NICE National Institute for Health and Care Excellence

Single Technology Appraisal

Gilteritinib for treating relapsed or refractory acute myeloid leukaemia [ID1484]

Committee Papers

NATIONAL INSTITUTE FOR HEALTH AND CARE EXCELLENCE

SINGLE TECHNOLOGY APPRAISAL

Gilteritinib for treating relapsed or refractory acute myeloid leukaemia [ID1484]

Contents:

The following documents are made available to consultees and commentators:

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

1. **Company submission** from Astellas

2. **Clarification questions and company responses**

3. **Patient group, professional group and NHS organisation submission** from:
   a. Leukaemia Care
   b. Royal College of Pathologists

4. **Expert personal perspectives** from:
   a. Dr Mike Dennis – clinical expert, nominated by Royal College of Pathologists and British Society for Haematology
   b. Professor Nigel Russell – clinical expert, nominated by NCRI-ACP-RCP
   c. Ms Charlotte Martin – patient expert, nominated by Leukaemia Care

5. **Evidence Review Group report** prepared by School of Health and Related Research

6. **Evidence Review Group – factual accuracy check**

7. **Technical engagement response** from Astellas

8. **Technical engagement responses from experts:**
   a. Ms Charlotte Martin – patient expert, nominated by Leukaemia Care
   b. Dr Mike Dennis – clinical expert, nominated by Royal College of Pathologists and British Society for Haematology
   c. Professor Nigel Russell – clinical expert, nominated by NCRI-ACP-RCP

9. **Technical engagement response from consultees and commentators:**
   a. NCRI-ACP-RCP-RCR

10. **Evidence Review Group critique of company response to technical engagement** prepared by School of Health and Related Research

11. Final Technical Report

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

Gilteritinib (XOSPATA™) for treating relapsed or refractory acute myeloid leukaemia ID1484

Document B

Company evidence submission

June 2019

File name	Version	Contains confidential information	Date
ID1484 Gilteritinib AML Company Submission ACIC	Final Redacted	Yes	23rd August 2019

Instructions for companies

This is the template for submission of evidence to the National Institute for Health and Care Excellence (NICE) as part of the single technology appraisal (STA) process. Please note that the information requirements for submissions are summarised in this template; full details of the requirements for pharmaceuticals and devices are in the user guide.

This submission must not be longer than 150 pages, excluding appendices and the pages covered by this template. If it is too long it will not be accepted.

Companies making evidence submissions to NICE should also refer to the NICE guide to the methods of technology appraisal and the NICE guide to the processes of technology appraisal.

In this template any information that should be provided in an appendix is listed in a box.

Highlighting in the template (excluding the contents list)

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.

Grey highlighted text in the footer does not work as an automatic form field, but serves the same purpose – as prompt text to show where you need to fill in relevant details. Replace the text highlighted in [grey] in the header and footer with appropriate text. (To change the header and footer, double click over the header or footer text. Double click back in the main body text when you have finished.)

Contents

Tables

Figures

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

B.1.1 Decision problem

The submission covers the full marketing authorisation of gilteritinib for this indication, namely, the treatment of adult patients who have relapsed or refractory acute myeloid leukaemia (AML) with a FLT3 mutation.

- This is relevant to NHS clinical practice; it would not be used in a wider population
- The evidence base on gilteritinib is limited to this population
- This population optimises the cost effectiveness of gilteritinib, because gilteritinib specifically targets FLT3 mutation positive AML, as identified by an appropriate diagnostic test
- This population reflects where gilteritinib provides the most clinical benefit
- Gilteritinib is not clinically effective in patients without the FLT3 mutation

Table 1 The Decision Problem

	Final scope issued by NICE	Decision problem addressed in the company submission	Rationale if different from the final NICE scope

Population	Adults with relapsed or refractory FLT3 mutation positive acute myeloid leukaemia	Adults with relapsed or refractory FLT3 mutation positive acute myeloid leukaemia	N/A
Intervention	Gilteritinib	Gilteritinib	N/A
Comparator(s)	Established clinical management without gilteritinib, for example: • Intermediate dose cytarabine (IDAC) • Fludarabine, cytarabine, granulocyte-colony stimulating factor (G-CSF) with idarubicin (FLAG-Ida) • Best supportive care • Hydroxycarbamide (for people who cannot have chemotherapy or stem cell transplant)	Established clinical management without gilteritinib including, but not limited to cytarabine or azacitidine based chemotherapy. For some patients, best supportive care may be their only option currently	The comparators used in the model are those included within the pivotal phase III trial (ADMIRAL). These were considered commonly used agents across the geographies for the trial
Outcomes	The outcome measures to be considered include: • overall survival • event-free survival • disease-free survival • response rates, including remission • stem cell transplant • adverse effects of treatment • health-related quality of life	The outcome measures to be considered include: • overall survival • event-free survival • disease-free survival • response rates, including remission • stem cell transplant • adverse effects of treatment • health-related quality of life	N/A

Special considerations including issues related to equity or equality		Gilteritinib represents an end-of-life (EoL) treatment based on NICE criteria: it is indicated in a population with a life expectancy less than 24 months and offers a survival extension of greater than 3 months. AML is an orphan condition, with an incidence of approximately 4.8 per 100,000. Relapsed or refractory patients are estimated to be 57% of these and FLT3 mutation occurs in approximately 30% of patients	Scope did not include such commentary, but did include the commentary around the life expectancy

B.1.2 Description of the technology being appraised

Gilteritinib is a tyrosine kinase-3 (FLT3) and AXL inhibitor. It is the first and only oral monotherapy shown to deliver over 9 months median OS (vs 5.6 months with salvage chemotherapy) in patients with relapsed or refractory FLT3 mutation positive AML.

Table 2 Technology Being Appraised

UK approved name and brand name	gilteritinib (XOSPATA™)
Mechanism of action	Gilteritinib is a tyrosine kinase-3 (FLT3) and AXL inhibitor. It is administered orally
Marketing authorisation/CE mark status	Gilteritinib does not currently have a marketing authorisation in the UK Astellas applied for a European Licence with the EMA on 28 February 2019 and expects CHMP recommendation in [redacted], with a licence granted in [redacted]. Gilteritinib is currently being assessed under Accelerated Assessment criteria
Indications and any restriction(s) as described in the summary of product characteristics (SmPC)	The expected indication of gilteritinib is for the treatment of adult patients who have relapsed or refractory acute myeloid leukaemia (AML) with a FLT3 mutation
Method of administration and dosage	Gilteritinib is administered as oral tablets, 120mg once daily (3 tablets x 40mg per tablet)
Additional tests or investigations	Patients undergo FLT3 mutation testing at the point of AML diagnosis as a prognostic and also to support the treatment decision for midostaurin. With the introduction of gilteritinib, relapsed or refractory patients should undergo FLT3 mutation testing to confirm their eligibility for treatment, regardless of their FLT3 mutation status at diagnosis. This represents an additional test associated with gilteritinib and has been incorporated in the economic analyses and budget impact assessment
List price and average cost of a course of treatment	The anticipated list price is [redacted] per 28-day pack It is assumed patients receive an average of [redacted] x 28 day dosing cycles ([redacted] days of treatment) This equates to an average cost of [redacted] per patient
Patient access scheme (if applicable)	Astellas has proposed a patient access scheme based on a simple discount of [redacted]. The anticipated PAS price is [redacted] per 28-day pack This equates to an average cost of [redacted] per patient

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

Acute myeloid leukaemia (AML) is a rare cancer of the blood and bone marrow. The incidence of AML varies by age, sex, and geography, with higher rates reported among the elderly, males, and individuals in developed Western nations.[1] Cancer Research UK (CRUK) estimates a crude incidence rate of 4.8 per 100,000 in the UK, with incidence increasing by age.[2] The European Medicines Agency has also acknowledged the orphan status of this disease.

AML is characterised by the uncontrolled proliferation and infiltration of bone marrow and blood by abnormally or poorly differentiated leukaemic blasts of the myeloid cell lineage.[14] Accumulation of leukaemic blasts in the bone marrow disrupts the formation of normal blood cells and platelets while their accumulation in the blood can disrupt circulation.[15] Leukaemic blasts also can infiltrate into other tissues and organs and impair their normal function.

AML is associated with poor outcomes with 5-year survival rates of 25%. AML is known for its tendency to progress rapidly and relapse.[1] Median survival is reported at two months or less in patients receiving supportive care alone,[3,4] suggesting the patient's condition deteriorates rapidly if remission is not attained. Complications arising from bone marrow failure, such as infection and haemorrhage, are the most common causes of mortality in AML.[1] Treatment choice and prognosis for AML depends on multiple factors including the extent of the disease, treatment history, patient age, symptoms, general state of health, as well as the disease risk group based on cytogenetic and molecular markers.[5–7] Although most treatment-eligible patients initially respond to therapy, the majority will experience disease relapse.[8–11] Prognosis for relapsed or refractory AML is especially poor, with less than 30% of patients still alive at 1 year post-relapse.[8]

FLT3 is a receptor tyrosine kinase involved in the normal development of haematopoietic stem cells and progenitor cells. Activating mutations in FLT3, are one of the most common class of recurring mutations in patients with AML and occur in around 30% of patients with AML.[6] Mutations in FLT3 are used to help stratify patients into prognostic risk groups and are suggestive of sensitivity to agents that inhibit

FLT3.[6,7,12,13] Patients who are FLT3 mutation positive have a poorer prognosis and a higher risk or relapse compared to patients with wild type AML,[12,14,15] Multiple studies have shown these patients have shorter remission duration and poorer survival outcomes than patients with wild-type FLT3.[12] The estimated median OS for patients with FLT3 mutations is 15.2 to 15.5 months compared to 19.3 to 28.6 months for patients with wild-type FLT3.[16]

There are two main types of mutation in FLT3. Internal tandem duplications (FLT3-ITD) are found in approximately 15% to 30% of the general AML population, and mutations in the FLT3 tyrosine kinase domain (FLT3-TKD) are found in approximately 5% to 10% of the general AML population.[17,18] FLT3-ITD mutations are associated with increased risk of relapse risk,[12,14,19] whilst the prognostic significance of FLT3-TKD is less clear.[12,14,15] Overall approximately 30% of AML patients carry the FLT3 mutation.[6]

Molecular genetic analysis is used to identify patients with the FLT3 mutation and is part of standard treatment guideline recommendations for characterising AML subgroups. Initial workups includes comprehensive medical history and physical examination, laboratory evaluations (including blood chemistry with complete blood count including platelets and differentials), bone marrow analysis with cytogenetics and evaluation of several molecular markers (e.g., FLT3), and HLA-typing (for patients eligible for allogeneic HSCT).[12]

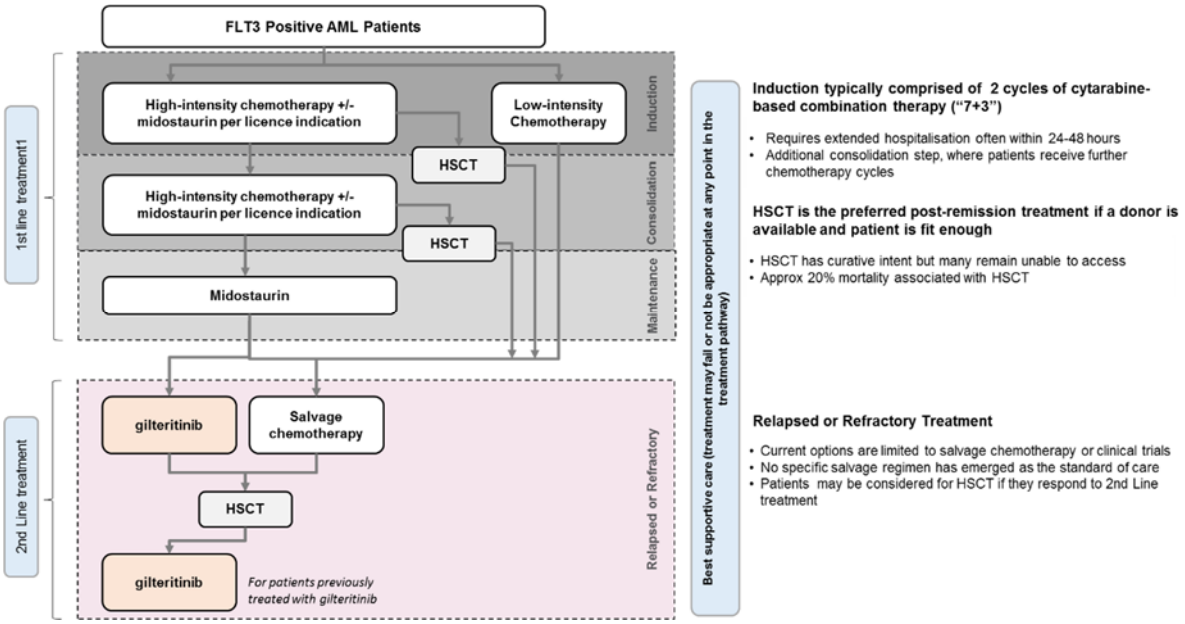
Testing for mutations in FLT3 is typically performed with PCR-based assays using DNA isolated from patient samples.[20] Tests that employ next-generation sequencing are an emerging option for identifying mutations in FLT3 and other molecular aberrations in samples from patients with AML, however this is considered a slower option with the potential for more false negatives and it is expected that PCR will remain the standard for FLT3 mutation testing until such factors are resolved.[21]

With the introduction of gilteritinib as a treatment option at the point a patient is diagnosed as relapsed or refractory, it would be recommended that patients are tested to ascertain their FLT3 mutation status. Only patients who test positive for the mutation would be eligible for gilteritinib. This would be conducted in all patients regardless of their previous FLT3 mutation status.

Clinical pathway of care

The clinical pathway of care for FLT3 AML is illustrated in Figure 1 below. It shows the four phases of care in AML: induction, consolidation, maintenance and treatment for relapse or refractory disease. The goal of treatment in AML is to induce remission and prevent relapse, with haemapoetic stem cell transplant (HSCT) being the only treatment with curative intent.[6,8,12,22] Despite the care pathway being well developed the outcomes from each of the treatments, except HSCT, remains poor, with median OS typically less than 6 months.

Figure 1 Clinical Pathway of Care



Abbreviations: AML: Acute myeloid leukaemia; HSCT: haematopoietic stem cell transplant

For newly diagnosed patients, current UK treatment guidelines, outlined in Table 3, generally recommend standard induction chemotherapy based on the 3+7 regimen followed by consolidation chemotherapy.[7,23–26] The 3+7 regimen is a combination of intravenous chemotherapy that includes 7 days of cytarabine and 3 days of an anthracycline, typically daunorubicin. If patients do not respond to the induction therapy, higher dosages may be administered. Some patients are not eligible for such treatment and receive low intensity chemotherapy with, for example LDAC or azacitidine, or best supportive care.

For patients who have achieved a complete response (CR) following induction consolidation therapy aims to prolong remission duration.[12] Standard consolidation therapy for AML often involves multiple cycles (typically 3-4) of high-dose cytarabine over 5 days in conjunction with additional chemotherapy drugs such as etoposide, daunorubicin or idarubicin.

As per its marketing authorisation, midostaurin (Rydapt) may also be used in combination with standard daunorubicin and cytarabine induction and high dose cytarabine consolidation chemotherapy, and as single agent maintenance therapy in patients in complete response, for adult patients with newly diagnosed FLT3 mutation positive AML.[27]

Patients with AML may receive a salvage chemotherapy regimen if they do not achieve a CR after induction chemotherapy or if they experience disease relapse after an initial response.[6,8,12] Although many patients who receive salvage chemotherapy are able to achieve a CR,[8] response durability is often limited and the treatment regimens are often associated with significant toxicity and mortality risk. Table 4 presents a list of salvage chemotherapy regimens used to treat patients with relapsed or refractory AML. There are no approved treatments specifically targeting FLT3 mutation positive relapsed or refractory AML and there is no standard of care given the paucity of effective treatment options. There is a need for treatment options which improve survival while being more tolerable and more convenient than intensive salvage chemotherapy.

Patients who respond to treatment and achieve a complete response may be considered eligible to receive haematopoietic stem cell transplant (HSCT) at any point in the treatment pathway. Where there is a suitable donor, HSCT is primarily an option for patients who are younger, sufficiently fit, and have high risk of disease relapse.

As an alternative to the above options, a significant number of patients enter clinical trials, with national protocols existing for such experimental treatments. This is again reflective of the paucity of effective treatment options in AML.

Table 3 Treatment Guidelines for Patients with Relapsed or Refractory FLT3 Mutation Positive AML

Guideline	Recommendations
British Society of Haematology (2006)[23]	**Relapsed AML:** Patients with relapsed disease should be stratified according to cytogenetics, age and length of CR1 to identify the best salvage approach
London Cancer (2015)[7]	**Relapsed AML:** Salvage chemotherapy or palliative care Patients entering a second CR should receive an allogenic SCT if possible For FLT3-ITD patients with relapsed AML and an identified donor, consider FLT3-inhibitors for initial therapy
Manchester Cancer (2015)[24]	**Refractory AML:** Salvage chemotherapy or allogeneic SCT Patients who are not suitable for allogeneic transplantation should be considered for investigational therapy of novel agents **Relapsed AML:** If CR duration was <6 months, palliative care or experimental therapy If CR duration >6months, consider salvage chemotherapy and SCT Patient who achieve CR should receive allogeneic SCT
NHS Birmingham (2011)[26]	**Relapsed AML:** Salvage chemotherapy to achieve a second CR should be considered if the patient has a transplant option available and is in good performance status

Table 4 Salvage Chemotherapy Regimens for Relapsed or Refractory AML

Regimen	Agents
HiDAC	High-dose cytarabine
FLAG	Fludarabine, cytarabine, G-CSF
FLAG-Ida	Fludarabine, cytarabine, G-CSF, idarubicin
FLA	Fludarabine, cytarabine
CLAG	Cladribine, cytarabine, G-CSF
CLAG-M	Cladribine, cytarabine, G-CSF, mitoxantrone
MEC	Mitoxantrone, etoposide, cytarabine
MEC-decitabine	Decitabine, mitoxantrone, etoposide, cytarabine
EMA-86	Mitoxantrone, cytarabine, etoposide
MAV	Mitoxantrone, cytarabine, etoposide
FLAD	Fludarabine, cytarabine, liposomal daunorubicin
FLAM	Flavopiridol, cytarabine, mitoxantrone
Hybrid FLAM	Flavopiridol, cytarabine mitoxantrone
Clofarabine cytarabine	Clofarabine, cytarabine
GCLAC	Clofarabine, cytarabine, G-CSF
HAA	Homoharringtonine, cytarabine, aclarubicin

B.1.4 Equality considerations

Astellas is not aware of any issues of equality in the management of AML in England and Wales and as such no equality considerations are considered in this submission.

B.2 Clinical effectiveness

B.2.1 Identification and selection of relevant studies

Two relevant studies have been identified: the phase I/II dose-escalating CHRYSALIS trial and the pivotal phase III ADMIRAL trial. CHRYSALIS (EudraCT 2014-002217-31) was an open-label, non-comparative study conducted in adults with relapsed or refractory AML. ADMIRAL (EudraCT 2015-000140-42) was conducted in adults with relapsed or refractory FLT3 mutation positive AML and compared gilteritinib with investigator's choice from specified salvage chemotherapies.

B.2.2 List of relevant clinical effectiveness evidence

The key evidence for the efficacy and safety of gilteritinib is drawn from CHRYSALIS (a phase I/II open-label, dose-escalating trial; EudraCT 2014-002217-31) and ADMIRAL (a phase III randomised open-label controlled trial; EudraCT 2015-000140-42). A systematic literature review was also conducted of relevant comparator studies, which is available in Appendix D.

CHRYSALIS

CHRYSALIS included patients with relapsed or refractory AML. The Full Analysis Set (FAS) included 249 patients and the Safety Analysis Set (SAF) included 252 patients. CHRYSALIS was a single-arm study in which patients were randomised to receive escalating doses of gilteritinib. The primary endpoints were safety, tolerability, and pharmacokinetics of gilteritinib.[28]

ADMIRAL

ADMIRAL included 371 patients with relapsed or refractory FLT3 mutation positive AML.[29] Patients were randomised in a 2:1 ratio to gilteritinib 120mg orally per day (administered as a single dose of 3 x 40mg tablets) and investigator's choice of pre-specified salvage chemotherapies (see Table 5). The co-primary endpoints were overall survival (OS), and the rate of complete remission (CR) and complete remission with partial haematological recovery (CR/CR_h).[29] The addition of the CR/CR_h endpoint

was based on guidance received from the FDA. ADMIRAL results have been presented at the American Association for Cancer Research (AACR) and European Haematology Association (EHA) annual meetings in 2019, and is pending full publication later in 2019.

