IV and subcutaneous administration were incurred at each treatment initiation. For oral chemotherapies, the cost was included per cycle.

Pre-medication

Ondansetron was set to be given to patients as pre-medication. It was assumed that patients receiving oral azacitidine would receive 8 mg of ondansetron twice a day for 5 days before the start of each on-treatment period. For midostaurin it was assumed that patients would receive 2.5 days of ondansetron twice a day each treatment cycle. Pre-medication drug costs were retrieved from the electronic market information tool (eMIT) 2020.[36]

Disease management cost

Rates of resource use for disease management were applied based on the QUAZAR AML-001 trial and further guided by expert opinion. Only the use of red blood cell and platelet transfusion was informed exclusively by UK clinical expert opinion. Resource use and treatment administration costs were obtained from the NHS reference cost 2019/2020.[37]

Best supportive care (BSC)

All patients, except for those in the RFS health state, were modelled to receive BSC. The calculation of BSC was largely based on UK expert opinion.[39] Dosing regimens were based on the respective SmPC. Acquisition costs were sourced from eMIT 2020[36] and the online BNF 2021.[38] Treatment administration costs were sourced from the NHS reference costs 2019/2020.[37] An overview of BSC resource use and associated costs can be found in CS Table B.3.22. and CS Table B.3.23.[1]

Table 4.8 summarises costing information for the medicine which was included in this submission.

Table 4.8: Medicine cost table

Drug name (type)	Admin route	Dose per tablet	Units per pack	Cost per pack (£) (list price)	Source
Intervention					
Oral azacitidine	Oral	300 mg	14	[redacted]	BMS data on file
FLT3 comparator					
Midostaurin	Oral	25 mg	56	£5,609.94	[38]
Premedication					
Ondansetron	Oral	8 mg	10	£0.93	[36]
Best supportive care					
Hydroxycarbamide	Oral	500 mg	100	£9.61	[36]
Ciprofloxacin	Oral	500 mg	10	£3.08	[36]
Posaconazole	Oral	100 mg	24	£175.32	[36]
Fluconazole	Oral	200 mg	7	£0.51	[36]
Tranexamic acid	Oral	500 mg	60	£7.98	[36]

Drug name (type)	Admin route	Dose per tablet	Units per pack	Cost per pack (£) (list price)	Source
Subsequent therapy					
Low dose cytarabine	Subcutaneous	100 mg	5	£22.52	[36]
Injectable azacitidine	Subcutaneous	100 mg	1	£220	[38]
Subsequent therapy: Salvage chemotherapy					
Daunorubicin	Intravenous	20 mg	10	£715	[38]
Cytarabine	Intravenous	500 mg	5	£22.38	[36]
Based on clarification response Table 52[9]					

Subsequent therapy

The share of patients that received subsequent therapies and the mix of subsequent therapies was informed by the QUAZAR AML-001 trial and validated by clinical advisors[39]. Cytarabine, injectable azacitidine, and salvage chemotherapy (daunorubicin and cytarabine) were considered as subsequent treatments. Salvage chemotherapy was assumed to consist of 3 days of daunorubicin and 7 days of cytarabine. Acquisition costs were sourced from eMIT 2020[36] and the online BNF.[38] An overview of subsequent treatment regimens, drug costs and treatment frequency are provided in CS Table B.3.24, CS Table B.3.25 and CS Table B.3.26.[1]

Treatment cost for HSCT was applied for the patient share which was modelled to receive HSCT. The cost for HSCT was taken from the NHS reference costs 2019/2020.[37]Table 4.9 gives an overview of the proportion of patients receiving HSCT as subsequent treatment and associated cost per treatment arm. Notably, a higher proportion of the watch and wait plus BSC group received HSCT than of the ITT population.

Table 4.9: HSCT use and cost

Parameter	Treatment	Proportion	Sources
Proportion of patients receiving stem cell transplant	Oral azacitidine	6.3%	QUAZAR CSR (Table 14.1.10.2)
	Watch and wait plus BSC	13.7%	QUAZAR CSR (Table 14.1.10.2)
	Oral azacitidine (FLT3)	6.3%	Assumed same as ITT
	Midostaurin (FLT3)	5.8%	[27]
Parameter	**Unit costs (£)**	**Source**	
Stem cell transplant	15,065.00	NHS reference costs 2019-2020. Peripheral blood stem cell transplant, autologous, 19 years and over. Code SA26A	
Based on CS table B.3.27[1] CS = Company Submission, ITT = intention to treat; NHS = National Health Service			

End-of-life cost

Upon death, end-of life costs were applied. In line with TA523, the cost was sourced from Nuffield 2014[40] and inflated to 2019/2020 based on the HCHS inflation index.[41] This resulted in end-of-life costs of £14,708.43.

Adverse Event costs

AE costs were applied for the following AEs: neutropenia, thrombocytopenia, anaemia, febrile neutropenia, diarrhoea, vomiting, nausea, and fatigue. An overview of AE frequency and associated cost can be found in CS Table B.3.28 and CS Table B.3.29.

In line with TA523,[34] costs for adult febrile neutropenia were assumed to be the same as those for AML with CC score 0-1. The average between in- and outpatient costs were applied to the model, where the weights were based on UK clinical expert opinion. Costs were informed by the NHS reference costs 2019/2020. Costs were applied in the first model cycle.

ERG comment: The main concerns of the ERG relate to a) lack of transparencies in the selection of evidence, and b) the use of different inflation indices for end-of-life costs.

a) There is uncertainty about the selected resource use estimates. The company did not perform a SLR on cost and resource use and relied heavily on expert opinion to inform resource use estimates. In response to clarification B6c and B13 the company provided a report of the expert opinion. The ERG remains unsure whether all resource use estimates are appropriate. Notably:
    i. On page 15 of the report of the expert opinion under "Resource use" it is stated that for midostaurin fewer than 21.8-22.7% would require red cell transfusions. The CS states that in the model, 21.8% of patients receiving midostaurin received red cell transfusions.
    ii. On page 16 of the report of the expert opinion under "RBC (red blood count) and platelet transfusion" it is stated that experts recommend 1 unit of RBC transfusion per cycle. The CS states that patients would, based on expert opinion, receive 2 units of transfusions per cycle.
    iii. On page 16 of the report of the expert opinion under "Subsequent therapies" it is stated that intensive chemotherapy could be removed as a subsequent treatment because patients would likely not be eligible. As a result, only decitabine has been removed from the subsequent treatment options while vidaza (injectable azacitidine) and salvage therapy (both are also chemotherapies) have remained as treatment options. It is unclear whether this is in line with expert opinion. If it is not in line with expert opinion, it is unclear why other chemotherapies were not also excluded.
    iv. On page 16 of the report of the expert opinion under "Subsequent therapies" it is stated that salvage chemotherapy may be more accurately described by applying therapies such as VenAza or FLAG-IDA and that the company should consider using half of the estimates shown to them. The company has implemented neither different therapies nor halved their estimates. It is unclear what guided the modelling decisions the company made in this case.

    Further clarification by the company and/or amendments to the model would be useful.

b) Section B.3.5.4.1 describes the calculation of end-of-life care costs. The Section describes that two different inflation indices were applied. It is unclear to the ERG why two different inflation indices were used.

5. COST EFFECTIVENESS RESULTS

5.1 Company's cost effectiveness results (updated in response to clarification)

The updated base-case cost effectiveness results (probabilistic) indicated that oral azacitidine (with PAS) is both more costly (additional costs of [redacted]) and more effective (incremental QALYs of [redacted]) than watch and wait plus BSC amounting to an ICER of £48,147 per QALY gained (Table 5.1). Moreover, the 95% percentiles for the probabilistic incremental costs and QALYs were ([redacted]) and ([redacted]) respectively. The probabilities of oral azacitidine being cost-effective, at thresholds of £20,000, £30,000, and £50,000 per QALY gained, compared to watch and wait plus BSC are 2%, 9% and 52% respectively.

Overall, the technology is modelled to affect QALYs by:

- Increased RFS, with an incremental of 0.854 years (80% of total incremental LYs) in the oral azacitidine arm (2.085 years) compared with watch and wait with BSC arm (1.232 years).
- Increased post-relapse survival, with an incremental of 0.211 years (20% of total incremental LYs) in the oral azacitidine arm (1.779 years) compared with watch and wait with BSC arm (1.568 years).

Overall, the technology is modelled to affect costs by:

- The higher drug costs (additional cost of [redacted], [redacted] of total incremental costs) and disease management costs (additional cost of [redacted]) in RFS on-treatment compared with watch and wait plus BSC.
- The lower disease management costs (reduced cost of [redacted]) in RFS off-treatment compared with watch and wait plus BSC.

For the FLT3 subgroup, midostaurin was dominated by oral azacitidine, the updated probabilistic ICER for oral azacitidine versus watch and wait plus BSC in this subgroup was £25,403 per QALY gained (Tables 5.1 and 5.2).

Table 5.1: Probabilistic base-case results with oral azacitidine PAS (updated in response to clarification)

Technology	Total costs (£)	Total LYG	Total QALYs	Incremental costs (£)	Incremental LYG	Incremental QALYs	ICER versus baseline (£/QALY)
Watch and wait plus BSC	[redacted]	2.815	[redacted]	-	-	-	-
Oral azacitidine	[redacted]	3.877	[redacted]	[redacted]	1.06	[redacted]	48,147

BSC = best supportive care; LYG = life years gained; ICER = incremental cost effectiveness ratio; PAS = Patient Access Scheme; QALY = quality adjusted life year

Table 5.2: Probabilistic results with oral azacitidine PAS for the FLT3 subgroup

Technology	Total costs (£)	Total LYG	Total QALYs	Incremental costs (£)	Incremental LYG	Incremental QALYs	ICER versus baseline (£/QALY)	Pairwise ICER versus oral azacitidine
Watch and wait plus BSC	[redacted]	2.69	[redacted]	-	-	-	-	25,403
Oral azacitidine	[redacted]	4.78	[redacted]	[redacted]	2.09	[redacted]	25,403	-
Midostaurin	[redacted]	3.58	[redacted]	[redacted]	0.89	[redacted]	272,290	Oral azacitidine is dominant

BSC = best supportive care; LYG = life years gained; PAS = Patient Access Scheme; ICER = incremental cost effectiveness ratio; QALY = quality adjusted life year

ERG comment: The main concern of the ERG relates to the lack of a fully incremental analysis for all comparators in the FLT3 subgroup. Contrary to the final scope issued by NICE, a full incremental analysis of oral azacitidine, midostaurin and watch and wait plus BSC was not performed for the FLT3 subgroup and was also not enabled as an option in the economic model.

5.2 Company's sensitivity analyses

The company performed and presented the results of PSA, DSA as well as scenario analyses. The parameters that had the greatest effect on the ICER based on the company's DSA were:

- Health state utility – RFS on treatment
- Health state utility – RFS off treatment
- Oral azacitidine relative dose intensity

CS scenarios that have the greatest impact on the ICER (not including scenarios related to discount rates and time horizon) were:

- Using the QUAZAR AML-001 Europe only population (decreased ICER to [redacted])
- Cure modelling with a 5-year cure point (decreased ICER to [redacted])
- Utility values based on Joshi 2019 for all health states (decreased ICER to [redacted])

ERG comment: The main concern of the ERG relates to the difference in results between the DSA and the PSA. Compared to the company's original deterministic analysis (ICER of £49,704 per QALY gained), the result of the original PSA based on 1,000 iterations was considerably lower (£45,130 per QALY gained). Upon request, the company provided a convergence plot and performed an updated PSA based on 5,000 iterations to assess the stability of the PSA results. The convergence plot based on 1,000 iterations demonstrated stable results, and the updated PSA result based on 5,000 iterations was similar to the updated deterministic base-case result. Hence, the ERG agrees that the PSA results are stable.

5.3 Model validation and face validity check

5.3.1 Face validity assessment

Guidance was sought from two clinical experts to ensure clinical validity by discussing in detail the model structure, inputs and key assumptions. Clinical experts' opinions were also sought on the clinical plausibility of the extrapolated survival functions to inform the final selections.

5.3.2 Technical verification

A checklist was used to manage quality control across different elements of the health economic model (reported in Table B.3.38 of the CS[1]). This included the execution of several stress tests on the model by testing the robustness of the model when using extreme values. In addition, in response to clarification question B25, the company also completed the TECH-VER checklist.[42]

5.3.3 Comparisons with other technology appraisals

Inclusion of comparators and clinical trials were cross-validated using published treatment guidelines for the management of AML including the ELN, published in 2017 which is the main guideline used in the UK[24], the BSCH published in 2006[25] and the European Society for Medical Oncology (ESMO) guideline published in 2013[43]. The company also compared other features of the economic analysis in Table B.3.1 of the CS.

5.3.4 Comparison with external data used to develop the economic model

No comparison with external data used to develop the economic model was performed.

5.3.5 Comparison with external data not used to develop the economic model

To the knowledge of the ERG, no comparison with external data (that was not used in the economic model) was performed.

ERG comment: The main concerns of the ERG relate to: a) internal validity; and b) cross validation with other technology appraisals.

a) The ERG considers that the internal validity of the company's model has been sufficiently established.
b) Upon request, the company provided detailed cross validation with other technology appraisals (TAs 399, 523, 545, 552, 642)[10, 11, 13, 33, 44] and the cost effectiveness analysis by Bewersdorf et al.[45] This prompted the company to perform further scenario analysis on alternative cost assumptions for nurse and haematologist visits (which increased the company's base-case ICER by £3,044). It was, however, unclear what informed the change in these costs (set to 40% in the original submission). In addition, the ERG considers it noteworthy that TAs 523, 545, and 642 included a HSCT health state. For TA545, the then ERG considered the model structure to be overly complex, however, it was unclear whether this related to the inclusion of the HSCT health state or to other health states as well (different refractory, relapse and post-HSCT states).

6. EVIDENCE REVIEW GROUP'S ADDITIONAL ANALYSES

6.1 Exploratory and sensitivity analyses undertaken by the ERG

Table 6.1 summarises the key issues related to the cost effectiveness categorised according to the sources of uncertainty as defined by Grimm et al. 2020.[46]

- Transparency (e.g., lack of clarity in presentation, description, or justification).
- Methods (e.g., violation of best research practices, existing guidelines, or the reference case).
- Imprecision (e.g., particularly wide confidence intervals, small sample sizes, or immaturity of data).
- Bias and indirectness (e.g., there is a mismatch between the decision problem and evidence used to inform it in terms of population, intervention/comparator and/or outcomes considered).
- Unavailability (e.g., lack of data or insight).

Identifying the source of uncertainty can help determine what course of action can be taken (i.e., whether additional clarifications, evidence and/or analyses might help to resolve the key issue). Moreover, Table 6.1 lists suggested alternative approaches, expected effects on the cost effectiveness, whether it is reflected in the ERG base-case as well as additional evidence or analyses that might help to resolve the key issues.

Based on all considerations in the preceding Sections of this ERG report, 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):[47]

- 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).

6.1.1 ERG base-case

Adjustments made by the ERG, to derive the ERG base-case (using the CS base-case as starting point) are listed below. Table 6.2 shows how individual adjustments impact the results plus the combined effect of all abovementioned adjustments simultaneously, resulting in the 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.

Fixing errors

No fixing errors were identified by the ERG.

Fixing violations

No fixing violations were identified by the ERG.

Matters of judgement

1. Consolidation therapy in the QUAZAR AML-001 trial not representative of UK clinical practice (Section 4.2.3)
   Use the consolidation subgroup instead of the ITT population.
2. The source for the calculation of the utility of the relapse health state. (Section 4.2.8)
   Calculate relapse utility based on Tremblay et al[22, 23, 48, 49] instead of from Joshi et al.[30]
3. Application of a disutility for patients receiving HSCT (Section 4.2.8)
   Remove the HSCT disutility.

6.1.2 ERG exploratory scenario analyses

The ERG performed the following exploratory scenario analyses to explore the impact of alternative assumptions conditional on the ERG base-case.

Exploratory scenario analyses

4. Post-HSCT utility increment for the proportion of patients treated with HSCT (Section 4.2.8)
   Apply RFS utility post HSCT for a duration of 1.67 years.

6.1.3 ERG subgroup analyses

The ERG adopted the same approach for the FLT3 subgroup as described in the ERG base-case and ERG exploratory scenario analyses above, except for using the ≥1 cycles of consolidation subgroup in the ERG base-case as this was not possible for the FLT3 subgroup.

Table 6.1: Overview of key issues related to the cost effectiveness

Key Issue	Section	Source of uncertainty	Alternative approaches	Expected impact on ICER[a]	Resolved in ERG base-case[b]	Required additional evidence or analyses
7) HSCT not appropriately reflected in the modelling	4.2.2	Methods	HSCT as a health state, or post HSCT utility increment.	+	No	Additional details for survival modelling of HSCT censored population. Evidence on post HSCT utility benefit.
8) QUAZAR trial not representative in terms of consolidation therapy	4.2.3	Bias and indirectness	Exclude patients with fewer than 2 cycles of consolidation therapy. Provide details for survival modelling of consolidation subgroup.	+/-	Partly	Evidence and scenario analysis based on the number of cycles of consolidation therapy in UK clinical practice.
9) Patient baseline characteristics not subgroup specific	4.2.3	Bias and indirectness	Use subgroup-specific patient baseline characteristics.	+/-	No	Updated model with updated patient baseline characteristics per subgroup.
10) Bias and lack of detail in survival analyses for the FLT3 subgroup	4.2.6	Bias and indirectness	Details for survival modelling of FLT3 subgroup. An analysis excluding patients without consolidation therapy for the FLT3 subgroup.	+/-	No	Details for survival modelling of FLT3 subgroup. An analysis excluding patients without consolidation therapy for the FLT3 subgroup.

Key Issue	Section	Source of uncertainty	Alternative approaches	Expected impact on ICER[a]	Resolved in ERG base-case[b]	Required additional evidence or analyses
11) Underestimation of AEs	4.2.7	Bias and indirectness	Further research is required to ascertain that AEs are not underestimated.	+	No	Further information on the AE duration and reoccurrence would help to resolve this issue.
12) Uncertainty in the choice of quality of life upon relapse	4.2.8	Bias and indirectness	Explore alternative sources for relapse utility. Investigate the impact of the relapse utility data as per the QUAZAR AML-001 trial.	+/-	Partly	Explore relapse utility as calculated based on the QUAZAR AML-001 trial data.
13) Lack of clarity about some resource use items	4.2.9	Transparency	Provide further justification, potentially updated analysis.	+/-	No	Further justification and updated analysis.
14) Lack of a fully incremental analysis for all comparators in the FLT3 subgroup	5.1	Methods	Perform a fully incremental analysis for all comparators in the FLT3 subgroup.	+/-	No	A fully incremental analysis for all comparators in the FLT3 subgroup.

AE = adverse event; ERG = Evidence Review Group; HSCT = Hematopoietic stem cell transplantation; ICER = incremental cost effectiveness ratio; UK = United Kingdom
[a]Likely conservative assumptions (of the intervention versus all comparators) are indicated by '-'; while '+/-' indicates that the bias introduced by the issue is unclear to the ERG and '+' indicates that the ERG believes this issue likely induces bias in favour of the intervention versus at least one comparator.
[b]Explored.

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

In Section 6.1 the ERG base-case was described, which was based on various changes compared to the company base-case. Table 6.2 shows how individual changes impact the results plus the combined effect of all changes simultaneously. The exploratory scenario analyses are presented in Table 6.3. These are all conditional on the ERG base-case. The analyses numbers in Tables 6.2 and 6.3 correspond to the numbers reported in Section 6.1. Finally, Table 6.4 provides the results of the subgroup analysis (described in Section 6.1.3). 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 6.2: ERG base-case

Technologies	Total costs (£)	Total QALYs	Incremental costs (£)	Incremental QALYs	ICER (£/QALY)
CS deterministic base-case					
Oral azacitidine	[redacted]	[redacted]			
w&w+BSC	[redacted]	[redacted]	[redacted]	[redacted]	48,660
Matter of judgement (key issue 9-consolidation subgroup)					
Oral azacitidine	[redacted]	[redacted]			
w&w+BSC	[redacted]	[redacted]	[redacted]	[redacted]	53,574
Matter of judgement (key issue 13-Relapse utility based on Tremblay)					
Oral azacitidine	[redacted]	[redacted]			
w&w+BSC	[redacted]	[redacted]	[redacted]	[redacted]	47,478
Matter of judgement (key issue 8-no temporary disutility for HSCT)					
Oral azacitidine	[redacted]	[redacted]			
w&w+BSC	[redacted]	[redacted]	[redacted]	[redacted]	48,729
Deterministic ERG base-case					
Oral azacitidine	[redacted]	[redacted]			
w&w+BSC	[redacted]	[redacted]	[redacted]	[redacted]	53,291
Probabilistic ERG base-case					
Oral azacitidine	[redacted]	[redacted]			
w&w+BSC	[redacted]	[redacted]	[redacted]	[redacted]	52,731

CS = Company Submission; ERG = Evidence Review Group; HSCT = hematopoietic stem cell transplantation; ICER = incremental cost effectiveness ratio; QALYs = quality-adjusted life years; w&w+BSC = watch & wait plus best supportive care

Table 6.3: Probabilistic scenario analyses (conditional on ERG base-case)

Technologies	Total costs	Total QALYs	Incremental costs	Incremental QALYs	ICER (£/QALY)
Probabilistic ERG base-case					
Oral azacitidine	[redacted]	[redacted]			

Technologies	Total costs	Total QALYs	Incremental costs	Incremental QALYs	ICER (£/QALY)
w&w+BSC	[redacted]	[redacted]	[redacted]	[redacted]	52,731
Scenario analysis (key issue 8-utility increment for HSCT)					
Oral azacitidine	[redacted]	[redacted]			
w&w+BSC	[redacted]	[redacted]	[redacted]	[redacted]	61,903
ERG = Evidence Review Group; ICER = incremental cost effectiveness ratio; QALYs = quality-adjusted life years; w&w+BSC = watch & wait plus best supportive care					

Table 6.4: ERG base-case FLT3 subgroup

Technologies	Total costs (£)	Total QALYs	Incremental costs (£)	Incremental QALYs	ICER versus baseline (£/QALY))	Pairwise ICER versus oral azacitidine
CS deterministic base-case						
Midostaurin	[redacted]	[redacted]	[redacted]	[redacted]	278,182	Oral azacitidine is dominant
Oral azacitidine	[redacted]	[redacted]	[redacted]	[redacted]	24,532	-
w&w+BSC	[redacted]	[redacted]	-	-		24,532
Matter of judgement (key issue 13-Relapse utility based on Tremblay)						
Midostaurin	[redacted]	[redacted]	[redacted]	[redacted]	244,739	Oral azacitidine is dominant
Oral azacitidine	[redacted]	[redacted]	[redacted]	[redacted]	24,547	-
w&w+BSC	[redacted]	[redacted]	-	-		24,547
Matter of judgement (key issue 8-no temporary disutility for HSCT)						
Midostaurin	[redacted]	[redacted]	[redacted]	[redacted]	278,898	Oral azacitidine is dominant
Oral azacitidine	[redacted]	[redacted]	[redacted]	[redacted]	24,548	-
w&w+BSC	[redacted]	[redacted]	-	-		24,548
Deterministic ERG base-case						
Midostaurin	[redacted]	[redacted]	[redacted]	[redacted]	245,293	Oral azacitidine is dominant

Technologies	Total costs (£)	Total QALYs	Incremental costs (£)	Incremental QALYs	ICER versus baseline (£/QALY))	Pairwise ICER versus oral azacitidine
Oral azacitidine	[redacted]	[redacted]	[redacted]	[redacted]	24,564	-
w&w+BSC	[redacted]	[redacted]	-	-		24,564
Probabilistic ERG base-case						
Midostaurin	[redacted]	[redacted]	[redacted]	[redacted]	236,519	Oral azacitidine is dominant
Oral azacitidine	[redacted]	[redacted]	[redacted]	[redacted]	25,275	
w&w+BSC	[redacted]	[redacted]				25,275
Probabilistic scenario analysis (key issue 8-utility increment for HSCT)						
Midostaurin	[redacted]	[redacted]	[redacted]	[redacted]	242,056	Oral azacitidine is dominant
Oral azacitidine	[redacted]	[redacted]	[redacted]	[redacted]	25,821	-
w&w+BSC	[redacted]	[redacted]	-	-		25,821

CS = company submission; ERG = Evidence Review Group; HSCT = hematopoietic stem cell transplantation; ICER = incremental cost effectiveness ratio; QALYs = quality-adjusted life years; w&w+BSC = watch & wait plus best supportive care

6.3 ERG's preferred assumptions

The estimated ERG base-case ICER (probabilistic), based on the ERG preferred assumptions highlighted in Section 6.1, was £52,731 per QALY gained. The probabilistic ERG base-case analyses indicated cost effectiveness probabilities of 2%, 7% and 43% at willingness to pay thresholds of £20,000, £30,000 and £50,000 per QALY gained. The most influential adjustment was using the consolidation subgroup instead of the ITT population. The ICER increased most in the scenario analysis assuming a post-HSCT utility increment for the proportion of patients treated with HSCT.

6.4 Conclusions of the cost effectiveness section

The company's cost effectiveness model complied with the NICE reference case. The most prominent issues highlighted by the ERG were 1) that the QUAZAR AML-001 trial was likely not representative of UK clinical practice in terms of consolidation therapy, 2) that HSCT was likely not appropriately reflected in the modelling, 3) that HRQoL estimates were biased due to uncertainty regarding the method of utility elicitation used, and the small sample size of the source used for the calculation of the utility of patients upon relapse, and 4) that the results of the FLT3 subgroup analyses were biased.

Firstly, the QUAZAR AML-001 trial was likely not representative of the UK clinical practice population, amongst others due to differences in consolidation therapy use. Although recommended after induction therapy in UK clinical practice[8], approximately 20% of patients in the QUAZAR AML-

001 trial were not treated with consolidation therapy. The ERG did not know the recommended number of cycles of consolidation therapy and its use in UK clinical practice, but it considered the consolidation subgroup (including patients that received at least 1 cycle of consolidation therapy) to be likely more appropriate than the QUAZAR AML-001 ITT population and adopted this subgroup in the ERG base-case. In addition, the company did not provide full details for the assessment of the NICE DSU TSD14 criteria to inform survival analyses for OS and RFS in the consolidation subgroup, and the ERG was therefore unable to assess the most appropriate curves for the modelling of OS and RFS. It adopted the company's modelling approach in its base-case but stressed that full details of the NICE DSU TSD14 criteria are necessary for a thorough assessment.

Secondly, HSCT was not included as a separate health state but was implicitly included in the modelling through the survival analysis and the application of costs and a temporary disutility associated with undergoing HSCT. The ERG considers that this way of handling HSCT in the model may cause biases, one because survival analysis of OS and RFS may be biased, and two because no benefit in HRQoL post HSCT was captured in the model. The company provided survival analyses on RFS and OS censored for HSCT, but these analyses lacked detail related to the assessment of the NICE DSU TSD14 criteria and a post HSCT HRQoL benefit was not explored. The ERG explored a post HSCT utility increment in a scenario analysis by applying a temporary return to the RFS utility after relapse for the proportion of patients undergoing HSCT, which substantially increased the ICER. This analysis, however, was suboptimal and the company should further explore this by incorporating an estimate of HSCT utility benefit in this population in the modelling. The impact of HSCT therefore remains an area of uncertainty in this model that does not appear to be appropriately explored.

Thirdly, the ERG questions the utility for the relapse health state based on a study by Joshi[30], in which the methodology used was not in line with the NICE reference case and the number of relapsed patients was small (n=23). The company argued that HRQoL was not measured upon relapse in the QUAZAR AML-001 trial, but these data were reported for a small number of patients in Appendix B.15 Table 78 of the clarification response and could have been explored as a scenario analysis in the model. Although relapse utilities from Stein[32] and Tremblay[30] were considered as alternatives, these utilities were also suboptimal as they were elicited in US populations and not using the EQ-5D. Although also suboptimal, in line with TA523 the ERG used Tremblay to calculate the relapse utility in their base-case, as the method of utility elicitation in this study was deemed more appropriate than Joshi.

Finally, the results of the FLT3 subgroup analyses were likely to be extremely biased due to limitations associated with the indirect comparison. Firstly, patient baseline characteristics should be updated in this analysis to reflect those of the subgroup. Secondly, as the survival analyses in this subgroup lacked transparency, a detailed description of the assessment of the NICE TSD DSU 14 criteria in the FLT3 subgroup should be provided. Furthermore, a fully incremental analysis for all comparators should be conducted. In addition, an analysis excluding patients without consolidation therapy would be potentially useful.

The updated CS base-case probabilistic and deterministic ICERs were £48,332 and £48,660 per QALY gained, respectively. For the FLT3 subgroup, midostaurin was dominated by oral azacitidine, and the probabilistic ICER for oral azacitidine versus watch and wait plus BSC in this subgroup was £25,403 per QALY gained. The estimated ERG base-case ICER (probabilistic), based on the ERG preferred assumptions highlighted in Section 6.1, was £52,731 per QALY gained. The most influential adjustment was using the consolidation subgroup instead of the ITT population. The ICER increased most in the scenario analysis assuming a post-HSCT utility increment for the proportion of patients

treated with HSCT. For the FLT3 subgroup, midostaurin was dominated by oral azacitidine in the ERG base-case and the probabilistic ICER for oral azacitidine versus watch and wait plus BSC was £25,275 per QALY gained.

There is large remaining uncertainty about the effectiveness and cost effectiveness of oral azacitidine, which can be partly resolved by the company by conducting further analyses. The appropriate number of cycles of consolidation therapy in UK clinical practice and the most appropriate curves for the modelling of OS and RFS in the consolidation subgroup are unknown. In addition, the current approaches (both in the CS and ERG base-case) to reflect HSCT in the modelling and to incorporate HRQoL, are likely biased. Results of the FLT3 subgroup are likely biased and updated baseline patient characteristics reflective of this subgroup are required, as well as a detailed description of survival analyses. Therefore, the ERG believes that the CS nor the ERG report contains an unbiased ICER of oral azacitidine compared with relevant comparators.

7. END OF LIFE

According to the company, this appraisal fulfils the end-of-life criteria as specified by NICE because:[1]

- patients who achieve CR/CRi after induction plus consolidation chemotherapy without undergoing maintenance treatment have a short life expectancy (median OS of patients in the placebo group, i.e., BSC of the QUAZAR AML-001 trial was 14.8 months); and
- there is sufficient evidence from the QUAZAR AML-001 study to indicate that oral azacitidine offers an extension to life of >3 months (prolongs median OS by 9.9 months, compared with placebo plus BSC).

ERG comment:

- The ERG's analyses do not indicate that the first end-of-life criteria (that life expectancy does not exceed 24 months) has been met. The ERG base-case for the ITT population results in 2.95 discounted life years for the comparator, and the ERG base-case for the FLT3 subgroup results in 2.73 discounted life years for watch and wait plus BSC and 3.60 discounted LYs for midostaurin. Both of these exceed the 24-month (two LYs) threshold specified by NICE.
- The ERG's analyses found uncertainty regarding whether the second criteria (that the treatment extend live by at least 3 months).
    - The ERG base-case for the ITT population results in 3.82 discounted LYs for oral azacitidine and 2.95 discounted LYs for the comparator, resulting in an 0.87 LY (10.4 month) gain for oral azacitidine.
    - The ERG base-case for the FLT3 subgroup results in 4.83 discounted LYs for oral azacitidine, 2.73 discounted LYs for watch and wait plus BSC and 3.60 discounted LYs for midostaurin. This implies a 2.1 LY gain compared with watch and wait plus BSC, and a 1.23 LY gain compared with midostaurin.
    - In Section 3 of this report, the ERG raised a number of problems with the evidence of clinical effectiveness upon which the claims that the second end-of-life criteria have been met.

8. REFERENCES

[1] Celgene Ltd. *Oral azacitidine for maintenance treatment of acute myeloid leukaemia after induction therapy [ID3892]: Document B (v2.0). Submission to National Institute of Health and Care Excellence. Single technology appraisal (STA)*, 2022 [accessed 11.1.22]

[2] National Institute for Health and Care Excellence. *Health Technology Appraisal: Oral azacitidine for maintenance treatment of acute myeloid leukaemia after induction therapy: final scope [Internet]*. London: NICE, 2021 [accessed 7.12.21] Available from: https://www.nice.org.uk/guidance/gid-ta10775/documents/final-scope

[3] Wei AH, Dohner H, Pocock C, Montesinos P, Afanasyev B, Dombret H, et al. Oral azacitidine maintenance therapy for acute myeloid leukemia in first remission. *N Engl J Med* 2020;383(26):2526-37.

[4] Celgene Ltd. CC-486-AML-001 (QUAZAR) Clinical Study Report [Data on file provided by the company]. 2019.

[5] U.S. Food and Drug Administration (FDA). *ONUREG (oral azacitidine) US Prescribing Information [Internet]*: FDA, 2020 [accessed 26.6.21] Available from: https://www.accessdata.fda.gov/drugsatfda_docs/label/2020/214120s000lbl.pdf

[6] European Medicines Agency (EMA). *Onureg: EPAR- Product information [Internet]*. Amsterdam: EMA, 2021 [accessed 29.11.21] Available from: https://www.ema.europa.eu/en/documents/product-information/onureg-epar-product-information_en.pdf

[7] European Medicines Agency (EMA). Onureg: azacitidine [Internet]. Amsterdam: EMA, n.d. [accessed 1.2.22]. Available from: https://www.ema.europa.eu/en/medicines/human/EPAR/onureg

[8] NHS. Treatment: acute myeloid leukaemia [Internet]. NHS, 2019 [accessed 9.2.22]. Available from: https://www.nhs.uk/conditions/acute-myeloid-leukaemia/treatment/

[9] Celgene Ltd. *Oral azacitidine for maintenance treatment of acute myeloid leukaemia after induction therapy [ID3892]: response to request for clarification from the ERG (v3.0)*, 2022 [accessed 10.2.22]. 249p.

[10] National Institute for Health and Care Excellence. *Liposomal cytarabine–daunorubicin for untreated acute myeloid leukaemia: NICE technology appraisal guidance 552 [Internet]* London: NICE, 2021 [accessed 27.1.22] Available from: https://www.nice.org.uk/guidance/ta552

[11] National Institute for Health and Care Excellence. *Gemtuzumab ozogamicin for untreated acute myeloid leukaemia: NICE technology appraisal guidance 545 [Internet]*. London: NICE, 2018 [accessed 7.12.21] Available from: https://www.nice.org.uk/guidance/ta545

[12] National Institute for Health and Care Excellence. *Azacitidine for the treatment of myelodysplastic syndromes, chronic myelomonocytic leukaemia and acute myeloid leukaemia: NICE technology appraisal guidance 218 [Internet]*. London: NICE, 2011 [accessed 27.1.22] Available from: https://www.nice.org.uk/guidance/ta218

[13] National Institute for Health and Care Excellence. *Midostaurin for untreated acute myeloid leukaemia: NICE technology appraisal guidance 523 [Internet]*. London: NICE, 2018 [accessed 29.11.21] Available from: https://www.nice.org.uk/guidance/ta523

[14] Schlenk RF, Weber D, Fiedler W, Salih HR, Wulf G, Salwender H, et al. Midostaurin added to chemotherapy and continued single-agent maintenance therapy in acute myeloid leukemia with FLT3-ITD. *Blood* 2019;133(8):840-51.

[15] Centre for Reviews and Dissemination. *Systematic Reviews: CRD's guidance for undertaking reviews in health care [Internet]*. York: University of York, 2009 [accessed 23.3.11] Available from: http://www.york.ac.uk/inst/crd/SysRev/!SSL!/WebHelp/SysRev3.htm

[16] Celgene Ltd. *Oral azacitidine for maintenance treatment of acute myeloid leukaemia after induction therapy [ID3892]: Appendices (v2.0). Submission to National Institute of Health and Care Excellence. Single technology appraisal (STA)*, 2022 [accessed 11.1.22]

[17] National Institute for Health and Care Excellence. *Methods for the development of NICE public health guidance (third edition): process and methods. Appendix I: Quality appraisal checklist - economic evaluations [Internet]*. London: NICE, 2012 Available from: https://www.nice.org.uk/process/pmg4/chapter/appendix-i-quality-appraisal-checklist-economic-evaluations

[18] Wei AH, Döhner H, Pocock C, Montesinos P, Afanasyev B, Dombret H, et al. The QUAZAR AML-001 maintenance trial: results of a phase III international, randomized, double-blind, placebo-controlled study of CC-486 (oral formulation of azacitidine) in patients with acute myeloid leukemia (AML) in first remission. *Blood* 2019;134(Suppl 2):LBA-3.

[19] U.S. Food and Drug Administration (FDA). *NDA multidisciplinary review and evaluation: Azacitidine (NDA 214120) [Internet]*: FDA, 2020 [accessed 7.12.21] Available from: https://www.accessdata.fda.gov/drugsatfda_docs/nda/2020/214120Orig1s000MultidisciplineR.pdf

[20] Phillippo D, Dias S, Palmer S, KR A, Welton N. NICE DSU Technical Support Document 18: Methods for population-adjusted indirect comparisons in submission to NICE. Appendix D: worked example of MAIC and STC [Internet]. 2016 [accessed 29.11.21]. Available from: http://nicedsu.org.uk/wp-content/uploads/2017/05/Population-adjustment-TSD-FINAL.pdf

[21] Higgins J, Thomas J, Chandler J, Cumpston M, Li T, Page M, et al. *Cochrane handbook for systematic reviews of interventions version 6.0 (updated July 2019) [Internet]*: Cochrane, 2019 [accessed 6.7.20] Available from: https://training.cochrane.org/handbook

[22] Tremblay G, Dolph M, Patel S, Brandt P, Forsythe A. Cost-effectiveness analysis for midostaurin versus standard of care in acute myeloid leukemia in the United Kingdom. *Cost Eff Resour Alloc* 2018;16:33.

[23] Tremblay G, Cariou C, Recher C, Dolph M, Brandt P, Blanc AS, et al. Cost-effectiveness of midostaurin in the treatment of newly diagnosed FLT3-mutated acute myeloid leukemia in France. *Eur J Health Econ* 2020;21:543–55.

[24] Dohner H, Estey E, Grimwade D, Amadori S, Appelbaum FR, Buchner T, et al. Diagnosis and management of AML in adults: 2017 ELN recommendations from an international expert panel. *Blood* 2017;129(4):424-47.

[25] 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.

[26] Panju AH, Danesh A, Minden MD, Kelvin DJ, Alibhai SM. Associations between quality of life, fatigue, and cytokine levels in patients aged 50+ with acute myeloid leukemia. *Support Care Cancer* 2009;17(5):539-46.

[27] Larson RA, Mandrekar SJ, Huebner LJ, Sanford BL, Laumann K, Geyer S, et al. Midostaurin reduces relapse in FLT3-mutant acute myeloid leukemia: the Alliance CALGB 10603/RATIFY trial. *Leukemia* 2021;35(9):2539-51.

[28] Dolan P. Modeling valuations for EuroQol health states. *Med Care* 1997;35(11):1095-108.

[29] Stein E, Xie J, Duchesneau E, Bhattacharyya S, Vudumula U, Ndife B, et al. Cost effectiveness of midostaurin in the treatment of newly diagnosed FLT3-mutated acute myeloid leukemia in the United States. *Pharmacoeconomics* 2019;37(2):239-53.

[30] Joshi N, Hensen M, Patel S, Xu W, Lasch K, Stolk E. Health state utilities for acute myeloid leukaemia: a time trade-off study. *Pharmacoeconomics* 2019;37(1):85-92.

[31] Matza LS, Deger KA, Howell TA, Koetter K, Yeager AM, Hogge D, et al. Health state utilities associated with treatment options for acute myeloid leukemia (AML). *J Med Econ* 2019;22(6):567-76.

[32] Stein EM, Yang M, Guerin A, Gao W, Galebach P, Xiang CQ, et al. Assessing utility values for treatment-related health states of acute myeloid leukemia in the United States. *Health Qual Life Outcomes* 2018;16(1):193.

[33] National Institute for Health and Care Excellence. *Gilteritinib for treating relapsed or refractory acute myeloid leukaemia: NICE technology appraisal guidance 642 [Internet]*. London: NICE, 2020 [accessed 29.11.21] Available from: https://www.nice.org.uk/guidance/ta642/

[34] NICE. Midostaurin for untreated acute myeloid leukaemia [TA523] [online]. Available from: https://www.nice.org.uk/guidance/ta523

[35] Szende A, Janssen B, Cabases Je. *Self-reported population health: an international perspective based on EQ-5D [Internet]*. Dordrecht (NL): Springer, 2014 [accessed 2.2.22]. Available from: https://pubmed.ncbi.nlm.nih.gov/29787044/

[36] Department of Health and Social Care. Drugs and pharmaceutical electronic market information tool (eMIT) [Database]. London: NHS, 2021 [accessed 2.2.22]. Available from: https://www.gov.uk/government/publications/drugs-and-pharmaceutical-electronic-market-information-emit

[37] NHS. *National Cost Collection 2019/20 Report [Internet] [Internet]*. London: NHS England and NHS Improvement, 2021 [accessed 29.11.21] Available from: https://www.england.nhs.uk/wp-content/uploads/2021/06/National-Cost-Collection-2019-20-Report-FINAL.pdf

[38] Joint Formulary Committee. *British National Formulary (BNF82)*. London: BMJ Group and Pharmaceutical Press, 2021.

[39] Bristol Myers Squibb. *ONUREG (Oral Azacitidine) in acute myeloid leukaemia: clinical expert meetings summary report [Document provided by the company]*, 2022 [accessed 21.1.22]

[40] Georghiou T, Bardsley M. *Exploring the cost of care at the end of life [Internet]*. London: Nuffield Trust 2014 [accessed 29.11.21] Available from: https://www.nuffieldtrust.org.uk/files/2017-01/end-of-life-care-web-final.pdf

[41] Curtis LA, Burns A. Unit costs of health & social care 2020 [Internet]. Canterbury: PSSRU, University of Kent, 2020 [accessed 7.12.21]. Available from: https://kar.kent.ac.uk/84818/

[42] Büyükkaramikli NC, Rutten-van Mölken M, Severens JL, Al M. TECH-VER: a verification checklist to reduce errors in models and improve their credibility. *Pharmacoeconomics* 2019;37(11):1391-408.

[43] Fey MF, Buske C. Acute myeloblastic leukaemias in adult patients: ESMO Clinical Practice Guidelines for diagnosis, treatment and follow-up. *Ann Oncol* 2013;24 (Suppl 6):vi138-43.

[44] National Institute for Health and Care Excellence. *Azacitidine for treating acute myeloid leukaemia with more than 30% bone marrow blasts: NICE technology appraisal guidance 399 [Internet]*. London: NICE, 2016 [accessed 29.11.21] Available from: https://www.nice.org.uk/guidance/ta399

[45] Bewersdorf JP, Patel KK, Huntington SF, Zeidan AM. Cost-effectiveness analysis of oral azacitidine maintenance therapy in acute myeloid leukemia. *Blood Adv* 2021;5(22):4686-90.

[46] Grimm SE, Pouwels X, Ramaekers BLT, Wijnen B, Knies S, Grutters J, et al. Development and Validation of the TRansparent Uncertainty ASsessmenT (TRUST) Tool for assessing uncertainties in health economic decision models. *Pharmacoeconomics* 2020;38(2):205-16.

[47] 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.

[48] Tremblay G, Dolph M, Ouagari KE, Brandt P, Forsythe A. Cost-effectiveness analysis of midostaurin (MIDO) with standard of care (SOC) for acute myeloid leukemia (AML) in Canada. *Value Health* 2018;21(Suppl 1):S28-9.

[49] Tremblay G, Dolph M, Patel S, Brandt P, Forsythe A. Cost-effectiveness analysis of midostaurin (MIDO) with standard chemotherapy (SOC) for acute myeloid leukemia (AML) in the united kingdom (UK). *Value Health* 2017;20 (9):A399.

in collaboration with:

Erasmus School of Health Policy & Management

Maastricht University

Oral azacitidine for maintenance treatment of acute myeloid leukaemia after induction therapy [ID3892]

ADDENDUM

Produced by Kleijnen Systematic Reviews (KSR) Ltd, in collaboration with Erasmus University Rotterdam (EUR) and Maastricht University Medical Centre (UMC)

Authors

Jeremy Howick, Reviews Manager, KSR Ltd, United Kingdom (UK)
Willem Witlox, Health Economist, Maastricht UMC, The Netherlands
Charlotte Ahmadu, Health Economist, KSR Ltd, UK
Sabine Grimm, Health Economist, Maastricht UMC, The Netherlands
Nigel Armstrong, Health Economics Manager, KSR Ltd, UK
Kevin McDermott, Systematic Reviewer, KSR Ltd, UK
Thomas Otten, Health Economist, Maastricht UMC, The Netherlands
Caro Noake, Information Specialist, KSR Ltd, UK
Robert Wolff, Managing Director, KSR Ltd, UK
Manuela Joore, Health Economist, Maastricht UMC
Jos Kleijnen, Founder and Owner, KSR Ltd, UK

Correspondence to Jeremy Howick, Kleijnen Systematic Reviews Ltd
Unit 6, Escrick Business Park
Riccall Road, Escrick
York, YO19 6FD
United Kingdom

Date completed 25/04/2022

Table 1: Updated ERG base-case

Technologies	Total costs (£)	Total QALYs	Incremental costs (£)	Incremental QALYs	ICER (£/QALY)
Updated CS deterministic base-case					
Oral azacitidine	[redacted]	[redacted]			
w&w+BSC	[redacted]	[redacted]	[redacted]	[redacted]	£32,718
Matter of judgement (key issue 9-consolidation subgroup)					
Oral azacitidine	[redacted]	[redacted]			
w&w+BSC	[redacted]	[redacted]	[redacted]	[redacted]	£41,238
Matter of judgement (key issue 13-Relapse utility based on Tremblay)					
Oral azacitidine	[redacted]	[redacted]			
w&w+BSC	[redacted]	[redacted]	[redacted]	[redacted]	£31,857
Matter of judgement (key issue 8-no temporary disutility for HSCT)					
Oral azacitidine	[redacted]	[redacted]			
w&w+BSC	[redacted]	[redacted]	[redacted]	[redacted]	£32,749
Deterministic ERG base-case					
Oral azacitidine	[redacted]	[redacted]			
w&w+BSC	[redacted]	[redacted]	[redacted]	[redacted]	£40,994
Probabilistic ERG base-case					
Oral azacitidine	[redacted]	[redacted]			
w&w+BSC	[redacted]	[redacted]	[redacted]	[redacted]	£40,768

CS = Company Submission; ERG = Evidence Review Group; HSCT = hematopoietic stem cell transplantation; ICER = incremental cost effectiveness ratio; QALYs = quality-adjusted life years; w&w+BSC = watch & wait plus best supportive care

Table 2: Updated and additional ERG scenario analyses

Technologies	Total costs	Total QALYs	Incremental costs	Incremental QALYs	ICER (£/QALY)
Probabilistic ERG base-case					
Oral azacitidine	[redacted]	[redacted]			
w&w+BSC	[redacted]	[redacted]	[redacted]	[redacted]	£40,768
Scenario analysis (key issue 8-utility increment for HSCT)					
Oral azacitidine	[redacted]	[redacted]			
w&w+BSC	[redacted]	[redacted]	[redacted]	[redacted]	£47,589
Scenario analysis dose extension [redacted]					
Oral azacitidine	[redacted]	[redacted]			
w&w+BSC	[redacted]	[redacted]	[redacted]	[redacted]	£41,349
Scenario analysis dose extension 30%					

Technologies	Total costs	Total QALYs	Incremental costs	Incremental QALYs	ICER (£/QALY)
Oral azacitidine	[redacted]	[redacted]			
w&w+BSC	[redacted]	[redacted]	[redacted]	[redacted]	£41,576
Scenario analysis dose extension 40%					
Oral azacitidine	[redacted]	[redacted]			
w&w+BSC	[redacted]	[redacted]	[redacted]	[redacted]	£41,845
Scenario analysis individual modelling of OS and RFS[1]					
Oral azacitidine	[redacted]	[redacted]			
w&w+BSC	[redacted]	[redacted]	[redacted]	[redacted]	£64,418

[1]For OS, the individual log-normal and individual generalised gamma were selected for oral azacitidine and w&w+BSC respectively. For RFS, the individual log-logistic was used for both arms.

ERG = Evidence Review Group; ICER = incremental cost effectiveness ratio; QALYs = quality-adjusted life years; w&w+BSC = watch & wait plus best supportive care

Table 3: FLT3 subgroup: updated ERG base-case and scenario analyses

Technologies	Total costs (£)	Total QALYs	Incremental costs (£)	Incremental QALYs	ICER versus baseline (£/QALY))	Pairwise ICER versus oral azacitidine
Updated CS deterministic base-case						
Midostaurin	[redacted]	[redacted]	[redacted]	[redacted]	£269,191	Oral azacitidine is dominant
Oral azacitidine	[redacted]	[redacted]	[redacted]	[redacted]	£19,063	
w&w+BSC	[redacted]	[redacted]				£19,063
Matter of judgement (key issue 13-Relapse utility based on Tremblay)						
Midostaurin	[redacted]	[redacted]	[redacted]	[redacted]	£237,034	Oral azacitidine is dominant
Oral azacitidine	[redacted]	[redacted]	[redacted]	[redacted]	£19,048	
w&w+BSC	[redacted]	[redacted]				£19,048
Matter of judgement (key issue 8-no temporary disutility for HSCT)						
Midostaurin	[redacted]	[redacted]	[redacted]	[redacted]	£269,861	Oral azacitidine

Technologies	Total costs (£)	Total QALYs	Incremental costs (£)	Incremental QALYs	ICER versus baseline (£/QALY))	Pairwise ICER versus oral azacitidine
						is dominant
Oral azacitidine	[redacted]	[redacted]	[redacted]	[redacted]	£19,076	
w&w+BSC	[redacted]	[redacted]				£19,076
Deterministic ERG base-case						
Midostaurin	[redacted]	[redacted]	[redacted]	[redacted]	£237,553	Oral azacitidine is dominant
Oral azacitidine	[redacted]	[redacted]	[redacted]	[redacted]	£19,061	
w&w+BSC	[redacted]	[redacted]				£19,061
Probabilistic ERG base-case						
Midostaurin	[redacted]	[redacted]	[redacted]	[redacted]	£228,820	Oral azacitidine is dominant
Oral azacitidine	[redacted]	[redacted]	[redacted]	[redacted]	£20,052	
w&w+BSC	[redacted]	[redacted]				£20,052
Probabilistic scenario analysis (key issue 8-utility increment for HSCT)						
Midostaurin	[redacted]	[redacted]	[redacted]	[redacted]	£233,871	Oral azacitidine is dominant
Oral azacitidine	[redacted]	[redacted]	[redacted]	[redacted]	£20,192	
w&w+BSC	[redacted]	[redacted]				£20,192

CS = company submission; ERG = Evidence Review Group; HSCT = hematopoietic stem cell transplantation; ICER = incremental cost effectiveness ratio; QALYs = quality-adjusted life years; w&w+BSC = watch & wait plus best supportive care

in collaboration with:

Erasmus School of Health Policy & Management

Maastricht University

Oral azacitidine for maintenance treatment of acute myeloid leukaemia after induction therapy [ID3892]

ERG addendum EU consolidation subgroup

Produced by Kleijnen Systematic Reviews (KSR) Ltd, in collaboration with Erasmus University Rotterdam (EUR) and Maastricht University Medical Centre (UMC)

Authors

Jeremy Howick, Reviews Manager, KSR Ltd, United Kingdom (UK)
Willem Witlox, Health Economist, Maastricht UMC, The Netherlands
Charlotte Ahmadu, Health Economist, KSR Ltd, UK
Sabine Grimm, Health Economist, Maastricht UMC, The Netherlands
Nigel Armstrong, Health Economics Manager, KSR Ltd, UK
Kevin McDermott, Systematic Reviewer, KSR Ltd, UK
Thomas Otten, Health Economist, Maastricht UMC, The Netherlands
Caro Noake, Information Specialist, KSR Ltd, UK
Robert Wolff, Managing Director, KSR Ltd, UK
Manuela Joore, Health Economist, Maastricht UMC
Jos Kleijnen, Founder and Owner, KSR Ltd, UK

Correspondence to Jeremy Howick, Kleijnen Systematic Reviews Ltd
Unit 6, Escrick Business Park
Riccall Road, Escrick
York, YO19 6FD
United Kingdom

Date completed 25/04/2022

Table 1: EU consolidation subgroup

Technologies	Total costs (£)	Total QALYs	Incremental costs (£)	Incremental QALYs	ICER (£/QALY)
ERG_1: EU consolidation subgroup					
Oral azacitidine	[redacted]	[redacted]			
w&w+BSC	[redacted]	[redacted]	[redacted]	[redacted]	£34,265
ERG_2: Matter of judgement - Relapse utility using Tremblay					
Oral azacitidine	[redacted]	[redacted]			
w&w+BSC	[redacted]	[redacted]	[redacted]	[redacted]	£33,881
ERG_3: Matter of judgement - Remove HSCT disutility					
Oral azacitidine	[redacted]	[redacted]			
w&w+BSC	[redacted]	[redacted]	[redacted]	[redacted]	£34,310
ERG deterministic base-case					
Oral azacitidine	[redacted]	[redacted]			
w&w+BSC	[redacted]	[redacted]	[redacted]	[redacted]	£33,925
ERG probabilistic base-case					
Oral azacitidine	[redacted]	[redacted]			
w&w+BSC	[redacted]	[redacted]	[redacted]	[redacted]	£33,809

CS = Company Submission; ERG = Evidence Review Group; HSCT = hematopoietic stem cell transplantation; ICER = incremental cost effectiveness ratio; QALYs = quality-adjusted life years; w&w+BSC = watch & wait plus best supportive care

Table 2: EU consolidation subgroup scenario analyses (probabilistic)

Technologies	Total costs (£)	Total QALYs	Incremental costs (£)	Incremental QALYs	ICER (£/QALY)
ERG probabilistic base-case					
Oral azacitidine	[redacted]	[redacted]			
w&w+BSC	[redacted]	[redacted]	[redacted]	[redacted]	£33,809
ERG_3: Scenario analysis - HSCT return to base-line					
Oral azacitidine	[redacted]	[redacted]			
w&w+BSC	[redacted]	[redacted]	[redacted]	[redacted]	£38,265
ERG_4: Scenario analysis – Dose extension - [redacted]					
Oral azacitidine	[redacted]	[redacted]			
w&w+BSC	[redacted]	[redacted]	[redacted]	[redacted]	£34,443
ERG_5: Scenario analysis - Dose extension - 30%					
Oral azacitidine	[redacted]	[redacted]			
w&w+BSC	[redacted]	[redacted]	[redacted]	[redacted]	£34,458
ERG_6: Scenario analysis - Dose extension - 40%					
Oral azacitidine	[redacted]	[redacted]			
w&w+BSC	[redacted]	[redacted]	[redacted]	[redacted]	£34,674
ERG_7: Scenario analysis - Individual survival modelling[1]					
Oral azacitidine	[redacted]	[redacted]			
w&w+BSC	[redacted]	[redacted]	[redacted]	[redacted]	£33,767

CS = Company Submission; ERG = Evidence Review Group; HSCT = hematopoietic stem cell transplantation; ICER = incremental cost effectiveness ratio; QALYs = quality-adjusted life years; w&w+BSC = watch & wait plus best supportive care

[1]OS: individual generalised gamma for both arms. RFS: individual log-logistic for both arms.

National Institute for Health and Care Excellence

Centre for Health Technology Evaluation

ERG report – factual accuracy check and confidential information check

Oral azacitidine for maintenance treatment of acute myeloid leukaemia after induction therapy [ID3892]

'*Data owners will be asked to check that confidential information is correctly marked in documents created by others in the technology appraisal process before release; for example, the technical report and ERG report.*' (Section 3.1.29, Guide to the processes of technology appraisals).

You are asked to check the ERG report to ensure there are no factual inaccuracies or errors in the marking of confidential information contained within it. The document should act as a method of detailing any inaccuracies found and how they should be corrected.

If you do identify any factual inaccuracies or errors in the marking of confidential information, you must inform NICE by **5pm on Tuesday 1 March** using the below 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.

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

Issue 1

Description of problem	Description of proposed amendment	Justification for amendment	ERG response
Page 12 of the report states "Increased post-relapse survival, with an incremental of 0.211 years (20% of total incremental LYs) in the oral azacitidine arm (1.779 years) compared with watch and wait with BSC arm (1.568 years)."	Please consider amending the text on page 12 to: "Increased post-relapse survival, with an incremental of **0.209** years (20% of total incremental LYs) in the oral azacitidine arm (**1.791** years) compared with watch and wait with BSC arm (**1.583** years)."	This amendment is based on the assumption that the intent of the ERG was to present probabilistic results using 1000 iterations.	No change made as ERG prefers presenting probabilistic results using 5,000 iterations.

Issue 2

Description of problem	Description of proposed amendment	Justification for amendment	ERG response
Page 12 of the report states "The higher drug costs (additional cost of [redacted], [redacted] of total incremental costs) and disease management costs (additional cost of [redacted]) in RFS on-treatment compared with watch and wait plus BSC." Page 12 of the report also states "The lower disease management costs (reduced cost of [redacted]) in RFS off-treatment compared with watch and wait plus BSC."	Please consider amending the text on page 12 to: The higher drug costs (additional cost of [redacted], [redacted] of total incremental costs) and disease management costs (additional cost of [redacted]) in RFS on-treatment compared with watch and wait plus BSC. Please also consider amending the text on page 12 to: "The lower disease management costs (reduced cost of [redacted]) in RFS off-treatment compared with watch and wait plus BSC."	These amendments are based on the assumption that the intent of the ERG was to present probabilistic results using 1000 iterations.	No change made as ERG prefers presenting probabilistic results using 5,000 iterations.