Table 5 Characteristics of the Relevant RCTs

Study	CHRYSALIS	ADMIRAL
Study design	Phase I/II, open-label, multicentre dose-escalation trial	Phase III, open-label, multicentre, randomised study that compared the efficacy and safety of gilteritinib to salvage chemotherapy
Population	Adults with relapsed or refractory AML	Adults with relapsed or refractory FLT3 mutation positive AML
Intervention(s)	Gilteritinib 20mg, 40mg, 80mg, 120mg, 200mg, 300mg and 450mg (N=252)	Gilteritinib 120mg/day orally (single dose of 3 x 40mg tablets) (N=247)
Comparator(s)	No comparator	Investigator's choice of salvage chemotherapy in the form of any one of the following regimens (N=124): • Low dose cytarabine (20mg twice-daily SC or IV injections for 10 days) • Azacitidine (75mg/m^2 daily SC or IV injections for 7 days) • MEC (mitoxantrone 8mg/m^2 per day, etoposide 100mg/m^2 per day, cytarabine 1,000mg/m^2 per day, all administered via IV injection for 5 days on days 1 through 5) • FLAG-Ida (fludarabine 30mg/m^2 per day and cytarabine 2,000mg/m^2 per day, both administered via IV injection for 5 days on days 2 through 6; G-CSF 300µg/m^2 per day administered via SC or IV injection for 5 days on days 1 through 5; idarubicin 10mg/m^2 per day administered via IV injection for 3 days on days 2 through 4)
Indicate if trial supports application for marketing authorisation	Yes	Yes

Study	CHRYSALIS	ADMIRAL
Indicate if trial used in the economic model	No	Yes
Rationale for use in the model	N/A	The ADMIRAL trial is the highest quality source supporting the efficacy and safety of gilteritinib relative to the comparators defined in the scoping document.
Reported outcomes specified in the decision problem	**Primary outcomes:** • Safety • Tolerability • Pharmacokinetics **Secondary outcomes:** • Complete remission (CR) rate • Composite complete remission (CR_c) rate • Best response • Duration of remission (DOR) • Overall survival (OS) • Leukaemia-free survival (LFS) • Drug-drug interactions • Pharmacodynamics	**Primary outcomes:** • Overall survival (OS) • Rate of complete remission and complete remission with partial haematological recovery (CR/CR_h) **Secondary outcomes:** • Event-free survival (EFS) • Complete remission (CR) rate • Leukaemia-free survival (LFS) • Duration of remission (DOR) • Complete remission with partial haematological recovery (CR_h) rate • Composite complete remission (CR_c) rate • Transfusion conversion rate; transfusion maintenance rate • Transplantation rate • Patient reported fatigue (Brief Fatigue Inventory [BFI]) • Adverse events (AEs) **Exploratory outcomes:** • Various including EQ-5D-5L

Abbreviations: AE: adverse event; AML: acute myeloid leukaemia; BFI: brief fatigue inventory; CR: complete remission; CR_c: Composite complete remission; CR_h: Complete remission with partial haematological recovery; DOR: duration of remission; EFS: event-free survival; FLAG-Ida: fludarabine, cytarabine and granulocyte colony-stimulating factor with idarubicin; FLT3: FMS-like tyrosine kinase-3; G-CSF: granulocyte colony stimulating factor; IV: intravenous; LFS: Leukaemia-free survival; MEC: mitoxantrone, etoposide and intermediate-dose cytarabine; SC: subcutaneous

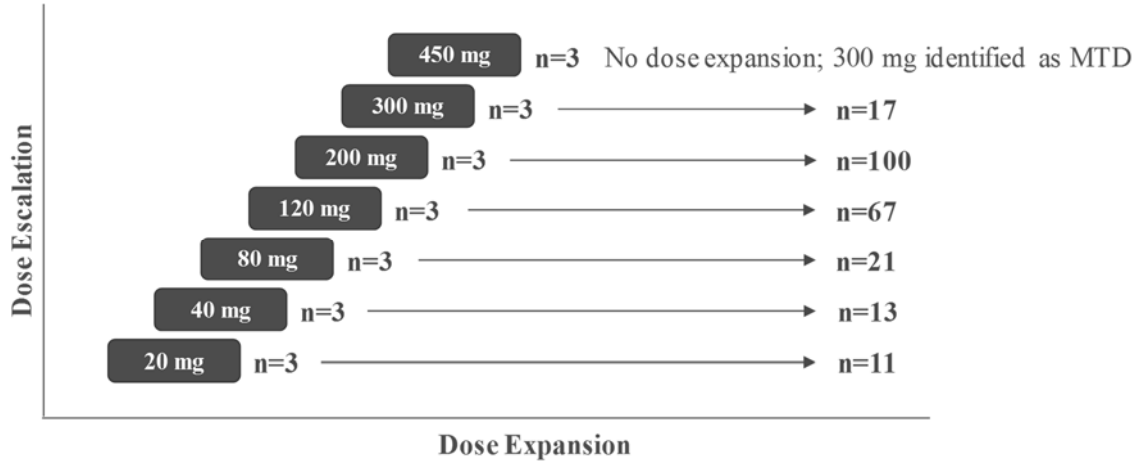
B.2.3 Summary of methodology of the relevant clinical effectiveness evidence

CHRYSALIS

CHRYSALIS was a multi-centre, open-label Phase I/II dose-escalation trial of gilteritinib in patients with relapsed or refractory AML. Patients received once-daily

oral gilteritinib in one of seven dose-escalation (n=23) or dose-expansion (n=229) cohorts in continuous 28-day cycles.[28]

Figure 2 CHRYSALIS Trial Design

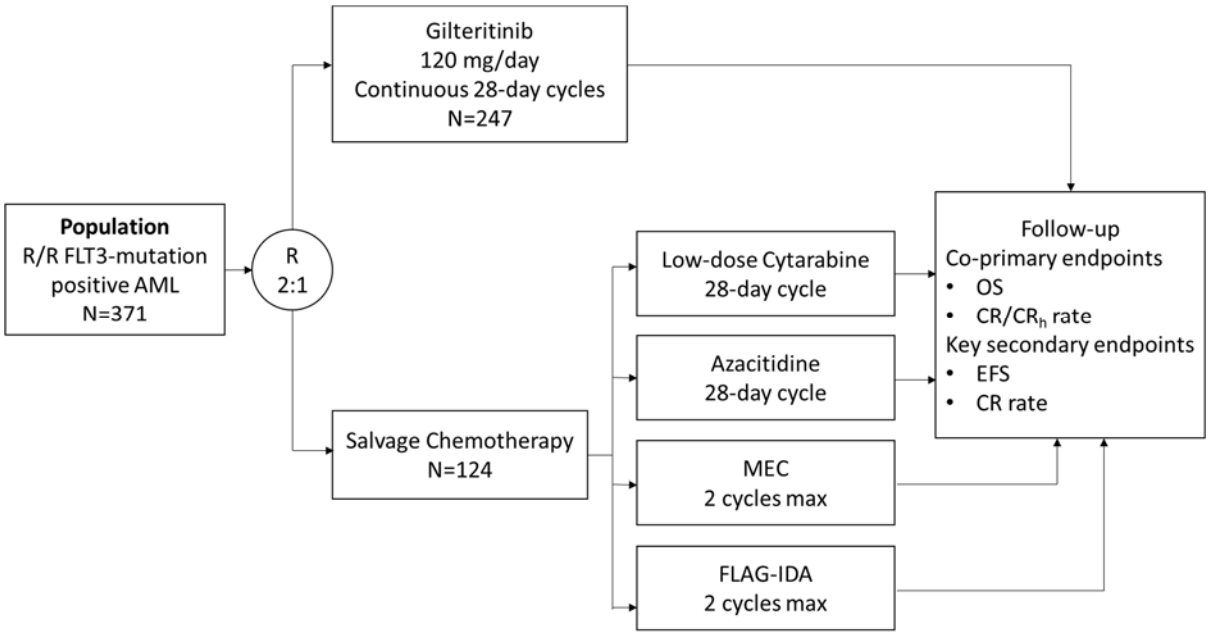


Abbreviations: MTD: maximum tolerated dose

ADMIRAL

ADMIRAL was a phase III, open-label, multicentre, randomised study comparing the efficacy and safety of gilteritinib to selected salvage chemotherapies in adults with relapsed or refractory FLT3 mutation positive AML.[29]

Figure 3 ADMIRAL Trial Design



Abbreviations: AML: acute myeloid leukaemia; CR: complete remission; CR_h: complete remission with partial haematologic recovery; EFS: event-free survival; FLAG Ida: fludarabine, cytarabine, granulocyte colony-stimulating factor, idarubicin; FLT3: FMS-like tyrosine kinase-3; MEC: mitoxantrone, etoposide, cytarabine; OS: overall survival; R/R: relapsed or refractory

Eligible patients were randomised in a 2:1 ratio to gilteritinib or salvage chemotherapy. The site investigator preselected a salvage chemotherapy regimen for each patient prior to randomisation; options for salvage chemotherapy were low-dose cytarabine (LoDAC); azacitidine; MEC induction chemotherapy (mitoxantrone, etoposide and intermediate-dose cytarabine); or FLAG-IDA induction chemotherapy (fludarabine, high-dose cytarabine, and granulocyte colony-stimulating factor [GCSF] with idarubicin). Randomisation was stratified by response to first-line AML therapy and preselected salvage chemotherapy.

Treatment continued until the patient met a treatment discontinuation criterion described below:

- Patient declined further study participation (i.e., withdrawal of consent)
- Patient was non-compliant with the protocol based on the investigator or medical monitor assessment
- Patient was found to have significantly deviated from any of the inclusion or exclusion criteria after enrolment (patients who had experienced clinical benefit may have remained in the study after discussion with the medical monitor).
- Patient developed intolerable or unacceptable toxicity
- Patient received any antileukaemic therapy other than the assigned treatment, with the exceptions of hydroxyurea for up to 2 weeks, prophylactic intrathecal chemotherapy or cranial irradiation, and donor lymphocyte infusion as part of a haematopoietic stem cell transplantation (HSCT) treatment plan
- Investigator/sub-investigator determined that continuation of the study treatment was detrimental to the patient
- Patient was lost to follow-up despite reasonable efforts by the investigator to locate the patient
- Patient was receiving MEC or FLAG-IDA and had non-response (NR) or progressive disease following cycle 1
- Patient was receiving LoDAC, azacitidine or gilteritinib and had progressive disease or NR and the patient, in the opinion of the investigator, was no longer deriving clinical benefit

- Patient was in the comparator group (chemotherapy) and went on to receive HSCT
- Female patient became pregnant
- Patient died

Patients receiving MEC or FLAG-IDA underwent one cycle of therapy and were assessed for response on or after day 15 as per institutional guidelines. If the bone marrow cellularity was ≥20% with ≥50% reduction in blasts, the patient received a second cycle of the same chemotherapy. If bone marrow cellularity was between 5% and 20%, the investigator decided whether the patient should receive another treatment cycle or be observed for recovery. If bone marrow cellularity was ≤5%, the patient was observed for recovery. Patients achieving CR, CR with incomplete haematologic recovery (CR_i) or CR with incomplete platelet recovery (CR_p) may have received a second cycle of chemotherapy at the investigator's discretion. Patients with no response (NR) or progressive disease following cycle 1 discontinued study treatment.

The study planned to randomise 369 patients in a 2:1 ratio to receive gilteritinib or salvage chemotherapy. It was expected that this sample size would be associated with 258 deaths for the analysis of OS. Actual enrolment was 371 patients, with 261 deaths observed by the data cut-off date of 17 September 2018.[29]

Further details of the CHRYSALIS and ADMIRAL trials are provided below.

Table 6 Summary of CHRYSALIS and ADMIRAL Methodology

STUDY	CHRYSALIS[28]	ADMIRAL[29]
Location	United States, Italy, France, Germany	107 sites in 14 countries: United States, United Kingdom, Germany, France, Spain, Italy, Belgium, Turkey, Poland, Canada, South Korea, Japan, Taiwan and Israel
Trial Design	Phase I/II, open-label, multicentre dose-escalation trial	Phase III, open-label, multicentre, randomised study
Eligibility criteria for participants	**Inclusion criteria:** • Age ≥18 years • Primary or secondary AML • ECOG performance status ≤2	**Inclusion criteria:** • Age ≥18 years • Refractory or relapsed AML (after first-line therapy with or without HSCT) • Confirmed FLT3 mutation

STUDY	CHRYSALIS[28]	ADMIRAL[29]
	• Refractory to at least 1 cycle of induction chemotherapy or relapsed after achieving remission with a previous drug • ALT or AST ≤2.5 X ULN • Serum creatinine ≤1.5 X ULN or eGFR >50 mL/min • Total bilirubin ≤1.5 X ULN **Exclusion criteria:** • APL • BCR-ABL-positive leukaemia • Malignant tumours other than AML or MDS • NYHA class 3 or 4 heart failure and those who had previously had NYHA class 3 or 4 heart failure, unless a screening ECG done within 3 months before study entry resulted in a LVEF of ≥45% • Long QTc syndrome • Active uncontrolled infections including hepatitis B or C and HIV • Pregnancy • Presence of grade ≥2 non-haematologic toxicities from prior AML treatment • Prior HSCT within 2 months of study treatment (Cycle 1, Day 1) • Persistent grade ≥2 non-haematologic toxicities related to HSCT • GvHD requiring treatment	• ECOG performance status ≤2 • Eligible for pre-selected salvage chemotherapy • ALT or AST ≤2.5 X ULN • Total bilirubin ≤1.5 X ULN • Serum creatinine ≤1.5 X ULN or eGFR >50 mL/min • Can receive oral therapy • Female patients must be either of non-child bearing potential or not pregnant at study initiation and not planning to become pregnant **Exclusion criteria:** • Acute promyelocytic leukaemia (APL) • Breakpoint cluster region-Abelson murine leukaemia • AML secondary to prior chemotherapy for other neoplasms (except for MDS) • History of another malignancy (unless disease free for ≥5 years) • Clinically active central nervous system leukaemia • Prior treatment with gilteritinib or other FLT3 inhibitors, with exception of sorafenib or midostaurin • Clinically significant abnormality of coagulation profile • Major surgery or radiation within 4 weeks of first study dose • NYHA class 3 or 4 heart failure and those who had previously had NYHA class 3 or 4 heart failure, unless a screening ECG done within 3 months before study entry resulted in a LVEF of ≥45% • Mean of triplicate QTcF >450 milliseconds • Long QT syndrome • Hypokalaemia or hypomagnesaemia • Active uncontrolled infections including hepatitis B or C and HIV, or other uncontrolled hepatic disorder • Active clinically significant GvHD or on treatment with systemic corticosteroids for GvHD
Settings and location where the data were collected	Data was collected from centres involved in the trial and reflected expected UK clinical practice	Data was collected from centres involved in the trial and reflected expected UK clinical practice
Trial drugs	Gilteritinib 20mg, 40mg, 80mg, 120mg, 200mg, 300mg and 450mg (N=252)	Arm I: Gilteritinib 120mg/day orally (single dose of 3 x 40mg tablets) for continuous 28-day cycles (N=247) Arm II: Salvage chemotherapy (N=124): Low dose cytarabine 20mg twice daily SC/IV for 7 days

STUDY	CHRYSALIS[28]	ADMIRAL[29]
		Azacitidine 75mg/m^2 daily SC/IV for 5 days MEC or FLAG-IDA IV for 5 days
Permitted and disallowed concomitant medication	Treatment with concomitant drugs that are strong inducers of CYP3A were prohibited. Treatment with concomitant drugs that are strong inhibitors or inducers of P-gp, and concomitant drugs that target serotonin 5HT2BR receptors or signma nonspecific receptor were to be avoided with the exception of drugs that are considered absolutely essential for the care of the subject. Treatment with concomitant drugs that are strong inhibitors of CYP3A were to be avoided with the exception of antibiotics, antifungals, and antivirals that are used as standard to prevent or treat infections If CYP3A inhibitors were used concomitantly, subjects were to be closely monitored for adverse events. Precaution was to be used with concomitant drugs that are known to prolong QT or QTc intervals. During the initial 15 days of treatment in expansion cohorts with DDI studies, moderate or strong CYP3A4 inhibitors were prohibited, unless required for treatment of active infections. In addition, during the initial 15 days of treatment for subjects assigned to Schedule 2E, MATE1 substrates were prohibited. Any other treatments of AML (including but not limited to chemotherapy, radiotherapy, surgery, immunotherapy or cellular therapy) are prohibited during therapy with the following exceptions: • Hydroxyurea up to 5gm daily for up to 2 weeks to keep the absolute blast count below 50,000 • Hematopoietic Stem Cell Transplants for patients with CRc or PR • Intrathecal Chemotherapy used as prophylaxis	*For gilteritinib arm only:* • Treatment with concomitant drugs that are strong inducers of CYP3A was prohibited • Treatment with concomitant drugs that are strong inhibitors or inducers of P-gp and concomitant drugs that target serotonin 5HT1R or 5HT2BR or sigma nonspecific receptor was to be avoided with the exception of drugs that were considered absolutely essential for the care of the patient • Treatment with concomitant drugs that are strong inhibitors of CYP3A should have been avoided with the exception of antibiotics, antifungals and antivirals that were used as standard of care to prevent or treat infections • Precaution should have been used in treatment of gilteritinib with concomitant drugs that are known to prolong QT or QTc intervals and drugs that are substrates of breast cancer resistance protein *For gilteritinib and salvage chemotherapy arm:* • Any other treatments of AML (including but not limited to chemotherapy, radiotherapy, surgery, immunotherapy or cellular therapy) were prohibited during therapy with the exception of hydroxyurea daily for up to 2 weeks to keep the absolute blast count below 50 x10^9/L and prophylactic intrathecal chemotherapy, cranial radiation, and donor lymphocyte infusion as part of the HSCT treatment plan
Primary outcomes	• Safety • Tolerability	• Overall survival (OS)

STUDY	CHRYSALIS[28]	ADMIRAL[29]
	• Pharmacokinetics	• Rate of complete remission and complete remission with partial haematological recovery (CR/CR_h)
Major secondary outcomes	• Complete remission (CR) rate • Composite complete remission (CR_c) rate • Best response • Duration of remission (DOR) • Overall survival (OS) • Leukaemia-free survival (LFS) • Drug-drug interactions • Pharmacodynamics	• Event-free survival (EFS) • Complete remission (CR) rate • Leukaemia-free survival (LFS) • Duration of remission (DOR) • Complete remission with partial haematological recovery (CR_h) rate • Composite complete remission (CR_c) rate • Transfusion conversion rate; transfusion maintenance rate • Transplantation rate • Patient-reported fatigue (Brief Fatigue Inventory [BFI]) • Adverse events
Pre-planned subgroups	Outcomes were analysed by dosing subgroup: 20mg, 40mg, 80mg, 120mg, 200mg, 300mg and 450mg	The following pre-planned subgroups were analysed for efficacy outcomes: • Age (<65 years, ≥65 years) • Sex (male, female) • Race (white, black or African American, Asian, other/missing) • Baseline ECOG (0-1, ≥2) • Region [North America, Europe (including Turkey and Israel), Asia] • Central FLT3 mutation type [FLT3-ITD alone, FLT3-TKD alone, FLT3-ITD and FLT3-TKD, others (unknown, missing, negative)] • Prior use of FLT3 inhibitor (yes, no) • Cytogenetic risk status (favourable, intermediate, unfavourable, other) • Response to first-line therapy per IRT (relapse within 6 months after allogeneic HSCT, relapse after 6 months after allogeneic HSCT, primary refractory without HSCT, relapse within 6 months after CR_c and no HSCT, relapse after 6 months after CR_c and no HSCT) • Pre-selected chemotherapy per IRT (high intensity, low intensity)

Abbreviations: AML: acute myeloid leukaemia; ALT: alanine aminotransferase; APL: acute promyelocytic leukaemia; AST: aspartate aminotransferase; BCR-ABL: Breakpoint cluster region-Abelson murine leukaemia viral oncogene homolog fusion protein (Philadelphia chromosome); BFI: brief fatigue inventory; CNS: central nervous system; CR: complete remission; CR_c: composite complete remission; CR_h: complete remission with partial haematological recovery; CYP3A: Cytochrome P 3A; ECG: electrocardiogram; ECOG: Eastern Cooperative Oncology Group; EFS: event-free survival; eGFR: estimated glomerular filtration rate; FLAG-Ida: fludarabine, cytarabine and granulocyte colony-stimulating factor with idarubicin; FLT3: FMS-like tyrosine kinase-3; FLT3-ITD: FMS-like tyrosine kinase-3-internal tandem duplication; FLT3-TKD: FMS-like tyrosine kinase-3-tyrosine kinase domain; GvHD: graft-versus-host disease; HIV: human immunodeficiency virus; HSCT: haematopoietic stem cell transplant; IRT: interactive response technology; IV: intravenous; LFS: leukaemia-free survival; LVEF: left ventricular ejection fraction; MEC: mitoxantrone, etoposide and intermediate-dose cytarabine; NYHA: New York Heart Association; OS: Overall survival; P-gp: P-glycoprotein; PRO: patient reported outcomes; QTcF: corrected QT interval by Fredericia; SC: subcutaneous; ULN: upper limit of normal; 5HT1R: 5-hydroxytryptamine-1 receptor; 5HT2BR: 5-hydroxytryptamine-2B receptor

Table 7 below presents the different definitions of response used to measure effectiveness.

Table 7 Definitions of Complete Response

Definition	Description
Complete Remission (CR)	For subjects to be classified as being in CR at a post-baseline visit, they must have bone marrow regenerating normal hematopoietic cells and achieve a morphologic leukemia-free state and must have an ANC ≥ 1 x 109/L and platelet count ≥ 100 x 109/L and normal marrow differential with < 5% blasts, and they will be RBC and platelet transfusion independent (defined as 1 week without RBC transfusion and 1 week without platelet transfusion) There should be no evidence of extramedullary leukemia
Complete Remission with Partial Hematologic Recovery (CRh)	At a post baseline visit, subjects will be classified as CRh if they have marrow blasts < 5%, partial hematologic recovery ANC ≥ 0.5 x 109/L and platelets ≥ 50 x 109/L, no evidence of extramedullary leukemia and cannot be classified as CR
Complete Remission with Incomplete Platelet Recovery (CRp)	For subjects to be classified as being in CRp at a post-baseline visit, they must achieve CR except for incomplete platelet recovery (< 100 x 10^9/L)
Complete Remission with Incomplete Haematologic Recovery (CRi)	For subjects to be classified as being in CRi at a post-baseline visit, they must fulfill all the criteria for CR except for incomplete hematological recovery with residual neutropenia < 1 x 109/L with or without complete platelet recovery. RBC and platelet transfusion independence is not required
Composite Complete Remission (CRc)	For subjects to be classified as being in CRc at a post-baseline visit, they must either achieve CR, CRp or CRi at the visit
Partial Remission (PR)	For subjects to be classified as being in PR at a post-baseline visit, they must have bone marrow regenerating normal hematopoietic cells with evidence of peripheral recovery with no (or only a few regenerating) circulating blasts and with a decrease of at least 50% in the percentage of blasts in the bone marrow aspirate with the total marrow blasts between 5% and 25%. A value of less or equal than 5% blasts is also considered a PR if Auer rods are present

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

CHRYSALIS

The primary objectives of CHRYSALIS were to assess the safety and tolerability of gilteritinib in patients with relapsed or refractory AML, including the maximum tolerated dose, and to determine the pharmacokinetic (PK) parameters of gilteritinib. The secondary objective was to assess efficacy outcomes of gilteritinib by dose.

Efficacy variables were assessed in term of the number of subjects (n), mean, standard deviation, median, minimum and maximum. Kaplan-Meier survival curves were used to display for time-to-event variables and median survival time was be estimated with 2-sided 95% confidence interval (CI). All statistical comparisons were

made using two sided tests at the α=0.05 significance level. All null hypotheses were of no treatment difference and all alternative hypotheses were two-sided. Data from each dose group were combined for the analysis of efficacy and safety endpoints. The sample size was not based on a statistical power calculation.

The Safety Analysis Set (SAF) was be used for the analyses of safety and biomarker variables. Full Analysis Set (FAS), Safety Analysis Set (SAF) and Per Protocol Set (PPS) were used for efficacy analysis. Pharmacokinetic Analysis Set (PKAS) was used for pharmacokinetic analyses. Pharmacodynamic Analysis Set (PDAS) was used for the analyses of pharmacodynamic data.[28]

ADMIRAL

ADMIRAL was powered to detect a difference between the gilteritinib arm and salvage chemotherapy arm for the co-primary efficacy endpoints, OS and CR/CR_h rate. The efficacy analyses were based on the intent-to-treat (ITT) study population which included all randomised patients. The Full Analysis Set (FAS) included all randomised patients with an FLT3 mutation. The Safety Analysis Set (SAF) included all patients who took at least one dose of study treatment (gilteritinib or salvage chemotherapy).

OS was analysed using a stratified log-rank test with strata to control for response to first-line AML therapy and preselected salvage chemotherapy. The CR/CR_h rate was analysed by calculating two-sided 95% exact CIs based on a binomial distribution for the gilteritinib and the salvage chemotherapy arms and checking for overlap.

The secondary endpoint of EFS was analysed using a stratified log-rank test in the same manner as OS. A hierarchical analysis was conducted whereby if the EFS endpoint was not met, the analyses of the subsequent secondary endpoints were considered as descriptive only. CR rate was analysed using the Cochran-Mantel-Haenszel (CMH) test. LFS and duration of remission were also analysed for patients who achieved remission using a stratified log-rank test. The number and percent of patients with CR_c and HSCT were summarised for each treatment arm with the exact 95% CI based on binomial distributions. Transfusion conversion rate and transfusion maintenance rate were calculated based on the transfusion status, with 95% CIs. The number and percentage of patients with transplantations were summarised for each

treatment arm with the exact 95% CI based on binomial distribution. The BFI global fatigue score was summarised using mean, SD, minimum, maximum and median values by treatment arm at each visit in the ITT.[29]

A safety analysis was also conducted, based on observed adverse events, clinical laboratory measurements, vital signs, ECG, ophthalmologic assessments and ECOG. Descriptive statistics were used to summarise the rates of these safety events. These events are summarised in Appendix F.

B.2.5 Quality assessment of the relevant clinical effectiveness evidence

Overall, CHRYSALIS and ADMIRAL were considered to be of high quality in terms of compliance to good clinical practice.[30] A summary of the quality assessment of CHRYSALIS and ADMIRAL is presented in Appendix D.

CHRYSALIS

CHRYSALIS was an open-label study in which the full analysis set included 249 patients, of which 191 had FLT3 mutations, and the safety set included 252 patients. The median treatment duration was 25.9 weeks and 300mg/day was identified as the maximum tolerated dose (MTD).[28] Detailed patient demographic and clinical characteristics are described in Table 8.

Table 8 Baseline Characteristics of Patients in CHRYSALIS

Parameter	20mg/day (n=16)	40mg/ day (n=16)	80mg/ day (n=24)	120mg/ day (n=70)	200mg/ day (n=103)	300mg /day (n=20)	450m g/day (n=3)
Median age, years (IQR)	65 (58 to 71)	62 (54 to 66)	62 (47 to 70)	60 (51 to 69)	64 (49 to 70)	64 (46 to 69)	64 (50 to 71)
Sex, n (%)							
Male	6 (38)	11 (69)	11 (46)	32 (46)	52 (51)	14 (70)	3 (100)
Female	10 (63)	5 (31)	13 (54)	38 (54)	51 (50)	6 (30)	0
Cytogenetic risk group, n (%)							
Favourable	0	0	2 (8)	1 (1)	3 (3)	0	1 (33)
Intermediate	13 (81)	5 (31)	11 (46)	42 (60)	64 (62)	8 (40)	0
Unfavourable	2 (13)	9 (56)	7 (29)	12 (17)	17 (17)	7 (35)	2 (67)
Median disease duration, months, (IQR)	10.6	7.1	16.8	9	8.3	7.3	6.3

Parameter	20mg/day (n=16)	40mg/ day (n=16)	80mg/ day (n=24)	120mg/ day (n=70)	200mg/ day (n=103)	300mg /day (n=20)	450m g/day (n=3)
	(7.2 to 16.1)	(5.1 to 11.7)	(8.3 to 29)	(4.7 to 16.6)	(3.9 to 13.8)	(2.7 to 16.5)	(3.5 to 11.9)
Prior HSCT							
0	11 (69)	13 (81)	15 (63)	49 (70)	71 (69)	18 (90)	2 (67)
1	4 (25)	2 (13)	9 (38)	20 (29)	29 (28)	2 (10)	1 (33)
≥2	1 (6)	1 (6)	0	1 (1)	3 (3)	0	0
Prior lines of therapy for AML, n (%)							
1	3 (19)	5 (31)	5 (21)	17 (24)	36 (35)	7 (35)	2 (67)
2	3 (19)	1 (6)	5 (21)	22 (31)	28 (27)	7 (35)	0
≥3	10 (63)	10 (63)	14 (58)	31 (44)	39 (38)	6 (30)	1 (33)
Prior treatment with TKI, n (%)	8 (50)	4 (25)	5 (21)	22 (31)	21 (20)	2 (10)	1 (33)
FLT3 mutation status, n (%)							
FLT3-ITD alone	12 (75)	6 (38)	10 (42)	47 (67)	79 (77)	8 (40)	0
FLT3-TKD (D835) alone	1 (6)	0	1 (4)	6 (9)	3 (3)	1 (5)	1 (33)
FLT3-ITD and FLT3-TKD	1 (6)	2 (13)	1 (4)	3 (4)	8 (8)	0	1 (33)

Abbreviations: AML: acute myeloid leukaemia; FLT3: FMS-like tyrosine kinase-3; HSCT: haematopoietic stem cell transplant; IQR: interquartile range; ITD: internal tandem duplication; TKD: tyrosine kinase domain; TKI: tyrosine kinase inhibitor

ADMIRAL

ADMIRAL was an open-label study with all outcome assessments based on the ITT principle. Randomisation in the trial was carried out appropriately, and baseline characteristics were well balanced across treatment groups. There is no evidence of any biological or genetic variation which may compromise the clinical validation or generalisability of this study to the population of England and Wales.

The mean patient age was [redacted] and [redacted] years in the gilteritinib and salvage chemotherapy arms, respectively, and the proportion of patients <65 years of age were [redacted] and [redacted], respectively. The median disease duration was [redacted] and [redacted] months in gilteritinib and salvage chemotherapy arms, respectively. Most patients had the FLT3-ITD mutation alone (87.0% and 91.1% of patients in the gilteritinib and salvage chemotherapy arms, respectively); the FLT3-TDK mutation was identified in 8.5% and 8.1% of patients, respectively; and both mutations were found in 2.8% of patients in gilteritinib arm but none of the patients in the salvage chemotherapy arm.

All patients in both treatment arms had previously received chemotherapy for the treatment of AML. Thirteen percent of patients in the gilteritinib arm and 11.3% patients in the salvage chemotherapy arm had previously received a FLT3 inhibitor (e.g. midostaurin), while 19.4% and 21.0% patients in the gilteritinib arm and salvage chemotherapy arm, respectively, had received HSCT. Response to first-line therapy was similar in both treatment arms: 39.4% of patients were characterised as primary refractory without HSCT and [redacted] experienced a relapse within 6 months after achieving CR_c without HSCT.

High intensity chemotherapy was given to [redacted] and [redacted] patients in the gilteritinib arm and salvage chemotherapy arm, respectively. Overall, cytogenetic risk was characterised as intermediate in most patients (73.0%), unfavourable in 10.0%, and favourable in only 1.3%.[29]

Table 9 Baseline Characteristics of Patients in ADMIRAL

Characteristic	Gilteritinib (N=247)	Salvage Chemotherapy (N=124)	Overall (N=371)
Age, years			
Mean (SD)	[redacted]	[redacted]	[redacted]
Median (range)	62.0 (20-84)	61.5 (19-85)	62.0 (19-85)
Age group, n (%)			
<65 years	[redacted]	[redacted]	[redacted]
≥65 years	[redacted]	[redacted]	[redacted]
Males, n (%)	116 (47.0)	54 (43.5)	170 (45.8)
Race, n (%)			
White	[redacted]	[redacted]	[redacted]
Asian	[redacted]	[redacted]	[redacted]
Black or African American	[redacted]	[redacted]	[redacted]
Native Hawaiian or other Pacific Islander	[redacted]	[redacted]	[redacted]
Other	[redacted]	[redacted]	[redacted]
Unknown	[redacted]	[redacted]	[redacted]
Missing	[redacted]	[redacted]	[redacted]
Baseline ECOG performance status			
0-1	[redacted]	[redacted]	[redacted]
2	[redacted]	[redacted]	[redacted]
Median disease duration, months (range)	[redacted]	[redacted]	[redacted]
FLT3 mutation status (central testing), n (%)			
FLT3-ITD alone	215 (87.0)	113 (91.1)	328 (88.4)
FLT3-TKD alone	21 (8.5)	10 (8.1)	31 (8.4)
FLT3-ITD and FLT3-TKD	7 (2.8)	0	7 (1.9)
Other (negative)	[redacted]	[redacted]	[redacted]

Characteristic	Gilteritinib (N=247)	Salvage Chemotherapy (N=124)	Overall (N=371)
Prior treatments received, n (%)			
Chemotherapy for AML	[redacted]	[redacted]	[redacted]
FLT3 inhibitor	32 (13.0)	14 (11.3)	46 (12.4)
HSCT	48 (19.4)	26 (21.0)	74 (19.9)
Response to first-line therapy			
Primary refractory without HSCT	98 (39.7)	48 (38.7)	146 (39.4)
Relapse within 6 months after CR_c and no HSCT	[redacted]	[redacted]	[redacted]
Relapse after 6 months after CR_c and no HSCT	[redacted]	[redacted]	[redacted]
Relapse within 6 months after allogeneic HSCT	[redacted]	[redacted]	[redacted]
Relapse after 6 months after allogeneic HSCT	[redacted]	[redacted]	[redacted]
Preselected Salvage Chemotherapy			
High intensity chemotherapy	[redacted]	[redacted]	[redacted]
Low intensity chemotherapy	[redacted]	[redacted]	[redacted]
Cytogenetic risk group, n (%)			
Favourable	4 (1.6)	1 (0.8)	5 (1.3)
Intermediate	182 (73.7)	89 (71.8)	271 (73.0)
Unfavourable	26 (10.5)	11 (8.9)	37 (10.0)
Other	35 (14.2)	23 (18.5)	58 (15.6)

Abbreviations: CR_c: composite complete remission; ECOG: Eastern Cooperative Oncology Group; FLT3: FMS-like tyrosine kinase-3; FLT3-ITD: FMS-like tyrosine kinase-3-internal tandem duplication; FLT3-TKD: FMS-like tyrosine kinase-3-tyrosine kinase domain; HSCT: haematopoietic stem cell transplant; SD: standard deviation

B.2.6 Clinical effectiveness results of the relevant trials

CHRYSALIS

In the full analysis set, the CR_c rate was 30%, median OS was 25 weeks, and median duration of response (DOR) was 17 weeks. Response was higher in patients who were FLT3 mutation positive, and in FLT3 mutation positive patients who received gilteritinib at a dose of ≥80mg/day. In the overall group, CR_c rate was 37%, median OS was 30 weeks, and median DOR was 20 weeks. In the group receiving ≥80mg/day gilteritinib, CR_c rate was 41%, median OS was 31 weeks, and DOR was 20 weeks. Of the 25 patients with resistance-associated D835 (TKD) point mutations who were treated with ≥80mg gilteritinib daily, 8 (32%) achieved a CR_i.[28]

A summary of outcomes reported from the CHRYSALIS trial is presented below.

Table 10 Key Efficacy Outcomes for CHRYSALIS

Outcomes	Full Analysis Set (n=249)	FLT3 mutation positive	
		All Patients (n=191)	Dosed at ≥80mg/day (n=169)
Response, %, (95% CI)			
CR	8 (5, 12)	9 (6, 15)	11 (6, 16)
CR_p	4 (2, 7)	5 (3, 9)	6 (3, 11)
CR_i	18 (14, 24)	22 (16, 29)	24 (18, 31)
CR_c	30 (25, 36)	37 (30, 44)	41 (33, 49)
Median DOR, weeks (95% CI)	17 (14, 29)	20 (14, 33)	20 (14, 33)
Median OS, weeks (95% CI)	25 (20, 30)	30 (23, 33)	31 (24, 59)

Abbreviations: CI: confidence interval; CR: complete remission; CR_c: composite complete remission; CR_i: complete remission with incomplete hematologic recovery; CR_p: complete remission with incomplete platelet recovery; DOR: duration of remission; FLT3: FMS-like tyrosine kinase-3; OS: overall survival

ADMIRAL

Gilteritinib demonstrated a statistically significant gains in efficacy in terms of the co-primary outcomes of OS and rate of CR/CR_h compared with salvage chemotherapy for the treatment of adults with relapsed or refractory FLT3 mutation positive AML.[29]

The co-primary and secondary outcomes of ADMIRAL are summarised in Table 11.

Median OS was longer in patients receiving gilteritinib (9.3 months; 95% CI: 7.7 to 10.7) compared with patients treated with salvage chemotherapy (5.6 months; 95% CI 4.7 to 7.3) (hazard ratio [HR]: 0.64 [95% CI: 0.49 to 0.83; p<0.001). After 1 year of follow-up, OS was more than doubled in patients receiving gilteritinib versus salvage chemotherapy (37.1% versus 16.7%). Six-month OS was also higher with gilteritinib (65.5%) versus salvage chemotherapy (48.9%). In a sensitivity analysis in which patients were censored at the time of HSCT, median OS was longer in the gilteritinib arm compared with the salvage chemotherapy arm (8.3 months versus 5.3 months; HR: 0.58 [95% CI: 0.43 to 0.76; p<0.0001). With respect to the second co-primary endpoint, the CR/CR_h rate was more than twice as high with gilteritinib versus salvage chemotherapy (34.0% versus 15.3%, p<0.001). The median gilteritinib dose in the trial was [redacted].

Gilteritinib prolonged median EFS compared with salvage chemotherapy (2.8 months [redacted] versus 0.7 months [redacted]), although the comparison was not statistically significant (p=0.0830). Given the hierarchical approach towards statistical testing, the remaining secondary endpoints are analysed descriptively. A subsequent analysis of modified EFS showed a significant benefit for gilteritinib (2.3 months versus 0.7 months; HR = 0.499 (95% CI 0.387-0.643, p<0.0001)). Modified event free survival was defined as failure to obtain a CR_c, with failures assigned as an event on randomisation, relapse or death form any cause including events at initiation of new anti-leukaemic treatments reported in long term follow-up. The CR rate was higher in the gilteritinib arm (21.1%) compared with the salvage chemotherapy arm (10.5%), as was the CR_c rate (54.3% versus 21.8%).

In the gilteritinib arm, the median duration of CR/CR_h, CR and LFS was 11.0 months, 14.8 months and 4.4 months, respectively. As the median duration of remission in the gilteritinib arm was not yet reached, the upper 95% CI for CR/CR_h and CR could not be reliably estimated. Likewise, because the majority of salvage chemotherapy patients finished the study by cycle 2 of treatment, the duration of exposure was short in the salvage chemotherapy arm, which led to limited follow-up of response and high censoring of the duration of CR and LFS. Therefore, the median duration of CR/CR_h, CR and LFS could not be reliably estimated in the salvage chemotherapy arm.

Among the 197 patients who were dependent on RBC and/or platelet transfusions at baseline, 68 patients became independent of RBC and platelet transfusions during any 56-day post-baseline period, corresponding to a transfusion conversion rate of 34.5% ([redacted]). Other transfusion-related outcomes are presented in Table 11. Finally, the haematopoietic stem cell transplantation (HSCT) rate was higher in the gilteritinib arm compared to salvage chemotherapy (25.5% vs 15.3%).

Seventy-eight gilteritinib patients required a dose increase from 120mg to 200mg, of whom [redacted] achieved CR/CR_h after the dose adjustment. Fifty-eight patients required a dose decrease from 120mg to 80mg, of whom [redacted] achieved CR/CR_h after the dose adjustment. The median daily gilteritinib dose in the trial was [redacted]

and the mean dose intensity was [redacted], calculated as the cumulative dose divided by duration of exposure.

Table 11 Key Efficacy Outcomes for ADMIRAL

Outcomes	Gilteritinib (N=247)	Salvage chemotherapy (N=124)
Overall Survival, median months (95% CI)	9.3 (7.7, 10.7)	5.6 (4.7, 7.3)
Patients achieving CR/CR_h (%)	84 (34.0%)	19 (15.3%)
Overall Survival Rate % (95% CI)		
6 months	65.5 (59.2, 71.1)	48.9 (39.3, 57.8)
12 months	37.1 (30.7, 43.6)	16.7 (9.9, 25.0)
24 months	19.0 (12.8, 26.0)	13.8 (7.5, 22.0)
Best response rate, n (%)		
CR	52 (21.1)	13 (10.5)
CR_p	19 (7.7)	0
CR_i	63 (25.5)	14 (11.3)
CR_c	134 (54.3)	27 (21.8)
CR_h	32 (13.0)	6 (4.8)
Duration of EFS, Median months (95% CI)	2.8 [redacted]	0.7 [redacted]
Duration of LFS, Median months (95% CI)	[redacted]	[redacted]
Duration of remission, Median months (95% CI)		
CR	[redacted]	[redacted]
CR_c	[redacted]	[redacted]
CR_h	[redacted]	[redacted]
CR/CR_h	[redacted]	[redacted]
Time to remission, Median months (95% CI)		
CR	[redacted]	[redacted]
CR_c	[redacted]	[redacted]
CR_h	[redacted]	[redacted]
CR/CR_h	[redacted]	[redacted]
Transplantation Rate, n (%)	63 (25.5)	19 (15.3)
Transfusion Conversion Rate*, n/N (%)	68/197(34.5%)	Not reported
Transfusion Maintenance Rate, n/N (%)**	[redacted]	Not reported
Change from baseline BFI Global Fatigue Score		
Cycle 1 day 8, Mean (SD)	[redacted]	[redacted]
Cycle 2 day 1, Mean (SD)	[redacted]	[redacted]

Abbreviations: BFI: brief fatigue inventory; CI: confidence interval; CR: complete remission; CR_c: composite complete remission; CR_i: complete remission with incomplete haematologic recovery; CR_h: complete remission with partial haematological recovery; CR_p: complete remission with incomplete platelet recovery; EFS: event-free survival; LFS: leukaemia-free survival; NE: not estimable; SD: standard deviation

***Transfusion Conversion Rate:** The number of subjects who were transfusion dependent at baseline period but become transfusion independent at post-baseline period divided by the total number of subjects who are transfusion dependent at baseline period.

****Transfusion Maintenance Rate:** The number of subjects who were transfusion independent at baseline period and still maintain transfusion independent at post-baseline period divided by the total number of subjects who are transfusion independent at baseline period.

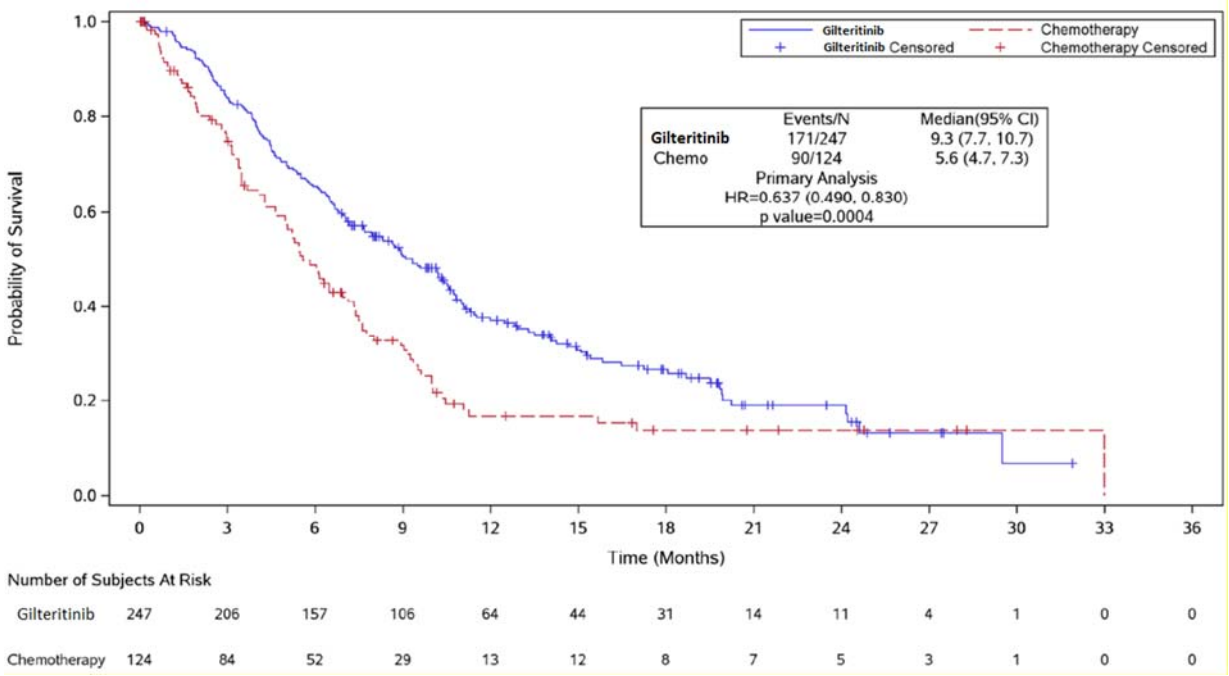
Table 12 Transfusion Rates in Patients Treated with Gilteritinib in ADMIRAL

Patient Group	Transfusion Conversion Rate*	Transfusion Maintenance Rate**
All patients	68 of 197 (34.5%)	[redacted]
Patients with post-baseline evaluable transfusion status	[redacted]	[redacted]

***Transfusion Conversion Rate:** The number of subjects who were transfusion dependent at baseline period but become transfusion independent at post-baseline period divided by the total number of subjects who are transfusion dependent at baseline period.