Issue 3

Description of problem	Description of proposed amendment	Justification for amendment	ERG response
Page 13 of the report states "Utility values based on Joshi 2019 for al health states (decreased ICER to [redacted])".	Please consider amending the text on page 27 to: "Utility values based on Joshi 2019 for all health states (decreased ICER to [redacted])".	Typographical error.	No change made. Utility values based on Joshi 2019 (RFS on and off treatment 0.89 and relapse 0.51) results in ICER reported by ERG.

Issue 4

Description of problem	Description of proposed amendment	Justification for amendment	ERG response
Page 13 (Table 1.2) states "Low dose cytarabine and subcutaneous azacitidine were not viewed by the company to be part of BSC."	Please consider amending the text on page 13 to: "Low dose **cytarabine** and subcutaneous azacitidine were not viewed by the company to be part of BSC."	Minor typographical errors.	Changes made
Page 13 (Table 1.2) also states "Additional evidence about the use of low dose cytarabine and subcutaneous azacitidine as BSC in this population, for example from independent clinical experts."	Please consider amending the text on page 13 to: "Additional evidence about the use of low dose **cytarabine** and subcutaneous azacitidine as BSC in this population, for example from independent clinical experts".		

Issue 5

Description of problem	Description of proposed amendment	Justification for amendment	ERG response
Page 28 of the report states "The company used a placebo comparator whereas placebo was not listed as a comparator in the final NICE scope"	Please consider amending the text on page 28 to: "The company used ~~a~~ placebo **with BSC** ~~was not listed~~ as a comparator ~~in the final NICE scope~~".	This text should be amended to reflect that Table 2.1 of the ERG report states that watch and wait with BSC is a comparator in the decision problem.	No change made; placebo is not mentioned in the final NICE scope.

Issue 6

Description of problem	Description of proposed amendment	Justification for amendment	ERG response
Page 28 of the report states "In their response to another question, the company notes that both experts confirmed that consolidation therapy is standard of care in UK clinical practice. This is a more sensible approach to determining standard of care (SoC).	Please consider removal of this text.	This text is factually incorrect as consolidation therapy is not the standard of care for patients in the AML maintenance setting. Rather, watch and wait with BSC is currently the standard of care for patients in maintenance.	No change made as the ERG has quoted what the company has said.

Issue 7

Description of problem	Description of proposed amendment	Justification for amendment	ERG response
Page 29 of the report states "While the company reports the OS independent of FLT3 status, this is not done for other efficacy or safety outcomes".	Please consider amending the text on page 29 to: "While the company reports the OS **and RFS** independent of FLT3 status, this is not done for ~~other efficacy or~~ safety outcomes".	A post-hoc analysis for OS and RFS was presented in section B.2.7.3 of the CS.	Amended to "While the company reports the OS and RFS independent of FLT3 status, this is not done for other efficacy or safety outcomes" (it was not done for HRQoL, for example)

Issue 8

Description of problem	Description of proposed amendment	Justification for amendment	ERG response
Page 29 of the report states "According to the company, this appraisal fulfils the end-of-life criteria as specified by NICE because the median survival of patients in the placebo group in the QUAZAR trial was 14.8 months, which is lower than 24 months, and because oral azacitidine plus BSC prolonged life by 9.9 months compared with placebo plus BSC".	Please consider amending the text on page 29 to: "According to the company, this appraisal fulfils the end-of-life criteria as specified by NICE because the median survival of patients in the placebo **plus BSC** group in the QUAZAR trial was 14.8 months, which is lower than 24 months, and because oral azacitidine plus BSC prolonged life by 9.9 months compared with placebo plus BSC".	In the QUAZAR AML-001 trial, patients in the control received placebo plus BSC. We request that this be amended to accurately reflect the control group of the trial.	Amended.

Issue 9

<table>
<tr><th>Description of problem</th><th>Description of proposed amendment</th><th>Justification for amendment</th><th>ERG response</th></tr>
<tr>
<td>Table 3.7 contains the following information regarding patient baseline characteristics:
<table>
<tr><th colspan="3">Reason ineligible for transplant[a], n (%)</th></tr>
<tr><td>Age</td><td>154 (65)</td><td>152 (65.)</td></tr>
<tr><td>Comorbidities</td><td>52 (22)</td><td>50 (21)</td></tr>
<tr><td>Performance Status</td><td>14 (6)</td><td>9 (4)</td></tr>
</table>
</td>
<td>Please amend this section of the table as outlined below:
<table>
<tr><th colspan="3">Reason ineligible for transplant[a], n (%)</th></tr>
<tr><td>Age</td><td>154 (65)</td><td>152 (65.)</td></tr>
<tr><td>Comorbidities</td><td>52 (22)</td><td>50 (21)</td></tr>
<tr><td>Performance Status</td><td>14 (6)</td><td>9 (4)</td></tr>
</table>
</td>
<td>Minor typographical error.</td>
<td>Amended</td>
</tr>
</table>

Issue 10

Description of problem	Description of proposed amendment	Justification for amendment	ERG response
Page 45 of the report states “Caveats of which were age (the trial limited its patient population to patients of age to ≥ 55 years), and cytogenetic risk (the study included patients with intermediate and poor cytogenetics whereas patients with favourable risk cytogenetics are less	Please consider amending the text on page 45 to: “Caveats of which were age (the trial limited its patient population to patients of age ≥ 55 years), and cytogenetic risk (the study included patients with intermediate and poor cytogenetics whereas patients with favourable risk cytogenetics are less likely to proceed to HSCT in first CR**)**”.	Minor typographical error.	Amended.

likely to proceed to HSCT in first CR".			

Issue 11

Description of problem	Description of proposed amendment	Justification for amendment	ERG response
Table 3.8 states that the sample size of the EU subgroup is 234.	Please consider amending the sample size to [redacted].	Incorrect sample size.	Amended to reflect company error in clarification response

Issue 12

Description of problem	Description of proposed amendment	Justification for amendment	ERG response
Figure 3.2 states the following: "Source: Table B.2.3 of the CS".	Please consider amending this to the following: "Source: **Figure** ~~Table~~ B.2.3 of the CS".	Typographical error.	Amended

Issue 13

Description of problem	Description of proposed amendment	Justification for amendment	ERG response
Page 73 of the report states "Discontinuation of study treatment because of AEs was reported for 13% of patients in the oral azacitidine group and 4% of patients in the placebo group. In the oral azacitidine group, AEs leading to treatment	Please consider removal of this text.	This text is repeated from the paragraph above.	Not a factual error (elaboration not repetition).

discontinuation reported by >1 patient in either treatment arm included nausea (2% versus 0%), diarrhoea (2% versus 0%), vomiting (1% versus 0%), abdominal pain (1% versus 0%), fatigue (1% versus 0%), and thrombocytopenia (0.4% versus 1%) for oral azacitidine versus placebo, respectively."			

Issue 14

Description of problem	Description of proposed amendment	Justification for amendment	ERG response
Page 75 of report states "Unmatched (n=472) and matched ([redacted]) results demonstrated that the HR is favourable for oral azacitidine when compared to midostaurin ([redacted]) and ([redacted]), respectively (see Table 3.24)."	Please consider amending the text to: "Unmatched (n=472) and matched ([redacted]) results demonstrated that the HR is favourable for oral azacitidine when compared to midostaurin ([redacted]) and ([redacted]), respectively (see Table 3.24)."	Minor typographical error.	Amended

Issue 15

Issue 16

Description of problem	Description of proposed amendment	Justification for amendment	ERG response
Page 81 of the report states “The aim was to identify published economic evaluations of interventions which address the decision problem. The eligibility criteria for the study selection were included in table B.5.25 of appendix G of the CS and appear relevant for the task at hand.”	Please consider amending the text to: “The aim was to identify published economic evaluations of interventions which address the decision problem. The eligibility criteria for the study selection were included in table B.5.**23** of appendix G of the CS and appear relevant for the task at hand.	Incorrect referencing.	Amended

Issue 17

Description of problem	Description of proposed amendment	Justification for amendment	ERG response
Page 82 of the report states “For HRQoL studies, inclusion criteria included direct utility values at baseline and utility increments or decrements by health state (using a number of different generic and disease specific HRQoL measures”.	Please consider amending the text to: “For HRQoL studies, inclusion criteria included direct utility values at baseline and utility increments or decrements by health state (using a number of different generic and disease specific HRQoL measures**)**”.	Minor typographical error.	Amended

Issue 18

Description of problem	Description of proposed amendment	Justification for amendment	ERG response
Page 82 of the report states "The nine included HRQoL studies exhibited considerable variation in study designs and utility values reported by health states. Utility values reported in the four economic studies (out of the nine included studies) were sourced from the literature but not included in the current review, as they did not align with the inclusion criteria".	Please consider amending the text to: "**Of the 20 studies included in the review, nine studies reported utility values by health state, which** ~~included HRQoL studies~~ exhibited considerable variation in study design~~s~~ and **reported** utility values ~~reported by health states~~. ~~Utility values reported in the five four economic studies (out of the nine included studies) were sourced from the literature but not included in the current review, as they did not align with the inclusion criteria~~".	In total, 20 studies were included in the HRQoL review, of which, nine studies reported utility values by health state. Of the nine studies reporting utility values by health state, five were economic studies and were included in the review. This is because the eligibility criteria outlined in Table B.5.29 of the CS did not restrict by study type, allowing for inclusion of economic evaluations.	Amended to reflect that it was the studies that reported utility values by health state

Issue 19

Description of problem	Description of proposed amendment	Justification for amendment	ERG response
Page 89 of the report states "The main concerns of the ERG relate to: a) exclusion of cytarabine and standard azacitidine and b) the appropriateness of the modelled subsequent treatments."	Please consider amending the text to: The main concerns of the ERG relate to: a) exclusion of cytarabine and ~~standard~~ **subcutaneous** azacitidine and b) the appropriateness of the modelled subsequent treatments."	Suggestion to revise wording to avoid confusion.	Amended
Page 89 of the report also states "The ERG questioned whether the exclusion of cytarabine and standard azacitidine was indeed	Please consider amending the text to: "The ERG questioned whether the exclusion of cytarabine and ~~standard~~ **subcutaneous** azacitidine was indeed appropriate, given that		

appropriate, given that these treatments were listed as comparators in the scope.”	these treatments were listed as comparators in the scope.”		

Issue 20

Description of problem	Description of proposed amendment	Justification for amendment	ERG response
Page 95 of the report states “This detail of the expert opinion, however, is not applied in the model as the duration of neutropenia in the model is also 1 week”.	Please consider amending the text to: “This detail of the expert opinion, however, is not applied in the model as the duration of neutropenia in the model is also 1 week. **Of note, a scenario analysis exploring a 4-week duration of all AEs was conducted**.”	Although a 1-week duration to all AEs was applied in the reference case, a scenario analysis with a 4-week duration of all AEs was conducted. This explores the uncertainty associated with this parameter and should be acknowledged.	Not a factual error.

Issue 21

Description of problem	Description of proposed amendment	Justification for amendment	ERG response
Page 95 of the report states: “Two other models were considered, one including the treatment arm and the other including a treatment arm and AEs. The same utility value (Model 3 intercept [redacted], SE [redacted]) was applied to RFS regardless of whether the patient was on- or	Please consider amending the text to: “Two other models were considered, one including the treatment arm and the other including a treatment arm and AEs. The same utility value (Model 3 intercept [redacted], SE [redacted]) was applied to RFS regardless of whether the patient was on- or off-treatment, as validated by expert opinion.”	An incorrect value for the SE was applied in the text of the report.	Not an ERG factual error Amended

off-treatment, as validated by expert opinion.			

Issue 22

Description of problem	Description of proposed amendment	Justification for amendment	ERG response
Table 4.7 contains the following information regarding health state utility values: **Health state** / **Utility value (base-case)** **RFS: on treatment** — 0.833 (0.012) **RFS: off treatment** — 0.833 (0.012) **Relapse** — 0.45 (0.46)	Please amend this section of the table as outlined below: **Health state** / **Utility value (base-case)** **RFS: on treatment** — 0.833 (**0.009**) **RFS: off treatment** — 0.833 (**0.009**) **Relapse** — 0.45 (0.46)	An incorrect value for the SE was applied for the RFS on treatment and RFS off treatment health states.	Not an ERG factual error Amended

Issue 23

Description of problem	Description of proposed amendment	Justification for amendment	ERG response
Page 98 of the report states “The ERG questions the choice of the source for the calculation of the utility of patients upon relapse. In	Please consider amending the text to: “The ERG questions the choice of the source for the calculation of the utility of patients upon relapse. In Section B.3.4.1 the company states	Incorrect table reference provided.	Not a factual error, ERG’s table reference aligns with

Section B.3.4.1 the company states that their study has not assessed relapse utility, however, Appendix B.15 Table 78 of the clarification response shows the EQ-5D measurement per cycle per treatment for a number of patients at a number of measurement points."	that their study has not assessed relapse utility, however, Appendix B.15 Table **75** of the clarification response shows the EQ-5D measurement per cycle per treatment for a number of patients at a number of measurement points."		latest version of clarification response document.

Issue 24

Description of problem	Description of proposed amendment	Justification for amendment	ERG response
Page 99 of the report states "The model assumed a relative dose intensity of [redacted] based on the QUAZAR AML-001 trial. The price for 14 doses of 300 mg was set at £[redacted]."	Please consider amending this text to: "The model assumed a relative dose intensity of [redacted] based on the QUAZAR AML-001 trial. The price for 14 doses of 300 mg was set at £[redacted]."	Transcription error.	Not an ERG factual error Amended

Issue 25

Description of problem	Description of proposed amendment	Justification for amendment	ERG response
Page 99 of the report states "The cost of midostaurin per unit was £100.18, with four units administered daily to a cost-per cycle of £11,219.88."	Please consider amending this text to: "The cost of midostaurin per unit was £100.18, with four units administered daily to a cost-per cycle of ~~£11,219.88~~ **£10,658.89**.	Transcription error. The updated model provided to the ERG incorporates the RDI for midostaurin. This leads to a final cost per cycle of £10,658.89.	Amended

Issue 26

Description of problem	Description of proposed amendment	Justification for amendment	ERG response
Page 103 of the report states "The updated base-case cost effectiveness results (probabilistic) indicated that oral azacitidine (with PAS) is both more costly (additional costs of [redacted]) and more effective (incremental QALYs of [redacted]) than watch and wait plus BSC amounting to an ICER of £44,714 per QALY gained (Table 5.1)."	Please consider amending this text to: "The updated base-case cost effectiveness results (probabilistic) indicated that oral azacitidine (with PAS) is both more costly (additional costs of [redacted]) and more effective (incremental QALYs of [redacted]) than watch and wait plus BSC amounting to an ICER of **£48,332** per QALY gained (Table 5.1)."	Transcription error. The updated model provided to the ERG has an ICER of £48,332 per QALY gained.	Amended

Issue 27

Description of problem	Description of proposed amendment	Justification for amendment	ERG response
Page 103 of the report states "Increased RFS, with an incremental of 0.854 years (80% of total incremental LYs) in the oral azacitidine arm (2.085 years) compared with watch and wait with BSC arm (1.232 years).	Please consider amending this text to: "Increased RFS, with an incremental of **0.853** years (80% of total incremental LYs) in the oral azacitidine arm (**2.088** years) compared with watch and wait with BSC arm (**1.235** years)."	These amendments are based on the assumption that the intent of the ERG was to present probabilistic results using 1000 iterations.	Not a factual error. No change made as ERG prefers presenting probabilistic results using 5,000 iterations
Page 103 of the report also states "Increased post-relapse survival, with an incremental of 0.211 years (20% of total	Please consider amending this text to: "Increased post-relapse survival, with an incremental of **0.209** years (20% of total incremental LYs) in the oral azacitidine arm		

incremental LYs) in the oral azacitidine arm (1.779 years) compared with watch and wait with BSC arm (1.568 years)."	(**1.791** years) compared with watch and wait with BSC arm (**1.583** years)."		

Issue 28

Description of problem	Description of proposed amendment	Justification for amendment	ERG response
Page 103 of the report states "The higher drug costs (additional cost of [redacted], [redacted] of total incremental costs) and disease management costs (additional cost of [redacted]) in RFS on-treatment compared with watch and wait plus BSC." Page 103 of the report also states "The lower disease management costs (reduced cost of [redacted]) in RFS off-treatment compared with watch and wait plus BSC."	Please consider amending this text to: "The higher drug costs (additional cost of [redacted], [redacted] of total incremental costs) and disease management costs (additional cost of [redacted]) in RFS on-treatment compared with watch and wait plus BSC. Please consider amending this text to: "The lower disease management costs (reduced cost of [redacted]) in RFS off-treatment compared with watch and wait plus BSC."	These amendments are based on the assumption that the intent of the ERG was to present probabilistic results using 1000 iterations.	Not a factual error. No change made as ERG prefers presenting probabilistic results using 5,000 iterations

Issue 29

Description of problem	Description of proposed amendment	Justification for amendment	ERG response
Page 104 (Table 5.2) outlines the probabilistic results with oral azacitidine PAS for the FLT3 subgroup. The values in Table 5.2 are incorrect.	Please consider populating this table using a 1000 iteration PSA as was done in Table 5.1.	The current values populating Table 5.2 align with probabilistic results using 5000 iterations, except for the values reported for midostaurin. Please consider updating these values if the intent was to present probabilistic results using 1000 iterations.	Not a factual error. Amended by populating Table 5.1 with 5,000 iterations PSA results in line with Table 5.2.

Issue 30

Description of problem	Description of proposed amendment	Justification for amendment	ERG response
Page 104 of the report states "Utility values based on Joshi 2019 for all health states (decreased ICER to [redacted])".	Please consider amending this text to: "Utility values based on Joshi 2019 for all health states (decreased ICER to [redacted])"	Transcription error.	Not a factual error. No change made. Utility values based on Joshi 2019 (RFS on and off treatment 0.89 and relapse 0.51) results in ICER reported by ERG.

Issue 31

Description of problem	Description of proposed amendment	Justification for amendment	ERG response
Page 105 of the report states "Inclusion of comparators and clinical trials were cross-validated using published treatment guidelines for the management of	Please consider amending this text to: "Inclusion of comparators and clinical trials were cross-validated using published treatment guidelines for the management of AML including the ELN, published in 2017	Inclusion of comparators were cross-validated using published treatment guidelines and were also informed by clinical experts in the UK.	Not a factual error, this section is about cross validation.

AML including the ELN, published in 2017 which is the main guideline used in the UK, the BSCH published in 2006 and the European Society for Medical Oncology (ESMO) guideline published in 2013."	which is the main guideline used in the UK, the BSCH published in 2006 and the European Society for Medical Oncology (ESMO) guideline published in 2013. **In addition, inclusion of comparators were also informed by clinical experts in the UK**".		

Issue 32

Description of problem	Description of proposed amendment	Justification for amendment	ERG response
Page 105 states "No comparison with external data used to develop the economic model was performed." Page 105 also states "5.3.5 Comparison with external data not used to develop economic model" Page 105 also states "To the knowledge of the ERG, no comparison with external data (that was not used in the economic model) was performed."	Please consider amending this text to: "External data from a SLR was used to develop the economic model." Please consider amending this text to: "5.3.5 Comparison with external data was used to develop the economic model". Please consider removal of this text.	The results of the economic SLR were used to inform and develop the economic model.	Not a factual error.

Issue 33

Description of problem	Description of proposed amendment	Justification for amendment	ERG response
Page 105 states “It was, however, unclear what informed the change in these costs (set to 40% in the original submission).”	Please consider removal of this text.	Cost assumptions for nurse and haematologist visits in scenario analyses were informed from NICE TA642 using PSSRU 2018.	Not a factual error.

Issue 34

Description of problem	Description of proposed amendment	Justification for amendment	ERG response
Page 108 and 109 (Table 6.1)	Please consider renumbering the key issues (ie, begin numbering with 1, rather than 8).	Minor typographical error.	Not a factual error. The numbers reflect those in the executive summary. They have been amended to reflect the accurate pointing to Key Issues.

Issue 35

Description of problem	Description of proposed amendment	Justification for amendment	ERG response
Page 113 of the report states “The company argued that HRQoL was not measured upon relapse in the QUAZAR AML-001 trial, but these data were reported for a small number of patients in Appendix B.15 Table 78 of the clarification response and could	Please consider amending this text to: “The company argued that HRQoL was not measured upon relapse in the QUAZAR AML-001 trial, but these data were reported for a small number of patients in Appendix B.15 Table ~~78~~ **75** of the clarification response and could have been explored as a scenario analysis in the model.”	Incorrect table reference provided.	Not a factual error, ERG’s table reference aligns with latest version of clarification response document.

have been explored as a scenario analysis in the model.”			

Technical engagement response form

Oral azacitidine for maintenance treatment of acute myeloid leukaemia after induction therapy [ID3892]

As a stakeholder you have been invited to comment on the evidence review group (ERG) report for this appraisal.

Your comments and feedback on the key issues below are really valued. The ERG report and stakeholders' responses are used by the appraisal committee to help it make decisions at the appraisal committee meeting. Usually, only unresolved or uncertain key issues will be discussed at the meeting.

Information on completing this form

We are asking for your views on key issues in the ERG report that are likely to be discussed by the committee. The key issues in the ERG report reflect the areas where there is uncertainty in the evidence, and because of this the cost effectiveness of the treatment is also uncertain. The key issues are summarised in the executive summary at the beginning of the ERG report.

You are not expected to comment on every key issue but instead comment on the issues that are in your area of expertise.

If you would like to comment on issues in the ERG report that have not been identified as key issues, you can do so in the 'Additional issues' section.

If you are the company involved in this appraisal, please complete the 'Summary of changes to the company's cost-effectiveness estimates(s)' section if your response includes changes to your cost-effectiveness evidence.

Please do not embed documents (such as PDFs or tables) because this may lead to the information being mislaid or make the response unreadable. Please type information directly into the form.

Do not include medical information about yourself or another person that could identify you or the other person.

We are committed to meeting the requirements of copyright legislation. If you want to include journal articles in your submission you must have copyright clearance for these articles. We can accept journal articles in NICE Docs. For copyright reasons, we will have to return forms that have attachments without reading them. You can resubmit your form without attachments, but it must be sent by the deadline.

Combine all comments from your organisation (if applicable) into 1 response. We cannot accept more than 1 set of comments from each organisation.

Please underline all confidential information, and separately highlight information that is submitted under 'commercial in confidence' in turquoise, all information submitted under 'academic in confidence' in yellow, and all information submitted under 'depersonalised data' in pink. If confidential information is submitted, please also send a second version of your comments with that information replaced with the following text: 'academic/commercial in confidence information removed'. See the Guide to the processes of technology appraisal (sections 3.1.23 to 3.1.29) for more information.

Deadline for comments by **5pm** on **Thursday 7 April 2022**. Please log in to your NICE Docs account to upload your completed form, as a Word document (not a PDF).

Thank you for your time.

We reserve the right to summarise and edit comments received during engagement, or not to publish them at all, if we consider the comments are too long, or publication would be unlawful or otherwise inappropriate.

Comments received during engagement are published in the interests of openness and transparency, and to promote understanding of how recommendations are developed. The comments are published as a record of the comments we received, and are not endorsed by NICE, its officers or advisory committees.

About you

Table 1 About you

Your name	[redacted]
Organisation name: stakeholder or respondent (if you are responding as an individual rather than a registered stakeholder, please leave blank)	**Bristol Myers Squibb Pharmaceuticals Ltd**
Disclosure Please disclose any past or current, direct or indirect links to, or funding from, the tobacco industry.	**None**

Key issues for engagement

All: Please use the table below to respond to the key issues raised in the ERG report.

Table 2 Key issues

Key issue	Does this response contain new evidence, data or analyses?	Response
Key issue 1: Low dose cytarabine and subcutaneous azacitidine are part of standard therapy according to NICE guidance yet were not viewed by the company to be part of BSC.	Yes	**Current treatment guidelines, real-world UK clinical practice, and clinical expert opinion confirm the exclusion of low dose cytarabine (LDAC) and subcutaneous (SC) azacitidine as relevant comparators.** The ERG cites NICE Technology Appraisal guidance (TA552, TA545, TA218) to justify the inclusion of LDAC and SC azacitidine as comparators in the oral azacitidine appraisal. However, these guidance documents are not specific to the maintenance setting and therefore, are not applicable to this oral azacitidine appraisal. More specifically, within its report the ERG focuses on consolidation stating that, "*At least some of these guidelines note that the treatments (including cytarabine) are recommended for consolidation therapy.*" Given that oral azacitidine is not licensed for consolidation therapy, the examples provided by the ERG are not relevant. **Treatment Guidelines** As stated in clarification question response B3.a, neither LDAC or SC azacitidine are recommended by NICE for the patient population eligible for maintenance treatment with oral azacitidine, nor is their use mentioned or endorsed for maintenance treatment in either the ELN (2017),[1] ESMO (2020),[2] or BSCH (2006) guidelines.[3] The QUAZAR population comprises patients in complete remission or complete remission with incomplete blood count recovery (CR/CRi) following intensive chemotherapy who are ineligible for

		transplant. Guidelines from ESMO (see **Appendix A**) illustrate that, among patients who are eligible for intensive chemotherapy, options for maintenance post response are limited to midostaurin for the subgroup with FLT3 mutations (**Figure A.1**).[2] The place in therapy for LDAC and SC azacitidine is front-line treatment for patients who are ineligible for intensive chemotherapy (**Figure A.2**), and in the R/R AML setting for patients who are ineligible for intensive chemotherapy (not shown; Figure 3 in ESMO guidelines).[2] This treatment paradigm is fully aligned with ELN guidelines which are widely followed in UK clinical practice.[1] Recent NICE Technology Appraisal guidance (e.g., TA765 [**Figure A.3**], TA523, etc.) also identify that LDAC and SC azacitidine are restricted to front line treatment of patients who are ineligible for intensive chemotherapy. **Real World UK Clinical Practice** The Haematological Malignancy Research Network (HMRN) is an ongoing population-based cohort of UK patients which was established to provide robust, generalisable data about leukaemias, lymphomas, myelomas, and related blood disorders to inform clinical practice and research. In order to match the QUAZAR population, a recent audit report by the HMRN sought to describe a subgroup of patients treated with induction therapy with or without consolidation who had achieved a CR or CRi, were 55 years or older, and who did not receive a stem cell transplant. In general, there was limited use of maintenance therapy in this subgroup of UK patients. Among [redacted] patients (total of subgroup 2, see Figure 1 of HMRN report in **Appendix G**), only [redacted] received SC azacitidine in the maintenance setting and [redacted] received LDAC (see Table 8 of HMRN report in **Appendix G**). This confirms that in real world UK clinical practice these treatments are not used as standard therapy in the maintenance setting. **Clinical Expert Opinion** Three independent clinical experts have also confirmed that LDAC and SC azacitidine are not part of standard therapy in the maintenance setting and are not relevant comparators. Please see clinical summaries in **Appendix F** for further details.
Key issue 2: Most patients in the QUAZAR trial received one dose or no doses of consolidation therapy, resulting in a	Yes	**Most patients in the QUAZAR trial received 1 or more cycles of consolidation therapy (pre-randomisation), reflecting real-world UK clinical practice.** As stated in clarification question response A12, the number of consolidation cycles was not driven by the trial protocol within the QUAZAR AML-001 study given patients were randomised post consolidation. Thus, the variation in the number of consolidation cycles reflects routine clinical

selection bias that could have exaggerated the benefits of oral azacitidine.		practice. Clinical experts confirmed that there is variability in the number of consolidation cycles and that up to 60% of UK patients are likely to receive only one dose or no dose in routine practice (please see clinical summaries in **Appendix F** for further details). Moreover, the HMRN report identified that among patients with a response to intensive induction chemotherapy in first line, [redacted] did not receive any cycles of consolidation (please see Table 4 of the HMRN report available in **Appendix G**). A subsequent analysis limited to the older population (patients ≥55 years; subgroup 2) showed that a much larger proportion of patients did not receive consolidation [redacted] reflecting that age is a factor in decision-making (see Appendix I of HMRN report available in **Appendix G**). This older subgroup was more representative of the QUAZAR population which also included older patients (median age in subgroup 2: [redacted] years vs. median age in QUAZAR: 68 years). Compared to subgroup 2 of the HMRN report [redacted], the proportions of patients in the QUAZAR study who did not receive consolidation were relatively low (20.0% in the ITT population and [redacted] in the EU subgroup; **Table 1**). These findings reflect the variability in the number of consolidation cycles used in clinical practice. **Table 1: Summary of Consolidation Therapy Usage in QUAZAR – ITT and EU** **ITT** / **EU** **Parameter** / **Oral azacitidine (N=238)** / **Placebo (N=234)** / **Total (N=472)** / **Oral azacitidine (N=167)** / **Placebo (N=147)** / **Total (N=314)** ***Received Consolidation Therapy Following Induction Therapy – n (%)*** **No** / 52 (22) / 42 (18) / 94 (20) / [redacted] **Yes** / 186 (78) / 192 (82) / 378 (80) / [redacted] **1 Cycle** / 110 (46) / 102 (44) / 212 (45) / [redacted] **2 Cycles** / 70 (29) / 77 (33) / 147 (31) / [redacted] **3 Cycles** / 6 (2.5) / 13 (6) / 19 (4) / [redacted] Abbreviations: EU = Europe; ITT = intention to treat. Data cutoff date: 15 Jul 2019 Source: ITT - Supplementary appendix to Wei et al. 2020; EU – see Table 6 of EU regional analysis (**Appendix H**)

		Subgroup analyses defined by number of consolidation cycles are presented in section B2.7.5 of the CS and demonstrate the treatment effect of oral azacitidine across subgroups. OS and RFS KM curves for these consolidation subgroups are presented in clarification question response A12. Compared to placebo, oral azacitidine was associated with consistent survival benefits (OS and RFS) regardless of the number of prior consolidation cycles (0, 1, and ≥2 cycles).[4]
Key issue 3: Few patients in the QUAZAR trial were recruited from UK sites, and there were relevant differences between the UK and analysed populations; this limits the generalisability to UK clinical practice.	Yes	**The economic model has been updated and now uses the EU-subgroup as the base case (rather than ITT) to improve generalisability to the UK setting.** Using the EU subgroup as the base case results in a deterministic ICER of £32,718/QALY. Further rationale to support the EU-subgroup as an alternative base case was previously provided in clarification question response A23**.** **This also aligns with feedback from the ERG report which states, "*The ERG however felt that the EU-subgroup would be more in line with what is expected to be seen in UK clinical practice*."** The generalisability of the QUAZAR trial's EU subgroup to the UK is supported by the findings of the HMRN report (**Appendix G**). To align with the QUAZAR trial eligibility criteria, the HMRN identified a cohort of patients treated with induction therapy with or without consolidation who had achieved a CR or CRi, were 55 years or older, and who did not receive a stem cell transplant. This cohort was referred to as subgroup 2. As shown in Appendix I of the HMRN report, the baseline characteristics of subgroup 2 were in line with the QUAZAR trial's EU subgroup, including the median age and sex distributions.
Key issue 4: HRQoL and fatigue were measured on day 1 of each 28-day cycle, when adverse	No	**To mitigate any risk that treatment-related AEs were not fully captured in the HRQoL measurements from the QUAZAR AML-001 trial, AE disutilities have been applied to the health state utility values in the base case.** BMS acknowledges that measuring HRQoL and fatigue only on day 1 of 28-day cycles may result in some treatment-related AEs not being captured. As stated in clarification question response A24,

events were less likely to arise.		the application of AE disutilities in the model ensured that the HRQoL impact for patients who experienced fatigue and other treatment-related AEs between measurement intervals would still be accounted for in the cost-effectiveness results. BMS also acknowledges the ERG's concern that, "*the reduction of number of clinical visits in a cycle from 2 (days 1 and 15) to 1, with the day 15 visit being optional and at the discretion of the investigator, beginning from cycle 25, might not have been appropriate in capturing the HRQoL outcome and AEs*." As HRQoL and fatigue were only measured on day 1 of each 28-day cycle, the change in the frequency of visits beginning from cycle 25 would not affect HRQoL and fatigue data collection.
Key issue 5: Randomisation of patients in RATIFY trial occurred at induction and not maintenance phase, potentially introducing a high risk of bias in any analysis at the maintenance phase.	No	**BMS agrees that randomisation of patients in the RATIFY trial occurred at induction and not the maintenance phase which may introduce bias into any analysis examining the maintenance phase. The RATIFY trial was not designed to determine the independent effect of maintenance therapy and the authors of RATIFY acknowledged this as a limitation.[5] However, an alternative ITC using data specific to the maintenance phase was not possible because the SLR did not find any other RCTs beyond RATIFY that provided evidence for midostaurin.** As stated in clarification question response A26, the study design of the QUAZAR AML-001 and RATIFY trials differed substantially. Although both trials included treatment with maintenance therapy, only QUAZAR AML-001 was prospectively designed to evaluate the efficacy of maintenance therapy in comparison with placebo. In contrast, the RATIFY trial was designed to assess the addition of midostaurin to induction and consolidation with standard chemotherapy versus chemotherapy alone. Although the RATIFY trial included a 12-month maintenance phase, patients were not re-randomised prior to the start of maintenance, so the independent effect of maintenance therapy could not be determined. Furthermore, not all patients who initiated treatment in the RATIFY trial went on to receive maintenance therapy.[5] While we recognise the limitations of indirect comparisons of oral azacitidine and midostaurin, the analysis conducted remains the most appropriate given the data available.

Key issue 6: The SLR eligibility criteria would not have identified the RATIFY trial; other midostaurin studies may also have been missed.	No	**The SLR eligibility criteria were designed to capture all relevant studies within AML, not just those with oral azacitidine as the comparator** As stated in clarification question response A25, the SLR eligibility criteria included midostaurin studies and captured the RATIFY trial. More specifically, the clinical SLR was designed to identify all trials with a relevant intervention (oral azacitidine) or comparator including midostaurin (RATIFY), placebo. It appears the ERG interpreted the SLR criteria to require Intervention AND Comparator. However, the actual criteria allowed for Intervention AND/**OR** Comparator and thus, the inclusion of studies was not restricted to trials evaluating both the intervention and comparator. Had such an approach been taken, only the QUAZAR AML-001 study would have been identified. In key issue 5, the ERG suggested, "*If possible, conducting an analysis using an RCT of midostaurin where patients were randomised [during the] maintenance phase."* However, since the SLR included midostaurin studies and did not find an RCT of midostaurin where patients were randomised in the maintenance phase, the alternative approach suggested by the ERG is unfortunately not possible.
Key issue 7: HSCT was not included as a separate health state but was implicitly included in the modelling through the survival analysis, increasing the likelihood of bias.	Yes	**HSCT was appropriately reflected in the model through the survival analysis, and not as a separate health state. The QUAZAR AML-001 trial did not collect sufficient data to allow modelling of HSCT as a separate health state, nor were these data available in the literature.** While patients who underwent HSCT in QUAZAR AML-001 were followed for survival, further parameters around HSCT and subsequent treatments were not recorded, including the proportion achieving successful HSCT and post-transplant outcomes. In models where HSCT has been included as a health state, the proportion of patients receiving HSCT tends to be substantially higher and HSCT is often administered during first CR rather than post-relapse. For example, in the RATIFY trial 59% of RYDAPT treated patients underwent HSCT.[5] Among these patients, 47.6% received HSCT during the first CR.[5] In contrast, subjects in the QUAZAR AML-001 trial who received another therapy (e.g., HSCT) for AML without documented relapse were censored on the date of the last bone marrow assessment, prior to receiving the other therapy.[6] Thus, the efficacy of these subsequent therapies did not contribute to RFS.[6] The OS hazard ratio [redacted] was maintained when censoring for HSCT ([redacted] QUAZAR AML-001, Table 14.2.1.5.4).[6] Further explanation is provided in

clarification question response B1. As described in clarification question responses A8 and A23, the most common reasons for transplant ineligibility in the QUAZAR population were age and comorbidities, reasons which would persist over time and most patients would remain unsuitable for transplant. However, a proportion were unsuitable due to donor availability, patient decision, or performance status. These criteria may change over time and therefore allowed some patients in the trial to become more suitable for HSCT.

A scenario analysis was conducted which included a utility benefit for patients on HSCT, to explore the long-term HRQoL impacts of HSCT. In the EU subgroup of the QUAZAR AML-001 trial, [redacted] of oral azacitidine patients and [redacted] of placebo plus BSC patients received HSCT (see **Appendix H** for further details). As in the company base case, the scenario analysis applied a utility decrement of -0.21 in the first 28-day cycle for all patients who received HSCT to capture the short-term impact of the procedure on HRQoL.[7] This was considered conservative given negative HRQoL impacts are often observed up to a year post-transplant.[8] However, unlike the base case, the scenario also considered a weighted average health state utility value for the relapse health state to account for the potential long-term HRQoL benefits of HSCT in those patients who went on to receive this as a subsequent treatment. Based on findings from the literature, the scenario analysis estimated that 38% of patients who received HSCT would develop chronic graft versus host disease with an associated utility of 0.37 (Bachier 2021;[9] Joshi 2019[8]). It was assumed that the remaining 62% of patients who received HSCT would not have long-term complications and thus, a utility of 0.94 was applied to represent the benefits of HSCT (Bachier 2021;[9] Joshi 2019[8]). All patients who did not receive HSCT were assumed to have a relapse utility value of 0.45 (Joshi 2019[8]). Using these inputs, a weighted average relapse utility value was calculated by treatment arm for the scenario analysis (**Table 1**). The results of the scenario analysis for the EU population are shown in **Table 2**. The ICER for the scenario analysis increased by 2.3% compared to the base case.

Table 1: Weighted average health state utility values for relapse in HSCT scenario analysis

Technology	Relapse Utility Value
Base Case – ITT	
Watch and wait plus BSC	0.45
Oral azacitidine	
Scenario Analysis – Long-term HRQoL Impacts of HSCT – ITT	

Watch and wait plus BSC	[redacted]
Oral azacitidine	[redacted]

Abbreviations: BSC = best supportive care; HRQoL = health related quality of life; HSCT = hematopoietic stem cell transplant; ITT = intention to treat.

Table 2: Deterministic results for scenario analysis of long-term HRQoL impacts of HSCT

Technology	Total costs (£)	Total LYG	Total QALYs	Inc. costs (£)	Inc. LYG	Inc. QALYs	ICER (£/QALY)
EU base case							
Watch and wait plus BSC	[redacted]	2.62	[redacted]	-	-	-	-
Oral azacitidine	[redacted]	3.97	[redacted]	[redacted]	1.35	[redacted]	32,718
EU base case – HSCT relapse scenario							
Watch and wait plus BSC	[redacted]	2.62	[redacted]	-	-	-	-
Oral azacitidine	[redacted]	3.97	[redacted]	[redacted]	1.35	[redacted]	33,472

Abbreviations: BSC = best supportive care; EU = Europe; HRQoL = health related quality of life; HSCT = hematopoietic stem cell transplant; ICER = incremental cost-effectiveness ratio; Inc. = incremental; LYG = life years gained; QALYs = quality-adjusted life years

As requested, additional details on the survival analyses are provided in **Appendix B**, including an overlay of the KM curves of ITT versus ITT with HSCT censored in one plot, the AIC/BIC fit for the individual distributions per treatment arm for the HSCT censored analysis, and all distributions in one plot. These analyses further reinforce that the assessment of curve fits for the ITT population censored for HSCT are aligned with the overall ITT population, with joint generalized gamma providing the optimal fit for OS and joint log-logistic providing the optimal fit for RFS.

<table>
<tr>
<td>Key issue 8: Some patients in QUAZAR trial received fewer cycles of consolidation therapy than is standard practice in the UK. This limits the applicability of the results to a UK setting.</td>
<td>Yes</td>
<td>Patients in the QUAZAR trial received sufficient consolidation therapy that was representative of UK clinical practice, so the results are applicable to the UK setting.

Please refer to the response to Key issue 2 for further details on the standard use of consolidation therapy in UK clinical practice.

In Appendix C, we have provided detailed assessments of the suitability of the survival models for the consolidation subgroup according to the NICE DSU TSD 14[10] criteria (Section 3), including log-cumulative hazard plots, evaluation of criterion 5 (OS/RFS gain pre- and post-extrapolation), and all distributions in one plot. Additional survival analyses in the consolidation subgroup aligned with the assessment for the ITT population, with joint generalised gamma providing the optimal fit for OS and joint log-logistic providing the optimal fit for RFS.</td>
</tr>
<tr>
<td>Key issue 9: Patient baseline characteristics in the model are not subgroup-specific (for example in the FLT3 subgroup, consolidation subgroup or Europe subgroup); patient baseline characteristics may not align with the subgroups being analysed.</td>
<td>Yes</td>
<td>Patient baseline characteristics in the model have been updated to be subgroup-specific and scenario analyses have been conducted. Overall, the use of subgroup-specific baseline characteristics had no material impact on the cost-effectiveness results.

Please see Table 1 for a list of baseline characteristics by subgroup.

Table 1: Patient baseline characteristics by subgroup
<table>
<tr><th>Parameter</th><th>ITT</th><th>EU</th><th>≥1 cycles of consolidation</th><th>FLT3</th></tr>
<tr><td>Average body weight (kg)</td><td>[redacted]</td><td>[redacted]</td><td>[redacted]</td><td>[redacted]</td></tr>
<tr><td>Body surface area (m^2)</td><td>[redacted]</td><td>[redacted]</td><td>[redacted]</td><td>[redacted]</td></tr>
<tr><td>Starting age (years)</td><td>67.9</td><td>[redacted]</td><td>[redacted]</td><td>[redacted]</td></tr>
<tr><td>Percent male</td><td>51.9%</td><td>[redacted]</td><td>[redacted]</td><td>[redacted]</td></tr>
<tr><td>Average height (cm)</td><td>[redacted]</td><td>[redacted]</td><td>[redacted]</td><td>[redacted]</td></tr>
</table>
</td>
</tr>
</table>

Abbreviations: EU = Europe; ITT = intention to treat.

Deterministic results for the EU and Consolidation subgroups when using subgroup-specific baseline characteristics (as compared to ITT) are shown below (**Table 2**).

Table 2: Deterministic results for the EU and Consolidation subgroup when using subgroup specific baseline characteristics

Technology	Total costs (£)	Total LYG	Total QALYs	Inc. costs (£)	Inc. LYG	Inc. QALYs	ICER (£/ QALY)
EU Subgroup (ITT baseline characteristics)							
Watch and wait plus BSC	[redacted]	2.63	[redacted]	-	-	-	-
Oral azacitidine	[redacted]	3.99	[redacted]	[redacted]	1.36	[redacted]	32,512
EU Subgroup (EU-specific baseline characteristics)							
Watch and wait plus BSC	[redacted]	2.62	[redacted]	-	-	-	-
Oral azacitidine	[redacted]	3.97	[redacted]	[redacted]	1.35	[redacted]	32,718
Consolidation Subgroup (ITT baseline characteristics)							
Watch and wait plus BSC	[redacted]	2.95	[redacted]	-	-	-	-
Oral azacitidine	[redacted]	3.82	[redacted]	[redacted]	0.87	[redacted]	41,554
Consolidation Subgroup (Consolidation subgroup baseline characteristics)							
Watch and wait plus BSC	[redacted]	2.96	[redacted]	-	-	-	-
Oral azacitidine	[redacted]	3.84	[redacted]	[redacted]	0.87	[redacted]	41,238

Abbreviations: BSC = best supportive care; EU = Europe; HSCT = hematopoietic stem cell transplant; ICER = incremental cost-effectiveness ratio; Inc. = incremental; ITT = intention to treat; LYG = life years gained; QALYs = quality-adjusted life years; RDI = relative dose intensity; ToT = time on treatment.

Deterministic results for the FLT3 subgroup when using subgroup-specific baseline characteristics (as compared to ITT) are shown below (**Table 3**).

Table 3: Deterministic results with oral azacitidine PAS for the FLT3 subgroup when using subgroup specific baseline characteristics

Tech.	Total costs (£)	Total LYG	Total QALYs	Inc. costs (£)	Inc. LYG	Inc. QALYs	ICER versus baseline (£/QALY)	Pairwise ICER versus oral azacitidine (£/QALY)
FLT3 Subgroup (ITT baseline characteristics)								
Watch and wait plus BSC	[redacted]	2.73	[redacted]	-	-	-	-	19,536
Oral azacitidine	[redacted]	4.83	[redacted]	[redacted]	2.10	[redacted]	19,536	-
Mido	[redacted]	3.60	[redacted]	[redacted]	0.87	[redacted]	278,182	Oral azacitidine is dominant
FLT3 Subgroup (FLT3 baseline characteristics)								
Watch and wait plus BSC	[redacted]	2.78	[redacted]	-	-	-	-	19,063
Oral azacitidine	[redacted]	4.95	[redacted]	[redacted]	2.18	[redacted]	19,063	-
Mido	[redacted]	3.67	[redacted]	[redacted]	0.89	[redacted]	269,191	Oral azacitidine is dominant

Abbreviations: BSC = best supportive care; ICER = incremental cost-effectiveness ratio; Inc. = incremental; ITT = intention to treat; LYG = life years gained; PAS = patient access scheme; QALYs = quality-adjusted life years; Tech. = technology.

Key issue 10: Survival analyses of the FLT3 subgroup are likely to be biased due to limitations	Yes	**BMS agrees that survival analyses of the FLT3 subgroup are likely to be biased due to inherent, unavoidable limitations associated with the indirect treatment comparison (ITC). However, the anchored Bucher ITC is the most appropriate option available to reduce this uncertainty.**

associated with the indirect comparison.		The anchored Bucher ITC is the most appropriate option because anchored comparisons relax the assumption of balanced prognostic variables between the trials that unanchored comparisons require and therefore provide a less biased effect estimate. As described in the submission, there were limitations in the estimates of comparative efficacy for the FLT3 subgroup related to data availability and trial design that likely introduced bias. The indirect treatment comparison provides the most robust estimate possible accounting for these limitations. In **Appendix D**, we have provided detailed assessments of the suitability of the survival models for the FLT3 subgroup according to the NICE DSU TSD 14[10] criteria (Section 3), including log-cumulative hazard plots, AIC/BIC statistics for individual models, plots showing all joint models in one plot, and evaluation of criterion 5. Additional survival analysis assessments for the FLT3 subgroup show parametric models provide a poor fit to the KM curves, likely due to the uncertainty associated with small sample size. Spline models for this subgroup were explored to address the uncertainty (Section B.2.9.4). Ultimately, clinician input aided in choosing clinically plausible survival extrapolations for the FLT3 subgroup. For clarity, BMS has previously provided the following efficacy and safety data for the FLT3 subgroup in the QUAZAR AML-001 trial: • OS and RFS (section B.2.7.3 of company submission) • AEs (Table B.3.8 of company submission)
Key issue 11: In the company's base-case analysis, only grade 3 and 4 AEs are applied with a maximum frequency of one and a duration of 1 week, which may underestimate the real impact of AEs.	Yes	**The frequency, duration, and disutility of adverse events were appropriately captured in the cost-effectiveness model with additional scenario analyses confirming the robustness of the base case model estimates.** In response to the initial concerns raised by the ERG during clarification questions, three additional scenario analyses were conducted: 1) addition of low-grade AEs (grade 1 and 2 occurring in ≥5% of patients; B12c), 2) additional AEs (grade 3 and 4 occurring in ≥2% of patients; B12b), and 3) applying AE disutilities over an average duration of 4 weeks vs. 1 week (base case) (B14b). All scenario analyses demonstrated a minimal impact on the ICER, with an increase of 2.6%, 0.08%, and 0.21%, respectively (**Table 1**).

In addition to these scenarios, the ERG questioned the validity of assuming a maximum frequency of 1 per patient per AE and requested a scenario analysis assuming that AEs occur in every on-treatment interval. As stated in the clarification question response to B14.c and B14.d, assuming a maximum frequency of 1 per patient per AE is a common and simplifying assumption which is consistent with the methods used in previous AML submissions. Further, the costs and disutilities are front-loaded in the model and thus are not impacted by discounting or reduced survival over time. This approach is considered conservative and may overestimate the real impact of AEs. Therefore, the suggested scenario was not conducted as it was not considered a valid application of AE frequencies and would over-estimate AE cost and disutilities in the model.

Table 1: Deterministic results for adverse event scenario analyses

Technology	Total costs (£)	Total LYG	Total QALYs	Inc. costs (£)	Inc. LYG	Inc. QALYs	ICER (£/QALY)
EU base case							
Watch and wait plus BSC	[redacted]	2.62	[redacted]	-	-	-	-
Oral azacitidine	[redacted]	3.97	[redacted]	[redacted]	1.35	[redacted]	32,718
EU base case – Grade 1 and 2 AEs included							
Watch and wait plus BSC	[redacted]	2.62	[redacted]	-	-	-	-
Oral azacitidine	[redacted]	3.97	[redacted]	[redacted]	1.35	[redacted]	33,559
EU base case – Grade 3 and 4 AEs occurring in ≥2% of the population included							
Watch and wait plus BSC	[redacted]	2.62	[redacted]	-	-	-	-
Oral azacitidine	[redacted]	3.97	[redacted]	[redacted]	1.35	[redacted]	32,744
EU base case – 4-week AE disutility duration (vs. 1-week in base case)							
Watch and wait plus BSC	[redacted]	2.62	[redacted]	-	-	-	-
Oral azacitidine	[redacted]	3.97	[redacted]	[redacted]	1.35	[redacted]	32,786

Abbreviations: AE = adverse event; BSC = best supportive care; EU = Europe; ICER = incremental cost-effectiveness ratio; inc. = incremental; LYG = life years gained; PAS = patient access scheme; QALYs = quality-adjusted life years.

Key issue 12: The current source of utility values may not accurately reflect the relapse utility.	Yes	**Relapse utilities were not uniformly reported in the QUAZAR AML-001 trial. Joshi 2019 remains the most valid source to inform relapse utility in the model.[8]** As stated in clarification question response B15.a.a-c, some patients in the QUAZAR AML-001 study who experienced a relapse with a bone marrow blast count between >5% and ≤15% may have received dose escalation. As these patients continued study treatment, utility data following relapse were collected. However, these utility data were severely limited because they were specific to patients with a moderately high bone marrow blast count (i.e., between >5% and ≤15%) and would not have captured patients with advanced disease (e.g. blast count >15%). Therefore, using a utility value derived from this small cohort of patients would be inappropriate and may overestimate the quality of life of relapsed patients. Joshi 2019 was selected to inform the relapse health state utility value in the base case analysis for several reasons: • The elicitation methodology used is preferred by NICE[11] (i.e., the composite time-trade off methodology is a choice-based method); • Utility values were sourced from individuals in the UK, increasing the applicability of the results; and • The utility value was clinically plausible Although the company acknowledges that there remains some uncertainty around the relapse utility value from Joshi 2019 (0.45)[8] due to the small sample size, when compared to other sources including Tremblay 2018 (0.55)[12] and Stein 2019 (0.62),[13] the reported value appears clinically plausible. Scenario analyses using relapse utility values from Tremblay 2018 and Stein 2019 were conducted to account for this uncertainty (**Table 1**) and showed minimal differences in the deterministic ICER compared to the EU subgroup base case (ICER (£/QALY) decreased by 2.6% and 4.3%, respectively).

Table 1: Deterministic results for scenario analyses of alternative health state relapse utility values

Technology	Total costs (£)	Total LYG	Total QALYs	Inc. costs (£)	Inc. LYG	Inc. QALYs	ICER (£/ QALY)
EU base case – Joshi 2019 relapse utility value							
Watch and wait plus BSC	[redacted]	2.62	[redacted]	-	-	-	-
Oral azacitidine	[redacted]	3.97	[redacted]	[redacted]	1.35	[redacted]	32,718
EU base case – Tremblay 2018 relapse utility value							
Watch and wait plus BSC	[redacted]	2.62	[redacted]	-	-	-	-
Oral azacitidine	[redacted]	3.97	[redacted]	[redacted]	1.35	[redacted]	31,857
EU base case – Stein 2019 relapse utility value							
Watch and wait plus BSC	[redacted]	2.62	[redacted]	-	-	-	-
Oral azacitidine	[redacted]	3.97	[redacted]	[redacted]	1.35	[redacted]	31,306

Abbreviations: BSC = best supportive care; EU = Europe; ICER = incremental cost-effectiveness ratio; Inc. = incremental; LYG = life years gained; QALYs = quality-adjusted life years.

Key issue 13: Some resource use estimates appear inconsistent with expert opinion and require further justification.	Yes	**The resource use estimates in the cost effectiveness model were consistent with UK clinical expert opinion.** However, to further explore how alternative resource use assumptions would impact results and address uncertainty, selected scenario analyses were conducted (ii, iii, iv). i. BMS maintains that the assumption that 21.8% of patients receiving midostaurin would receive red cell transfusions is appropriate. On page 15 of the report of expert opinion under "Resource Use" it is stated that, "*For midostaurin, fewer patients than 21.8%-22.7% would require red cell transfusions*

but mainly because they are younger. If an older population would be eligible, then it is comparable." Since the QUAZAR AML-001 trial was limited to patients aged 55 years and older, BMS considers the latter to apply to the analysed population and believes 21.8% to be a valid estimate.

ii. A scenario analysis was conducted assuming 1 unit of RBC transfusion per cycle. The results are presented below (**Table 1**). The impact on the ICER was marginal (<1% decrease).

Table 1: Deterministic results for scenario analysis assuming 1 unit of RBC transfusion per cycle

Technology	Total costs (£)	Total LYG	Total QALYs	Inc. costs (£)	Inc. LYG	Inc. QALYs	ICER (£/ QALY)
EU base case							
Watch and wait plus BSC	[redacted]	2.62	[redacted]	-	-	-	-
Oral azacitidine	[redacted]	3.97	[redacted]	[redacted]	1.35	[redacted]	32,718
EU base case - 1 unit of RBC							
Watch and wait plus BSC	[redacted]	2.62	[redacted]	-	-	-	-
Oral azacitidine	[redacted]	3.97	[redacted]	[redacted]	1.35	[redacted]	32,481

Abbreviations: BSC = best supportive care; EU = Europe; ICER = incremental cost-effectiveness ratio; Inc. = incremental; LYG = life years gained; QALYs = quality-adjusted life years; RBC = red blood count.

iii/iv. A scenario analysis was conducted to align the distribution of subsequent therapies more closely with clinical expert opinion (see **Appendix F**). Additional therapies, including VenAza and FLAG-IDA, were considered in this analysis. The difference between inputs in the base case and scenario analysis are shown in **Table 2**. The results of the scenario analysis are shown in **Table 3**. The impact on the ICER was marginal (~2% increase).

Table 2: Distribution of subsequent therapies in base case vs. scenario analysis

Technology	Low-Dose Cytarabine	Azacitidine IV	Salvage chemo: 3+7	Salvage chemo: FLAG-IDA	VenAza	Total
Base Case						
Watch and wait plus BSC	10.7%	15.4%	33.8%	-	-	59.9%
Oral azacitidine	14.3%	8.4%	26.1%	-	-	48.8%
Scenario						
Watch and wait plus BSC	0%	12.5%	10%	20%	27.5%	70.0%
Oral azacitidine	0%	12.5%	10%	20%	27.5%	70.0%

Abbreviations: BSC = best supportive care; IV = intravenous.

Table 3: Deterministic results for scenario analysis assuming different distribution of subsequent therapies

Technology	Total costs (£)	Total LYG	Total QALYs	Inc. costs (£)	Inc. LYG	Inc. QALYs	ICER (£/QALY)
EU base case							
Watch and wait plus BSC	[redacted]	2.62	[redacted]	-	-	-	-
Oral azacitidine	[redacted]	3.97	[redacted]	[redacted]	1.35	[redacted]	32,718
EU base case - Alternative subsequent therapy distribution							
Watch and wait plus BSC	[redacted]	2.62	[redacted]	-	-	-	-

		Oral azacitidine | [redacted] | 3.97 | [redacted] | [redacted] | 1.35 | [redacted] | 33,404 Abbreviations: BSC = best supportive care; EU = Europe; ICER = incremental cost-effectiveness ratio; Inc. = incremental; ITT = intention to treat; LYG = life years gained; QALYs = quality-adjusted life years.
Key issue 14: Treatment effectiveness in the FLT3 subgroup was analysed for the different comparisons separately; preventing comparison of oral azacitidine, midostaurin, watch and wait plus BSC.	No	**BMS agrees that the ERG's implementation of the fully incremental analysis in the FLT3 subgroup was the most appropriate approach.** The fully incremental analysis (Table B.3.37. of CS and Table 5.2 of ERG report) showed that oral azacitidine was dominant over midostaurin. When comparing against watch and wait plus BSC in the FLT3 subgroup, treatment with oral azacitidine led to a cost per QALY gained of £19,063.

Additional issues

All: Please use the table below to respond to additional issues in the ERG report that have not been identified as key issues. Please do **not** use this table to repeat issues or comments that have been raised at an earlier point in this appraisal (for example, at the clarification stage).