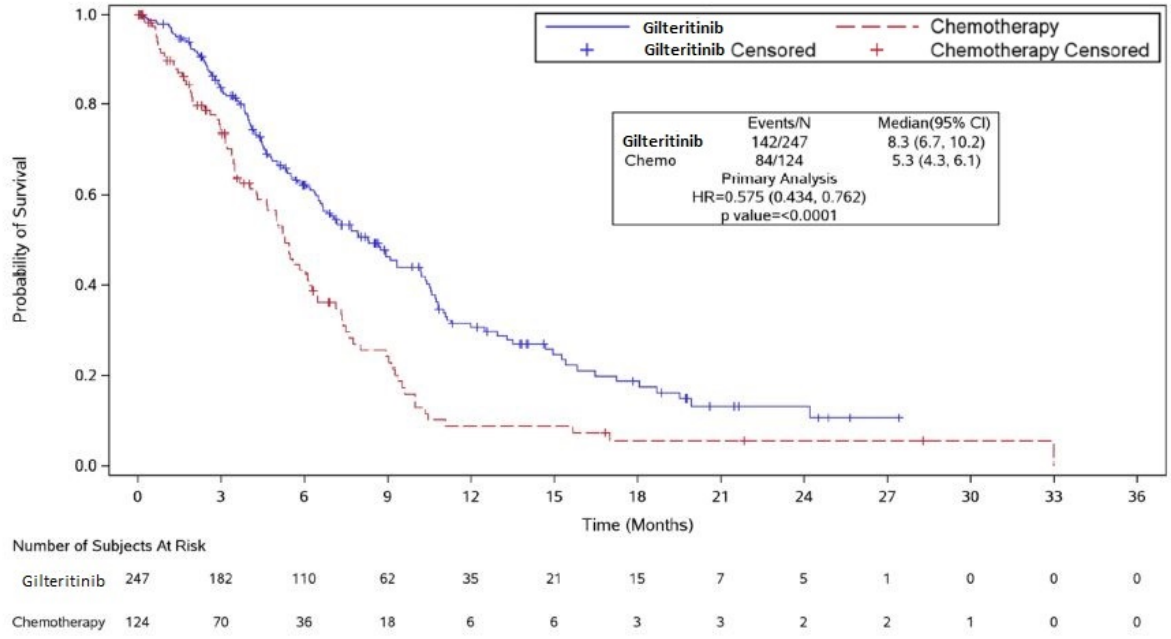
****Transfusion Maintenance Rate:** The number of subjects who were transfusion independent at baseline period and still maintain transfusion independent at post-baseline period divided by the total number of subjects who are transfusion independent at baseline period.

Figure 4 Kaplan-Meier Plot of Overall Survival by Treatment Arm



Abbreviations: CI: confidence interval; HR: hazard ratio; OS: overall survival

Figure 5 Kaplan-Meier Plot of OS Censored at HSCT by Treatment Arm



Abbreviations: CI: confidence interval; HR: hazard ratio; NE: not estimable

Figure 6 Kaplan-Meier Plot of EFS by Treatment Arm

[redacted]

Abbreviations: CI: confidence interval; HR: hazard ratio; NE: not estimable

There are two unusual characteristics of the above KM curve. Firstly, the sharp drop on Day 1 for both treatment curves can be explained as follows: in the ADMIRAL trial,

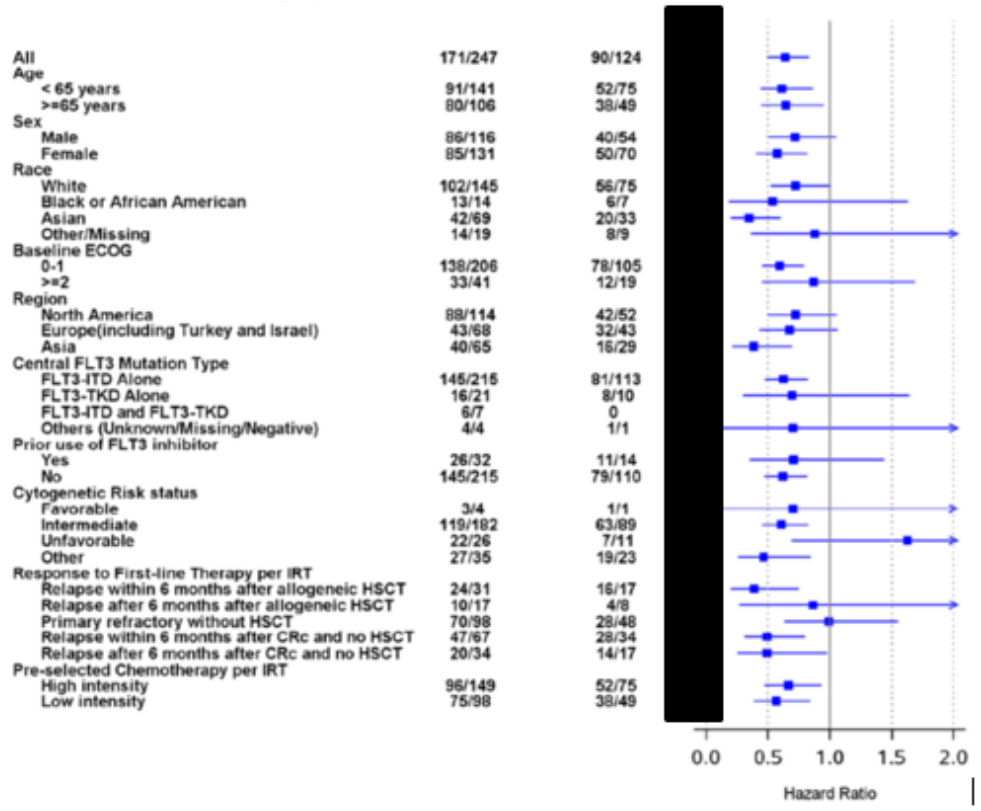
EFS was defined as the time from the date of randomisation until the date of documented relapse, treatment failure or death. Any event recorded in the first month of treatment should be recorded as a Day 0 event to account for variable patient follow up times. Secondly, the short duration of the salvage chemotherapy treatments reflects their fixed cycles of treatment and leads to a short curve.

Median EFS was 2.8 months ([redacted]) in the gilteritinib arm and 0.7 months ([redacted]) in the salvage chemotherapy arm. Although there was a trend towards increased EFS duration in the gilteritinib arm, the EFS endpoint did not meet the pre-specified criteria for statistical significance ([redacted]). A robust comparison of EFS between the gilteritinib and salvage chemotherapy arms could not be conducted due to the large number of censoring events in the first month in the salvage chemotherapy arm.

B.2.7 Subgroup analysis

Pre-planned subgroup analyses were performed on the OS endpoint in ADMIRAL, as summarised in Figure 7 below. A consistent benefit for gilteritinib was observed, although the numbers are small in some patient groups (note that the upper 95% CI crossed 1.0 for some groups). OS was improved with gilteritinib versus salvage chemotherapy in patients who had relapsed within 6 months following HSCT. A survival benefit for gilteritinib was also observed in other subgroups according to response to first-line therapy. When OS was evaluated according to racial group, the most pronounced survival advantage was seen in Asian patients ([redacted]). There was also a survival advantage for gilteritinib among White patients ([redacted]). Too few Black or African American patients were included for meaningful analysis of this subgroup. There is no clinical evidence about the biological or genetic difference which could inhibit the generalisability of these results on the population of England and Wales.[29] The population of ADMIRAL has been validated by several UK clinicians that this reflects their caseload.

Figure 7: Forest Plot for Subgroup Analysis of OS in ITT Population



B.2.8 Meta-analysis

CHRYSALIS was a dose escalation Phase I/II RCT and only one relevant Phase III RCT (ADMIRAL) was identified. Thus, a meta-analysis was not required or performed.

B.2.9 Indirect and mixed treatment comparisons

ADMIRAL directly compares gilteritinib with the relevant comparators defined in the scoping document. Therefore, an indirect or mixed treatment comparison was not required or performed.

B.2.10 Adverse reactions

CHRYSALIS

Gilteritinib was generally well tolerated. The most common treatment-emergent AEs reported in CHRYSALIS included diarrhoea (37%), anaemia (34%), fatigue (33%), elevated AST (26%), and elevated ALT (19%). The most common Grade 3/4 AEs were

febrile neutropenia (39%), anaemia (24%), thrombocytopenia (13%), sepsis (11%), and pneumonia (11%).[28]

A detailed breakdown of treatment-emergent AEs is presented in Appendix F.

ADMIRAL

Gilteritinib was generally well tolerated. Almost all patients experienced at least one TEAE. When adjusted by patient-year of exposure, the incidence of all types of TEAE was lower in the gilteritinib arm than the salvage chemotherapy arm. Drug-related TEAEs occurred at a rate of 16.6 per patient-year in the gilteritinib group compared with 47.2 per patient-year in the salvage chemotherapy group; corresponding values for drug-related TEAEs leading to death were 0.1 and 0.7 per patient-year. A total of [redacted] of patients receiving gilteritinib died ([redacted] deaths per patient-year) compared with [redacted] of patients receiving salvage chemotherapy ([redacted] deaths per patient-year).[29]

Table 13 Summary of Key Safety Events from ADMIRAL

	Gilteritinib		Salvage Chemotherapy	
	n (%) patients (N=246)	Number of events/PY (PY=121.7)	n (%) patients (N=109)	Number of events/PY (PY=11.9)
TEAE	[redacted]	[redacted]	[redacted]	[redacted]
Drug-related TEAE	[redacted]	[redacted]	[redacted]	[redacted]
Serious TEAE	[redacted]	[redacted]	[redacted]	[redacted]
Drug-related serious TEAE	[redacted]	[redacted]	[redacted]	[redacted]
TEAE leading to death	[redacted]	[redacted]	[redacted]	[redacted]
Drug-related TEAE leading to death	[redacted]	[redacted]	[redacted]	[redacted]
TEAE leading to withdrawal of treatment	[redacted]	[redacted]	[redacted]	[redacted]
Drug-related TEAE leading to withdrawal of treatment	[redacted]	[redacted]	[redacted]	[redacted]
NCI-CTCAE Grade 3 or higher TEAE	[redacted]	[redacted]	[redacted]	[redacted]
Drug-related Grade 3 or higher TEAE	[redacted]	[redacted]	[redacted]	[redacted]
Death	[redacted]	[redacted]	[redacted]	[redacted]

Abbreviations: NCI-CTCAE: National Cancer Institute-Common terminology criteria for adverse events; PY: Patient-year; TEAE: Treatment emergent adverse events

The most frequently reported TEAEs in the gilteritinib arm were anaemia (47.2%), febrile neutropenia (46.7%) and pyrexia (42.7%). Increased levels of alanine aminotransferase and aspartate aminotransferase were observed in 41.9% and 40.2% patients, respectively. Serious TEAEs occurred in [redacted] of patients receiving

gilteritinib ([redacted] events per patient-year) and [redacted] of patients receiving salvage chemotherapy ([redacted] events per patient-year). The most frequent serious TEAEs (≥5.0%) in the gilteritinib arm were febrile neutropenia ([redacted]), acute myeloid leukaemia ([redacted]), pyrexia ([redacted]), pneumonia ([redacted]), sepsis ([redacted]), acute kidney injury ([redacted]), lung infection ([redacted]) and ALT increased ([redacted]). The most frequent serious TEAEs (≥5.0%) in the salvage chemotherapy arm were febrile neutropenia ([redacted]) and sepsis ([redacted]). The most frequent drug-related serious TEAEs in the gilteritinib arm were febrile neutropenia ([redacted]), ALT increased ([redacted]), AST increased ([redacted]), pneumonia ([redacted]) and anaemia ([redacted]). The most frequent TEAEs leading to withdrawal of treatment in the gilteritinib arm were AML ([redacted]), lung infection ([redacted]) and AST increased ([redacted]).[29]

A more detailed breakdown of TEAEs and SAEs is provided in Appendix F.

B.2.11 Ongoing studies

The clinical evidence supporting the use of gilteritinib to treat adults with relapsed or refractory FLT3 mutation positive AML is based on the phase I/II CHRYSALIS trial and the phase III ADMIRAL randomised controlled trial. Several other studies are currently ongoing that will provide evidence of the long-term safety of gilteritinib and comparisons of gilteritinib with different therapies and its role in earlier lines of therapy in FLT3 mutation positive AML.

Table 14 Gilteritinib Clinical Development Program in FLT3 Mutation Positive AML

Study	Country	Study Number	Study Design	Status
Ongoing Monotherapy Studies				
Relapsed or Refractory				
Phase III Mono Asia	Asia, Russia	2215-CL-0303	Phase III study of gilteritinib for first-line treatment of patients with R/R AML	Study results expected: [redacted]
Maintenance Therapy Post-1L Induction Chemotherapy				
Phase III Mono Maintenance Post-HSCT (GOSSAMER)	North America, EU, Asia, Central and South America, rest of the world	2215-CL-0302	Phase III, double-blinded, placebo-controlled study of gilteritinib maintenance therapy for patients with AML and FLT3-ITD mutations in their first CR following 1L induction/consolidation therapy	Study results expected: [redacted]
Maintenance Therapy Post-allogeneic HSCT				

Study	Country	Study Number	Study Design	Status
Phase III Mono Maintenance (MORPHO)	North America, EU, Asia	2215-CL-0304	Phase III, double-blind, placebo-controlled study of gilteritinib maintenance therapy for patients with AML and FLT3-ITD mutations in their first CR following allogeneic HSCT	Study results expected: [redacted]
Ongoing Combination Therapy Studies				
Newly Diagnosed, Intensive Chemotherapy-eligible				
Phase III (HOVON)	US	NCT03836209 PrE0905	Phase III study of gilteritinib vs. midostaurin in combination with standard chemotherapy of daunorubicin and cytarabine during induction and high-dose cytarabine during consolidation in patients with FLT3 acute myeloid leukaemia (AML)	Planned
Phase I Combo	US	2215-CL-0103	Phase I study of gilteritinib in combination with induction and consolidation chemotherapy in patients with newly diagnosed AML	Ongoing
Phase I Combo JP	Japan	2215-CL-0104	Phase I study of gilteritinib in combination with induction and consolidation chemotherapy in Japanese patients with newly diagnosed AML	Ongoing
Newly Diagnosed, Intensive Chemotherapy-ineligible				
Phase IIb/III Mono and Combo (LACEWING)	North America, Europe, Asia	2215-CL-0201	Phase IIb/III, 3-arm study of gilteritinib, the combination of gilteritinib plus azacitidine, or azacitidine alone in HIC-ineligible newly diagnosed patients who have AML with FLT3 mutations	Ongoing; study results expected: [redacted]

Abbreviations: 1L = first-line; AML = acute myeloid leukaemia; CR = complete remission; FLT3 = FMS-like tyrosine kinase 3; HIC = high-intensity chemotherapy; HSCT = haematopoietic stem cell transplant; NA = Not available

B.2.12 Innovation

Gilteritinib is the first and only oral monotherapy shown to deliver over 9 months median OS (vs 5.6 months with salvage chemotherapy) in patients with relapsed or refractory FLT3 mutation positive AML. No standard of care has been established for the treatment of relapsed or refractory FLT3 mutation positive AML, and gilteritinib thus has the potential to become first-line therapy in in FLT3 mutation positive relapsed or refractory AML. The innovative mechanism of gilteritinib, targeting both FLT3-ITD and FLT3-TKD, means that patients are less likely to acquire FLT3 resistance

mutations (e.g., D835) than earlier generation multi-kinase inhibitors such as midostaurin.[29]

B.2.13 Interpretation of clinical effectiveness and safety evidence

The prognosis for patients with AML is poor. Mutations of the FLT3 gene occur in approximately 30% of AML cases. The two most common types are FLT3-ITD which occurs in 15-30% of cases and FLT3-TKD which occurs 5-10% of cases.[17,18] The prognosis of AML in patients with FLT3 mutation is poor compared with AML patients without this mutation. The estimated median OS for patients with FLT3 mutations is 15.2 to 15.5 months compared to 19.3 to 28.6 months for patients with wild-type FLT3.[16] In addition, prognosis worsens further for those with relapsed or refractory disease.[8] There are currently no effective treatment options for these patients and as such there is no established standard of care.

Gilteritinib is a targeted therapy for the FLT3 mutation. Regulatory approval is being sought for the treatment of patients with relapsed or refractory FLT3 mutation positive AML as a once-daily oral monotherapy.

In the phase I/II CHRYSALIS study, gilteritinib at doses between 80mg/day and 300mg/day were generally well tolerated and achieved a CR_c rate of 41%, with median OS of 31 weeks and median DOR of 20 weeks.[28] In the phase III ADMIRAL trial, gilteritinib doubled the 1-year OS rate versus salvage chemotherapy (37.1% versus 16.7%), and was associated with a significantly longer median duration of OS (9.3 months versus 5.6 months; $p<0.001$). A total of 34.0% of patients in gilteritinib arm achieved CR/CR_h compared with 15.3% patients in the salvage chemotherapy arm ($p<0.001$). Gilteritinib prolonged EFS to 2.8 months versus only 0.7 months with salvage chemotherapy, this result was not statistically significant. Because of the hierarchical statistical testing employed, analysis of other secondary outcomes was descriptive only. Finally, almost double the number of patients receiving gilteritinib went on to receive HSCT versus salvage chemotherapy (25.5% vs 15.3%).[29]

Gilteritinib was generally well tolerated, with a lower incidence of TEAEs than salvage chemotherapy when adjusted by patient-year, with a discontinuation rate of 10%. There were no significant differences observed in the efficacy or safety of gilteritinib between patients younger or older than 65 years.

Thus, gilteritinib is a clinically effective treatment option which is generally well tolerated, and represents an option for patients who currently have a poor prognosis and no effective treatments.

Table 15 End-of-life Criteria

Criterion	Data available
The treatment is indicated for patients with a short life expectancy, normally less than 24 months	Median survival is reported to be two months or less in patients receiving supportive care alone [3,4] and the pivotal ADMIRAL phase III trial showed the median overall survival in the comparator salvage chemotherapy arm was 5.6 months.
There is sufficient evidence to indicate that the treatment offers an extension to life, normally of at least an additional 3 months, compared with current NHS treatment	The pivotal ADMIRAL phase III trial showed median overall survival in the gilteritinib arm was 9.3 months vs 5.6 months in the salvage chemotherapy arm, a gain of 3.7 months.

B.3 Cost effectiveness

This section outlines the cost-effectiveness analysis (CEA) model of gilteritinib in adult patients with relapsed or refractory FLT3 mutation positive AML. This section includes detailed descriptions on methods and model inputs; including model structure, model assumptions, extrapolation approach, inputs, scenario and sensitivity analyses, and validation methods. The model and cost inputs are based on NHS England and the Personal Social Service (PSS) perspective. Table 16 provides a detailed summary of the content included in this section.

Table 16 Contents of the Economic Dossier

Topic	Content included
Published cost-effectiveness studies (**B3.1**)	Overview of published cost-effectiveness studies for gilteritinib are given in Appendix G; published cost effectiveness for comparator treatments are given in Appendix I
Economic analysis (**B3.2**)	• Patient population, time horizon and discount rate are in line with the NICE reference case and reflect an NHS and PSS perspective

Topic	Content included
	• The intervention technology and its comparators; salvage chemotherapies (azacitidine, FLAG-IDA, MEC, LDAC), BSC, and a weighted comparator, are described • The model structure of the CEA is discussed and justified
Clinical parameters and variables (**B3.3**)	Efficacy inputs: • Clinical inputs for gilteritinib and comparators derived from ADMIRAL phase III RCT • Extrapolation method considered for efficacy measures • Rationale for the selected approaches and assumptions
Measurement and valuation of health effects (**B3.4**)	Utility inputs: • Trial-based utility measures using the UK EQ-5D tariff • Rationale for the selected approaches and assumptions
Cost and healthcare resource use identification, measurement and valuation (**B3.5**)	Cost inputs: • Resource use derived from ADMIRAL phase III RCT and, where necessary, the literature • Unit prices from English data sources, particularly NHS reference prices and PSSRU data
Summary of base-case analysis inputs and assumptions (**B3.6**)	Details of the base-case inputs and assumptions
Base-case results (**B3.7**)	Base-case results in terms of cost per quality-adjusted life year (QALY) gained
Sensitivity analyses (**B3.8**)	Results of the deterministic and probabilistic sensitivity analyses of the base case and consideration of alternative scenarios
Subgroup analyses (**B3.9**)	Cost-effectiveness results for pre-specified subgroups
Validation (**B3.10**)	Internal and external validation of cost-effectiveness model
Interpretation and conclusions of economic evidence (**B3.11**)	Interpretation of the economic evidence and conclusions from the cost-effectiveness analysis

B.3.1 Published cost-effectiveness studies

No published cost-effectiveness analyses are available for gilteritinib.

B.3.2 Economic analysis

A *de novo* model has been constructed to assess the cost-effectiveness of gilteritinib relative to common salvage chemotherapies alternatives and best supportive care (BSC) for the management of adult patients with relapsed or refractory FLT3 mutation positive AML. The analysis is conducted in terms of cost per quality-adjusted life year (QALY) gained from the perspective of NHS England and the Personal Social Service (PSS). Only direct health care costs are considered in the base case.

A lifetime horizon is considered to comprehensively capture the expected costs and health outcomes of patients over their remaining lifetime from the initiation of the treatment. In the base-case, both costs and effectiveness are discounted at 3.5%

annually. During the modelled time horizon, costs and effectiveness are estimated for each treatment arm included in the model. The following cost components are considered: drug and drug administration costs, adverse event (AE) costs, subsequent hematopoietic stem cell transplantation (HSCT) costs, medical costs associated with health states, post-progression treatment costs, and FLT3 mutation testing and terminal care costs. Effectiveness measures include life years (LYs) and quality-adjusted life years (QALYs). The incremental cost-effectiveness ratios (ICERs) of gilteritinib vs. each comparator are evaluated in terms of the incremental cost per QALY gained and the incremental cost per LY gained.

Table 17 Key Features of the Cost-effectiveness Model

Features	Description
Target population	Adult patients with R/R FLT3 mutation positive AML
Perspective	NHS England and PSS perspective
Time horizon	Lifetime
Model structure and cycle length	Decision-tree structure followed by partitioned survival models with monthly cycles
Intervention and Comparators	Gilteritinib (intervention)
	Azacitidine
	FLAG-IDA: Fludarabine + cytarabine + granulocyte colony stimulating factor + idarubicin
	MEC: Mitoxantrone + etoposide + cytarabine
	LDAC: Low-dose cytarabine
	BSC: best supportive care (referring to supportive care only without any active treatments)
	Weighted comparator (based on ADMIRAL trial, base case comparator)
Model components	• Treatment efficacy • Health-state utilities • Drug and drug administration costs • AEs associated with initial treatments and corresponding costs and disutilities • Subsequent HSCT following initial treatments and corresponding costs and disutilities • Medical costs associated with health states, FLT3 mutation testing and terminal care • Post-progression treatment costs
Model outputs	• Total and incremental effectiveness o Life years (LYs) o Quality-adjusted life years (QALYs) • Total and incremental costs o Treatment costs (drug acquisition and administration costs) o Treatment-associated adverse event (AE) costs o HSCT costs o Medical costs (including health states, FLT3 testing and terminal care costs) o Post-progression treatment costs • ICERs o Incremental cost per LY gained o Incremental cost per QALY gained

Abbreviations: AE, adverse event; AML, acute myeloid leukaemia; BSC, best supportive care; FLT3, FMS-like tyrosine kinase 3; HSCT, haematopoietic stem cell transplantation; ICER, incremental cost-effectiveness ratio; LY, life year; NHS, National Health Service; PSS, Personal Social Service; QALY, quality-adjusted life year; R/R, relapsed or refractory; UK, United Kingdom

B.3.2.1 Patient population

The patient population considered in the economic evaluation consists of adult patients with relapsed or refractory FLT3 mutation positive AML. This patient population is consistent with the anticipated indication for gilteritinib and corresponds to the patient population evaluated in the pivotal phase III ADMIRAL trial.[29] Leading UK clinicians have reviewed the ADMIRAL patient population and confirmed that it is reflective of their typical caseload.