Table 3 Additional issues from the ERG report

Issue from the ERG report	Relevant section(s) and/or page(s)	Does this response contain new evidence, data or analyses?	Response
Additional issue 1: End of life	Section 7, page 115 of the ERG report	No	**BMS acknowledges the uncertainty regarding whether end of life (EOL) criteria apply to oral azacitidine based on the extrapolated mean OS estimates from the model. However, the trial data clearly show that the majority of patients do not live beyond 2 years.** BMS maintains that in the EU subgroup of the QUAZAR study (2020 data), median OS in the BSC arm was only [redacted] months with only [redacted] of patients alive at 24 months which meets NICE End of Life criteria 1 - life expectancy < 24 months.[14] Furthermore, in the EU subgroup oral azacitidine was associated with an incremental increase in OS of [redacted] months compared to BSC which meets NICE End of Life criteria 2 - extension of life by at least 3 months compared to current NHS treatments. Similarly, in the ITT population of the QUAZAR study (2019 data), median OS in the BSC arm was only 14.8 months with only 37% of patients alive at 24 months. In the ITT population, oral azacitidine was

			associated with an incremental increase in OS of 9.9 months compared to BSC. Most patients with AML who achieve a CR/CRi after induction chemotherapy will experience disease relapse. It is clear that a substantial unmet need remains for a well-tolerated and easily administered AML maintenance treatment that significantly prolongs survival among patients with AML who are in remission after intensive chemotherapy without compromising HRQoL.
Additional issue 2: Clarification on duration of oral azacitidine treatment effect	Section 3.2.5.1, page 49/50 (and tables 3.2, 3.3) of the ERG report	No	The ERG suggests a rapid diminishing of treatment effect after three years compared to placebo. However, this is based on the July 2019 data cut-off (Figure 3.2). The extension phase of QUAZAR (Sept 2020 cut-off) provides a more mature data cut (Figure 3.3) and demonstrates a persistent treatment effect beyond three years.
Additional issue 3: Clarification on placebo and best supportive care (BSC)	Section 2.3.1, page 28 of the ERG report	No	Clarification question response A5.b states that the placebo arm within the QUAZAR study is considered to represent 'watch and wait'. Throughout the treatment period of the trial, patients in both arms were permitted to receive best supportive care (BSC) in addition to either oral azacitidine or placebo (see Table B.2.3 of CS for definition of concomitant therapy). Therefore, placebo reflects a 'watch and wait plus BSC' strategy.
Additional issue 4: Design of QUAZAR	Section 3.2.1, page 38 of the ERG report	No	To clarify, QUAZAR AML-001 was designed to account for the possibility of up to 4 cycles of consolidation treatment in line with treatment

AML-001 trial to account for consolidation			guidelines (e.g., ELN 2017) which recommend, for younger patients with favourable risk genetics, 2-4 cycles of consolidation treatment. This upper limit of up to 4 cycles of consolidation does not indicate that 4 cycles is the standard of care, particularly when considering the QUAZAR population was comprised of older patients (≥55 years) with intermediate- or poor-risk cytogenetics at AML diagnosis.
Additional issue 5: Clarification on the definition of documented relapse	Section 3.2.5.3, page 54 of the ERG report	No	Please refer to Table B.2.5 of the company submission for the definition of documented relapse: *Documented relapse was defined as the earliest date of any of the following (according to IWG for AML criteria):* • *≥5% BM blasts from the central pathology report;* • *The appearance of >0% blasts in the peripheral blood with a later BM confirmation (BM blasts ≥5%) within 100 days; or* • *At least two peripheral blasts ≥5% within 30 days*
Additional issue 6: HRQoL tables	Section 3.2.5.5, page 56 of the ERG report	Yes	Please see **Appendix I** for the requested HRQoL results tables (Tables 14.3.6.2.1-14.3.6.2.4).

Summary of changes to the company's cost-effectiveness estimate(s)

Company only: If you have made changes to the base-case cost-effectiveness estimate(s) in response to technical engagement, please complete the table below to summarise these changes. Please also provide sensitivity analyses around the revised base case. If there are sensitivity analyses around the original base case which remain relevant, please re-run these around the revised base case.

Table 4 Changes to the company's cost-effectiveness estimate

Key issue(s) in the ERG report that the change relates to	Company's base case before technical engagement	Change(s) made in response to technical engagement	Impact on the company's base-case incremental cost-effectiveness ratio (ICER)
Key issue 3: Few UK patients and questionable generalisability to UK NHS setting Key issue 9: Subgroup specific patient baseline characteristics Additional issue 1: End of Life	The original base-case population was the ITT population from the QUAZAR AML-001 trial. All model inputs were based on this population.	To address uncertainty around the generalisability of the cost-effectiveness estimates in the ITT population to the UK NHS setting, the base-case was updated to model the EU subgroup of the QUAZAR AML-001 trial. The updated base-case uses EU-specific inputs including baseline characteristics, OS 2020, RFS 2019, time on treatment, relative dose intensity, and proportion of patients undergoing HSCT.	The updated base-case results in a deterministic ICER (cost/QALY) of £32,718. This is a 14.6% decrease from the deterministic ICER in the ITT base-case (£38,293).

Sensitivity analyses around revised base case

Figure 1 Tornado plot of deterministic sensitivity analysis: impact on incremental costs – (PAS) price

Abbreviations: AE = adverse event; AZA = azacitidine; incr. = incremental; RDI = relative dose intensity; SCT = stem stell transplant; RBC = red blood-cell count.

Figure 2 Tornado plot of deterministic sensitivity analysis: impact on incremental QALYs – (PAS) price

Abbreviations: AE = adverse event; AZA = azacitidine; QALYs = quality adjusted life years; RFS = relapse-free survival; SCT = stem cell transplant.

Figure 3 Tornado plot of deterministic sensitivity analysis: ICER – (PAS) price

Abbreviations: AZA = azacitidine; ICER = incremental cost effectiveness ratio; QALYs = quality adjusted life years; RDI = relative dose intensity; RFS = relapse-free survival; SCT = stem cell transplant.

Table 5 Deterministic results from the scenario analyses – PAS price

	PAS price				
	Incremental costs	**Incremental QALYs**	**Incremental LYs**	**ICER**	**% change in ICER**
Base case	[redacted]	[redacted]	**1.35**	**32,718**	**-**
Discount rate scenarios					
Discount rate. Costs: 0%, QALYs: 0%	[redacted]	[redacted]	1.66	30,337	-7.28%
Discount rate. Costs: 6%, QALYs: 6%	[redacted]	[redacted]	1.19	34,239	+4.65%
Discount rate. Costs: 0%, QALYs: 6%	[redacted]	[redacted]	1.19	40,583	+24.04%
Discount rate. Costs: 6%, QALYs: 0%	[redacted]	[redacted]	1.66	25,594	-21.77%
Time horizon scenarios					
Time horizon: 10 years	[redacted]	[redacted]	1.07	36,338	+11.06%
Time horizon: 15 years	[redacted]	[redacted]	1.25	33,974	+3.84%
Time horizon: 20 years	[redacted]	[redacted]	1.33	33,105	+1.18%
Vial sharing					
Include	[redacted]	[redacted]	1.35	32,793	+0.23%
Survival model : Extrapolation OS					
Cure model	[redacted]	[redacted]	1.51	32,230	-1.49%
Hybrid model	[redacted]	[redacted]	1.32	32,846	+0.39%
Joint log-normal model	[redacted]	[redacted]	1.41	32,312	-1.24%
Adverse events scenarios					

	PAS price				
	Incremental costs	Incremental QALYs	Incremental LYs	ICER	% change in ICER
AE disutilities doubled	[redacted]	[redacted]	1.35	32,741	+0.07%
Utility					
Joshi 2019	[redacted]	[redacted]	1.35	30,865	-5.66%

Abbreviations: AE = adverse event; ICER = incremental cost effectiveness ratio; LY = life year; OS = overall survival; QALY = quality adjusted life year; RFS = relapse-free survival.

References

[1]Dohner H, Estey E, Grimwade D, Amadori S, Appelbaum FR et al. (2017) Diagnosis and management of AML in adults: 2017 ELN recommendations from an international expert panel. *Blood* 129 (4): 424-447.

[2]Heuser M, Ofran Y, Boissel N, Brunet Mauri S, Craddock C et al. (2020) Acute myeloid leukaemia in adult patients: ESMO Clinical Practice Guidelines for diagnosis, treatment and follow-up. *Ann Oncol* 31 (6): 697-712.

[3]Milligan DW, Grimwade D, Cullis JO, Bond L, Swirsky D et al. (2006) Guidelines on the management of acute myeloid leukaemia in adults. *Br J Haematol* 135 (4): 450-474.

[4]Wei A, Roboz GJ, Dombret H, Döhner H, Schuh AC et al. (2020) CC-486 Improves Overall Survival (OS) and Relapse-Free Survival (RFS) for Patients with Acute Myeloid Leukemia (AML) in First Remission after Intensive Chemotherapy (IC), Regardless of Amount of Consolidation Received: Results from the Phase III QUAZAR AML-001 Maintenance Trial. *Blood* 136 38-40.

[5]Stone RM, Mandrekar SJ, Sanford BL, Laumann K, Geyer S et al. (2017) Midostaurin plus Chemotherapy for Acute Myeloid Leukemia with a FLT3 Mutation. *N Engl J Med* 377 (5): 454-464.

[6]Celgene (2019) Data on file. CC-486-AML-001 (QUAZAR) Clinical Study Report.

[7]Matza LS, Deger KA, Howell TA, Koetter K, Yeager AM et al. (2019) Health state utilities associated with treatment options for acute myeloid leukemia (AML). *J Med Econ* 22 (6): 567-576.

[8]Joshi N, Hensen M, Patel S, Xu W, Lasch K et al. (2019) Health State Utilities for Acute Myeloid Leukaemia: A Time Trade-off Study. *Pharmacoeconomics* 37 (1): 85-92.

[9]Bachier CR, Aggarwal SK, Hennegan K, Milgroom A, Francis K et al. (2021) Epidemiology and Treatment of Chronic Graft-versus-Host Disease Post-Allogeneic Hematopoietic Cell Transplantation: A US Claims Analysis. *Transplant Cell Ther* 27 (6): 504.e501-504.e506.

[10]Latimer NR (2013). *NICE Decision Support Unit Technical Support Documents*. eds). National Institute for Health and Care Excellence (NICE)

Copyright © 2013 National Institute for Health and Clinical Excellence, unless otherwise stated. All rights reserved.: London.

[11]Papaioannou D, Brazier J, Paisley S (2010). *NICE Decision Support Unit Technical Support Documents*. eds). National Institute for Health and Care Excellence (NICE)

Copyright © 2010 National Institute for Health and Clinical Excellence, unless otherwise stated. All rights reserved.: London.

[12]Tremblay G, Dolph M, Patel S, Brandt P, Forsythe A (2018) Cost-effectiveness analysis for midostaurin versus standard of care in acute myeloid leukemia in the United Kingdom. *Cost Eff Resour Alloc* 16 33.

[13]Stein E, Xie J, Duchesneau E, Bhattacharyya S, Vudumula U et al. (2019) Cost Effectiveness of Midostaurin in the Treatment of Newly Diagnosed FLT3-Mutated Acute Myeloid Leukemia in the United States. *Pharmacoeconomics* 37 (2): 239-253.

[14]Wei AH, Döhner H, Pocock C, Montesinos P, Afanasyev B et al. (2019) The QUAZAR AML-001 Maintenance Trial: Results of a Phase III International, Randomized, Double-Blind, Placebo-Controlled Study of CC-486 (Oral Formulation of Azacitidine) in Patients with Acute Myeloid Leukemia (AML) in First Remission. *Blood* 134 (Supplement_2): LBA-3-LBA-3.

Oral azacitidine for maintenance treatment of acute myeloid leukemia after induction therapy [ID3892]

Deterministic and Probabilistic Analyses

May 31, 2022

1.0 Base Case Results (EU Subgroup)

Table 1.1: Base Case Model Results with Oral AZA Discount ([redacted])

Technology	Total costs (£)	Total LYG	Total QALYs	Incremental costs (£)	Incremental LYG	Incremental QALYs	ICER versus baseline (£/QALY)
Deterministic							
Watch and wait plus BSC	[redacted]	2.62	[redacted]	-	-	-	-
Oral azacitidine	[redacted]	3.97	[redacted]	[redacted]	1.35	[redacted]	32,718
Probabilistic (5000 iterations)							
Watch and wait plus BSC	[redacted]	2.64	[redacted]	-	-	-	-
Oral azacitidine	[redacted]	3.99	[redacted]	[redacted]	1.35	[redacted]	32,480

Abbreviations: BSC = best supportive care; ICER = incremental cost-effectiveness ratio; LYG = life year gained; QALY = quality-adjusted life year.

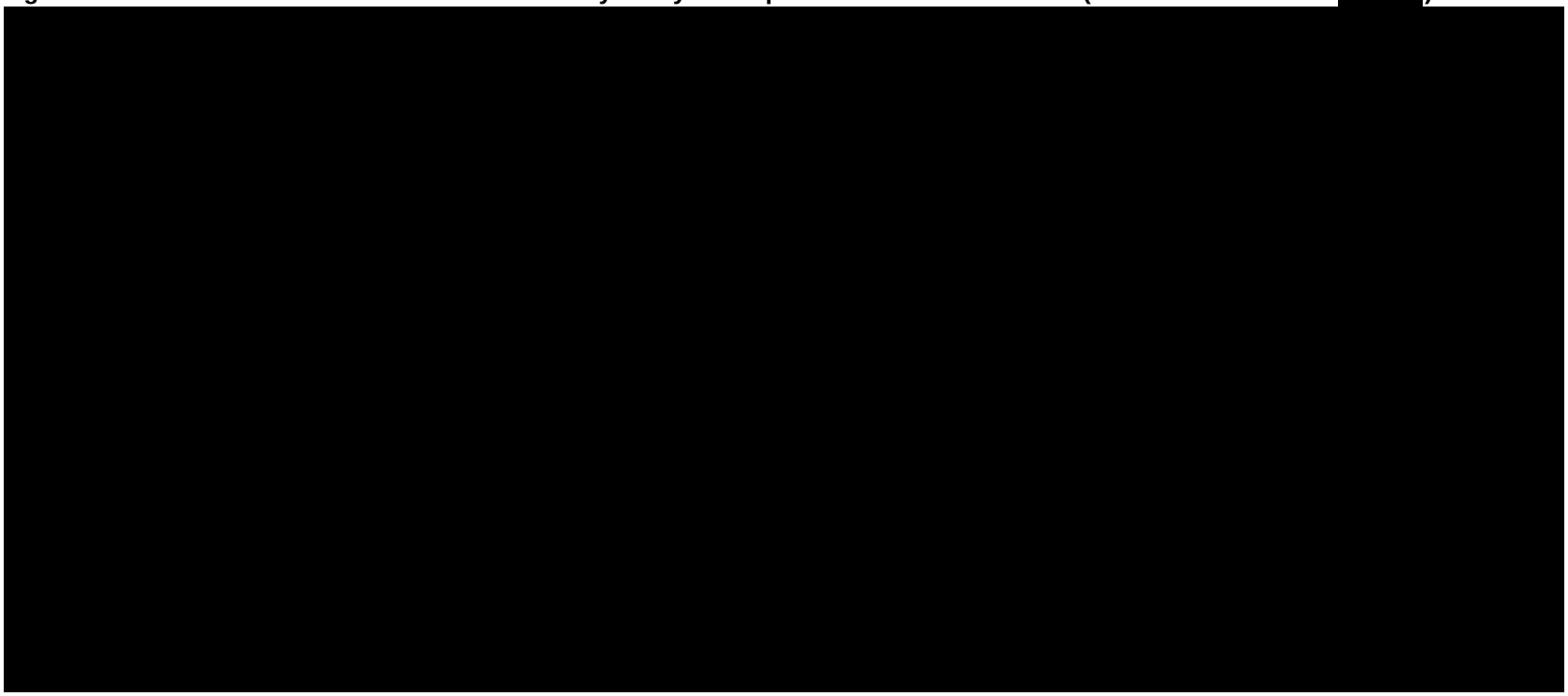
Figure 1.1: Scatterplot of Incremental Costs and QALYs (Oral AZA Discount = [redacted])

Abbreviations: AZA = azacitidine; QALY = quality-adjusted life year.

Figure 1.2: Cost-effectiveness Acceptability Curve (Oral AZA Discount = [redacted])

Abbreviations: AZA = azacitidine; CEAC = cost-effectiveness acceptability curve.

Figure 1.3: Tornado Plot of Deterministic Sensitivity Analysis: Impact on Incremental Costs (Oral AZA Discount = [redacted])



Abbreviations: AZA = azacitidine; RBC = red blood cell transfusion; RDI = relative dose intensity; SCT = stem cell transplant.

Figure 1.4: Tornado Plot of Deterministic Sensitivity Analysis: Impact on Incremental QALYs (Oral AZA Discount = [redacted])

Abbreviations: AE = adverse event; AZA = azacitidine; QALY = quality-adjusted life year; RFS = relapse-free survival; SCT = stem cell transplant.

Figure 1.5: Tornado Plot of Deterministic Sensitivity Analysis: ICER (Oral AZA Discount = [redacted])

Abbreviations: AZA = azacitidine; ICER = incremental cost-effectiveness ratio; RDI = relative dose intensity; RFS = relapse-free survival; SCT = stem cell transplant.

2.0 Scenario Analyses Results

Table 2.1: Deterministic and Probabilistic Results from Scenario Analyses

	Deterministic Results					Probabilistic Results				
	Incremental Costs	**Incremental QALYs**	**Incremental LYs**	**ICER**	**ICER Change (%)**	**Incremental Costs**	**Incremental QALYs**	**Incremental LYs**	**ICER**	**ICER Change (%)**
Base case	[redacted]	[redacted]	**1.35**	**32,718**	**-**	[redacted]	[redacted]	**1.35**	**32,480**	**-**
Discount rate										
Discount rate. Costs: 0%, QALYs: 0%	[redacted]	[redacted]	1.66	30,337	-7.28%	[redacted]	[redacted]	1.66	30,107	-7.31%
Discount rate. Costs: 6%, QALYs: 6%	[redacted]	[redacted]	1.19	34,239	+4.65%	[redacted]	[redacted]	1.19	34,000	+4.68%
Discount rate. Costs: 0%, QALYs: 6%	[redacted]	[redacted]	1.19	40,583	+24.04%	[redacted]	[redacted]	1.19	40,296	+24.06%
Discount rate. Costs: 6%, QALYs: 0%	[redacted]	[redacted]	1.66	25,594	-21.77%	[redacted]	[redacted]	1.66	25,403	-21.79%
Time horizon										
Time horizon: 10 years	[redacted]	[redacted]	1.07	36,338	+11.06%	[redacted]	[redacted]	1.06	36,096	+11.13%
Time horizon: 15 years	[redacted]	[redacted]	1.25	33,974	+3.84%	[redacted]	[redacted]	1.24	33,725	+3.83%
Time horizon: 20 years	[redacted]	[redacted]	1.33	33,105	+1.18%	[redacted]	[redacted]	1.32	32,849	+1.14%
Vial sharing										
Include	[redacted]	[redacted]	1.35	32,793	+0.23%	[redacted]	[redacted]	1.35	32,553	+0.22%
Survival model: Extrapolation OS										
Cure model	[redacted]	[redacted]	1.51	32,230	-1.49%	[redacted]	[redacted]	1.50	31,995	-1.49%
Hybrid model	[redacted]	[redacted]	1.32	32,846	+0.39%	[redacted]	[redacted]	1.31	32,624	+0.44%
Joint log-normal model	[redacted]	[redacted]	1.41	32,312	-1.24%	[redacted]	[redacted]	1.42	32,283	-0.61%

	Deterministic Results					Probabilistic Results				
	Incremental Costs	**Incremental QALYs**	**Incremental LYs**	**ICER**	**ICER Change (%)**	**Incremental Costs**	**Incremental QALYs**	**Incremental LYs**	**ICER**	**ICER Change (%)**
Adverse events										
AE disutilities doubled	[redacted]	[redacted]	1.35	32,741	+0.07%	[redacted]	[redacted]	1.35	32,503	+0.07%
Utility										
Joshi 2019	[redacted]	[redacted]	1.35	30,865	-5.66%	[redacted]	[redacted]	1.35	30,629	-5.70%

Abbreviations: AE = adverse event; ICER = incremental cost-effectiveness ratio; LY = life years; OS = overall survival; QALY = quality-adjusted life year.

3.0 Subgroup Analysis

Table 3.1: Model Results with Oral AZA Discount ([redacted]), *FLT3 Subgroup*

Technology	Total costs (£)	Total LYG	Total QALYs	Incremental costs (£)	Incremental LYG	Incremental QALYs	ICER versus baseline (£/QALY)	Pairwise ICER versus oral azacitidine
Deterministic								
Watch and wait plus BSC	[redacted]	2.78	[redacted]	-	-	-	-	19,063
Oral azacitidine	[redacted]	4.95	[redacted]	[redacted]	2.18	[redacted]	19,063	-
Rydapt	[redacted]	3.67	[redacted]	[redacted]	0.89	[redacted]	269,191	Oral azacitidine is dominant
Probabilistic (5000 iterations)								
Watch and wait plus BSC	[redacted]	2.73	[redacted]	-	-	-	-	19,878
Oral azacitidine	[redacted]	4.90	[redacted]	[redacted]	2.17	[redacted]	19,878	-
Rydapt	[redacted]	3.65	[redacted]	[redacted]	0.92	[redacted]	256,807	Oral azacitidine is dominant

Abbreviations: BSC = best supportive care; ICER = incremental cost-effectiveness ratio; LYG = life year gained; QALY = quality-adjusted life year.

Figure 3.1: Scatterplot of Incremental Costs and QALYs (Oral AZA Discount = [redacted]), *FLT3* Subgroup

Abbreviations: AZA = azacitidine; FLT3 = FMS-like tyrosine kinase 3; QALY = quality-adjusted life year.

Figure 3.2: Cost-effectiveness Acceptability Curve (Oral AZA Discount = [redacted]), *FLT3* Subgroup

Abbreviations: AZA = azacitidine; CEAC = cost-effectiveness acceptability curve; FLT3 = FMS-like tyrosine kinase 3.

Figure 3.3: Tornado Plot of Deterministic Sensitivity Analysis: Impact on Incremental Costs (Oral AZA Discount = [redacted]), *FLT3* Subgroup

Abbreviations: AEs = adverse events; AZA = azacitidine; FLT3 = FMS-like tyrosine kinase 3; RDI = relative dose intensity; SCT = stem cell transplant.

Figure 3.4: Tornado Plot of Deterministic Sensitivity Analysis: Impact on Incremental QALYs (Oral AZA Discount = [redacted]), *FLT3* Subgroup

Abbreviations: AE = adverse event; AZA = azacitidine; FLT3 = FMS-like tyrosine kinase 3; QALY = quality-adjusted life year; RFS = relapse-free survival; SCT = stem cell transplant.

Figure 3.5: Tornado Plot of Deterministic Sensitivity Analysis: ICER (Oral AZA Discount = [redacted]), *FLT3* Subgroup

Abbreviations: AZA = azacitidine; FLT3 = FMS-like tyrosine kinase 3; ICER = incremental cost-effectiveness ratio; RDI = relative dose intensity; RFS = relapse-free survival; SCT = stem cell transplant.

NICE National Institute for Health and Care Excellence

Clinical expert statement and technical engagement response form

Oral azacitidine for maintenance treatment of acute myeloid leukaemia after induction therapy [ID3892]

Thank you for agreeing to comment on the evidence review group (ERG) report for this appraisal, and for providing 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. The ERG report and stakeholder responses are used by the appraisal committee to help it make decisions at the appraisal committee meeting. Usually, only unresolved or uncertain key issues will be discussed at the meeting.

Information on completing this form

In part 1 we are asking for your views on this technology. The text boxes will expand as you type.

In part 2 we are asking for your views on key issues in the ERG report that are likely to be discussed by the committee. The key issues in the ERG report reflect the areas where there is uncertainty in the evidence, and because of this the cost effectiveness of the treatment is also uncertain. The key issues are summarised in the executive summary at the beginning of the ERG report (section 1.1). You are not expected to comment on every key issue but instead comment on the issues that are in your area of expertise.

A clinical perspective could help either:

- resolve any uncertainty that has been identified OR
- provide missing or additional information that could help committee reach a collaborative decision in the face of uncertainty that cannot be resolved.

In part 3 we are asking you to provide 5 summary sentences on the main points contained in this document.

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. Please type information directly into the form.

Do not include medical information about yourself or another person that could identify you or the other person.

We are committed to meeting the requirements of copyright legislation. If you want to include **journal articles** in your submission you must have copyright clearance for these articles. We can accept journal articles in NICE Docs. For copyright reasons, we will have to return forms that have attachments without reading them. You can resubmit your form without attachments, but it must be sent by the deadline.

Combine all comments from your organisation (if applicable) into 1 response. We cannot accept more than 1 set of comments from each organisation.

Please underline all confidential information, and separately highlight information that is submitted under 'commercial in confidence' in turquoise, all information submitted under 'academic in confidence' in yellow, and all information submitted under 'depersonalised data' in pink. If confidential information is submitted, please also send a second version of your comments with that information replaced with the following text: 'academic/commercial in confidence information removed'. See the Guide to the processes of technology appraisal (sections 3.1.23 to 3.1.29) for more information.

Please note, part 1 can be completed at any time. We advise that **part 2** is completed after the expert engagement teleconference (if you are attending or have attended). At this teleconference we will discuss some of the key issues, answer any specific questions you may have about the form, and explain the type of information the committee would find useful.

Deadline for comments by **5pm** on **Thursday 7 April**. Please log in to your NICE Docs account to upload your completed form, as a Word document (not a PDF).

Thank you for your time.

We reserve the right to summarise and edit comments received during engagement, or not to publish them at all, if we consider the comments are too long, or publication would be unlawful or otherwise inappropriate.

NICE National Institute for Health and Care Excellence

Comments received during engagement are published in the interests of openness and transparency, and to promote understanding of how recommendations are developed. The comments are published as a record of the comments we received, and are not endorsed by NICE, its officers or advisory committees.

NICE National Institute for Health and Care Excellence

Part 1: Treating acute myeloid leukaemia and current treatment options

Table 1 About you, aim of treatment, place and use of technology, sources of evidence and equality

1. Your name	Professor Charles Craddock
2. Name of organisation	University of Birmingham
3. Job title or position	Professor of Haemato-oncology, University of Birmingham, Director BMT Unit, University Hospitals Birmingham NHS Trust
4. Are you (please tick all that apply)	☐ An employee or representative of a healthcare professional organisation that represents clinicians? ☒ A specialist in the treatment of people with acute myeloid leukaemia? ☐ A specialist in the clinical evidence base for acute myeloid leukaemia or technology? ☐ Other (please specify):
5. Do you wish to agree with your nominating organisation's submission? (We would encourage you to complete this form even if you agree with your nominating organisation's submission)	☒ Yes, I agree with it ☐ No, I disagree with it ☐ I agree with some of it, but disagree with some of it ☐ Other (they did not submit one, I do not know if they submitted one etc.)
6. If you wrote the organisation submission and/or do not have anything to add, tick here. (If you tick this box, the rest of this form will be deleted after submission)	☐ Yes
7. Please disclose any past or current, direct or indirect links to, or funding from, the tobacco industry.	
8. What is the main aim of treatment for acute myeloid leukaemia?	In younger patients who are treated with intensive chemotherapy it is A) Cure and failing this prolongotion of Overall Survival

(For example, to stop progression, to improve mobility, to cure the condition, or prevent progression or disability)	
9. 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)	Increase in overall survival of more than 6 months
10. In your view, is there an unmet need for patients and healthcare professionals in acute myeloid leukaemia?	Yes
11. How is acute myeloid leukaemia currently treated in the NHS? • Are any clinical guidelines used in the treatment of the condition, and if so, which? • Is the pathway of care well defined? Does it vary or are there differences of opinion between professionals across the NHS? (Please state if your experience is from outside England.) • What impact would the technology have on the current pathway of care?	There is a clear standard of care as articulated by ELN 2017 and ESMO Guidelines. This represents a Europe wide consensus CC486 represents an important new strategy to improve survival of patients with AML who achieve a complete response (CR) after intensive chemotherapy but are unable to proceed to a potentially curative allogeneic stem cell transplant (allo-SCT) either on grounds of age, lack of a donor, co-morbidities or patient preference. Consequently CC486 maintenance in patients who have achieved a CR after induction chemotherapy and cannot proceed to an allograft represents a step change in therapeutic option and also addresses the fundamentally important issue that a significant number of allo-mandatory patients with AML in CR1 cannot proceed to a transplant because of lack of donor availability. This is a majo0r issue of equity of access to curative therapy and disproprtionately affects patients from ethnic minorities..
12. Will the technology be used (or is it already used) in the same way as current care in NHS clinical practice? • How does healthcare resource use differ between the technology and current care? • In what clinical setting should the technology be used? (for example, primary or secondary care, specialist clinic)	Yes Easy to implement and does not require development of new treatment pathways or investment in infrastructure

• What investment is needed to introduce the technology? (for example, for facilities, equipment, or training)	
13. Do you expect the technology to provide clinically meaningful benefits compared with current care? • Do you expect the technology to increase length of life more than current care? • Do you expect the technology to increase health-related quality of life more than current care?	Yes Yes
14. Are there any groups of people for whom the technology would be more or less effective (or appropriate) than the general population?	See answer to Point 11. The new technology is of particular relevance to patients for particular ethnic groups who through no fault of their own do not have access to a potentially curative allogeneic transplant because they do not have a suitable donor
15. 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)	No
16. Will any rules (informal or formal) be used to start or stop treatment with the technology? Do these include any additional testing?	No
17. 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. See answers to Question 11 and 14. This technology provides an effective treatment for patients who are currently denied effective treatment because of lack of donor availability. This represents a major advance and addresses a very

• Do the instruments that measure quality of life fully capture all the benefits of the technology or have some been missed? For example, the treatment regimen may be more easily administered (such as an oral tablet or home treatment) than current standard of care	substantial inequity in access to effective treatment for patients from very sizeable ethnic groups
18. 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? • Is the technology a 'step-change' in the management of the condition? • Does the use of the technology address any particular unmet need of the patient population?	Yes. It is the first effective maintenance therapy ever to be developed in Acute Myeloid Leukaemia
19. How do any side effects or adverse effects of the technology affect the management of the condition and the patient's quality of life?	Cc486 is in my extensive experience with this agent extremely well tolerated in the great majority of patients
20. Do the clinical trials on the technology reflect current UK clinical practice? • If not, how could the results be extrapolated to the UK setting? • What, in your view, are the most important outcomes, and were they measured in the trials? • If surrogate outcome measures were used, do they adequately predict long-term clinical outcomes? • Are there any adverse effects that were not apparent in clinical trials but have come to light subsequently?	The pivotal study by Wei et al is an accurate and fair representation of UK practice The Wei et al study is a very robust RCT and importantly demonstrated an improvement in OS-the most robust endpoint in AML trials
21. Are you aware of any relevant evidence that might not be found by a systematic review of the trial evidence?	No

22. Are you aware of any new evidence for the comparator treatment(s) since the publication of NICE technology appraisal guidance TA523?	There is an update on QUAZAR 001 in this week's Blood Roboz et al 2022 139(4) 2145-2155
23. How do data on real-world experience compare with the trial data?	Accurate reflection
24. NICE considers whether there are any equalities issues at each stage of an appraisal. Are there any potential equality issues that should be taken into account when considering this condition and this treatment? Please explain if you think any groups of people with this condition are particularly disadvantaged. Equality legislation includes people of a particular age, disability, gender reassignment, marriage and civil partnership, pregnancy and maternity, race, religion or belief, sex, and sexual orientation or people with any other shared characteristics. Please state if you think this appraisal could • exclude any people for which this treatment is or will be licensed but who are protected by the equality legislation • lead to recommendations that have a different impact on people protected by the equality legislation than on the wider population • lead to recommendations that have an adverse impact on disabled people. Please consider whether these issues are different from issues with current care and why.	Please see above This is a major issue for this appraisal

More information on how NICE deals with equalities issues can be found in the NICE equality scheme. Find more general information about the Equality Act and equalities issues here.	