B.3.2.2 Intervention technology and comparators

Gilteritinib is a tyrosine kinase-3 (FLT3) and AXL inhibitor. It is the first and only oral monotherapy shown to deliver over 9 months median OS (vs 5.6 months with salvage chemotherapy) in patients with relapsed or refractory FLT3 mutation positive AML. In the ADMIRAL phase III open-label trial, gilteritinib was compared to the following treatment alternatives, which have been confirmed by leading UK clinicians as reflective of UK clinical practice:

- Low dose cytarabine (20mg twice-daily SC or IV injections for 10 days)
- Azacitidine (75mg/m^2 daily SC or IV injections for 7 days)
- MEC (mitoxantrone 8mg/m^2 per day, etoposide 100mg/m^2 per day, cytarabine 1,000mg/m^2 per day, all administered via IV injection for 5 days on days 1 through 5)
- FLAG-Ida (fludarabine 30mg/m^2 per day and cytarabine 2,000mg/m^2 per day, both administered via IV injection for 5 days on days 2 through 6; G-CSF 300µg/m^2 per day administered via SC or IV injection for 5 days on days 1 through 5; idarubicin 10mg/m^2 per day administered via IV injection for 3 days on days 2 through 4)
- Best supportive care (BSC). Based on findings from a global chart review study,[31] BSC is one of the most commonly used options for the treatment of relapsed or refractory FLT3 mutation positive AML. It is recommended by the European Leukemia Net (ELN)[6] and European Society for Medical Oncology (ESMO) guidelines[5] as a potential treatment option for relapsed or refractory AML patients who are not fit for intensive therapy or HSCT. BSC here refers to supportive medication or procedures that do not include any active anti-leukemic treatments. Typical management strategies may include hydroxyurea, blood transfusions, growth factors, and anti-infective treatments

These comparators were included in the model on the basis of their relevance in the treatment of relapsed or refractory FLT3 mutation positive AML and because they were the active comparators in the ADMIRAL phase III trial therefore provide head-to-head evidence relative to gilteritinib.

As the sample size of ADMIRAL was powered to detect a difference of clinical benefits between the gilteritinib arm and the overall salvage chemotherapy arm, the efficacy inputs for the individual comparators in the model are based on the pooled efficacy data reported for the overall salvage chemotherapy arm in the trial. The sample size for each individual salvage chemotherapy arm is limited (ranging from 17 to 42 patients). Drug and administration costs and hospitalisation costs differ between different salvage chemotherapy arms.

B.3.2.3 Model structure

B.3.2.3.1 Justification of the selected model structure

The model structure is comprised of a decision-tree component to stratify patients based on their transplantation status, followed by two separate three-state partitioned survival models to predict the long-term survival status of the target patient population conditional on their transplantation status. The decision tree is implemented to account for the lag between the beginning of treatment and any eventual HSCT. A one-month cycle length is considered. The model has been developed in Microsoft Excel.

Partitioned survival analysis is the most commonly utilised decision modelling approach for appraisals of advanced and metastatic cancer interventions and is well-accepted by health technology assessment (HTA) bodies.[32] The partitioned survival model structure eliminates the need to generate assumptions for the transition of patients between health states and allows for the direct use of ADMIRAL Kaplan-Meier (K-M) curves to estimate the proportion of patients in each clinical state.

The strengths of partitioned survival analysis are derived from the direct correspondence between EFS and OS, time-to-event endpoints reported for the ADMIRAL trial. This correspondence makes the approach intuitive and transparent. As observed mortality during ADMIRAL[29] reached more than 90%, the survival data can be considered mature in terms of OS and EFS. [redacted] of the 247 patients randomised to gilteritinib ([redacted]) remained on treatment at the data cutoff date. This limits any survival curve extrapolation bias and facilitates a partitioned survival structure. Other modelling alternatives such as Markov or semi-Markov (state-transition) structures require further assumptions to estimate transition probabilities,

cannot incorporate time-varying transition probabilities for all considered health states and have stricter individual-level data requirements to use treatment arms not directly evaluated in ADMIRAL.

Receiving a hematopoietic stem cell transplant (HSCT) is a key intermediate clinical event ("landmark") that is prognostic of patient outcomes. For relapsed or refractory AML, HSCT is a key clinical event because it is the only established intervention that has plausible curative potential for the target population based on the current treatment landscape. Several studies have investigated the relationship between HSCT status and the long-term survival outcomes among relapsed or refractory AML patients.[33–35] These studies have consistently found that HSCT appears to be an important prognostic determinant of OS, highlighting the importance of stratifying by transplant status. In ADMIRAL, OS for patients who proceeded to HSCT ranged from 6.5 to 39 months in the relapsed or refractory setting compared to 1.5 to 11.9 months in patients who did not receive HSCT. This is consistent with other studies that have found a favourable impact of HSCT on survival outcomes.[36,37] Patients with FLT3 mutation positive AML receiving HSCT showed significantly improved OS and relapse-free survival compared to FLT3 mutation positive patients managed with chemotherapy only.

In addition, a high percentage of AML patients receive HSCT following complete remission which is associated with distinct outcomes, costs and utilities. A typical partitioned survival model with only three states cannot provide the granularity to separately model the clinical benefit associated with the HSCT treatment. Therefore, a model structure that introduces a structural link between HSCT and the overall treatment benefit is required to appropriately reflect the treatment pathway and associated outcomes and costs for the target population. Such an approach has been used in a number of previous HTA submissions for other hematological disorders.[38–40]

A three-state model (alive and event-free, alive and post-event, and death) was selected to describe the clinical pathway for each subgroup of patients as it is the most commonly-used model structure for advanced cancer therapies. The three-state structure allows us to directly apply the primary clinical endpoints evaluated in the

ADMIRAL clinical trial (EFS and OS) to the economic evaluation and is reflective of the typical clinical pathway of AML patients. Patients under event-free survival are expected to have better quality of life and utilise less healthcare resources compared to those who experience progression, relapse or treatment failure. By separating out these patients, distinct utilities and medical costs can be assigned to each specific health state. In addition, less data (efficacy, utilities, costs, etc.) and fewer assumptions are required to populate the model than for structures with a greater number of states, e.g. four-state models accounting for response status. For inputs that have been directly sourced from the ADMIRAL trial, these assumptions allow us to directly apply clinical trial data to the economic evaluation.

To account for the fact that the patient must first survive during a certain "waiting time" to receive a stem cell transplant, a decision-tree component was added to the model. With the decision-tree modelling approach, all patients begin in the "alive and event-free without HSCT state" following treatment initiation. A proportion of patients, corresponding to the HSCT transplant rate in ADMIRAL, transition to the "alive and event-free with HSCT state" after the average time to HSCT observed in ADMIRAL has elapsed. The addition of the decision-tree should align the model structure more closely to the actual clinical pathway and capture the benefit of treatments before stem cell transplantation.

The current model structure incorporates feedback from the NICE PRIMA scientific review service to separately model efficacy based on the HSCT status and is consistent with prior HTA submissions in advanced or relapsed or refractory haematological cancers.[38–40] There has also been previous HTA submissions in 'newly diagnosed' AML (midostaurin and azacitidine) but these submissions are not directly applicable to relapsed or refractory FLT3 mutation positive indication due to the different survival expectation for newly diagnosed vs. relapsed or refractory FLT3 mutation positive patients.[41,42] However, critiques related to survival extrapolation approaches, and to the modelling of subsequent treatments in prior AML submissions, have been incorporated during the development of the current model.

The structure of the model, as well as the inputs and assumptions (per B.3.2.3.2 below), have been guided by lessons from prior submissions, as discussed during the Decision Problem meeting. These are outlined in Table 18 below.

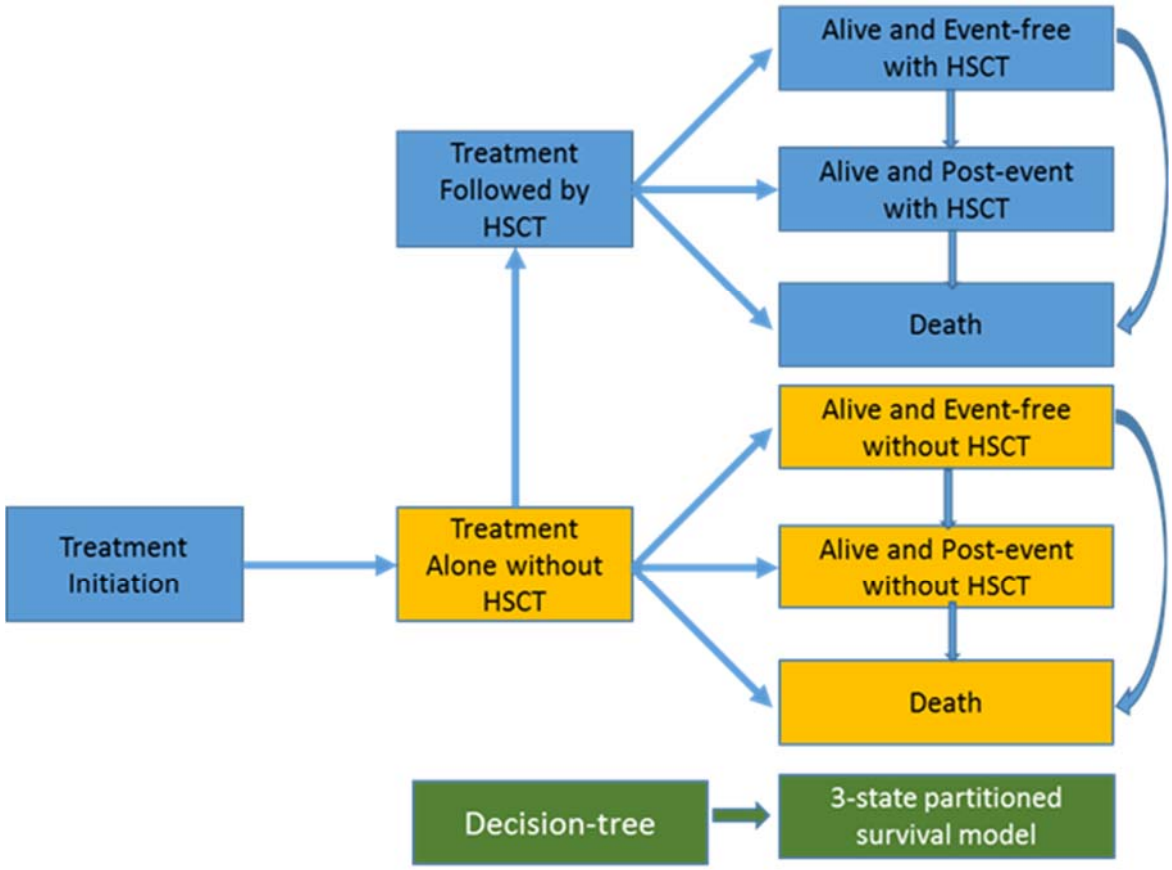
Table 18 Lessons from Previous Appraisals (Decision Problem Meeting)

Previous appraisal	Issue	How addressed in previous appraisals and committee's preferred assumptions	Astellas approach and rationale
Midostaurin (TA523)	Model structure vs clinical practice	ERG attempted correction	Structure based on NICE PRIMA advice and UK clinical experts
	Utility in relapse	Took mid-point of literature and company submission	Based on EQ5D collection from pivotal trial
	Post transplant costs	Costs excluded	Included in base case with toggle to exclude
	Relapse state costs	Lower cost applied for limited time before entering (new) cured health state	HealthCare Resource Use (HCRU) as observed in pivotal trial applied
	Post HSCT Standard Mortality Ratio (SMR)	Uncertainty remained, but SMR=2 applied	SMR=2 applied
	Cure point	Committee prefer to use the latest point at which the data showed a levelling out effect	Cure point aligned with flattening of KM curves from a range of publications (describing comparable population survival)

B.3.2.3.2 Description of model transitions and health states

In the model, all patients begin in the "treatment alone without HSCT" state following treatment initiation, and a proportion of patients transition to the "with HSCT" states after the average time to HSCT based on the ADMIRAL trial. Patients with and without subsequent HSCT are modelled separately from this point. A model schematic is presented in Figure 8.

Figure 8 Decision-tree and Partitioned Survival Model Structure



In partitioned survival analysis, within each group (HSCT or “no HSCT”), the EFS and OS curves do not represent mutually exclusive state membership estimates. The OS curve includes all alive individuals, regardless of these being in the “event-free” or “post-event” states. Since there are multiple health states in which alive patients can reside, state membership is derived from the areas under the survival curves as follows. Evidently, the proportion of “alive and event-free” individuals is directly provided by the value of the EFS curve. The proportion of subjects who are dead is 1 minus the value of the OS curve at a given time point. The proportion of “alive and post-event” individuals is, by necessity, the difference between the OS and EFS curves. Such proportions represent the marginal probabilities for each clinical state. In the model, the EFS and OS curves have been fitted to the K-M data following a parametric survival modelling approach (see section B3.3).

Patients who did not receive HSCT:

Patients who did not receive subsequent HSCT transition between the following three states: EFS without HSCT, post-event without HSCT and death. Treatment-specific efficacy inputs were used to inform transitions between the states. We have:

- **EFS without HSCT:** Patients are defined to be in this state if they have not received subsequent HSCT and have not yet experienced an event (i.e. relapse, progression or treatment failure), or death. Patients in this health state follow the EFS curve estimated among relapsed or refractory AML patients who have not received HSCT
- **Post-event without HSCT**: Patients are defined to be in this state if they have not received subsequent HSCT and have relapsed, progressed or experienced treatment failure. The proportion of patients in this health state is equal to the difference between the proportion of patients who are alive (the OS curve) without HSCT and the proportion of patients who are event-free and alive (the EFS curve) without HSCT
- **Death**: The absorbing state. The proportion of patients in this health state is estimated as 1 minus the OS curve for patients who did not receive HSCT

Patients who did receive HSCT

After the average time to HSCT observed in ADMIRAL[29] has elapsed, patients receiving subsequent HSCT transition between the following three states: EFS with HSCT, alive and post-event with HSCT, and death. The same efficacy inputs were considered for all patients regardless of the initial treatment.

- **EFS with HSCT:** Patients are defined to be in this state if they have received subsequent HSCT and have not yet experienced an event (i.e., relapse, progression or treatment failure), or death. Patients in this health state follow the EFS curve estimated among relapsed or refractory AML patients receiving HSCT
- **Post-event with HSCT**: Patients are defined to be in this state if they have received subsequent HSCT and have relapsed, progressed, or experienced treatment failure. The proportion of patients in this health state is set equal to the difference between the proportion of patients who are alive (the OS curve) with HSCT and the proportion of patients who are event-free and alive (the EFS curve) with HSCT

- **Death**: The absorbing state. The proportion of patients in this health state is estimated as 1 minus the OS curve for patients who received HSCT

Overall population

Finally, the curves for each subgroup of patients (HSCT or "no HSCT") are weighted by transplant (or non-transplant) rates to give marginal state membership probabilities for the overall population. A monthly model cycle has been applied to estimate the proportion of patients in each heath state over time. During each cycle, patients "transition" between the defined health states, with death being the absorbing state. "Alive and event-free" patients can either transition to the "alive and post-event" state, transition to the "death" state or remain in the same health state. "Alive and post-event" patients can either transition to the "death" state or remain in the same health state. The transition back to complete remission has not been modelled but the model does consider outcomes for "cured" longer-term survivors.

B.3.3 Clinical parameters and variables (efficacy inputs)

Efficacy inputs for the model include OS and EFS by HSCT status (i.e. OS without HSCT, OS with HSCT, EFS without HSCT, and EFS with HSCT), long-term survival, and OS benefits associated with post-HSCT maintenance therapy.

The efficacy inputs for OS and EFS without HSCT are assumed to be different across the treatment arms. Based on the ADMIRAL trial data, gilteritinib has been shown to significantly improve OS among patients who did not receive subsequent HSCTF.[29] Estimates of survival for the gilteritinib and pooled salvage chemotherapy arms were based on individual patient data (IPD) from the intent-to-treat (ITT) data set without HSCT from the ADMIRAL trial. As there were small sample sizes associated with the individual salvage chemotherapy treatments (the number of patients treated with azacitidine, FLAG-IDA, MEC, and LDAC was 32, 42, 33, and 17, respectively), these comparators were pooled to estimate an overall salvage chemotherapy survival curve. This is similar to the approach taken in the submission of midostaurin, NICE TA399.[41]

BSC efficacy inputs are based on data from relevant clinical trial publications. The efficacy input sources and the methodology used to predict OS and EFS without HSCT inputs are described in detail in sections **B.3.3.2** and **B.3.3.3**, respectively.

The ADMIRAL trial was not powered to demonstrate differences in post-HSCT survival based on prior treatment. Hence, we assume conservatively that patients who received HSCT across all treatment arms share the same efficacy inputs. ADMIRAL was also not designed to reliably inform the long-term survival in patients who received subsequent HSCT (for example, post-HSCT patient numbers are small and follow-up is not mature enough). Therefore, these inputs have been derived from relevant publications with similar patient populations. Specifically, Evers et al., 2018 was used to derive long-term OS following HSCT.[43] As Evers does not report EFS, this input has been derived from OS assuming a constant cumulative HR, following the method of prior NICE submissions.[40,44,45] The efficacy input sources and the methodology used to predict OS and EFS inputs for patients with subsequent HSCT are described in detail in section **B.3.3.4**.

Incremental analysis in health economic evaluation require estimating the difference in mean survival times between different arms, based on the relevant areas under the EFS and OS survival curves. This requires a statistical extrapolation beyond the time horizon of the trial. To estimate the mean survival outcome times, a parametric approach to survival analysis was followed, as recommended by NICE.[46] When parametric models are fitted to each treatment arm, NICE guidelines suggest fitting the same distribution to both treatments, as different distributions allow for very different shapes and assumptions.

Alternative parametric models/distributions have been proposed for the extrapolation of survival curves. These include the exponential, Weibull, Gompertz, log-logistic, log-normal and generalised gamma distributions. In the base-case analysis, the efficacy inputs for OS and EFS without HSCT for gilteritinib were predicted using parametric survival models estimated using the ADMIRAL trial data. For survival without HSCT, the model supports applying a proportional hazards assumption if appropriate survival curves are chosen (i.e. is an exponential, Weibull or Gompertz model), in which case a single hazard ration can be applied to the gilteritinib curve to derive a comparator curve. Alternatively, individual parametric models can be fitted to OS and EFS, separately for each treatment.

OS with HSCT is predicted for all treatment arms using parametric models fitted to data from the relevant literature. As previously noted, the follow-up period and patient numbers add substantial uncertainty to the post-HSCT OS derived from ADMIRAL and therefore published curves are more appropriate. As Evers et al.[43] does not report any EFS measures, and as a single HR is used to derive the EFS inputs for HSCT patients, the candidate distributions to model OS with HSCT are only proportional hazards (PH) compatible models (exponential, Weibull and Gompertz).

The parametric survival models or HRs are used to inform OS and EFS until year 3. Afterwards, all patients who remain alive are assumed as cured from the disease and follow survival linked to the general population (SMR rates are explored in scenario analyses).

The role of varying survival analysis extrapolation assumptions is limited in this submission given the maturity of the ADMIRAL data. Seventy-five to eighty-five percent of patients did not receive HSCT group, and mortality in this cohort was 90-95% by data cut-off.[29] As such, extrapolation is only applied to a minority of survivors and only up to the three-year cure point. Table 19 provides a summary of data sources and extrapolation methods used for all the efficacy inputs in the base-case. Section **B.3.3.1** details the criteria used to evaluate the fit of the survival models. The methodology for parametric extrapolation and the long-term survival assumptions are described in more detail in sections **B.3.3.2** to **B.3.3.6**.

Table 19 Summary of Efficacy Data Sources and Base-case Extrapolation Approach

Efficacy inputs	Treatments	Extrapolation methods	Data sources
OS without HSCT	Gilteritinib	Parametric survival model (log-logistic)	ADMIRAL trial[29]
	Salvage chemotherapy comparators (i.e., azacitidine, FLAG-IDA, MEC, and LDAC)	Parametric survival model (log-logistic)	ADMIRAL trial[29]
	BSC	HR with gilteritinib as reference	ADMIRAL trial[29] Sarkozy et al.[47]
EFS without HSCT	Gilteritinib	Parametric survival model (log-logistic)	ADMIRAL trial[29]
	Salvage chemotherapy comparators (i.e., azacitidine, FLAG-IDA, MEC, and LDAC)	Parametric survival model (log-logistic)	ADMIRAL trial[29]
	BSC	All patients were assumed to start in post-event state	N/A

Efficacy inputs	Treatments	Extrapolation methods	Data sources
OS with HSCT	All treatments	Parametric survival model (Gompertz)	Evers et al.[43]
EFS with HSCT	All treatments	EFS estimated applying HR to OS	Evers et al.[43] Ustun et al.[48]

Abbreviations: BSC, best supportive care; EFS, event-free survival; FLAG-IDA, the combination therapy of fludarabine, cytarabine, granulocyte colony stimulating factor, and idarubicin; HR, hazard ratio; HSCT, haematopoietic stem cell transplantation; LDAC, low-dose cytarabine; MEC, the combination therapy of mitoxantrone, etoposide, and cytarabine; N/A, not applicable; OS, overall survival

B.3.3.1 Extrapolation of data and curve fitting

Following the systematic survival model selection process recommended by NICE DSU TSD14,[46] a range of methods, when appropriate, have been used to assess the suitability of parametric survival models for all efficacy inputs.