Part 2: Technical engagement questions for clinical experts

We welcome your comments on the key issues below, but you may want to concentrate on issues that are in your field of expertise. If you think an issue that is important to clinicians or patients has been missed in the ERG report, please also advise on this in the space provided at the end of this section.

The text boxes will expand as you type. Your responses to the following issues will be considered by the committee and may be summarised and presented in slides at the appraisal committee meeting.

For information: the professional organisation that nominated you has also been sent a technical engagement response form (a separate document) which asks for comments on each of the key issues that have been raised in the ERG report. These will also be considered by the committee.

Table 2 Issues arising from technical engagement

Key issue 1: Low dose cytarabine and subcutaneous azacitidine are part of standard therapy according to NICE guidance yet were not viewed by the company to be part of best supportive care.	Agree-these are not appropriate comparators
Key issue 2: Most patients in the QUAZAR trial received one dose or no doses of consolidation therapy, resulting in a selection bias that could have exaggerated the benefits of oral azacitidine.	The population studied was representative of clinical reality and represents an important area of unmet need

Key issue 3: Few patients in the QUAZAR trial were recruited from UK sites, and there were relevant differences between the UK and analysed populations; this limits the generalisability to UK clinical practice.	No relevant. The overall patient population studied is a fair representation of UK practice
Key issue 4: Health-related quality of life and fatigue were measured on day 1 of each 28-day cycle, when adverse events were less likely to arise.	Disagree. CC486 is very well tolerated in my extensive experience
Key issue 5: Randomisation of patients in RATIFY trial occurred at induction and not maintenance phase, potentially introducing a high risk of bias in any analysis at the maintenance phase.	Disagree. I don't believe the design of the RATIFY allows any confident comments to eb made on the benefits of midostaurin maintenance in this patient population
Key issue 6: The systematic literature review eligibility criteria would not have identified the RATIFY trial; other midostaurin studies may also have been missed.	Not an important issue
Key issue 7: Haematopoietic stem cell transplantation (HSCT) was not included as a separate health state but was implicitly included in the modelling through the survival analysis, increasing the likelihood of bias.	Not qualified to comment

Key issue 8: Some patients in QUAZAR trial received fewer cycles of consolidation therapy than is standard practice in the UK. This limits the applicability of the results to a UK setting.	Disagree. See above
Key issue 9: Patient baseline characteristics in the model are not subgroup-specific (for example in the FLT3 subgroup, consolidation subgroup or Europe subgroup); patient baseline characteristics may not align with the subgroups being analysed.	Not important
Key issue 10: Survival analyses of the FLT3 subgroup are likely to be biased due to limitations associated with the indirect comparison.	Not qualified to comment
Key issue 11: In the company's base-case analysis, only grade 3 and 4 AEs are applied with a maximum frequency of one and a duration of 1 week, which may underestimate the real impact of AEs.	Disagree. See above. CC486 is remarkably well tolerated maintenance therapy-if anything BMS over-estimate the side-effects associated with maintenace
Key issue 12: The current source of utility values may not accurately reflect the relapse utility.	Not qualified

Key issue 13: Some resource use estimates appear inconsistent with expert opinion and require further justification.	Disagree
Key issue 14: Treatment effectiveness in the FLT3 subgroup was analysed for the different comparisons separately; preventing comparison of oral azacitidine, midostaurin, watch and wait plus best supportive care.	See above
Key issue 15 (not numbered in the ERG report): The ERG's analysis did not find that oral azacitidine meets NICE's criteria to be considered a life-extending treatment at the end of life (see section 2.5 - page 29 and section 7 - page 115 of the report).	Disagree
Are there any important issues that have been missed in ERG report?	

Part 3: Key messages

In up to 5 sentences, please summarise the key messages of your statement:

First maintenance therapy to improve OS in AML. This is an important paradigmatic shift
Improves OS in adults with AML in CR after induction chemotherapy who are unable to proceed to allogeneic stem cell transplantation
CC486 maintenance therapy is very well tolerated
Addresses fundamental issues of lack of equity to transplant for patients from ethnic minorities
Decisions about whether to proceed to transplant in fit odler patients are very complex and include patient choice, social circumstances-including carer responsibilities- and CC486 therefore represents an important new treatment option to be discussed in patients who for complex reasons may ultimately not proceed to transplant

Thank you for your time.

Your privacy

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

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

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

NICE National Institute for Health and Care Excellence

Clinical expert statement and technical engagement response form

Oral azacitidine for maintenance treatment of acute myeloid leukaemia after induction therapy [ID3892]

Thank you for agreeing to comment on the evidence review group (ERG) report for this appraisal, and for providing 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. The ERG report and stakeholder responses are used by the appraisal committee to help it make decisions at the appraisal committee meeting. Usually, only unresolved or uncertain key issues will be discussed at the meeting.

Information on completing this form

In part 1 we are asking for your views on this technology. The text boxes will expand as you type.

In part 2 we are asking for your views on key issues in the ERG report that are likely to be discussed by the committee. The key issues in the ERG report reflect the areas where there is uncertainty in the evidence, and because of this the cost effectiveness of the treatment is also uncertain. The key issues are summarised in the executive summary at the beginning of the ERG report (section 1.1). You are not expected to comment on every key issue but instead comment on the issues that are in your area of expertise.

A clinical perspective could help either:

- resolve any uncertainty that has been identified OR
- provide missing or additional information that could help committee reach a collaborative decision in the face of uncertainty that cannot be resolved.

In part 3 we are asking you to provide 5 summary sentences on the main points contained in this document.

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. Please type information directly into the form.

Do not include medical information about yourself or another person that could identify you or the other person.

We are committed to meeting the requirements of copyright legislation. If you want to include **journal articles** in your submission you must have copyright clearance for these articles. We can accept journal articles in NICE Docs. For copyright reasons, we will have to return forms that have attachments without reading them. You can resubmit your form without attachments, but it must be sent by the deadline.

Combine all comments from your organisation (if applicable) into 1 response. We cannot accept more than 1 set of comments from each organisation.

Please underline all confidential information, and separately highlight information that is submitted under 'commercial in confidence' in turquoise, all information submitted under 'academic in confidence' in yellow, and all information submitted under 'depersonalised data' in pink. If confidential information is submitted, please also send a second version of your comments with that information replaced with the following text: 'academic/commercial in confidence information removed'. See the Guide to the processes of technology appraisal (sections 3.1.23 to 3.1.29) for more information.

Please note, part 1 can be completed at any time. We advise that **part 2** is completed after the expert engagement teleconference (if you are attending or have attended). At this teleconference we will discuss some of the key issues, answer any specific questions you may have about the form, and explain the type of information the committee would find useful.

Deadline for comments by **5pm** on **Thursday 7 April**. Please log in to your NICE Docs account to upload your completed form, as a Word document (not a PDF).

Thank you for your time.

We reserve the right to summarise and edit comments received during engagement, or not to publish them at all, if we consider the comments are too long, or publication would be unlawful or otherwise inappropriate.

Comments received during engagement are published in the interests of openness and transparency, and to promote understanding of how recommendations are developed. The comments are published as a record of the comments we received, and are not endorsed by NICE, its officers or advisory committees.

NICE National Institute for Health and Care Excellence

Part 1: Treating acute myeloid leukaemia and current treatment options

Table 1 About you, aim of treatment, place and use of technology, sources of evidence and equality

1. Your name	Manoj Raghavan
2. Name of organisation	University Hospitals Birmingham NHS Trust and University of Birmingham
3. Job title or position	Consultant Haematologist and Clinical Senior Lecturer
4. Are you (please tick all that apply)	☐ An employee or representative of a healthcare professional organisation that represents clinicians? ☒ A specialist in the treatment of people with acute myeloid leukaemia? ☒ A specialist in the clinical evidence base for acute myeloid leukaemia or technology? ☐ Other (please specify):
5. Do you wish to agree with your nominating organisation's submission? (We would encourage you to complete this form even if you agree with your nominating organisation's submission)	☒ Yes, I agree with it ☐ No, I disagree with it ☐ I agree with some of it, but disagree with some of it ☐ Other (they did not submit one, I do not know if they submitted one etc.)
6. If you wrote the organisation submission and/or do not have anything to add, tick here. (If you tick this box, the rest of this form will be deleted after submission)	☐ Yes
7. Please disclose any past or current, direct or indirect	None

links to, or funding from, the tobacco industry.	
8. What is the main aim of treatment for acute myeloid leukaemia? (For example, to stop progression, to improve mobility, to cure the condition, or prevent progression or disability)	The main aim is to improve the overall survival of patients with acute myeloid leukaemia. For patients who are fit enough for intensive, potentially curative chemotherapy, this will involve preventing relapse after complete remission is obtained. For patients who are not fit enough for intensive therapy but can tolerate a non-intensive form of chemotherapy that is unlikely to be curative, the aim is still to improve overall survival but also to give an improved quality of life for that duration.
9. 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)	Treatment response to induction chemotherapy, whether intensive or non-intensive implies a complete remission (CR), defined as a reduction to less than 5% blast cells (leukaemia cells) and a return to normal blood counts. The aim is then to maintain the CR with further chemotherapy or allogeneic transplant. Sometimes the blood counts do not return to normal termed CR with incomplete blood counts (CRi), but the leukaemia is in remission (less than 5%).
10. In your view, is there an unmet need for patients and healthcare professionals in acute myeloid leukaemia?	The risk of relapse for patients is substantial. For the small number of patients with good risk genetic abnormalities in their leukaemia, the risk of relapse is less than 30% at 2 years. For those with intermediate and high risk genetics the risk is much higher. Intermediate risk patients have between 40 and 70% and high risk it is in excess of 70% and usually inevitable. The number of patients in the higher risk groups increases with age. Allogeneic stem cell transplant is the most effective consolidation for patients with higher risk disease. It is offered to patients if fit enough (generally but not always under the age of 70) to reduce the risk of relapse. However, there needs to be a donor available. Full siblings of donors have a one in four chance of being a fully matched donor, as long as they are fit and willing to do so. For Caucasians there is approximately 80% chance of finding a volunteer unrelated donor (VUD), but this is far low lower for other ethnic groups who are under represented on the global stem cell registries. There are higher risk alternative donors from cord stem cell banks or using haploid matched related donors (a partially matched first degree relation), but many people will still not have a donor option.

	In addition, many of patients above the age of 60 cannot undergo such an intensive procedure due to their other medical conditions or fitness. The median age of the incidence of AML is between 70-75 years. The options for treatment at relapse are limited in the older age group; the vast majority will only have palliative chemotherapy or conservative management with transfusions and antibiotic treatment alone. The risk benefit of intensive chemotherapy is often unknown at the start of treatment. For many it is best to give the benefit of the doubt and go ahead. Once the patent is in remission they must then decide with guidance of the treating physicians whether they wish to go ahead with an allogeneic stem cell transplant if that is recommended. However these are balanced decisions; there is a considerable morbidity and mortality from transplant to be weighed against the risk of relapse with chemotherapy alone. The use of oral azacitidine opens a further option to reduce the risk of relapse that has considerably less risk of morbidity compared to transplant. Individuals will weigh up these risks and come to their own conclusions – there is no one right answer. In summary, there is an unmet clinical need for patients to reduce the risk of relapse for unfit patients in the older, frailer age group and for ethnic minority patients who do not have donor options.
11. How is acute myeloid leukaemia currently treated in the NHS? • Are any clinical guidelines used in the treatment of the condition, and if so, which? • Is the pathway of care well defined? Does it vary or are there differences of opinion between professionals across the NHS? (Please state if your experience is from outside England.)	The treatment of acute myeloid leukaemia can be either intensive or non-intensive depending on the fitness of the patient. NICE have produced a pathway outlining the treatment options for patients with AML, shortly to be withdrawn (Myeloid leukaemia - NICE Pathways). Regional guidelines are published (Pan London: Haemato-oncology guidelines (rmpartners.nhs.uk), West Midlands: Final Guidance for Acute Myeloid Leukaemia Treatment in Adults in the West Midlands v18 clean.pdf (wmcanceralliance.nhs.uk)) The pathway is well defined. Consensus in treatment has led from the wide use of protocols from the National Cancer Research Institute (NCRI) AML trials that have been widely followed around the UK. Intensive treatment is aimed to be curative, with induction chemotherapy followed by consolidation with chemotherapy and/or an allogeneic stem cell transplant depending on the risk stratification of the patient. The chemotherapy requires a few weeks' hospital admission with life threatening adverse effects. There are different indications for the variations of induction chemotherapy, depending on the risk stratification of the disease. For good risk patients who enter CR after induction chemotherapy, patients a further 3 courses of chemotherapy. For patients with a FLT3 mutation, they will receive induction chemotherapy with a targeted drug, midostaurin. They will receive consolidation chemotherapy with midostaurin and if deemed high risk

• What impact would the technology have on the current pathway of care?	they will proceed to an allogeneic transplant. If not, after completion of the 4th course of chemotherapy they will have monthly midostaurin maintenance treatment for 12 months. Non-intensive chemotherapy protocols are palliative aiming to increase the overall survival of the patient and improve their quality of life. The two options are monthly azacitidine (Aza) subcutaneous injections, either monotherapy for patients with leukaemia blasts cells less than 30%, or monthly cycles of venetoclax with azacitidine (VenAza). If patients respond (either CR or CRi for VenAza, or at least stable disease for Aza monotherapy) then treatment is continued monthly until disease progression. The technology will not have an impact on non-intensive pathway, but for patients who are not able to have transplant consolidation and do not have a FLT3 mutation; it will lead to them having monthly azacitidine oral treatment until disease progression, or intolerance. In conventional care, patients would normally attend clinic monthly for the first 12 months to monitor for relapse, with their visits then becoming less frequent over time. However over the period of the pandemic many patients are having blood tests and telephone reviews so the time spent attending hospital for follow up is less. For many oral chemotherapies, if patients are tolerating the medication, we are also arranging blood tests and performing telephone reviews, sending out medication by courier or post. While there is a need for review of the patients prior to each course of treatment, it is now possible to do this in a less onerous way for the patients.
12. Will the technology be used (or is it already used) in the same way as current care in NHS clinical practice? • How does healthcare resource use differ between the technology and current care? • In what clinical setting should the technology be used? (for example, primary or secondary care, specialist clinic)	As outlined above, other than for midostaurin for lower risk FLT3 mutated AML, maintenance chemotherapy is not routinely prescribed for patients with AML. Those with FLT3 mutations who receive midostaurin (a tyrosine kinase inhibitor) for up to 12 months following completion of consolidation chemotherapy if they do not proceed to allogeneic stem cell transplant. These agents are always prescribed in secondary or tertiary care by a specialist haematologist familiar with the use of chemotherapy. Haematology clinics are generally equipped with facilities for a point of care blood count. The technologies for phone clinics are broadly available across the NHS. There will be an increase in the number of follow up AML patients being prescribed and monitored for potential adverse effects. The frequency of visits will not change with time although it may be possible for other healthcare staff to be involved in routine prescription of maintenance chemotherapy e.g. specialist oncology pharmacists.

• What investment is needed to introduce the technology? (for example, for facilities, equipment, or training)	
13. Do you expect the technology to provide clinically meaningful benefits compared with current care? • Do you expect the technology to increase length of life more than current care? • Do you expect the technology to increase health-related quality of life more than current care?	The Quazar trial would suggest that patients who are unable or unwilling to have a transplant would have a longer overall survival compared to those simply on a watch and wait pathway. This is as long as patients are able to tolerate the treatments. My experience of subcutaneous azacitidine is that most are able to do so, but there are some adverse effects e.g. tiredness, constipation. From the Quazar trial, compared with placebo, there were increased numbers of patients with gastro-intestinal toxicity (nausea, vomiting, diarrhoea and constipation) and fatigue (https://doi.org/10.1186/s13045-021-01142-x). Although most were grade 1 or 2, given the drug needs to be taken chronically this could affect the patient's quality of life. The paper also mentions neutropenia and thrombocytopenia but in general haematologists can manage these problems. The fatigue and quality of life scores were not significantly different between patients on treatment and placebo. (https://www.nejm.org/doi/suppl/10.1056/NEJMoa2004444/suppl_file/nejmoa2004444_appendix.pdf) There is no standard treatment at relapse and most patients would deteriorate rapidly requiring transfusions and having life threatening infections.
14. Are there any groups of people for whom the technology would be more or less effective (or appropriate) than the general population?	As outlined above, patients for whom it is not possible to find a suitable stem cell donor will benefit from treatment with oral azacitidine. There are reduced numbers of donors of Asian and African origin on the international stem cell donor panels. In the absence of a related donor, it can be very difficult to find a suitable donor for these patients. Oral azacitidine offers an opportunity for them to improve their survival.
15. Will the technology be easier or more difficult to use for patients or healthcare professionals than current care? Are	As outlined above, haematology clinics are well versed in the use of oral chemotherapy agents in the outpatient settings and would be staffed to implement this. There may be an increased resource in terms of the number of patients returning to clinic 4 weekly having chemotherapy, but this may be ameliorated with telephone clinics and the use of specialist pharmacists.

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)	
16. Will any rules (informal or formal) be used to start or stop treatment with the technology? Do these include any additional testing?	Treatment will be stopped if the leukaemia recurs. There may be occasions where patients cannot tolerate the treatment. Neither of these situations would require additional testing.
17. 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? • Do the instruments that measure quality of life fully capture all the benefits of the technology or have some been missed? For example, the treatment regimen may be more easily administered	No, I would expect the increase in overall survival should increase the QALY calculation.

(such as an oral tablet or home treatment) than current standard of care	
18. 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? • Is the technology a 'step-change' in the management of the condition? • Does the use of the technology address any particular unmet need of the patient population?	Outside of FLT3 mutated AML, no post-consolidation maintenance chemotherapy regimens have shown an improvement in overall survival, although some have shown an improvement in disease free survival or relapse free survival. The QUAZAR trial is the first to show an overall survival advantage for maintenance therapy. Its use of oral azacitidine, unlike parenteral azacitidine; delivering this form of chemotherapy is practical because it is an oral tablet and seems reasonably well tolerated, giving a reasonable quality of life. The oral form has these advantages and is likely more effective by delivering a higher pharmacokinetic dose of the drug. It gives an alternative pathway and improved survival for patients without a transplant option, either because • they are from an ethnic group for which there are few donors on the international registries • they are not fit enough to go ahead with an allogeneic stem cell transplant • The patient does not wish to go ahead with an allogeneic stem cell , because of the risks of morbidity or mortality (or for any other reason)
19. How do any side effects or adverse effects of the technology affect the management of the condition and the patient's quality of life?	Compared with placebo, there were increased numbers of patients with gastro-intestinal toxicity (nausea, vomiting, diarrhoea and constipation) and fatigue (https://doi.org/10.1186/s13045-021-01142-x). Although most are grade 1 or 2, given the drug needs to be taken chronically this could affect the patient's quality of life. The paper also mentions neutropenia and thrombocytopenia but in general haematologists can manage these problems reasonably well. The fatigue and quality of life scores were not significantly different between patients on treatment and placebo (https://www.nejm.org/doi/suppl/10.1056/NEJMoa2004444/suppl_file/nejmoa2004444_appendix.pdf
20. Do the clinical trials on the technology reflect current UK clinical practice?	Yes they do. The trial reflects how patients are treated in the UK using intensive chemotherapy can therefore be applied in the maintenance setting in the same way as the trial.

• If not, how could the results be extrapolated to the UK setting? • What, in your view, are the most important outcomes, and were they measured in the trials? • If surrogate outcome measures were used, do they adequately predict long-term clinical outcomes? • Are there any adverse effects that were not apparent in clinical trials but have come to light subsequently?	The most important outcome is that maintenance chemotherapy extended **overall survival** significantly without adversely affecting quality of life for pateitns who cannot or do not wish to have an allogeneic stem cell transplant.
21. Are you aware of any relevant evidence that might not be found by a systematic review of the trial evidence?	No I am not. There is a comprehensive review of maintenance therapy in AML across all modalities here: Marcos de Lima, Gail J. Roboz, Uwe Platzbecker, Charles Craddock, Gert Ossenkoppele, AML and the art of remission maintenance, Blood Reviews, Volume 49, 2021,100829, ISSN 0268-960X, https://doi.org/10.1016/j.blre.2021.100829 .
22. Are you aware of any new evidence for the comparator treatment(s) since the publication of NICE technology appraisal guidance TA523?	No I am not. TA523 refers specifically to patients with FLT3 mutated AML. Approximately one third of cases of de-novo AML will have this mutation and be eligible for midostaurin maintenance if they do not proceed to transplant. The majority of patients with AML would be unsuitable for this maintenance treatment.

23. How do data on real-world experience compare with the trial data?	I do not know of any published real world data, especially as this is a new treatment.
24. NICE considers whether there are any equalities issues at each stage of an appraisal. Are there any potential equality issues that should be taken into account when considering this condition and this treatment? Please explain if you think any groups of people with this condition are particularly disadvantaged. Equality legislation includes people of a particular age, disability, gender reassignment, marriage and civil partnership, pregnancy and maternity, race, religion or belief, sex, and sexual orientation or people with any other shared characteristics. Please state if you think this appraisal could • exclude any people for which this treatment is or will be licensed but who	Patients from ethnic minorities have less access to allogeneic stem cell transplant as there are fewer donors from these ethinic groups on the international registries. The require donors with matched HLA types ebven if an alternative cord stem cell donor is thought appropriate. This treatment allows acces to treatment that can prolong their survival in the absence of donor availability. There is a need to make sure it is available to all those who not eligible for transplant including ethnic minorities.

are protected by the equality legislation • lead to recommendations that have a different impact on people protected by the equality legislation than on the wider population • lead to recommendations that have an adverse impact on disabled people. Please consider whether these issues are different from issues with current care and why. More information on how NICE deals with equalities issues can be found in the NICE equality scheme. Find more general information about the Equality Act and equalities issues here.	

Part 2: Technical engagement questions for clinical experts

We welcome your comments on the key issues below, but you may want to concentrate on issues that are in your field of expertise. If you think an issue that is important to clinicians or patients has been missed in the ERG report, please also advise on this in the space provided at the end of this section.

The text boxes will expand as you type. Your responses to the following issues will be considered by the committee and may be summarised and presented in slides at the appraisal committee meeting.

For information: the professional organisation that nominated you has also been sent a technical engagement response form (a separate document) which asks for comments on each of the key issues that have been raised in the ERG report. These will also be considered by the committee.

Table 2 Issues arising from technical engagement

Key issue 1: Low dose cytarabine and subcutaneous azacitidine are part of standard therapy according to NICE guidance yet were not viewed by the company to be part of best supportive care.	In the UK, very little use is now made of subcutaneous cytarbine as as its efficacy is poor (15% CR rate in newly dignosed patients on UK AML11 trial). Subcutaneous azacitidine is available only to patients with less than 30% blasts. Both are used for patients who are unsuitable for intensive therapy, so would not be appropriate for maintenance treatment as they follow a different treatment pathway.
Key issue 2: Most patients in the QUAZAR trial received one dose or no doses of consolidation therapy, resulting in a selection bias that could have exaggerated the benefits of oral azacitidine.	Further doses of consolidation therapy could have improved the outcome of patients. However the data from the forest plot (figure 3 from the NEJM paper) suggest that overall survival is improved regardless of receiving consolidation or not (only 3 patients received 3 courses of consolidation, so not significant). The recent paper on measurable residual disease is consistent with this showing higher MRD was not a factor in shortening survival in the study. Therefore treatment with fewer consolidationis unlikely to have exaggerated the overall survival outcomes (Roboz GJ, Ravandi F, Wei AH, Dombret H, Thol F, Voso MT,

	Schuh AC, Porkka K, La Torre I, Skikne B, Zhong J, Beach CL, Risueño A, Menezes DL, Ossenkoppele G, Döhner H. Oral azacitidine prolongs survival of patients with AML in remission independently of measurable residual disease status. Blood. 2022 Apr 7;139(14):2145-2155. doi: 10.1182/blood.2021013404. PMID: 34995344.).
Key issue 3: Few patients in the QUAZAR trial were recruited from UK sites, and there were relevant differences between the UK and analysed populations; this limits the generalisability to UK clinical practice.	The treatment algorithm follows the pathway of patients treated in the UK. In general patients would receive more courses of consolidation in the UK as that is standard practice (between 2 or 3 courses). It is directly applicable to standard UK practice.
Key issue 4: Health-related quality of life and fatigue were measured on day 1 of each 28-day cycle, when adverse events were less likely to arise.	It is appropriate to measure the quality of life data after completion of each course of treatment as it is the overall impact on the patient's quality of life across the month that is most important. Patients will have days when they feel better or worse but a broader measure acroos each course is appropriate.
Key issue 5: Randomisation of patients in RATIFY trial occurred at induction and not maintenance phase, potentially introducing a high risk of bias in any analysis at the maintenance phase.	From an intention to treat perspective, randomising at induction may be appropriate. With induction treatment there is a significant mortality risk and so patients may drop out. I am not sure this would bias the treatment in one direction. If patients entered maintenance early because they could not tolerate further intensive therapy this may cause a bias. For the same reason they would not be able to proceed to transplant – in reality these would be the patients deemed appropriate for maintenance treatment.
Key issue 6: The systematic literature review eligibility criteria would not have identified the RATIFY trial; other midostaurin studies may also have been missed.	I do not know in which way the studies would not have met the eligibility criteria.

Key issue 7: Haematopoietic stem cell transplantation (HSCT) was not included as a separate health state but was implicitly included in the modelling through the survival analysis, increasing the likelihood of bias.	I am uncertain as to how this biases the model. Six patients on the oral azacitidine arm subsequently underwent an allogenic stem cell transplant, so discontinued. Thirty-two discontinued and underwent transplantation on the placebo arm. I cannot see the data, but the presumption would be that these patients relapsed and underwent transplant in second remission (but it could have a choice come off the trial while if first complete remission).
Key issue 8: Some patients in QUAZAR trial received fewer cycles of consolidation therapy than is standard practice in the UK. This limits the applicability of the results to a UK setting.	If the data suggests that patients can have an improved overall survival with fewer courses of consolidation then this would change the management of patients in the UK. From a patient perspective this would reduce the amount of time spent in hospital and reduce the risk of life threatening complications. From a financial perspective this would reduce the costs of inpatient stay and need for high dose antibiotic and anti-fungal treatment, transfusions of blood and platelets and potential intensive care stay.
Key issue 9: Patient baseline characteristics in the model are not subgroup-specific (for example in the FLT3 subgroup, consolidation subgroup or Europe subgroup); patient baseline characteristics may not align with the subgroups being analysed.	
Key issue 10: Survival analyses of the FLT3 subgroup are likely to be biased due to limitations associated with the indirect comparison.	From a clinical perspective, if a patient is started on midostaurin because they have a FLT3 mutation, they are unlikely to switch to azacitidine. The pathway for treatment with midostaurin includes maintenance if the patient does not proceed to transplant and that should be route taken. The vast majority of patients with a FLT3 mutation should be started on midostaurin with intensive chemotherapy at induction. The only circumstance that may not occur would be intolerance of midostaurin. Therefore the comparison is not of clinical relevance (applies to issue 9 also).

Key issue 11: In the company's base-case analysis, only grade 3 and 4 AEs are applied with a maximum frequency of one and a duration of 1 week, which may underestimate the real impact of AEs.	It may have a bigger implication for the quality of life scores if patients are significantly debilitated for a week or two. Analysis of QOL in comparison to grade3/4 AEs may be helpful?
Key issue 12: The current source of utility values may not accurately reflect the relapse utility.	Cannot comment.
Key issue 13: Some resource use estimates appear inconsistent with expert opinion and require further justification.	Would need to know which aspects are inconsistent.
Key issue 14: Treatment effectiveness in the FLT3 subgroup was analysed for the different comparisons separately; preventing comparison of oral azacitidine, midostaurin, watch and wait plus best supportive care.	See key issue 10.
Key issue 15 (not numbered in the ERG report): The ERG's analysis did not find that oral azacitidine meets NICE's criteria to be considered a life-extending treatment at the end of life (see section 2.5 - page 29 and section 7 - page 115 of the report).	That may be reasonable. Patients will be in remission and well. They are not at end of life.

Are there any important issues that have been missed in ERG report?	No

Part 3: Key messages

In up to 5 sentences, please summarise the key messages of your statement:

In patients who receive intensive chemotherapy for AML and do not have a FLT3 mutation:

1. Oral azacitidine meets the clinical need for patients to reduce the risk of relapse for patients not fit for allogeneic stem cell transplant.
2. It meets the need for patients who do not have a matched related or unrelated donor.
3. This is particularly so for those ethnic minority patients with few donors of similar ethnicity on the marrow registries.
4. It gives a choice to patients who are concerned about the morbidity and mortality associated with an allogeneic stem cell transplant.
5. It is a treatment that can be delivered in home setting and if tolerated may cause the least disruption to patients' lives.

Thank you for your time.

Your privacy

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

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

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

NICE National Institute for Health and Care Excellence

Patient expert statement and technical engagement response form

Oral azacitidine for maintenance treatment of acute myeloid leukaemia after induction therapy [ID3892]

Thank you for agreeing to give us your views on this treatment and its possible use in the NHS.

Your comments and feedback on the key issues below are really valued. You can provide a unique perspective on conditions and their treatment that is not typically available from other sources. The evidence review group (ERG) report and stakeholder responses are used by the appraisal committee to help it make decisions at the appraisal committee meeting. Usually, only unresolved or uncertain key issues will be discussed at the meeting.

Information on completing this form

In part 1 we are asking you about living with acute myeloid leukaemia or caring for a patient with acute myeloid leukaemia. The text boxes will expand as you type.

In part 2 we are asking for your views on key issues in the ERG report that are likely to be discussed by the committee. The key issues in the ERG report reflect the areas where there is uncertainty in the evidence, and because of this the cost effectiveness of the treatment is also uncertain. The key issues are summarised in the executive summary at the beginning of the ERG report (section 1.1).

A patient perspective could help either:

- resolve any uncertainty that has been identified OR

- provide missing or additional information that could help committee reach a collaborative decision in the face of uncertainty that cannot be resolved.

You are not expected to comment on every key issue but instead comment on the issues that are in your area of expertise. We have given guidance on the issues in which we expect this to be the case and advice on what you could consider when giving your response.

In part 3 we are asking you to provide 5 summary sentences on the main points contained in this document.

Help with completing this form

If you have any questions or need help with completing this form please email the public involvement (PIP) team at pip@nice.org.uk (please include the ID number of your appraisal in any correspondence to the PIP team).

Please use this questionnaire with our hints and tips for patient experts. You can also refer to the Patient Organisation submission guide. **You do not have to answer every question** – they are prompts to guide you. There is also an opportunity to raise issues that are important to patients that you think have been missed and want to bring to the attention of the committee.

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. Please type information directly into the form.

We are committed to meeting the requirements of copyright legislation. If you want to include **journal articles** in your submission you must have copyright clearance for these articles. We can accept journal articles in NICE Docs. For copyright reasons, we will have to return forms that have attachments without reading them. You can resubmit your form without attachments, but it must be sent by the deadline.

Your response should not be longer than 15 pages.

Please note, **part 1** can be completed at any time. We advise that **part 2** is completed after the expert engagement teleconference (if you are attending or have attended). At this teleconference we will discuss some of the key issues, answer any specific questions you may have about the form, and explain the type of information the committee would find useful.

Deadline for comments by **5pm** on **Thursday 7 April**. Please log in to your NICE Docs account to upload your completed form, as a Word document (not a PDF).

Thank you for your time.

We reserve the right to summarise and edit comments received during engagement, or not to publish them at all, if we consider the comments are too long, or publication would be unlawful or otherwise inappropriate.

Comments received during engagement are published in the interests of openness and transparency, and to promote understanding of how recommendations are developed. The comments are published as a record of the comments we received, and are not endorsed by NICE, its officers or advisory committees.

Part 1: Living with this condition or caring for a patient with acute myeloid leukaemia

Table 1 About you, acute myeloid leukaemia, current treatments and equality

1. Your name	Martin Burr
2. Are you (please tick all that apply)	x A patient with acute myeloid leukaemia? x A patient with experience of the treatment being evaluated? ☐ A carer of a patient with acute myeloid leukaemia? ☐ A patient organisation employee or volunteer? ☐ Other (please specify):
3. Name of your nominating organisation	Leukaemia Care
4. Has your nominating organisation provided a submission? (please tick all options that apply)	☐ No (please review all the questions and provide answers when possible) x Yes, my nominating organisation has provided a submission ☐ I agree with it and **do not wish to** complete a patient expert statement x Yes, I authored / was a contributor to my nominating organisations submission ☐ I agree with it and **do not wish to** complete this statement x I agree with it and **will be** completing
5. How did you gather the information included in your statement? (please tick all that apply)	x I am drawing from personal experience ☐ I have other relevant knowledge or experience (for example, I am drawing on others' experiences). Please specify what other experience: x I have completed part 2 of the statement **after attending** the expert engagement teleconference

	☐ I have completed part 2 of the statement **but was not able to attend** the expert engagement teleconference ☐ I have not completed part 2 of the statement
6. What is your experience of living with acute myeloid leukaemia? **If you are a carer (for someone with acute myeloid leukaemia) please share your experience of caring for them**	It was a frightening and worrying time. Apart from the immediate concerns of feeling extremely weak and unwell, there were the major concerns raised by a bleak prognosis and worries about the future, not only for myself but for my wife and children. I had been living with the condition polycythemia vera for twelve years which had been controlled through regular treatment. I was taken ill with stomach pain and admitted to hospital. My blood tests were abnormal and my haematology consultant decided to refer me to a colleague who specialised in myloproliferative neoplasms. The weeks spent waiting for an appointment were difficult as I was feeling very unwell but had no idea what might be the cause. I had not been warned that it was possibly leukaemia. When I eventually saw the consultant and more tests were carried out, I was given the diagnosis of AML. The consultant said I was too ill to withstand a stem cell transplant and that alternative treatment could give me up to eight months of life. The impact on me and my wife was enormous. To be told I had what I considered a terminal condition and probably would not see the year out sent both of us into a state of shock. We then had to go home and work out how to tell our daughters – one studying for A Levels and one at university. It was difficult as we didn't want to upset them too much but didn't want them to think I would soon be well again. There followed a period of hospital appointments and admissions to improve my health to the point where I was told I would be a candidate for a stem cell transplant but that I would only have a five per cent chance of surviving. Although slim, this was still a chance for survival and I tried to focus on it. During these months our lives were disrupted as I was unable to work or pursue my regular leisure activities and it was difficult to maintain any sort of social life with such a bleak future ahead. I had no energy, felt ill, tired easily and was living a restricted

	life. At the same time I was trying to get all my financial affairs in order to ensure my family was secure. News that my one brother was a perfect match for a stem cell donor gave us cause for optimism. Then the rigorous process of preparing me for transplant began. This involved numerous hospital appointments and long periods as an in patient on an isolation ward. I was subjected to increasing intensive cycles of chemotherapy to destroy my immune system. The side effects were numerous and it was difficult for my family when they came to visit to see me looking so unwell, completely hairless and attached to numerous machines. My wife and the daughter still living at home travelled up every evening after work/school, a journey involving two trains and taking at least an hour each way. This was a particularly stressful time for my wife as she was working, worrying about me and trying to maintain a normal family life for our daughter while spending three hours a day on hospital visits. Our whole summer was focused on the coming stem cell transplant. We tried to ignore the small chance of success and take a positive attitude. The transplant itself came almost as an anticlimax after all those months of preparation. It was in the form of an intravenous infusion, just like the blood and plasma I had been receiving regularly. But we felt optimistic, and although I spent some weeks in hospital with an infection, things looked brighter. So when I got the news a few months later that the transplant was failing, it was probably the most devastating moment in the whole process. This news came just before Christmas. My wife and I tried to give our daughters a normal family Christmas, thinking that it would be my last one and that I would only be around for a few months. But we did not tell our family or friends. Then, in the new year my consultant was able to turn round the failing transplant and kick start it into working, despite his warnings that I should not get too optimistic. From that point, things improved, though there was still a long way to go before I could live anything like the life I had enjoyed before. I was weak and at risk of infection. So I was unable to go to places like the theatre, concert halls or galleries. I was told not to eat out in restaurants. Even preparing food at home was fraught with difficulties. For example, I was told not to use ground pepper as the peppercorns had been dried under unknown conditions so might put me at risk of infection. Numerous

	hospital visits combined with my weakened physical condition to make a return to work difficult. The whole period, lasting for considerably longer than a year, was an emotional rollercoaster for me and my family. There were a few highs and plenty of lows, periods of optimism and longer periods of despair. I think the whole period left us all traumatised and this emotional impact lasted a lot longer than the physical issues.
7a. What do you think of the current treatments and care available for acute myeloid leukaemia on the NHS? **7b. How do your views on these current treatments compare to those of other people that you may be aware of?**	a) I cannot fault the care that was made available to me. Once I was offered a stem cell transplant, every effort was made to ensure I was strong enough to withstand this before the conditioning was started to prepare me for it. Then, when the transplant was found not to be working after three months, the consultant to whom I was assigned did everything he could to turn it around despite having been told, as I later found out, that there was very little hope of success. It was unknown territory for him, but he tried a cocktail of oral drugs and subcutaneous azacitidine which worked to the point where within a few months I was in complete remission. He then maintained me on gradually reducing cycles of azacitidine for the next three years to prevent relapse. I opted for a stem cell transplant because although my chances of surviving were small, there was no alternative. Of course I am glad I did so, in retrospect. But had there been an option for a treatment like oral azacitidine to give me quality of life without the high-risk transplant I might have been tempted to go down this route.. If the need for a transplant can be avoided by adding oral azacitidine to the end of existing treatment, then this would be financially beneficial for the NHS. b)During my treatment, the only other AML patients I met were either pre or post stem cell transplant. I cannot recollect any of them having anything other than a positive view of the treatment they were given. But I presume that, like me, they had not been offered an alternative. It was a similar situation during my 34 cycles of subcutaneous azacitidine. I heard no mention of an oral version being available but from discussions with other patients I cannot think of any who would not have preferred the oral version.

8. If there are disadvantages for patients of current NHS treatments for acute myeloid leukaemia (for example, how oral azacitidine is given or taken, side effects of treatment, and any others) please describe these	Having a stem cell transplant had a massive impact on my life for a considerable period of time. For eight months leading up to the transplant, I spent most of my time either as an in-patient or as an out-patient attending various clinics while the haematology team brought me to the point where I was well enough to withstand a transplant and then conditioned me ready for it. My regular life was virtually put on hold as I felt very weak and ill and then spent the later stages in quarantine as my immune system was gradually reduced to zero. It took me a long time to recover after the transplant, particularly as it initially failed. I had to isolate due to my compromised immune system, there were many hospital appointments and an emergency admission with sepsis which I had been told to expect. Although I was not in pain I remained a semi-invalid for many months. Initially I was on seven days of treatment once a month. Owing to my compromised immune system I was unable to use public or hospital transport. I was also too unwell to drive. So I had to travel to and from the hospital by minicab, which cost me in excess of £400 during each monthly cycle. Having the azacitidine subcutaneously meant I spent a lot of time at the hospital for one week every month. On the first day of the cycle, I had to leave home around 6am to get to the hospital for an early blood test. This was to give time for the results to be assessed by the haematology team and then allow sufficient time for the pharmacy to make up the prescription. It would then be ready for the chemotherapy day team to administer it around 5pm at the earliest. After that I would have to order a minicab and would be lucky to get home by 7.30pm. That is a long day for anyone to be waiting round in a hospital, let alone someone still feeling weak and tired after the transplant. On days two three and five of the cycle I would be away from home for seven hours. Day four would require another blood test resulting in another long day. Then at the weekends, the chemo day unit was closed so I had to have the azacitidine administered in one of the haematology wards, waiting until the busy nursing team had time to do it. As my immune system recovered, I was eventually allowed to travel to the hospital by public transport but the monthly cycles still meant I was spending a considerable time at the hospital for one week in every four. This was on top of the other regular

	hospital follow-up appointments. I suffered from graft versus host disease following the transplant, leading to numerous appointments in the respiratory, dermatology and ophthalmology departments. It was decided I needed extracorporeal photopheresis (ECP) to combat my GvHD. This involved two afternoons of treatment every month. These numerous hospital appointments, which all stemmed from the transplant, had a major impact on my ability to return to work. I was fortunate enough to have an understanding employer who allowed me to work from home when I felt able and in between hospital appointments. But working in a rapidly-changing environment I often had difficulty fitting in and catching up. It also caused issues for members of the team reporting to me as I was not always available for them.
9a. If there are advantages of oral azacitidine over current treatments on the NHS please describe these. For example, the effect on your quality of life, your ability to continue work, education, self-care, and care for others? **9b. If you have stated more than one advantage, which one(s) do you consider to be the most important, and why?** **9c. Does oral azacitidine help to overcome or address any of the listed disadvantages of current treatment that you have described in question 8? If so, please describe these**	a)In my case, being prescribed oral rather than intravenous azacitidine would have had a massive impact on the quality of my life in terms of time spent in hospital visits, my availability to work and on my personal finances. Treatment with oral azacitidine as an alternative to a stem cell transplant would have greatly reduced the number of hospital appointments, time spent in hospital and time spent dealing with the side effects. It would have had a much reduced impact on the quality of my life in terms of pain, ability to work effectively and changes to my lifestyle. The most significant physical side effect of subcutaneous azacitidine was an extremely sore abdomen. Having two injections in the abdomen of the viscous liquid daily for seven days meant the area was almost permanently sore and inflamed. This would have been avoided with oral azacitidine. b) Although the financial impact and impact on my ability to work effectively were significant, it would have been in the quality of life where the advantage of oral over subcutaneous azacitidine would have been greatest. It would have meant not putting my life on virtual hold for a whole week every month and being unable to make any arrangements or commitments for 25 per cent of my time. At the time it seemed to me I spent almost my whole time either at hospital or travelling to or from it.

	c)If I had been offered a choice of maintenance treatment rather than a transplant, it would have significantly reduced time spent at hospital, both as an in and out-patient because there would be no need for the lengthy and painful pre-conditioning leading up to the transplant. It would also have avoided all the complications due to a compromised immune system and treatments for the GvHD effects on my eyes, liver, skin and lungs. If I had been given oral azacitidine following my transplant rather than the subcutaneous version, it would have meant far fewer hospital visits and avoided having to set aside one week in four for the treatment. This would have reduced money spent on minicabs, enabled me to work more regularly and improved the quality of my life by allowing me to spend more time with friends and family and pursuing my leisure interests.
10. If there are disadvantages of oral azacitidine over current treatments on the NHS please describe these. For example, are there any risks with oral azacitidine? If you are concerned about any potential side effects you have heard about, please describe them and explain why	I can think of no disadvantages of oral azacitidine compared with the subcutaneous version. At the time I was receiving treatment I was unaware there was an oral version. If I had been, I would have requested it. In terms of the disadvantages of oral azacitidine compared with a stem cell transplant, in my opinion the latter would be preferable as it would offer a permanent cure for patients in a position to undergo it..
11. Are there any groups of patients who might benefit more from oral azacitidine or any who may benefit less? If so, please describe them and explain why Consider, for example, if patients also have other health conditions (for example difficulties with mobility, dexterity or cognitive impairments) that affect the suitability of different treatments	Patients who cannot afford to pay for private transportation to and from hospital would benefit from oral azacitidine. If they are too immunocompromised to use hospital or public transport yet cannot afford the cost of a minicab, they would not have access to subcutaneous azacitidine treatment which seems to me an unfair financial inequality. Many people, by the time they need azacitidine, could be in a position where they are receiving reduced or even no wages. Oral azacitidine makes the treatment available to everyone on an equal basis whatever their financial position. Other groups of patients who might benefit from oral azacitidine compared with a stem cell transplant could be those with underlying health conditions which make a

	transplant unviable; people with responsibilities which would prevent them spending a considerable time in hospital then an extended recovery period before they can resume their normal life, such as single parents and carers; and patients who are offered a transplant but told there is only a slim chance of its success and do not want to take the risk.
12. Are there any potential equality issues that should be taken into account when considering acute myeloid leukaemia and oral azacitidine? Please explain if you think any groups of people with this condition are particularly disadvantaged Equality legislation includes people of a particular age, disability, gender reassignment, marriage and civil partnership, pregnancy and maternity, race, religion or belief, sex, and sexual orientation or people with any other shared characteristics More information on how NICE deals with equalities issues can be found in the NICE equality scheme Find more general information about the Equality Act and equalities issues here.	The main equality issue of oral azacitidine compared with the subcutaneous version is financial. The cost of funding their own transport to and from hospital while their immune system is compromised can make subcutaneous azacitidine inaccessible to patients on a reduced or non-existent income. Also, the time commitment in hospital attendance required during a seven day subcutaneous azacitidine cycle means people are not able to work for most of one week in every four. This can also make it inaccessible to patients who cannot afford such a significant drop in earnings. A stem cell transplant requires a massive time commitment, both leading up to and post transplant. The patient will not be able to meet their regular commitments for a long period of time. There will be many people in a wide range of circumstances or stages of their life who cannot afford to do this – single parents of young children or carers of an elderly relative, for example. Such people may have to refuse the offer of a stem cell transplant. In the name of equality, these groups should not be denied treatment and oral azacitidine would be a viable alternative to a transplant.
13. Are there any other issues that you would like the committee to consider?	

NICE National Institute for Health and Care Excellence

Part 2: Technical engagement questions for patient experts

Issues arising from technical engagement

The issues raised in the ERG report are listed in table 2. We welcome your comments on the issues, but you do not have to provide a response to every issue, such as the ones that are technical, that is, cost effectiveness-related issues. We have added a comment to the issues where we consider a patient perspective would be most relevant and valuable. If you think an issue that is important to patients has been missed in the ERG report, please let us know in the space provided at the end of this section.

For information: the patient organisation that nominated you has also been sent a technical engagement response form (a separate document) which asks for comments on each of the key issues that have been raised in the ERG report, the patient organisation responses will also be considered by the committee.

Table 2 Issues arising from ERG report

Key issue 1: Low dose cytarabine and subcutaneous azacitidine are part of standard therapy according to NICE guidance yet were not viewed by the company to be part of best supportive care. *Are low dose cytarabine and subcutaneous azacitidine part of best supportive care in the population being considered in this appraisal? Should*	I would regard subcutaneous azacitidine an inferior comparator against the oral version. It requires many more hospital visits at a time when the patient's immune system could be severely compromised. This can result in difficulty getting to the hospital and having to self fund transport costs as public and hospital transport is not an option. The numerous hospital visits limit a patient's capacity to work and reduce the quality of their life for one week in every four. Also, with a compromised immune system, even being in a hospital environment for extended periods can be a health risk. Taking oral azacitidine avoids all these issues.

they be considered as comparators to oral azacitidine? **We consider patient perspectives may particularly help to address this issue.**	
Key issue 2: Most patients in the QUAZAR trial received one dose or no doses of consolidation therapy, resulting in a selection bias that could have exaggerated the benefits of oral azacitidine. *How many cycles of consolidation therapy would be expected in this population following induction treatment?* **We consider patient perspectives may particularly help to address this issue.**	
Key issue 3: Few patients in the QUAZAR trial were recruited from UK	

sites, and there were relevant differences between the UK and analysed populations; this limits the generalisability to UK clinical practice. *Is the QUAZAR trial population generalisable to those who would receive oral azacitidine as maintenance therapy in UK clinical practice?* **We consider patient perspectives may particularly help to address this issue.**	
Key issue 4: Health-related quality of life and fatigue were measured on day 1 of each 28-day cycle. This may have missed adverse events, given that patients would have been off oral azacitidine treatment for 14 days prior to the measurement. *Do you consider that the effect of treatment-related adverse events would have been captured in the measurement of health-related quality of life or would they be underestimated*	

for oral azacitidine (compared to placebo/standard of care)? **We consider patient perspectives may particularly help to address this issue.**	
Key issue 5: Randomisation of patients in RATIFY trial occurred at induction and not maintenance phase, potentially introducing a high risk of bias in any analysis at the maintenance phase.	
Key issue 6: The systematic literature review eligibility criteria would not have identified the RATIFY trial; other midostaurin studies may also have been missed.	
Key issue 7: Haematopoietic stem cell transplantation (HSCT) was not included as a separate health state but was implicitly included in the modelling through the survival analysis, increasing the likelihood of bias.	
Key issue 8: Some patients in QUAZAR trial received fewer cycles of consolidation therapy than is standard practice in the UK. This limits the	

applicability of the results to a UK setting.	
Key issue 9: Patient baseline characteristics in the model are not subgroup-specific (for example in the FLT3 subgroup, consolidation subgroup or Europe subgroup); patient baseline characteristics may not align with the subgroups being analysed.	
Key issue 10: Survival analyses of the FLT3 subgroup are likely to be biased due to limitations associated with the indirect comparison.	
Key issue 11: In the company's base-case analysis, only grade 3 and 4 AEs are applied with a maximum frequency of one and a duration of 1 week, which may underestimate the real impact of AEs.	
Key issue 12: The current source of utility values may not accurately reflect the relapse utility.	
Key issue 13: Some resource use estimates appear inconsistent with expert opinion and require further justification.	

Key issue 14: Treatment effectiveness in the FLT3 subgroup was analysed for the different comparisons separately; preventing comparison of oral azacitidine, midostaurin, watch and wait plus best supportive care.	
Key issue 15 (not numbered in the ERG report): The ERG's analysis did not find that oral azacitidine meets NICE's criteria to be considered a life-extending treatment at the end of life (see section 2.5 - page 29 and section 7 - page 115 of the report). *Does oral azacitidine meet NICE's criteria to be considered a life-extending treatment at the end of life, that is:* • *is life expectancy without oral azacitidine for the population being considered in this appraisal normally less than 24 months?* • *does treatment with oral azacitidine extend life by at least*	

an additional 3 months compared to current treatment?	
Are there any important issues that have been missed in ERG report?	

Part 3: Key messages

In up to 5 sentences, please summarise the key messages of your statement:

- Oral azacitidine is far more beneficial for the patient that the subcutaneous version in terms of expense and time.
- Oral azacitidine removes the inequality inherent in the subcutaneous version as it is less financially burdensome to all and greatly reduces the time commitment.
- A stem cell transplant is not a viable option for certain groups of people depending on their lifestyle, commitments and health.
- Oral azacitidine can solve this issue of inequality by offering a viable alternative.
- From a patient's perspective, oral azacitidine is the preferential option compared with the subcutaneous version and as an alternative to a stem cell transplant it empowers the patient with choice.

Thank you for your time.

Your privacy

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

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

For more information about how we process your personal data please see NICE's privacy notice.

NICE National Institute for Health and Care Excellence

Technical engagement response form

Oral azacitidine for maintenance treatment of acute myeloid leukaemia after induction therapy [ID3892]

As a stakeholder you have been invited to comment on the evidence review group (ERG) report for this appraisal.

Your comments and feedback on the key issues below are really valued. The ERG report and stakeholders' responses are used by the appraisal committee to help it make decisions at the appraisal committee meeting. Usually, only unresolved or uncertain key issues will be discussed at the meeting.

Information on completing this form

We are asking for your views on key issues in the ERG report that are likely to be discussed by the committee. The key issues in the ERG report reflect the areas where there is uncertainty in the evidence, and because of this the cost effectiveness of the treatment is also uncertain. The key issues are summarised in the executive summary at the beginning of the ERG report.

You are not expected to comment on every key issue but instead comment on the issues that are in your area of expertise.

If you would like to comment on issue.

s in the ERG report that have not been identified as key issues, you can do so in the 'Additional issues' section.

If you are the company involved in this appraisal, please complete the 'Summary of changes to the company's cost-effectiveness estimates(s)' section if your response includes changes to your cost-effectiveness evidence.

Please do not embed documents (such as PDFs or tables) because this may lead to the information being mislaid or make the response unreadable. Please type information directly into the form.

Do not include medical information about yourself or another person that could identify you or the other person.

We are committed to meeting the requirements of copyright legislation. If you want to include journal articles in your submission you must have copyright clearance for these articles. We can accept journal articles in NICE Docs. For copyright reasons, we will have to return forms that have attachments without reading them. You can resubmit your form without attachments, but it must be sent by the deadline.

Combine all comments from your organisation (if applicable) into 1 response. We cannot accept more than 1 set of comments from each organisation.

Please underline all confidential information, and separately highlight information that is submitted under 'commercial in confidence' in turquoise, all information submitted under 'academic in confidence' in yellow, and all information submitted under 'depersonalised data' in pink. If confidential information is submitted, please also send a second version of your comments with that information replaced with the following text: 'academic/commercial in confidence information removed'. See the Guide to the processes of technology appraisal (sections 3.1.23 to 3.1.29) for more information.

Deadline for comments by **5pm** on **Thursday 7 April 2022**. Please log in to your NICE Docs account to upload your completed form, as a Word document (not a PDF).

Thank you for your time.

We reserve the right to summarise and edit comments received during engagement, or not to publish them at all, if we consider the comments are too long, or publication would be unlawful or otherwise inappropriate.

Comments received during engagement are published in the interests of openness and transparency, and to promote understanding of how recommendations are developed. The comments are published as a record of the comments we received, and are not endorsed by NICE, its officers or advisory committees.

NICE National Institute for Health and Care Excellence

About you

Table 1 About you

Your name	Novartis
Organisation name: stakeholder or respondent (if you are responding as an individual rather than a registered stakeholder, please leave blank)	Stakeholder
Disclosure Please disclose any past or current, direct or indirect links to, or funding from, the tobacco industry.	

Key issues for engagement

All: Please use the table below to respond to the key issues raised in the ERG report.