Specifically, model fit has been evaluated based on the following criteria:

- **Akaike information criterion** (**AIC)/Bayesian information criterion (BIC) tests:** The AIC and the BIC provide useful statistical tests of the relative fit of different parametric survival models. Such criteria attempt to estimate the out-of-sample prediction error of each model without external data or further model fits. These tests weight the improved fit of models with the potentially inefficient use of additional parameters. Lower AIC and BIC values indicate better (complexity-adjusted) goodness-of-fit to the data

- **Visual inspection:** visual inspection evaluates visually how well a parametric survival model fits the observed K-M. The parametric survival model that most closely follows the K-M curve could be considered that with the best fit

- **Examination of the log-cumulative hazard plots:** Log-cumulative hazard plots illustrate whether the hazards observed in the clinical trial over time are likely to be non-monotonic, monotonic or constant. Since different parametric survival models incorporate different hazard functions (e.g., the exponential implies a constant hazard, the Gompertz implies a monotonic hazard, etc.), the observed hazard plots are used to evaluate whether the parametric survival models have hazard functions with suitable and clinically plausible shapes

- **Testing the proportional hazards assumption:** The proportional hazards (PH) assumption has been evaluated when hazard ratios are utilised and applied to a

base survival curve to compare different comparator arms with the same reference case. In most cases, one HR is applied to the entire modelled period. In this scenario, the Schoenfeld residual test has been used to test the proportional hazards assumption and ensure that the treatment effect is proportional over time

- **Clinical input and external validation:** We sought clinical inputs and validated the model prediction with external data sources. Inputs were sought from clinical experts related to the clinical validity of estimated long-term survival and cure assumptions. Additionally, external data sources with comparable patient populations and interventions, but longer follow-up, were used for validation against the projected survival curves from the model

The methodology for parametric extrapolation, the long-term survival assumptions, and the selection of best-fit survival models are described in more detail in sections B.3.3.2 to B.3.3.6.

B.3.3.2 Overall survival without HSCT

OS data for gilteritinib and salvage chemotherapy were derived using individual patient data from the ADMIRAL phase III trial based on the ITT population who did not receive subsequent HSCT. Specifically, 184 patients and 105 patients randomised to the gilteritinib and salvage chemotherapy arms did not receive HSCT during the trial follow-up and have been included in the OS analysis. For gilteritinib, standard parametric models have been used to fit an OS curve and to extrapolate overall survival estimates. For the comparator arm, the model allows for independent parametric survival models or the application of a hazard ratio to the gilteritinib OS data.

The following parametric functions were considered: exponential, Weibull, Gompertz, log-logistic, log-normal, and generalised gamma distributions. Following the survival model selection process specified in section **B.3.3.1**, the following criteria were considered to select the parametric survival model with the best fit:

Information criterion (AIC/BIC)

Goodness-of-fit criteria based on the Akaike information criterion (AIC) and the Bayesian information criterion (BIC) were estimated for each parametric model to evaluate model fit based on statistical test results (Table 20 and Table 21).

Table 20 Summary of Goodness-of-fit Statistics for gilteritinib - OS without HSCT

Treatment	Efficacy inputs	Parametric Function	AIC	BIC
Gilteritinib	OS without HSCT	Exponential	943.394	946.609
		Weibull	938.745	945.175
		Log-logistic	**930.703**	**937.133**
		Log-normal	933.216	939.646
		Gompertz	944.713	951.143
		Generalised gamma	933.209	942.854

Abbreviations: OS, overall survival; HSCT, haematopoietic stem cell transplantation; AIC, Akaike information criterion; BIC, Bayesian information criterion

Table 21 Summary of Goodness-of-fit Statistics for Salvage Chemotherapy - OS without HSCT

Treatment	Efficacy inputs	Parametric Function	AIC	BIC
Salvage Chemotherapy	OS without HSCT	Exponential	485.933	488.587
		Weibull	485.728	491.036
		Log-logistic	**482.731**	**488.039**
		Log-normal	485.307	490.615
		Gompertz	487.889	493.197
		Generalised gamma	483.885	491.847

Abbreviations: OS, overall survival; HSCT, haematopoietic stem cell transplantation; AIC, Akaike information criterion; BIC, Bayesian information criterion

AIC and BIC suggest that the log-logistic was the best fitting model in both arms of the trial.

Visual inspection

The predicted curves were plotted against observed survival, shown in Figure 9 and Figure 10. These show limited differentiation between the curves but in general the log-logistic and log-normal curves appear to provide the best fit in both arms.

Figure 9 Parametric Models for OS - gilteritinib

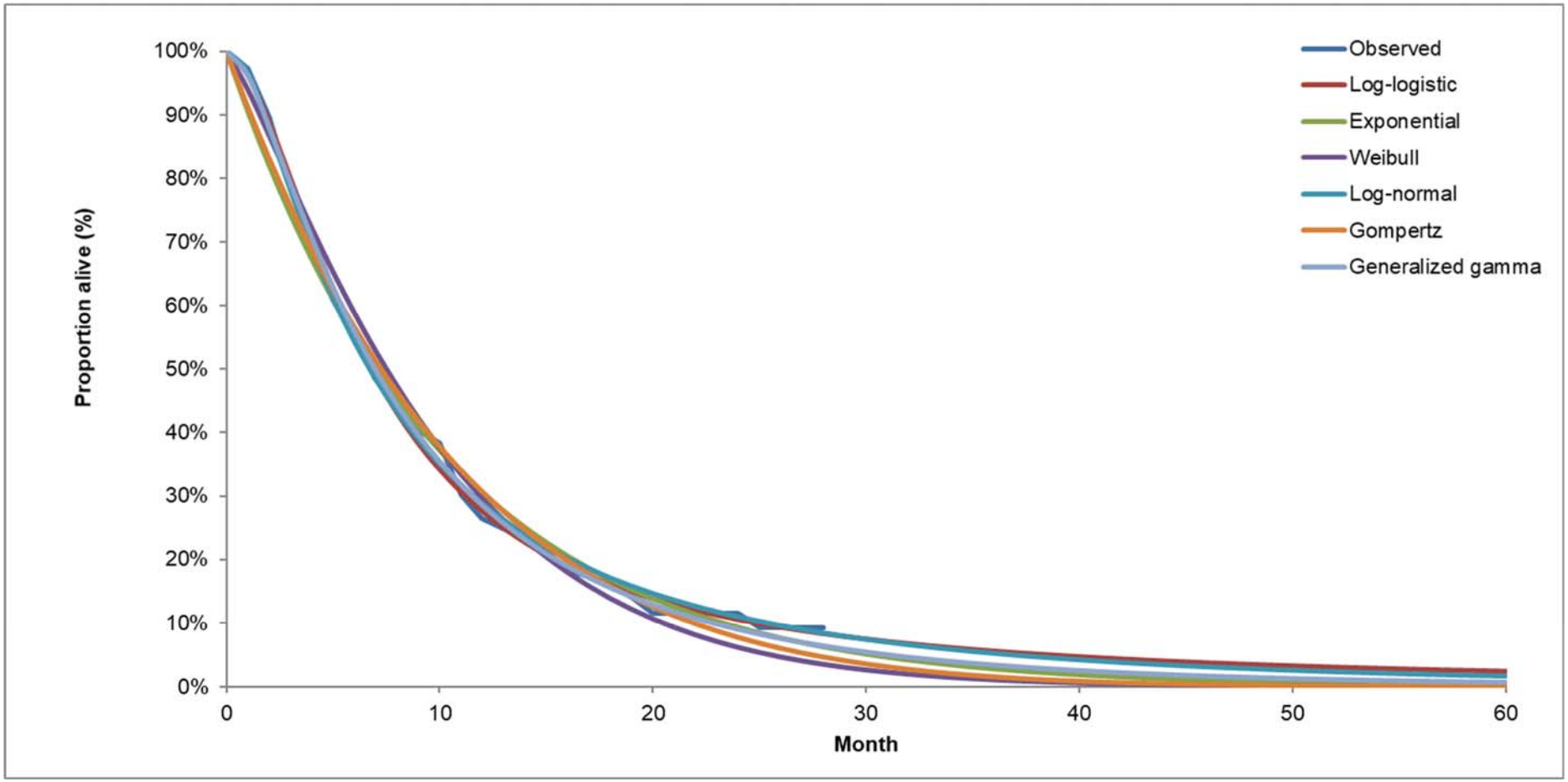
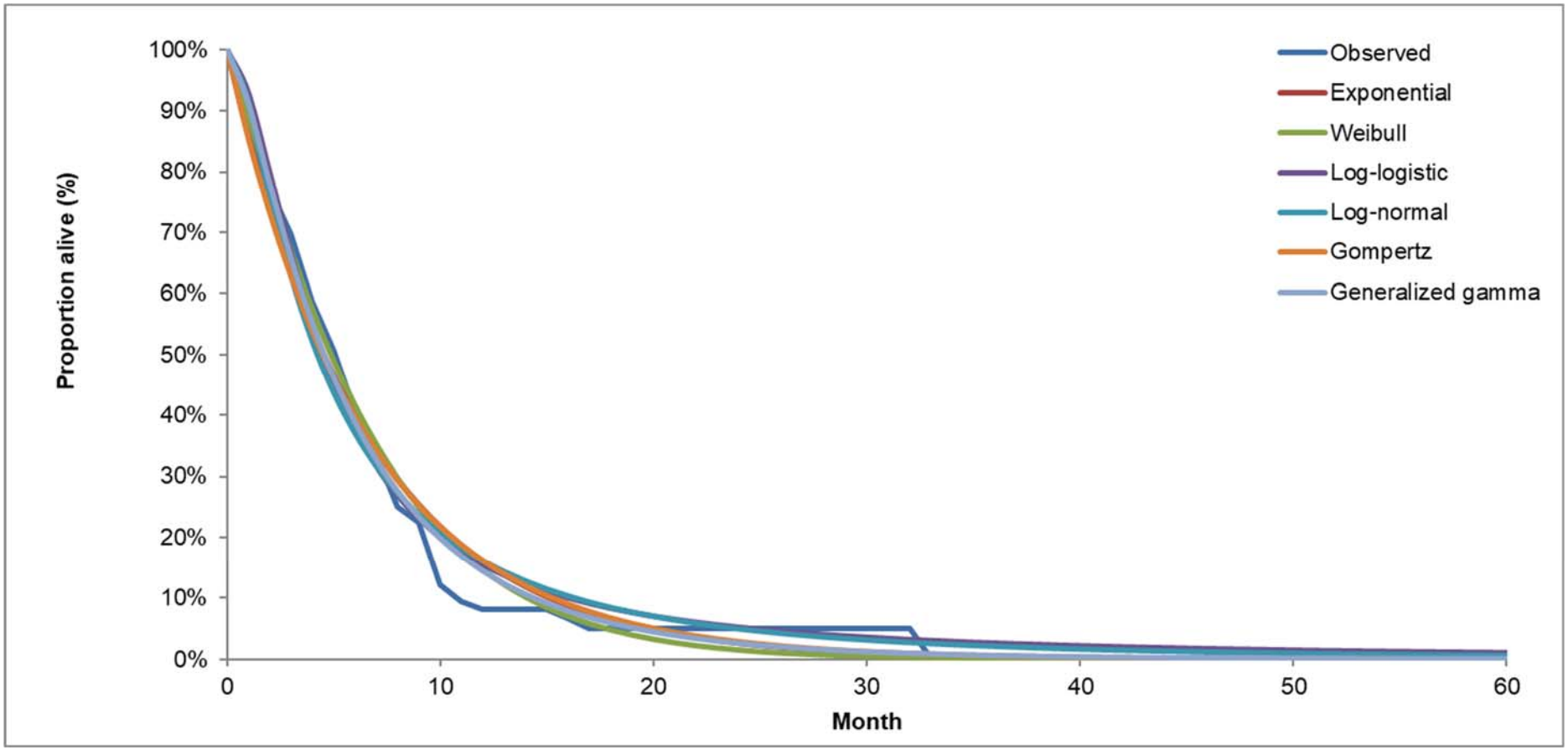


Figure 10 Parametric Models for OS - Salvage Chemotherapy



Log cumulative hazard plots

A log cumulative hazard plot was generated for gilteritinib and salvage chemotherapy based on observed OS among the ITT population without HSCT. In the plot (Figure 11), the gradients of the plots do not appear to be constant over time, therefore a log-logistic, log-normal, or generalized gamma model with non-monotonic hazard appear more suitable. In addition, the plot shows that the hazards are reasonably proportional between the two treatment arms, indicating that the PH assumption would likely hold between the gilteritinib arm and the salvage chemotherapy arm.

Figure 11 Log Cumulative Hazard Plots of gilteritinib and Salvage Chemotherapy - OS without HSCT

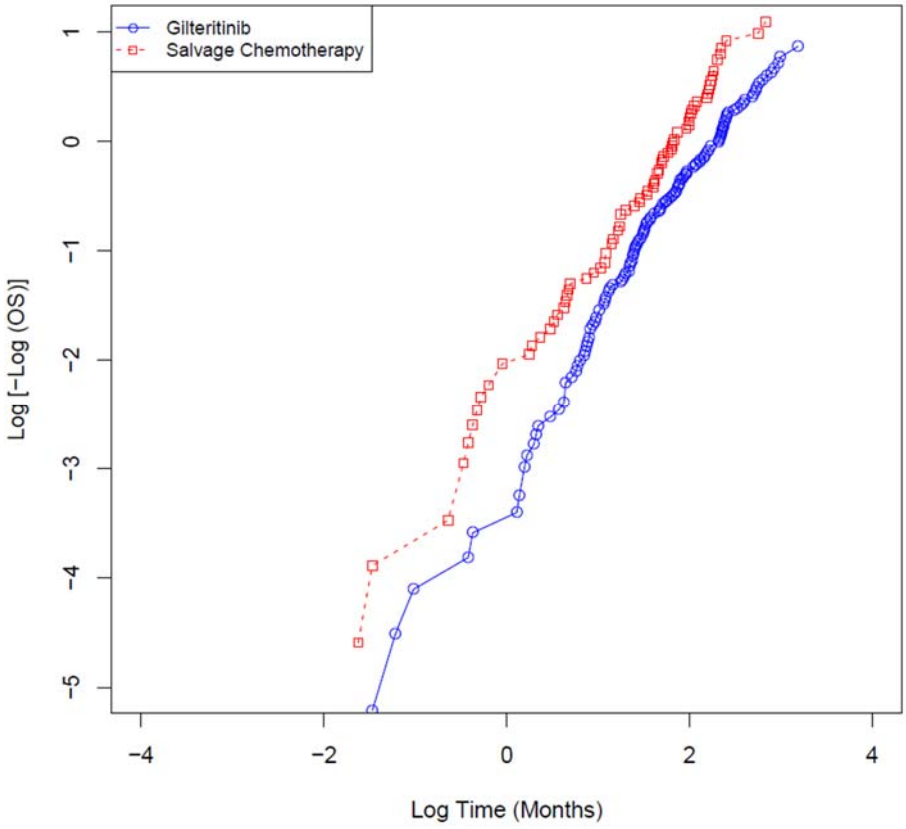
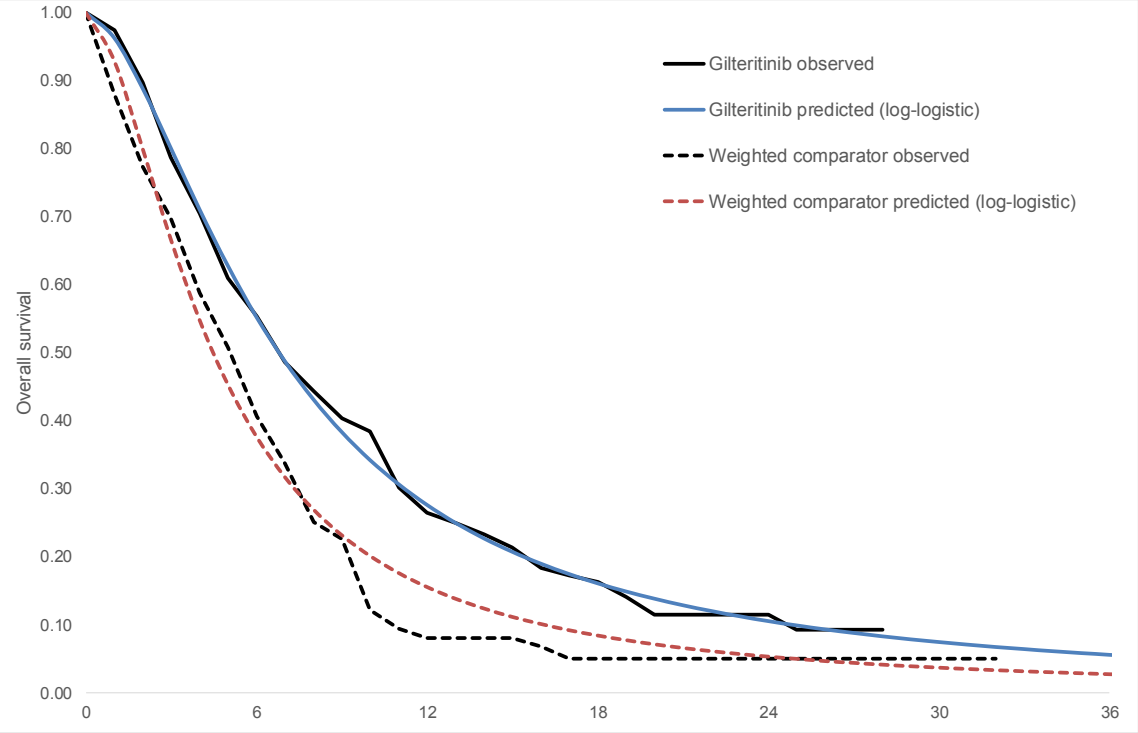


Figure 11 also shows that the proportional hazards assumption appears to be satisfied between the gilteritinib and the salvage chemotherapy arms, but the PH-compatible models (exponential, Weibull and Gompertz) did not fit the individuals arm well, particularly in terms of AIC and BIC.

The log-logistic model was selected as the best-fitting parametric model for both gilteritinib and salvage chemotherapy arms based on the following considerations: 1) it has the lowest AIC and BIC values among all survival models; 2) it demonstrates good fit against the observed K-M curves upon visual inspection; 3) the log cumulative hazard plots for both treatments indicate non-monotonic hazard profiles over time, which are consistent with the underlying hazard assumptions of the log-logistic model. Therefore, the base case uses log-logistic survival curves fitted separately to the gilteritinib and comparator survival data. These curves are plotted against observed OS in Figure 12.

Figure 12 Observed vs Predicted Overall Survival, gilteritinib and Weighted Comparator



As BSC was not included as a comparator in the ADMIRAL trial, OS inputs for the BSC comparator are based on relevant publications of relapsed or refractory AML. A targeted literature review was conducted to identify studies that reported efficacy of BSC in a comparable patient population to the target population. Sarkozy et al.[47] was selected as the most relevant publication because it evaluated efficacy of BSC in a comparable population (i.e. AML patients in first relapse including patients with and without FLT3 mutation positive) and included a large sample size of patients who received BSC (N=124).

As Sarkozy et al. [47] do not report a survival curve for BSC, a hazard ratio was applied to estimate the relatively effectiveness of this comparator with gilteritinib as the reference arm. The application of a HR to a log-logistic survival curve is not the methodologically preferred approach but it has been adopted here due to 1) limited BSC data, and 2) the superior fit of the log-logistic model to the gilteritinib data and the relatively poor fit of the PH-compatible models. The HR was calculated by comparing the ratio between the median OS reported for LDAC and BSC in Sarkozy et al.[47] and the HR comparing salvage chemotherapy with gilteritinib in the ADMIRAL[29] trial. This HR estimate is been based on a naïve comparison and therefore potential differences in patient baseline characteristics between the ADMIRAL trial and Sarkozy et al.[47] may not be fully accounted for. Table 22 presents a summary of the hazard ratios used for OS without HSCT for BSC vs. gilteritinib. As BSC is unlikely to be the comparator of greatest interest, we feel this pragmatic approach is justifiable.

Table 22 Summary of HRs for OS without HSCT for Comparators vs. gilteritinib

Treatment	HR vs. gilteritinib	Source
BSC	2.86	Sarkozy et al.[47]

Abbreviations: BSC, best supportive care; HR, hazard ratio

In the base-case model, the log-logistic was been chosen to inform OS for gilteritinib and salvage chemotherapy among patients without HSCT based on superior AIC/BIC goodness-of-fit, visual inspection, and the examination of the log-cumulative hazard plot. For BSC, OS has been informed based on the application of HRs with gilteritinib as the reference arm.

B.3.3.3 EFS without HSCT

As for OS, EFS data for gilteritinib and salvage chemotherapy were derived using IPD from the ADMIRAL trial for the ITT population who did not receive HSCT.[29] In the ADMIRAL trial, EFS was defined as the time from the date of randomisation until the date of documented relapse, treatment failure or death. Given the short duration of the salvage chemotherapy treatments, most treatment failures in this arm occurred within the first month of treatment. This was represented by setting the event date of treatment failure to study Day 1. As shown in Figure 13, this definition leads to an

immediate drop in the EFS curves for both gilteritinib and salvage chemotherapy: [redacted] of patients in the gilteritinib arm and [redacted] of patients in the salvage chemotherapy arm experienced a treatment failure event in the ADMIRAL trial.

Figure 13 Kaplan-Meier Plot of EFS by Treatment Arm

The non-parametric shape of the EFS curves complicates fitting parametric distributions to the data. To address this, the patients with treatment failures were assumed to be in the post-event state from cycle 0, and those without treatment failure events were used to fit parametric survival models for EFS prediction.

The same parametric functions were tested as with OS (see Section B.3.3.2) – exponential, Weibull, Gompertz, log-logistic, log-normal, and generalised gamma – and the same model selection process was followed.

For gilteritinib, the log-logistic model was selected as the best-fitting model based on the following considerations: 1) it has the lowest AIC and BIC values (see Table 23); 2) it demonstrated good fit against the observed curves based on visual inspection (see Figure 14); 3) the log cumulative hazard plot indicates a non-monotonic hazard

pattern over time (see Figure 15), which is consistent with the underlying hazard assumptions of the log-logistic model.

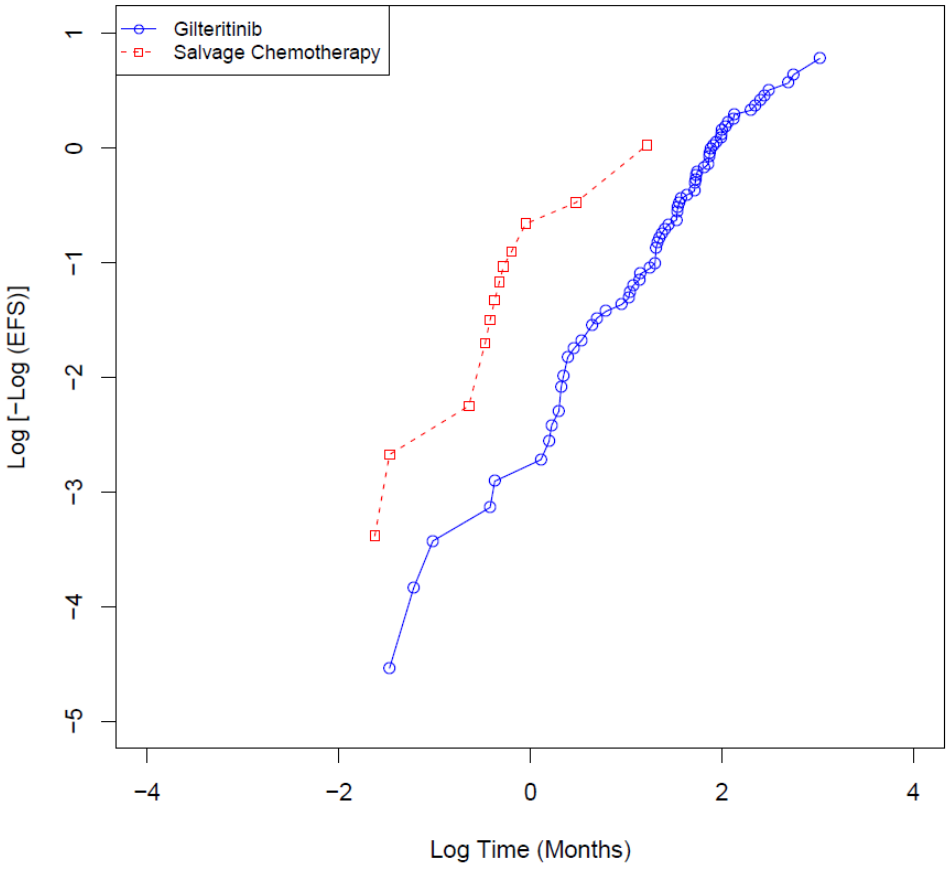
Table 23 Summary of Goodness of Fit Statistics for gilteritinib - EFS without HSCT

Treatment	Efficacy inputs	Parametric Function	AIC	BIC
Gilteritinib	EFS without HSCT	Exponential	452.601	455.186
		Weibull	454.232	459.402
		Log-logistic	**445.292**	**450.462**
		Log-normal	447.616	452.786
		Gompertz	453.179	458.349
		Generalized gamma	449.379	457.134

Abbreviations: AIC, Akaike information criterion; BIC, Bayesian information criterion; EFS, event-free survival; HSCT, haematopoietic stem cell transplantation

Figure 14 Parametric Models for EFS – gilteritinib

Figure 15 Log Cumulative Hazard Plots of gilteritinib and Salvage Chemotherapy - EFS without HSCT



For salvage chemotherapy, the generalised gamma model was the best-fitting model based on the AIC criterion, and the log-normal model was the best-fitting model based on the BIC criterion. However, NICE guidance suggests that the same distribution should be used to model each treatment for a given endpoint, as different distributions allow for very different shapes and assumptions.[46] For this reason, and for parsimony with respect to the OS endpoint without HSCT modelled using a log-logistic, a log-logistic distribution was also selected to model salvage chemotherapy based on the following considerations: 1) it gives a reasonably good fit based on AIC and BIC criteria (see Table 24); 2) it demonstrated good clinical plausibility and consistency; 3) visual inspection suggested a reasonable correspondence with observed EFS (see Figure 16); and 4) the log cumulative hazard plot indicates a non-monotonic hazard profile over time (see Figure 17), which is consistent with the underlying hazard assumptions of the log-logistic model.

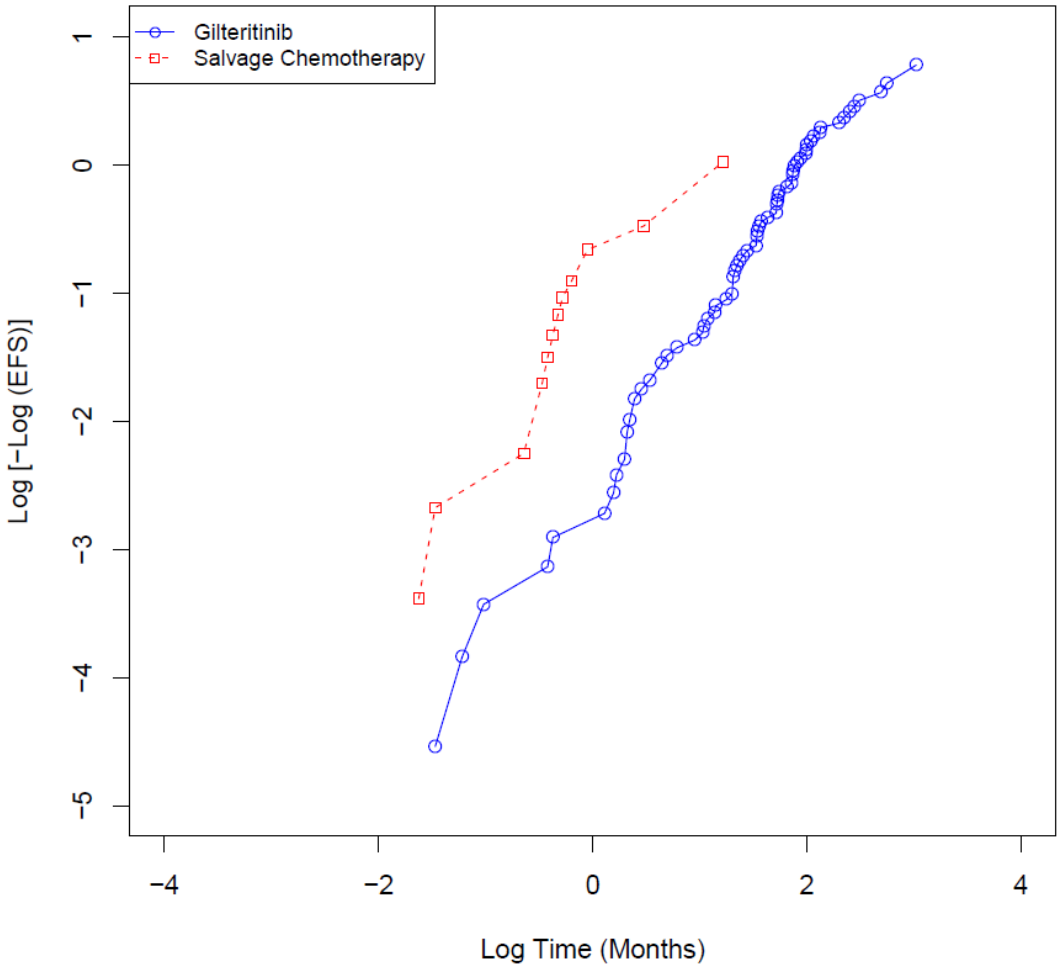
Table 24 Summary of Goodness of Fit Statistics for Salvage Chemotherapy - EFS without HSCT

Treatment	Efficacy inputs	Parametric Function	AIC	BIC
Salvage Chemotherapy	EFS without HSCT	Exponential	63.366	65.477
		Weibull	65.331	69.553
		Log-logistic	62.459	66.681
		Log-normal	61.140	65.362
		Gompertz	64.110	68.332
		Generalized gamma	**60.870**	**67.203**

Abbreviations: AIC, Akaike information criterion; BIC, Bayesian information criterion; EFS, event-free survival; HSCT, haematopoietic stem cell transplantation

Figure 16 Parametric Models for EFS - Salvage Chemotherapy

Figure 17 Log Cumulative Hazard Plots of gilteritinib and Salvage Chemotherapy - EFS without HSCT



Patients receiving BSC were considered to have low likelihood to achieve response based on clinical inputs. Consequently, all patients in the BSC comparator have been considered to start from the post-event state from cycle 0.

B.3.3.4 HSCT rate

According to the treatment guidelines for AML, HSCT is typically recommended for relapsed or refractory AML patients who achieve a second CR with chemotherapy.[6,8,12,22,23] In the base case, HSCT rates are based on observed HSCT rates in ADMIRAL.[29] The HSCT rate with BSC is assumed to be zero.

Table 25 Summary of HSCT Rates by Treatment in the Base-Case Model

Treatment	Proportion of patients (%)	Source

Gilteritinib	25.5	ADMIRAL trial[29]
Azacitidine	15.3	ADMIRAL trial[29] pooled salvage chemotherapy
FLAG-IDA	15.3	ADMIRAL trial[29] pooled salvage chemotherapy
MEC	15.3	ADMIRAL trial[29] pooled salvage chemotherapy
LDAC	15.3	ADMIRAL trial[29] pooled salvage chemotherapy
BSC	0.0	Assumption of no HSCT in BSC arm

Abbreviations: FLAG-IDA, the combination therapy of fludarabine, cytarabine, granulocyte colony stimulating factor, and idarubicin; MEC, the combination therapy of mitoxantrone, etoposide, and cytarabine; LDAC, low-dose cytarabine; BSC, best supportive care

A scenario is presented where it is assumed that only patients who achieve a second CR are eligible for HSCT (second CR: gilteritinib 21.1% vs salvage chemotherapy 10.5%).[29] Therefore, the base case should be considered highly conservative, as the difference in transplant rates is likely to be greater between the treatment options. This has been confirmed by an analysis of the British Society of Blood and Bone Marrow Transplantation database which shows 8.3% of all relapsed or refractory AML patients treated with salvage chemotherapy received a HSCT in 2017.[49] Some UK centres are known to apply this strict HSCT criteria.

B.3.3.5 OS and EFS with HSCT

All patients who received subsequent HSCT have been assumed to have the same clinical outcomes, regardless of the prior treatment received. In the ADMIRAL trial, 63 patients randomised to gilteritinib and 19 patients randomised to salvage chemotherapy received HSCT and have available OS and EFS data. The follow-up for these patients was limited: the median follow-up post-HSCT survival was [redacted] months with [redacted] having survival data beyond year 1, and only [redacted] patients ([redacted]) having data beyond year 2. Therefore, the literature were considered more robust to inform OS and EFS for patients with HSCT.

A targeted literature review was conducted to identify publications reporting post-HSCT OS and EFS data from a patient population comparable to that of the ADMIRAL trial. Ten publications were identified and selected, taking into consideration the following elements: 1) patient population comparable to the model target population; 2) relevant OS and EFS measures reported in the form of K-M curves; 3) large and mature enough sample sizes to reduce uncertainty in the survival extrapolation. All the studies evaluated for inclusion are presented in Table 26. None of the identified

publications focused on relapsed or refractory FLT3 mutation positive AML patients, but there is evidence that FLT3 mutation status does not impact survival outcomes post-HSCT.[50] Therefore, publications that evaluated survival for relapsed or refractory AML patients overall were considered more appropriate as a proxy to inform post-HSCT survival for the target population.

Evers et al.[43] was selected as the most relevant data source because it included the largest sample size and had the longest follow-up time. Since Evers et al. report OS stratified by remission status prior to HSCT, OS for patients with second CR (CR2) has been selected to inform OS with HSCT in the base-case. This is based on evidence from the ADMIRAL trial where the majority ([redacted]) of patients who received HSCT had achieved a composite complete remission (CRc).[29] The DSA explores considering different post-HSCT survival rates based on CR status. The model is not sensitive to this assumption, as described in section B.3.8.1. In addition, Ustun et al.[48] was tested as an alternative efficacy input source in the sensitivity analysis (section B.3.8.1), to evaluate the potential impact of *FLT3* mutation status on post-HSCT survival.

Data were extracted from the published K-M curves reported in Evers et al.[43] using the digitisation software Engauge. Pseudo patient-level data were derived based on the K-M data using the algorithm outlined in Guyot et al.[51] Information on "numbers at risk" and "number of events" has been incorporated into the reconstruction of IPD where available. The same set of standard parametric survival models described in section B.3.3.2 have been used to project survival estimates for OS with HSCT. Following the survival model selection process specified in section B.3.3.1, the following criteria have been considered to select the best-fitting parametric survival model:

- Goodness-of-fit criteria, such as AIC and BIC have been estimated for each parametric model to evaluate model fit based on statistical test results (Table 27)
- An overlay of the reconstructed K-M curves and the curves of each parametric survival model are presented in Figure 18 for visual inspection

- Log cumulative hazard plots have been generated for gilteritinib and salvage chemotherapy based on the K-M curve to assess if the hazards are constant, monotonic, or non-monotonic over time, see Figure 19.

Table 26 Clinical Data Source for Post-HSCT Survival Input

Source	Patient Population	Sample Size	Follow-up	OS at year 3	OS Definition
Evers[43]	R/R AML patients after 1st line treatment	Overall: 498 CR2: 128	Maximum follow-up: Over 11 years Median follow-up: 6.5 years	46%	From HSCT
Steckel[52]	R/R AML patients after 1-2 lines of chemotherapy	Overall: 292 Relapsed after 1st line: 51	Maximum follow-up reported: 5 years Median follow-up: 4.8 years	41%	From HSCT
Fong[53]	Relapsed AML patients	58	Maximum follow-up: 59 months Median follow-up: 6.7 months	55%	From treatment start date
Frazer[54]	AML patients in CR2 after relapse	55	NR	46%	From HSCT
Schmid[55]	R/R AML patients after one or multiple lines of chemotherapy	103	Maximum follow-up: 5.6 years Median follow-up: 2.1 years	32%	From HSCT
Oran[56]	FLT3 mutation positive AML patients	48	NR	54%	From diagnosis date
Schlenk[57]	FLT3 mutation positive AML patients in CR1	93	Median follow-up: 5.9 years	49%	From date of entry into the study
Song[58]	FLT3 mutation positive AML patients	262	NR	38%	From HSCT
Ustun[59]	Pooled FLT3 mutation positive AML patients in CR1 and CR2	284	Median follow-up: 3 years	50%	From HSCT
Deol[50]	Pooled FLT3 mutation positive AML patients in CR1 and CR2	158	Maximum follow-up: 5.4 years Median follow-up: 3.1 years	49%	From HSCT

Abbreviations: AML, acute myeloid leukaemia; CI, confidence interval; CR1, first complete remission; CR2, second complete remission; FLT3 mutation positive, FMS-like tyrosine kinase 3-mutated; HSCT, hematopoietic stem cell transplant; NR, not reported; OS, overall survival; R/R, refractory or relapsed; URD: unrelated donors

The generalised gamma model had the lowest AIC and BIC among all survival models. However, it does not have a proportional hazards parametrisation and is incompatible with the application of a single HR to derive EFS with HSCT. Instead, the Gompertz model, the best-fitting PH distribution, is used to model EFS with HSCT. The Gompertz model has been selected based on the following considerations: 1) it has the lowest AIC and BIC values among PH-compatible survival models; 2) it demonstrates reasonable fit against the observed curves based on visual inspection; 3) the log

cumulative hazard plot indicates a non-monotonic hazard pattern over time, which is consistent with the underlying hazard assumptions of the Gompertz model.

In the base-case model, the Gompertz model is used to predict OS with HSCT until the end of year 3. Afterwards, the model assumes that all patients who remain alive are cured from AML and follow SMR-adjusted general population mortality. More details about long-term survival are presented in section B.3.3.6.

Table 27 Summary of Goodness of Fit Statistics for All Treatments - OS with HSCT

Treatment	Efficacy inputs	Parametric Function	AIC	BIC
All treatments	OS with HSCT	Exponential	942.452	945.304
		Weibull	849.022	854.726
		Log-logistic	835.345	841.049
		Log-normal	828.555	834.259
		Gompertz	840.883	846.587
		Generalised gamma	**800.352**	**808.908**

Abbreviations: AIC, Akaike information criterion; BIC, Bayesian information criterion; OS, overall survival

Figure 18 Parametric Models for OS with HSCT

Figure 19 Log cumulative Hazard Plots of OS with HSCT - Evers et al., 2018

EFS data were not available in Evers et al.[43] to inform inputs for EFS with HSCT. In the absence of such data, the EFS curve has been derived from the available OS curve using an approach consistent with prior NICE submissions.[40,44,45] It has been assumed that, before year 3, the cumulative hazard function for EFS is proportional to the cumulative hazard function for OS. The ratio between post-HSCT EFS and post-HSCT OS has been modelled based on Ustun et al.[48] using data reported for the unrelated donor (URD) group. To estimate an overall cumulative HR between OS and EFS for patients with HSCT, the ratio was first estimated as the natural log of OS probability divided by the natural log of EFS probability at monthly intervals until year 5. The overall cumulative HR between OS and EFS was then calculated as the average of cumulative HRs at all monthly intervals. This assumption is justifiable based on clinical inputs and on the basis of evidence that EFS is highly correlated with OS.[60] In this model, the proportional relationship between EFS and OS is assumed to continue up to year 3. After year 3, the cumulative survival probabilities of EFS are assumed to remain constant until they reach OS. EFS is assumed to be less than or equal to OS at all time points.

B.3.3.6 Long-term survival

In the base case, the end of year 3 is assumed as the cure point for the target population. After year 3, those patients who remained alive and event-free are considered cured. This time point is consistently cited in the AML literature and HTA submissions as a clinically important landmark, after which there is a minimal risk of relapse,[41,61–63] and has been validated by clinical opinion leaders. This assumption is further validated by an observed plateau after year 3 in the OS curves of relapsed or refractory AML patients published in the literature.[62,64] The survival for cured patients has been modelled using general population mortality based on the 2012 UK life table, with an SMR adjustment to account for the higher mortality risk of the target population. Similar to the prior NICE submission of midostaurin,[41] the SMR in the base case is set to a multiple of two, validated by clinical expert opinion. It is worth highlighting that this assumption reduces some of the long-term uncertainties arising in the extrapolation of data beyond the reported follow-up of ADMIRAL. The estimated survival rate is applied to all patients who remain alive from year 3 onwards in the model. To evaluate the uncertainty around these assumptions, extensive sensitivity analyses have been performed on the long-term survival assumptions. These have included varying the cure assumption and applying a different SMR adjustment to the general population mortality. More details are described in section B.3.8.1.

B.3.3.7 Overall survival benefits associated with post-HSCT maintenance therapy

In the ADMIRAL trial, patients could continue receiving gilteritinib following HSCT, with [redacted] of patients continued therapy post-HSCT.[29] In the base case model, the costs and potential benefits associated with the post-HSCT treatment are considered to capture more comprehensively the expected cost and benefit for the target patient population in clinical practice. Because the post-HSCT survival data in the ADMIRAL trial are immature and unstable in the long-term, OS benefits associated with post-HSCT gilteritinib maintenance therapy are informed by the application of a hazard ratio. Specifically, a HR of [redacted] has been estimated using OS data from patients receiving gilteritinib maintenance after HSCT in the ADMIRAL trial and OS data of CR2 patients from Evers et al., 2018. The estimated HR has then been applied to the predicted OS of the proportion of patients receiving gilteritinib maintenance therapy

after HSCT. A scenario that does not consider costs or benefits of post-HSCT gilteritinib maintenance therapy is evaluated in the DSA (section B.3.8.1).

B.3.4 Measurement and valuation of health effects (utility inputs)

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

The EuroQol Group-5 Dimension-5 Level Instrument (EQ-5D-5L) was used to measure patients' health related quality of life in the ADMIRAL trial.[29] The instrument was administered at cycle 1 day 1 pre-dose, cycle 2 day 1 (± 2 days) and all subsequent cycles' day 1 (± 2 days) as well as during the pre-HSCT/end of treatment visit and the 30-day follow-up visit. During long-term follow up, patients were contacted by site personnel via telephone to provide responses to the questionnaire. The number of patients who provided EQ-5D scores at baseline, during cycle 1-5, during cycle 6-10, cycle 11 onwards, end of treatment, 30-day follow-up visit, and other times were 279, 221, 106, 46, 158, 54, 105 and 59, respectively.

Descriptive statistics on the EQ-5D values generated using the ADMIRAL trial data have been calculated according to the following categories, which correspond to different model health states:

- **EQ-5D measures for EFS without HSCT:** any EQ-5D assessment corresponding to patients in the EFS state before receiving HSCT, i.e., on or after the treatment start date and before the date of HSCT, relapse, treatment failure or death. The definition of EFS is consistent with that used in the ADMIRAL trial protocol

- **EQ-5D measures for post-event without HSCT:** any EQ-5D assessment corresponding to patients in the post-event state before receiving HSCT. The post-event state is defined as that after relapse or treatment failure

- **EQ-5D measurements for EFS with HSCT:** any EQ-5D assessment undertaken when the patient is in the EFS state after receiving HSCT, i.e., on or after the subsequent HSCT date and before the date of relapse, treatment failure or death

- **EQ-5D measurements for post-event with HSCT:** any EQ-5D assessment undertaken when the patient is in the post-event state after receiving HSCT. The post-event state is defined as that after relapse or treatment failure

EQ-5D-5L utility scores were calculated based on the individual dimension scores and mapped to EuroQol Group-5 Dimension-3 Level (EQ-5D-3L) scores using the UK preference-weights. The data were mapped back to the 3L tool using the crosswalk method, with the mapping function from van Hout et al.[65] This analysis did not impute values for missing evaluations. Descriptive statistics on the mapped EQ-5D-3L utility values and the total sample size for the above health state categories are shown in Table 28. A generalised estimating equation (GEE) model was developed to estimate patient utility scores with a robust variance estimator to account for correlation within patients' repeated assessments. The results are presented in Figure 20.

Table 28 Descriptive Statistics on EQ-5D Utility Values in the ADMIRAL Trial

Health states	N patients[a]	N assessments	Mean	SD
EFS without HSCT	[redacted]	[redacted]	[redacted]	[redacted]
Post-event without HSCT	[redacted]	[redacted]	[redacted]	[redacted]
EFS with HSCT	[redacted]	[redacted]	[redacted]	[redacted]
Post-event with HSCT	[redacted]	[redacted]	[redacted]	[redacted]

Abbreviations: EFS, event-free survival; HSCT, haematopoietic stem cell transplantation; SD, standard deviation
a. The same patient can have multiple health states at different visits. The statistics presented here reflect the number of patients with at least one assessment with the specified health state

Figure 20 EQ-5D Utility Mapping by Health States

B.3.4.2 Health-related quality-of-life data used in the CEA

B.3.4.2.1 Health state utilities

Utility values in the model are assumed to be dependent on the health states and independent of the treatment arms. The utility inputs considered in the model are summarised in Table 29. In the base-case model, patients who remain alive after year 3 are considered long-term survivors. The utility associated with long-term survivorship is assumed to equal that of patients in the "EFS with HSCT" state. Similar assumptions were considered in prior NICE submissions of midostaurin.[41]

There is no published literature that reports utility inputs for relapsed or refractory FLT3 mutation positive AML patients, and there are few publications in the general AML literature that use consistent methods to derive utilities for the specific health states

considered in the model. Joshi et al.[66] recruited 210 participants and used the composite time trade off approach to estimate utility values for 10 different health states for newly diagnosed AML. Many of the health states are comparable with those considered in the current CEA model. These inputs are largely like the utility estimates from the ADMIRAL trial data, although Joshi et al., 2019 predicted greater differences between pre-event and post-event states. For example, treatment failure/relapse/refractory is associated with a utility value of 0.51 based on Joshi et al., 2019; whereas post-event states from the ADMIRAL trial analysis have utilities ranging from [redacted] to [redacted]. Long-term follow-up after SCT > 1 year is associated with a utility value of 0.94 based on Joshi et al.;[66] the estimate from the ADMIRAL trial analysis is [redacted] (EFS with HSCT state).[29] Overall, EQ-5D analysis based on the ADMIRAL trial[29] data estimated lower utility for the EFS states and predicted smaller differences between EFS and post-event states compared to Joshi et al.[66] Alternative utility inputs based on Joshi et al. are explored in the sensitivity analyses.