Table 2 Key issues

Key issue	Does this response contain new evidence, data or analyses?	Response
Key issue 1: Low dose cytarabine and subcutaneous azacitidine are part of standard therapy according to NICE guidance yet were not viewed by the company to be part of BSC.	Yes/No	Please provide your response to this key issue, including any new evidence, data or analyses
Key issue 2: Most patients in the QUAZAR trial received one dose or no doses of consolidation therapy, resulting in a selection bias that could have exaggerated the benefits of oral azacitidine.	Yes/No	Please provide your response to this key issue, including any new evidence, data or analyses
Key issue 3: Few patients in the QUAZAR trial were recruited from UK sites, and there were relevant differences between the UK and analysed populations; this limits	Yes/No	Please provide your response to this key issue, including any new evidence, data or analyses

the generalisability to UK clinical practice.		
Key issue 4: HRQoL and fatigue were measured on day 1 of each 28-day cycle, when adverse events were less likely to arise.	Yes/No	Please provide your response to this key issue, including any new evidence, data or analyses
Key issue 5: Randomisation of patients in RATIFY trial occurred at induction and not maintenance phase, potentially introducing a high risk of bias in any analysis at the maintenance phase.	Yes/No	We do not believe that midostaurin maintenance is an appropriate comparator for this appraisal as explained in the Additional issue 1. If midostaurin is considered a comparator, we further do not believe that the economic model's conclusion that maintenance oral azacitidine is more effective compared with midostaurin in FLT-3 patients is supported by the evidence. In particular, the ITC comparing oral azacitidine and midostaurin maintenance for the subgroup of FLT-3 patients is not appropriate. An anchored comparison is used to compare results from the post-hoc analysis of the RATIFY trial and a post-hoc analysis of the QUAZAR AML-001 trial using the placebo group from each study as a common comparator (to connect these two studies). Anchored comparisons are used to respect randomisation and assume that treatment effect modifiers are balanced between treatment arms within studies[1] but also in-between studies which is not the case here. As highlighted by the ERG, patients in the RATIFY trial were not randomised at the start of maintenance introducing significant bias. Perhaps more importantly, in addition to the lack of randomisation and potential imbalances in patients' characteristics (and treatment effect modifiers within the post-hoc analysis of the RATIFY trial), all patients on midostaurin maintenance in the post-hoc analysis of the RATIFY trial received prior midostaurin as part of their induction and consolidation treatment, while those in the placebo group did not receive prior midostaurin. Therefore, the two groups in the post-hoc analysis of the RATIFY trial are two separate single arm cohorts (1) those previously treated with midostaurin at induction and consolidation receiving midostaurin maintenance and (2) those not previously treated with midostaurin at induction and consolidation and receiving placebo maintenance. Treating this post-hoc analysis as a RCT is

		therefore incorrect as the population in the midostaurin and placebo group are not directly comparable. It should also be noted that the FLT-3 subgroup in the QUAZAR AML-001 trial is also a post-hoc analysis and therefore randomisation is broken introducing bias in any estimate of the treatment effect of oral azacitidine compared with placebo if patients' characteristics are not balanced (which is not reported). Furthermore none of the patients in the QUAZAR AML-001 received prior induction and consolidation midostaurin, making the population not comparable to that on midostaurin in the post-hoc analysis of the RATIFY trial. In addition to this, comparing naively the post-hoc analysis of the RATIFY trial and that of the QUAZAR AML-001 trial is problematic as there are clear imbalances in treatment effects modifiers between these two studies that are not accounted for. All patients in the post-hoc analysis of the RATIFY trial who underwent maintenance treatment received 4 cycles of consolidation, while most patients in the QUAZAR AML-001 trial received less than 2 cycles of consolidation, with no patients receiving 4 cycles of consolidation. This is problematic as in the ITT population of the QUAZAR AML-001 trial the HR for OS for the full population (including those who did and did not receive consolidation) was 0.69 (0.56 – 0.86). However, when considering only patients with 2 or more consolidation cycles, the HR (0.75; 95% CI: 0.5 – 1.11) is no longer significant and the curve for oral azacitidine and placebo join by month 36 (Figure 3.9; ERG report). The HR is also no longer significant when considering patients with 1 or more cycle of consolidation (0.75; 95%CI: 0.55 – 1.02). Therefore, comparing naively the HRs from the post-hoc analysis of the RATIFY and QUAZAR AML-001 trial is not appropriate as there are some clear treatment effect modifiers and that most patients in the QUAZAR trial did not receive consolidation, overestimating the treatment effect for oral azacitidine compared with midostaurin.
Key issue 6: The SLR eligibility criteria would not have identified the RATIFY trial; other midostaurin	Yes/No	Two additional studies examined the use of midostaurin as a maintenance treatment: • AMLSG 16-10 - Phase 2, open-label, nonrandomized study of midostaurin in combination with induction and consolidation chemotherapy and as

studies may also have been missed.		single agent maintenance therapy following consolidation with alloSCT or chemotherapy in patients (aged 18-70 years) with newly diagnosed FLT3-ITD+ AML • Radius: A Phase 2, Randomized Trial of Standard of Care with or without Midostaurin to Prevent Relapse Following Allogeneic Hematopoietic Stem Cell Transplant in Patients with FLT3-Itd-Mutated Acute Myeloid Leukemia (AML)
Key issue 7: HSCT was not included as a separate health state but was implicitly included in the modelling through the survival analysis, increasing the likelihood of bias.	Yes/No	
Key issue 8: Some patients in QUAZAR trial received fewer cycles of consolidation therapy than is standard practice in the UK. This limits the applicability of the results to a UK setting.	Yes/No	Please provide your response to this key issue, including any new evidence, data or analyses
Key issue 9: Patient baseline characteristics in the model are not subgroup-specific (for example in the FLT3 subgroup, consolidation subgroup or Europe subgroup); patient baseline characteristics may not align with the subgroups being analysed.	Yes/No	Please provide your response to this key issue, including any new evidence, data or analyses

Key issue 10: Survival analyses of the FLT3 subgroup are likely to be biased due to limitations associated with the indirect comparison.	Yes/No	We do not believe that midostaurin maintenance is an appropriate comparator for this appraisal as explained in additional issue 1. If midostaurin is considered a comparator, as highlighted in our response to Key issue 5, we further do not believe that the economic model's conclusion that maintenance oral azacitidine is more effective compared with midostaurin is supported by the evidence. The ITC presented comparing oral azacitidine and midostaurin maintenance in the subgroup of FLT-3 patients is not appropriate because (a) patients in the RATIFY trial were not randomised at the start of maintenance introducing significant bias, (b) the two arms in the post-hoc analysis of the RATIFY trial are two separate single arm cohorts (those previously treated or not with midostaurin at induction and consolidation), (c) randomisation is broken in the post-hoc analysis of the FLT-3 subgroup of the QUAZAR AML-001 trial and not patients received prior midostaurin and (d) the clear imbalances in treatment effects modifiers between the two studies that are not accounted for, notably the number of consolidation cycles. All patients in the post-hoc analysis of the RATIFY trial who underwent maintenance treatment receive 4 cycles of consolidation while most patients in the QUAZAR AML-001 trial received less than 2 cycles of consolidation. This is problematic as in the ITT population of the QUAZAR AML-001 trial the HR for OS for the full population (including who and did not receive consolidation) was 0.69 (0.56 – 0.86). However, when considering only patients with 2 or more consolidation cycles, the HR (0.75; 95% CI: 0.5 – 1.11) is no longer significant and the curve for oral azacitidine and placebo join together by month 36 (Figure 3.9; ERG report). Therefore, comparing naively, the HRs from the post-hoc analysis of the RATIFY and QUAZAR AML-001 trial is not appropriate as there are some clear treatment effect modifiers and that most patients in the QUAZAR trial did not receive consolidation, overestimating the treatment effect for oral azacitidine compared with midostaurin. Furthermore, details for the survival analyses for the FLT-3 subgroup are also limited. Larson et al (2021)[2] reported in the post-hoc analysis of the RATIFY trial that 10 patients had started maintenance therapy (7 on midostaurin and 3 in placebo) prior to receiving SCT. As SCT is the only curative option in AML

		estimate for the long-term survival for midostaurin should be consistently greater compared with that of placebo due to the higher rate of SCT in the midostaurin group compared with placebo. In contrast, in the FLT3 subgroup in the post-hoc QUAZAR AML-001 trial patients on oral azacitidine had a lower rate of SCT compared with placebo with SCT being the only curative option. In conclusion, we do not believe that the evidence supports the conclusion that oral azacitidine is more effective compared with midostaurin because of the different number of consolidation treatments received, differences between studies and that STC is the only curative option (relative to placebo, more patients in the post-hoc analysis of RATIFY received SCT compared with oral azacitidine in the post-hoc analysis of the QUAZAR AML-001).
Key issue 11: In the company's base-case analysis, only grade 3 and 4 AEs are applied with a maximum frequency of one and a duration of 1 week, which may underestimate the real impact of AEs.	Yes/No	
Key issue 12: The current source of utility values may not accurately reflect the relapse utility.	Yes/No	
Key issue 13: Some resource use estimates appear inconsistent with expert opinion and require further justification.	Yes/No	We do not believe that midostaurin maintenance is an appropriate comparator for this appraisal as explained in additional issue 1. If midostaurin is considered a comparator, we believe that the higher costs post-relapse for patients on midostaurin maintenance compared with those on oral azacitidine is not supported by the evidence. Subsequent treatments distribution in patients receiving placebo in the QUAZAR AML-001 are used for midostaurin with no justification.

		Significantly more patients with midostaurin are assumed to receive salvage chemotherapy compared with oral azacitidine (36.1% vs. 23.3%) which is associated with a very high cost. This led to an arbitrary higher cost post-relapse for midostaurin compared with oral azacitidine overestimating total costs for maintenance midostaurin. It is also unclear if the higher proportion of patients who received post-relapse treatment in the placebo arm in the QUAZAR AML-001 trial (which is used to inform the distribution for midostaurin without justification) is due to the higher number of relapses during the observed period of the trial which therefore bias the estimate for midostaurin against oral azacitidine.
Key issue 14: Treatment effectiveness in the FLT3 subgroup was analysed for the different comparisons separately; preventing comparison of oral azacitidine, midostaurin, watch and wait plus BSC.	Yes/No	Please provide your response to this key issue, including any new evidence, data or analyses
Key issue 15 (not numbered in the ERG report): The ERG's analysis did not find that oral azacitidine meets NICE's criteria to be considered a life-extending treatment at the end of life (see section 2.5 - page 29 and section 7 - page 115 of the report).	Yes/No	Please provide your response to this key issue, including any new evidence, data or analyses

Additional issues

All: Please use the table below to respond to additional issues in the ERG report that have not been identified as key issues. Please do **not** use this table to repeat issues or comments that have been raised at an earlier point in this appraisal (for example, at the clarification stage).

Table 3 Additional issues from the ERG report

Issue from the ERG report	Relevant section(s) and/or page(s)	Does this response contain new evidence, data or analyses?	Response

Additional issue 1: Midostaurin as a relevant comparator for the FLT-3 mutation subgroup	Section 2.3.2 of ERG report	No	We believe that midostaurin maintenance should not be a comparator in this appraisal. While we recognise that midostaurin is included in the comparator list in the NICE final scope, midostaurin maintenance is only permitted in people that receive induction and consolidation therapy with midostaurin as stated in the license reproduced below: "*Rydapt is indicated in combination with standard daunorubicin and cytarabine induction and high-dose cytarabine consolidation chemotherapy, and for patients in complete response followed by Rydapt single agent maintenance therapy, for adult patients with newly diagnosed acute myeloid leukaemia (AML) who are FLT3 mutation-positive (see section 4.2)*". The license states that midostaurin maintenance can only be used as a standalone treatment in patients with FLT-3 mutation that had midostaurin as induction and consolidation therapy. The license further suggests that following complete response on induction and consolidation treatment with midostaurin, single agent maintenance midostaurin should be offered. The license does not permit the use of midostaurin only as a maintenance treatment without the induction and consolidation phase with midostaurin. There is also no evidence of oral azacitidine maintenance in patients with FLT-3 mutation who

			responded to both 1-2 cycles of induction and four cycles of consolidation midostaurin.

Additional issue 2: Treatment costs	Section 4.2.9 of ERG report	No	We do not believe that midostaurin maintenance is an appropriate comparator for this appraisal as explained in additional issue 1. If midostaurin is considered a comparator, we do not believe that drug acquisition costs are calculated correctly for the FLT-3 subgroup. Treatment costs for midostaurin have been incorrectly calculated and overestimated compared with oral azacitidine. According to the license (and as given in the RATIFY trial), midostaurin maintenance is given for a maximum of 12 cycles until relapse 1) **Incorrect use of the median as a proxy for the mean time on treatment**. Midostaurin is given for a maximum of 12 cycles until relapse therefore the median is not a good proxy for the mean. This can be seen by comparing RFS from the post-hoc analysis of the RATIFY trial and the median value used. At 10-month RFS is around 70%. 2) **The time on treatment for midostaurin cannot be greater than RFS**: Due to the ITC, RFS predicted in the economic model for midostaurin do not align with that reported in the post-hoc analysis of the RATIFY trial (and therefore the treatment duration). This can be seen when comparing RFS in Figure B.3.33 from the company submission and RFS from the post-hoc analysis of the RATIFY trial. At 10 months, RFS in the post-hoc analysis of the RATIFY trial is about 70% but that predicted by the model is considerably lower as shown in Figure B.3.33. Therefore, the

			current approach to costing for midostaurin is incorrect and overestimate the number of cycles of treatment. A more realistic approach to costing is to use RFS from the economic model and cap costs after 336 days (12 cycles of 28 days) as patients are treated until relapse. It should be noted that this would also overestimate costs as discontinuation due to other reasons than relapse is not accounted for (for example adverse events), 3) **The mean RDI should be used for midostaurin**. The median RDI is used for midostaurin but the mean RDI is used for oral azacitidine. For consistency the mean RDI for midostaurin (89.8%) should be used and is reported the Australian public assessment report (page 169),[3] 4) **Fix dosing for midostaurin for the FLT-3 subgroup (not extended dosing).** It is unclear if extended dosing for oral azacitidine has been included in the economic model for the FLT-3 subgroup comparison against midostaurin. Dohner et al (2021)[4] reported that 51 patients (21%) treated with oral azacitidine in the QUAZAR AML-001 trial were assigned to ≥ 21 day/cycle schedule, with a median number of cycles of 2.0 (range 1- 45) and 43% receiving ≥ 3 cycles of 21d dosing. It is unclear from the ERG report if this has been included for the FLT3 subgroup. If not, there is therefore a mismatch between the clinical effectiveness (which include the

			effect of extended dosing) and costs of oral azacitidine. This would underestimate costs for oral azacitidine compared with treatment that are given daily like midosaturin for the FLT-3 subgroup.
Additional issue N: Insert additional issue			[INSERT / DELETE ROWS AS REQUIRED]

Summary of changes to the company's cost-effectiveness estimate(s)

Company only: If you have made changes to the base-case cost-effectiveness estimate(s) in response to technical engagement, please complete the table below to summarise these changes. Please also provide sensitivity analyses around the revised base case. If there are sensitivity analyses around the original base case which remain relevant, please re-run these around the revised base case.

Table 4 Changes to the company's cost-effectiveness estimate

Key issue(s) in the ERG report that the change relates to	Company's base case before technical engagement	Change(s) made in response to technical engagement	Impact on the company's base-case incremental cost-effectiveness ratio (ICER)
Insert key issue number and title as described in the ERG report	Briefly describe the company's original preferred assumption or analysis	Briefly describe the change(s) made in response to the ERG report	Please provide the ICER resulting from the change described (on its own), and the change from the company's original base-case ICER.
Insert key issue number and title as described in the ERG report	…	…	[INSERT / DELETE ROWS AS REQUIRED]
Company's base case following technical engagement (or revised base case)	Incremental QALYs: [QQQ]	Incremental costs: [£££]	Please provide company revised base-case ICER

Sensitivity analyses around revised base case
[PLEASE DESCRIBE HERE]

1. Phillippo DM, A.A., Dias S, Palmer S, Abrams KR, Welton NJ., *Methods for Population-Adjusted Indirect Comparisons in Health Technology Appraisal.* Med Decis Making, 2018. **38**(2): p. 200-211.
2. Larson RA, M.S., Huebner LJ, Sanford BL, Laumann K, Geyer S, Bloomfield CD, Thiede C, Prior TW, Döhner K, Marcucci G, Voso MT, Klisovic RB, Galinsky I, Wei AH, Sierra J, Sanz MA, Brandwein JM, de Witte T, Niederwieser D, Appelbaum FR, Medeiros BC, Tallman MS, Krauter J, Schlenk RF, Ganser A, Serve H, Ehninger G, Amadori S, Gathmann I, Döhner H, Stone RM., *Midostaurin reduces relapse in FLT3-mutant acute myeloid leukemia: the Alliance CALGB 10603/RATIFY trial.* Leukemia. , 2021. **35**(9): p. 2539-2551.
3. *Australian Public Assessment Report (AusPAR) for Midostaurin*, in *Australian Government - Department of Health - Therapeutic Good Administration*. 2019.
4. Dohner H, W.A., Montesinos P, Dombret H, Ravandi F, Sayar H et al., *Escalated dosing schedules of CC-486 for patients experiencing first acute myeloid leukemia (AML) relapse: Results from the phase III QUAZAR AML-001 maintenance trial.* 2020: Meeting Abstract | 2020 ASCO Annual Meeting I; Abstract 7513.

in collaboration with:

Erasmus School of Health Policy & Management

Maastricht University

Oral azacitidine for maintenance treatment of acute myeloid leukaemia after induction therapy [ID3892]

ADDENDUM: Critique of the company's response to Technical Engagement

Produced by Kleijnen Systematic Reviews (KSR) Ltd, in collaboration with Erasmus University Rotterdam (EUR) and Maastricht University Medical Centre (UMC)

Authors
Jeremy Howick, Reviews Manager, KSR Ltd, United Kingdom (UK)
Willem Witlox, Health Economist, Maastricht UMC, The Netherlands
Charlotte Ahmadu, Health Economist, KSR Ltd, UK
Sabine Grimm, Health Economist, Maastricht UMC, The Netherlands
Nigel Armstrong, Health Economics Manager, KSR Ltd, UK
Kevin McDermott, Systematic Reviewer, KSR Ltd, UK
Thomas Otten, Health Economist, Maastricht UMC, The Netherlands
Caro Noake, Information Specialist, KSR Ltd, UK
Robert Wolff, Managing Director, KSR Ltd, UK
Manuela Joore, Health Economist, Maastricht UMC
Jos Kleijnen, Founder and Owner, KSR Ltd, UK

Correspondence to Jeremy Howick, Kleijnen Systematic Reviews Ltd
Unit 6, Escrick Business Park
Riccall Road, Escrick
York, YO19 6FD
United Kingdom

Date completed 25/04/2022

Company's response to technical engagement

The purpose of this addendum is to provide a critique of the new evidence submitted by the company as part of their response to the technical engagement (TE) report.

In their response to technical engagement, the company submitted responses to the key issues raised in the Technical Report written by the National Institute for Health and Care Excellence (NICE) technical team, and some additional evidence relevant to these issues.

1. ***Key issue 1: Low dose cytarabine and subcutaneous azacitidine are part of standard therapy according to NICE guidance yet were not viewed by the company to be part of BSC.***

The ERG reiterates that subcutaneous azacitidine and cytarabine were included in the final NICE scope. It is true that only midostaurin is recommended as an option in the ESMO guidelines and only after initial treatment with midostaurin. However, the ESMO guidelines cited by the company state: "Maintenance treatment with subcutaneous azacitidine in older AML patients who obtained CR after induction and consolidation treatment improved disease free survival but not OS in a randomised study." (p.704) With respect to the real world HMRN data, it is also true that very few patients received maintenance treatment, but at least some did, and this was subcutaneous azacitidine. The ERG notes that a comprehensive audit of current practice is required. Notwithstanding the clinical expert opinion cited by the company, the ERG also notes the need for independent clinical experts. The ERG therefore maintains that subcutaneous azacitidine might still be a valid comparator for this population and this remains a key issue.

2. ***Key issue 2: Most patients in the QUAZAR trial received one dose or no doses of consolidation therapy, resulting in a selection bias that could have exaggerated the benefits of oral azacitidine.***

The ERG acknowledges some apparent contradiction between the NHS website (https://www.nhs.uk/conditions/acute-myeloid-leukaemia/treatment/), which suggests everyone gets consolidation and the HMRN that suggests [redacted] do not receive it. The ERG suspects that this is at least partly because HMRN includes patients diagnosed a long time ago (from 2004), which might include period of change in clinical practice. The more recent (2020) ESMO guidelines state: "*As soon as patients achieve CR/CRi after 1 or 2 induction cycles, they should proceed to consolidation treatment [II, B].*" The ERG therefore reiterates that consolidation is expected, and that the relevant population is the consolidation subgroup. The ERG also acknowledges the uncertainty surrounding the optimal number of rounds of consolidation therapy. Therefore, this remains a key issue.

3. ***Key issue 3: Few patients in the QUAZAR trial were recruited from UK sites, and there were relevant differences between the UK and analysed populations; this limits the generalisability to UK clinical practice.***

The company updated its base-case using the EU-subgroup, stating that this improves the generalisability to the UK setting. As the company states, the ERG did suggest that the EU-subgroup would be more in line with what is expected to be seen in UK clinical practice. However, this was relative to the ITT population and notwithstanding the ERG's preference for the consolidation subgroup. It might be informative to perform an analysis using a subgroup of the EU-subgroup who received at least one cycle of consolidation therapy: this should be accompanied with an updated NICE TSD 14 criteria assessment. This therefore remains a key issue.

4. ***Key issue 4: HRQoL and fatigue were measured on day 1 of each 28-day cycle, when adverse events were less likely to arise.***

No compelling new arguments or evidence provided. Hence the ERG perspective as described in the ERG report remains unchanged.

5. ***Key issue 5: Randomisation of patients in RATIFY trial occurred at induction and not maintenance phase, potentially introducing a high risk of bias in any analysis at the maintenance phase.***

The company agree with the limitation identified by the ERG. In turn, the ERG agrees with the company that "the analysis conducted remains the most appropriate given the data available." Nevertheless, the timing of the randomisation implies that there is additional uncertainty around the efficacy and safety estimates of oral azacitidine versus midostaurin.

6. ***Key issue 6: The SLR eligibility criteria would not have identified the RATIFY trial; other midostaurin studies may also have been missed.***

Notwithstanding the incorrect specification of azacitidine as the only intervention, the ERG acknowledges the company's clarification of their inclusion criteria. The ERG also acknowledges the technical engagement response form from Novartis claiming that two additional studies examined the use of midostaurin as a maintenance treatment:

- AMLSG 16-10 - Phase 2, open-label, nonrandomized study of midostaurin in combination with induction and consolidation chemotherapy and as single agent maintenance therapy following consolidation with alloSCT or chemotherapy in patients (aged 18-70 years) with newly diagnosed FLT3-ITD+ AML
- Radius: A Phase 2, Randomized Trial of Standard of Care with or without Midostaurin to Prevent Relapse Following Allogeneic Hematopoietic Stem Cell Transplant in Patients with FLT3-Itd-Mutated Acute Myeloid Leukemia (AML)

However, the AMLSG 16-10 study is not confined to midostaurin in the maintenance phase and the maintenance phase follows HSCT.[1] The Radius trial does compare midostaurin in the maintenance phase, but again only post-HSCT.[2] Therefore, the ERG considers that the results of neither trial are informative.

7. ***Key issue 7: HSCT was not included as a separate health state but was implicitly included in the modelling through the survival analysis, increasing the likelihood of bias.***

Although the ERG would have preferred to see an updated model structure including HSCT as a health state, the evidence provided in Figures B1 and B2 of Appendix B illustrated that the impact of HSCT on survival analyses of OS and RFS was likely minor. The long-term impact of HSCT on quality of life is, however, still unclear. The company conducted a scenario analysis in which a weighted average relapse utility value was calculated by treatment arm increasing the ICER by 2.3%. However, the ERG questions whether the long-term benefits of HSCT are fully explored in this scenario, as it was assumed that 38% of patients receiving HSCT would develop chronic graft versus host disease and the lower relapse utility for these patients likely outweighed the potential HSCT utility benefit.

8. ***Key issue 8: Some patients in QUAZAR trial received fewer cycles of consolidation therapy than is standard practice in the UK. This limits the applicability of the results to a UK setting.***

As described in response to key issue 3, the ERG prefers using the consolidation subgroup (i.e. the population that received at least one cycle of consolidation therapy) in its base-case. For this subgroup, the company provided updated assessments of the suitability of the survival models according to the NICE DSU TSD 14 criteria in Appendix C. The company stated that the "additional survival analyses in the consolidation subgroup aligned with the assessment for the ITT population, with joint generalised gamma providing the optimal fit for OS and joint log-logistic providing the optimal fit for RFS". The ERG agrees that using an AFT model with a treatment covariate (i.e. joint generalised gamma and joint log-logistic for OS and RFS respectively) is acceptable when log-cumulative hazard plots are not parallel but relatively straight. However, the ERG would like to highlight that using individual parametric models without a treatment covariate (the alternative acceptable approach) has a substantial impact on the ICER. For OS, based on the visual and statistical fit to the data as provided in Appendix C, the individual log-normal model was the best fitting model for oral azacitidine and the individual generalised gamma model was the best fitting model for placebo. For RFS, the best fitting model for oral azacitidine was the individual log-logistic model and for placebo this was the individual Gompertz model. However, the ERG agrees with the company that the tail of the Gompertz extrapolation may not be clinically plausible and considers the log-logistic (2nd best statistical fit) to be more appropriate for the modelling of RFS in the placebo arm. Individually modelling OS and RFS using the curves above increases the ERG base-case ICER from £53,291 to £83,279 per QALY gained. It should be noted that the company did not provide expert opinion to assess the clinical plausibility of the extrapolated curves in Appendix C.

9. ***Key issue 9: Patient baseline characteristics in the model are not subgroup-specific (for example in the FLT3 subgroup, consolidation subgroup or Europe subgroup); patient baseline characteristics may not align with the subgroups being analysed.***

The ERG is satisfied with the company's additional scenario analyses showing that the use of subgroup-specific baseline characteristics had no material impact on the cost-effectiveness results. However, no updated economic model was provided and hence these scenario analyses could not be replicated by the ERG.

10. ***Key issue 10: Survival analyses of the FLT3 subgroup are likely to be biased due to limitations associated with the indirect comparison.***

As stated in the ERG report, bias due to limitations associated with the ITC seems not to be resolvable. Although the ERG appreciates that the company provided updated assessments of the suitability of the survival models in the FLT3 subgroup according to the NICE DSU TSD 14 criteria in Appendix D, the ERG considers any approach or chosen model likely to be biased. For OS, the company selected the individual generalised gamma in its base-case ([redacted]) based on expert opinion. The ERG explored using the individual log-normal ([redacted]) which increased the deterministic ERG base-case (£24,564 per QALY gained) to £27,173 per QALY gained. For RFS, the company explored spline models due to the poor fit in the individual standard parametric models. The company used the 1 internal knot, odds linear predictor model based on statistical fit and clinical plausibility, which the ERG agrees may be more suitable than using any of the standard parametric models.

11. ***Key issue 11: In the company's base-case analysis, only grade 3 and 4 AEs are applied with a maximum frequency of one and a duration of 1 week, which may underestimate the real impact of AEs.***

No compelling new arguments/evidence provided. Hence the ERG perspective as described in the ERG report remains unchanged.

12. ***Key issue 12: The current source of utility values may not accurately reflect the relapse utility.***

No compelling new arguments/evidence provided. Hence the ERG perspective as described in the ERG report remains unchanged.

13. ***Key issue 13: Some resource use estimates appear inconsistent with expert opinion and require further justification.***

Although it is not clear to the ERG whether the scenario analyses of the company are 100% consistent with expert opinion, the impact of this key issue on the ICER is likely minor.

14. ***Key issue 14: Treatment effectiveness in the FLT3 subgroup was analysed for the different comparisons separately; preventing comparison of oral azacitidine, midostaurin, watch and wait plus BSC.***

The company in its technical engagement response wrongly stated that a fully incremental analysis was performed. No compelling new arguments/evidence provided. Hence the ERG perspective as described in the ERG report remains unchanged.

Additional issues

1. ***Additional issue 1: End of life***

No new evidence has been submitted and the ERG therefore maintains its position regarding whether the end-of-life criteria has been met that is stated in the ERG report (section 7 of ERG report).

2. ***Additional issue 2: Clarification on duration of oral azacitidine treatment effect***

No compelling new arguments/evidence provided. Hence the ERG perspective as described in the ERG report remains unchanged.

3. ***Additional issue 3: Clarification on placebo and best supportive care (BSC)***

No new evidence has been submitted and the ERG maintains it's concerns regarding the use of the placebo comparator (section 2.3.1 of ERG report).

4. ***Additional issue 4: Design of QUAZAR AML-001 trial to account for consolidation***

No compelling new arguments/evidence provided. Hence the ERG perspective as described in the ERG report remains unchanged.

5. ***Additional issue 5: Clarification on the definition of documented relapse***

The ERG acknowledges the clearer definition of documented relapse.

6. ***Additional issue 6: HRQoL tables***

The ERG acknowledges the receipt of additional HRQoL tables

Conclusion

The ERG acknowledges the responses from the company and agrees that Key Issue 9 has been mostly resolved.

Key Issue 10 appears to be not resolvable, and the approach taken by the company is the best option available, and Key Issue 13 is likely to only have a minor impact on the ICER.

The ERG reiterates that Key Issues 1, 2, 3, 4, 5, 6, 7, 8, 11, 12, and 14 have not been resolved and introduce considerable uncertainty regarding the apparent benefits, safety, and cost-effectiveness of oral azacitidine. The ERG takes all of these to be fundamental as they might have a considerable impact on the estimates of benefit, harm, and cost-effectiveness of oral azacitidine.

References

[1] University of Ulm, Novartis Pharmaceuticals. *Protocol in Acute Myeloid Leukemia With FLT3-ITD. NCT01477606. In: ClinicalTrials.gov [Internet]. Bethesda (MD): National Library of Medicine (US)*, 2020 [accessed 22.4.22] Available from: https://clinicaltrials.gov/ct2/show/NCT01477606

[2] Maziarz RT, Levis M, Patnaik MM, Scott BL, Mohan SR, Deol A, et al. Midostaurin after allogeneic stem cell transplant in patients with FLT3-internal tandem duplication-positive acute myeloid leukemia. *Bone Marrow Transplant* 2021;56(5):1180-9.