The utility inputs derived from the ADMIRAL trial analyses are also similar to those considered in prior NICE submissions of FLT3 mutation positive AML. In the prior midostaurin submission of newly diagnosed patients with FLT3 mutation positive AML, the utility values for patients in CR or post SCT recovery ranged between 0.810 to 0.830,[41] which are similar to the utility inputs derived from the ADMIRAL trial.[29] In addition, in the midostaurin submission, the NICE committee considered 0.780 as the most plausible utility value for the relapsed state as patients who relapse from initial treatment can still achieve subsequent remission.[41] In the current model, the utilities of the post-event states are lower since the patients have already experienced two relapses.

B.3.4.2.2 Subsequent HSCT disutility

For patients who underwent subsequent HSCT, an additional disutility of [redacted] is considered based on the disutility value of SCT procedures reported in Joshi et al.[66] As Joshi et al. do not report an estimate of duration associated with the reported disutility estimate, the disutility associated with subsequent HSCT is assumed to last 6 months in the base case, based on the input of clinical experts.

B.3.4.2.3 Age-related utility decrements

The model considers additional age-related decrements as the modelled population becomes older over the modelled time horizon. The age decrements have been calculated as the relative ratio of the general population utility at the modelled age compared to the utility at the starting age ([redacted] years old based on the average age of patients in the ADMIRAL trial).[29] The calculated age adjustments were then applied to the estimated health state utility values when estimating the total QALYs. The inputs for age-related utility have been derived from Janssen et al.,[67] which describes the health utilities of healthy populations by different age groups. These utilities have been derived using the EQ-5D index population norms based on UK time-trade-off value sets.

B.3.4.2.4 Adverse event disutilities

For the base-case model, AE-specific utilities were estimated for any grade 3 and above AEs and have been added to each treatment arm as one-time utility decrements at the beginning of the model. AE rate inputs have been obtained from the ADMIRAL trial data.[29] The assumptions and inputs for AE rates are described in more detail in section **B.3.5.3**. The disutility inputs considered for individual AEs are reported in Table 29. Based on these inputs, the calculated AE disutilities for gilteritinib, salvage chemotherapies and BSC are -0.211, -0.143 and 0, respectively.

Table 29 Utilities and Disutilities

Parameter	Utility/disutility inputs	Source
Health state utilities (base-case)		
EFS without HSCT	[redacted]	ADMIRAL trial[29]
Post-event without HSCT	[redacted]	
EFS with HSCT	[redacted]	
Post-event with HSCT	[redacted]	
AML long-term survivors	[redacted]	Assumed equal to EFS with HSCT
Health state utilities (sensitivity analysis)		
EFS without HSCT	0.89	Joshi et al.[66]
Post-event without HSCT	0.51	
EFS with HSCT	0.94	
Post-event with HSCT	0.51	
AML long-term survivors	0.94	
Subsequent HSCT disutility		

Parameter	Utility/disutility inputs	Source
One-time HSCT disutility for patients with subsequent HSCT	-0.21	Joshi et al.[66]
Age-related utilities		
Age 55-64	0.799	Janssen et al.[67]
Age 65-74	0.779	
Age 75+	0.726	
AE disutilities		
Anaemia	-0.119	Swinburn et al.[68]
Dyspnoea	-0.050	Doyle et al.[69]
Elevated alanine aminotransferase	0.000	Assumed no disutility for abnormal lab tests
Elevated aspartate aminotransferase	0.000	Assumed no disutility for abnormal lab tests
Elevated blood phosphocreatine kinase	0.000	Assumed no disutility for abnormal lab tests
Fatigue	-0.115	Lloyd et al.[70], 2006
Febrile neutropenia	-0.150	Lloyd et al.[70], 2006
Hyperglycaemia	0.000	Assumed no disutility for abnormal lab tests
Hypertension	-0.153	Swinburn et al.[68]
Hypokalaemia	0.000	Assumed no disutility for abnormal lab tests
Hyponatraemia	0.000	Assumed no disutility for abnormal lab tests
Hypophosphatemia	0.000	Assumed no disutility for abnormal lab tests
Hypotension	-0.153	Assumed equal to hypertension
Leucopenia	-0.090	Assumed equal to neutropenia
Neutropenia	-0.090	Nafees et al.[71]
Neutrophil count decreased	0.000	Assumed no disutility for abnormal lab tests
Platelet count decreased	0.000	Assumed no disutility for abnormal lab tests
Pneumonia	-0.153	Assumed maximum disutility of all other AEs
Progressive acute myeloid leukaemia	[redacted]	ADMIRAL trial[29]
Sepsis	-0.090	Assumed equal to neutropenia
Thrombocytopenia	-0.090	Assumed equal to neutropenia
White blood cell count decreased	0.000	Assumed no disutility for abnormal lab tests

Abbreviations: AE, adverse event; AML, acute myeloid leukaemia; EFS, event-free survival; HSCT, haematopoietic stem cell transplantation

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

The model considers drug and drug administration costs, AE costs, subsequent HSCT costs, medical costs associated with health states, FLT3 mutation testing costs, terminal care costs, and post-progression treatment costs. The resource use specific to the gilteritinib and salvage chemotherapy arms has been obtained from the ADMIRAL trial. The cost and resource use for BSC has been obtained from literature to the extent possible. The inputs and assumptions are described in detail below.

B.3.5.1 Intervention and comparators' costs and resource use

This subsection includes drug acquisition costs and administration costs. In the model, drug and drug administration costs for each treatment are incurred during the modelled time horizon and are applied to all patients as a one-time input at the beginning of the model regardless of whether patients were treated with subsequent HSCT or not. The one-time input is derived from the distribution of treatment durations observed in the ADMIRAL trial. Mean observed treatment duration from the ADMIRAL trial has been used to estimate the overall drug and administration costs for gilteritinib and salvage chemotherapy regimens. For salvage chemotherapy arms, regimen-specific exposure time has been used because the treatment discontinuation rule varies across different salvage regimens (see details in section B.3.2.2). For high intensity regimens (FLAG-IDA and MEC), only 1-2 cycles were administered to patients. For low intensity regimens (LDAC and azacitidine), patients could continuously receive treatment until a lack of clinical benefit, intolerance, or a protocol-defined discontinuation criterion was met.

The ADMIRAL trial data are sufficiently mature at the time of data cut-off to capture majority of the expected treatment use. In the trial, all patients receiving salvage chemotherapies discontinued the treatment and only [redacted] of patients receiving gilteritinib remained on the initial treatment at data cut-off date. For gilteritinib, [redacted] of patients with HSCT re-initiated gilteritinib treatment 30 to 90 days post-HSCT.[29] In the base case analysis, the treatment exposure time associated with gilteritinib maintenance therapy is considered in the estimation of the mean treatment duration. The use of the treatment duration observed in the trial is a common approach used in

economic models and can accurately capture duration of treatment use in accordance with the clinical trial observation, and in accordance with the efficacy data. The BSC arm is assumed to incur no drug and drug administration costs.

Drug acquisition costs have been calculated as a function of unit drug cost, dosing, relative dose intensity, and treatment duration. Unit drug costs for brand-name active treatments have been retrieved from the Monthly Index of Medical Specialties (MIMS),[72] while unit drug costs for treatments that are available as generics have been obtained from the electronic Market Information Tool (eMIT) from the Commercial Medicines Unit of the NHS.[73] Unit drug prices for gilteritinib are based on the proposed list price for a 28-day pack ([redacted] for 84 tablets). A Patient Access Scheme (PAS) simple confidential discount of [redacted] has been proposed and applied in the model.

Dosing schedule, relative dose intensity, and treatment duration have been derived from the ADMIRAL trial. Drug administration costs have been calculated as a function of unit administration cost and administration frequency. Unit administration costs have been derived from NHS reference costs 2017-2018 based on the route of administration for each treatment.[74] The administration frequency is based on the dosing schedule from the ADMIRAL trial.[29]

Table 30 summarises the drug and drug administration costs for each treatment in the model, calculated based on dosing schedule and unit costs.

Table 30 Drug and Drug Administration Costs

Treatment	Dosing schedule	Per 28-day dosing cycle		Dosing cycles (N)	Relative dose intensity (%)	Total drug and admin costs (2018 GBP)	Source (Dosing schedule, dosing cycles, relative dose intensity; unit drug cost; unit admin cost)
		Drug cost (2018 GBP)	Admin cost (2018 GBP)				
Gilteritinib	120 mg daily	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	ADMIRAL trial[29]; assumptions; Chemotherapy Regimens Clinical Coding Standards and Guidance OPCS-4 (2017)[75]
Azacitidine	75 mg/m^2 daily, Day 1-7	£4,537.50	£1,573.79	2.24	101.0	£13,698.02	ADMIRAL trial[29]; MIMS[72]; Chemotherapy Regimens Clinical Coding Standards and Guidance OPCS-4 (2017) [75]
FLAG-IDA	G-CSF: 300 µg/m^2 daily, Day 1-5 Fludarabine: 30 mg/m^2 daily, Day 2-6 Cytarabine: 2000 mg/m^2 daily, Day 2-6 Idarubicin: 10 mg/m^2 daily, Day 2-4	£1,849.50	£1,418.51	1.02	G-CSF: 87.0 Fludarabine: 98.8 Cytarabine: 98.6 Idarubicin: 98.7	£3,335.71	ADMIRAL trial[29]; MIMS[72] for idarubicin and eMIT[73] for other treatments; Chemotherapy Regimens Clinical Coding Standards and Guidance OPCS-4 (2017) [775]
MEC	Mitoxantrone: 8 mg/m^2 daily, Day 1-5 Etoposide: 100 mg/m^2 daily, Day 1-5 Cytarabine: 1000 mg/m^2 daily, Day 1-5	£456.18	£1,185.28	1.13	Mitoxantrone: 105.5 Etoposide: 105.7 Cytarabine: 106.0	£1,848.99	ADMIRAL trial[2]; eMIT[73]; Chemotherapy Regimens Clinical Coding Standards and Guidance OPCS-4 (2017) [775]
LDAC	Cytarabine: 20 mg twice daily, Day 1-10	£77.32	£2,327.64	1.68	90.1	£4,048.06	

Abbreviations: Admin, administration; eMIT, electronic Market Information Tool; FLAG-IDA, the combination therapy of fludarabine, cytarabine, granulocyte colony-stimulating factor, and idarubicin; GBP, Great Britain Pound; G-CSF, granulocyte colony-stimulating factor; LDAC, low-dose cytarabine; MEC, the combination therapy of mitoxantrone, etoposide, and cytarabine; MIMS, Monthly Index of Medical Specialties

B.3.5.2 Health-state unit costs and resource use

The medical costs associated with health states in the model are considered separately based on whether patients are in the EFS or post-event state and whether patients receive or do not receive subsequent HSCT. In addition, it is assumed that patients considered cured after year 3 do not incur as much resource use, and therefore incur health state costs associated to long-term AML survivorship. In the base-case model, the medical costs associated with long-term survivors are assumed to be equal to those incurred by the “EFS with HSCT” state. The total medical cost inputs for each health state are summarised in Table 31.

For patients not receiving subsequent HSCT, medical costs associated with each health state include costs of outpatient visits, emergency department (ED) visits, hospitalisations, diagnostic procedures, lab tests, and blood transfusions. Different resource use frequencies have been considered for the "EFS without HSCT" state and the "post-event without HSCT" state. The frequencies of different resource use are based on information collected in a retrospective chart review study of relapsed or refractory FLT3 mutation positive AML patients in Europe.[76] The chart review study collected data from 93 patients with relapsed or refractory FLT3 mutation positive AML from the following European countries: UK, France, Germany, Italy, Spain and the Netherlands. The study collected data regarding AML-related inpatient admissions, intensive care unit, ED, outpatient visits, diagnostic procedures, lab tests, and blood transfusions. Separate inputs were collected during the event-free period and the post-event period.

The unit costs for outpatient visits have been derived from Unit Costs of Health and Social Care 2018 reported by the Personal Social Services Research Unit (PSSRU).[77] For ED visits, hospitalisations, diagnostic procedures, and lab tests, unit costs have been derived from the NHS reference costs 2017-2018.[74] The unit costs for blood transfusions have been obtained from the NHS blood and transplant price list.[78] Monthly resource use and costs for each state of patients not receiving subsequent HSCT are summarised in Table 32.

Different data sources have been considered to inform the frequencies of hospitalisation for patients in “EFS without HSCT” and “post-event without HSCT”

states. In the ADMIRAL trial,[29] frequencies of hospitalisation were collected before patients progressed, relapsed or experienced treatment failure events. The monthly lengths of stay for inpatient and intensive care unit admissions have been derived directly from the trial for the “EFS without HSCT” state. Separate estimates were derived for gilteritinib, low-dose chemotherapies and high-dose chemotherapies as the observed utilisation differs across these treatments. Monthly hospitalisation duration has been estimated based on ADMIRAL trial data among patients without HSCT.[29] All observed hospitalisation during the EFS period has been summarised as the mean hospitalisation or ICU days per month. Patients with no observed hospitalisation event during the period are included in the analysis and considered to have zero hospitalisation days. Hospitalisation inputs for the "EFS without HSCT" state are summarised in Table 33.

Data related to hospitalisation frequencies were sparsely collected in the ADMIRAL trial after patients experienced an event. Consequently, the sample size was insufficient to estimate robust inputs of hospital stay for patients in the “post-event without HSCT” health state. The frequencies of hospitalisation in this health state have been estimated based on information collected in the retrospective chart review study of relapsed or refractory FLT3 mutation positive AML patients in Europe,[76] and have been assumed to be the same across all treatment arms. Hospitalisation inputs for the “post-event without HSCT” state are summarised in Table 32 along with the other resource use categories collected in the chart review study.

For patients receiving subsequent HSCT, medical costs associated with the “EFS with HSCT” and “post-event with HSCT” states are based on literature and assumptions. For “EFS with HSCT”, the categories of relevant services and the monthly frequency of each service category are based on resource utilisation for the “SCT treatment” health state as reported in Tremblay et al.[79] These inputs were collected via a clinician survey that included data from 7 clinician respondents and represent the expected resource use for patients with AML receiving HSCT. The unit cost inputs are based on PSSRU 2018.[77] The related inputs are summarised in Table 34. As Tremblay et al.[79] do not differentiate between EFS and post-event for patients receiving HSCT, the medical cost inputs for the “post-event with HSCT” state are assumed to be equal to those in the “post-event without HSCT” state.

Table 31 Medical Costs Associated with Health States

Health state	Monthly costs (2018 GBP)	Source (Resource use; unit cost)
EFS without HSCT		
Gilteritinib	[redacted]	ADMIRAL trial,[29] AML chart review study[76]; PSSRU 2018[77]; NHS Reference Costs 2017-18[74]; NHS Blood and Transplant Price List 2017-18[78]
Azacitidine	£2,519.69	
FLAG-IDA	£5,583.83	
MEC	£5,583.83	
LDAC	£2.519.69	
BSC	£1,829.59	
Post-event without HSCT	£2,742.51	AML chart review study[76]; PSSRU 2018[77]; NHS Reference Costs 2017-18[77]; NHS Blood and Transplant Price List 2017-18[78]
EFS with HSCT	£170.02	Tremblay et al.[79]; PSSRU 2018[77]
Post-event with HSCT	£2,742.51	Assumed to be equal to post-event without HSCT state
Long-term survivors	£170.02	Assumed to be equal to EFS with HSCT state

Abbreviations: BSC, best supportive care; EFS, event-free survival; FLAG-IDA, the combination therapy of fludarabine, cytarabine, granulocyte colony stimulating factor, and idarubicin; GBP, Great Britain Pound; HSCT, haematopoietic stem cell transplantation; LDAC, low-dose cytarabine; MEC, the combination therapy of mitoxantrone, etoposide, and cytarabine

B.3.5.3 Adverse reaction unit costs and resource use

AE costs have been calculated for gilteritinib and salvage chemotherapy based on AE rates and unit costs regardless of whether patients were treated with subsequent HSCT or not. No AE costs have been considered for BSC. The AE rate inputs were obtained from the ADMIRAL trial data. As, to ensure a large enough sample size, gilteritinib was compared to salvage chemotherapy overall in the ADMIRAL trial, AE rate inputs for individual chemotherapy regimens are assumed to be the same as the overall inputs in the salvage chemotherapy arm. Grade 3/4 AEs have been included in the model if they affected over 5% of patients receiving any treatment. The costs associated with each of the AEs were derived from NHS reference costs 2017-2018 and literature (Table 35).[74]

Table 32 Monthly Resource use and costs - EFS without HSCT and post-event without HSCT (excluding hospitalisation inputs for EFS without HSCT)

Resource	Unit cost (2018 GBP)	Monthly frequency for EFS	Monthly frequency for post-event	NHS reference code (if available)	Source (Monthly frequency; unit cost)
Outpatient use					
Haematologist visits	£108.00	2.63	2.79	-	AML chart review study[76]; PSSRU 2018[77]
Nurse visits	£37.00	2.77	3.05	-	AML chart review study[76]; PSSRU 2018[77]
General practitioner visits	£93.75	1.67	1.56	-	AML chart review study[76]; PSSRU 2018[77]
ED use					
ED visits	£202.15	0.27	0.58	VB01Z, VB04Z, VB05Z, VB07Z, VB08Z	AML chart review study[76]; NHS Reference Costs 2017-18[77]
Hospitalisations					
Hospitalisation days	£396.30	Summarised in Table 33	2.13	SA25G, SA25H, SA25J, SA25K, SA25L, SA25M	AML chart review study[76]; NHS Reference Costs 2017-18[77]
ICU	£1,049.23		0.22	XC01Z - XC07Z	
Diagnostic procedures					
Imaging procedures	£42.30	0.71	0.57	RD20A, RD21A, RD22Z, RD20A, RD21A, RD22Z, DAPF	AML chart review study[76]; NHS Reference Costs 2017-18[77]
Lab tests					
Bone marrow biopsy	£519.82	1.07	0.32	SA33Z	AML chart review study[76]; NHS Reference Costs 2017-18[77]
Lumbar puncture	£519.82	0.18	0.16	SA33Z	
Blood transfusions					
Red blood cells	£128.99	1.73	2.41	BC001	AML chart review study[76]; NHS Blood and Transplant Price List[78]
Platelets	£208.68	1.50	1.82	BC044, BC045	
Plasma	£28.46	0.56	0.90	BC080	

Abbreviations: AML, acute myeloid leukaemia; ED, emergency department; EFS, event-free survival; GBP, Great Britain Pound; HSCT, haematopoietic stem cell transplantation; ICU, intensive care unit; NHS, National Health Service

Table 33 Monthly hospitalisation costs by regimen - EFS without HSCT

Treatments	Unit cost (2018 GBP)	Monthly frequency for EFS
Source	NHS Reference cost 2017-18[74]	ADMIRAL trial[29]
Gilteritinib		
ICU days	£1,049.23	[redacted]
Hospitalisation days (day case)	£396.30	[redacted]
Both ICU and hospitalisation days	£722.76	[redacted]
Low-dose chemotherapies		
ICU days	£1,049.23	0.00
Hospitalisation days (day case)	£396.30	1.03
Both ICU and hospitalisation days	£722.76	0.39
High-dose chemotherapies		
ICU days	£1,049.23	0.54
Hospitalisation days (day case)	£396.30	6.59
Both ICU and hospitalisation days	£722.76	0.80

Notes: Low-dose chemotherapies include azacitidine and LDAC; High-dose chemotherapies include FLAG-IDA and MEC
Abbreviations: ED, emergency department; EFS, event-free survival; FLAG-IDA, the combination therapy of fludarabine, cytarabine, granulocyte colony stimulating factor, and idarubicin; FLT3, fms-like tyrosine kinase 3; GBP, Great Britain Pound; LDAC, low-dose cytarabine; ICU, intensive care unit; MEC, the combination therapy of mitoxantrone, etoposide, and cytarabine

Table 34 Monthly resource use and costs - EFS with HSCT

Service category	Unit cost (2018 GBP)	Monthly frequency (minutes)	Source (Monthly frequency; unit cost)
Follow-up visit	£1.68 per minute	101	Tremblay et al.[79]; PSSRU 2018[77]

Abbreviation: EFS, event-free survival; GBP, Great Britain Pound; HSCT, haematopoietic stem cell transplantation

Table 35 Adverse Event Rates and Costs

Grade 3+ AEs ≥5%	Gilteritinib	Azacitidine	FLAG-IDA	MEC	LDAC	NHS reference code (if available)	Unit cost (2018 GBP)	Source for unit cost
Source for AE rates	ADMIRAL trial[29]							
Anaemia	40.7%	30.3%	30.3%	30.3%	30.3%	-	£211.73	Pixantrone NICE submission[80]
Dyspnoea	[redacted]	2.8%	2.8%	2.8%	2.8%	DZ27S, DZ27T, DZ27U	£422.41	NHS Reference Costs 2017-18[74]
Elevated alanine aminotransferase	13.8%	4.6%	4.6%	4.6%	4.6%	GC01E, GC01F	£550.36	
Elevated aspartate aminotransferase	14.6%	1.8%	1.8%	1.8%	1.8%	-	£550.36	Assumed equal to elevated alanine aminotransferase
Elevated blood phosphocreatine kinase	[redacted]	0.0%	0.0%	0.0%	0.0%	-	£550.36	
Fatigue	[redacted]	1.8%	1.8%	1.8%	1.8%	-	£91.68	Pixantrone NICE submission[80]
Febrile neutropenia	45.9%	36.7%	36.7%	36.7%	36.7%	-	£1,775.72	
Hyperglycaemia	[redacted]	8.3%	8.3%	8.3%	8.3%	KB02G, KB02H, KB02J, KB02K	£379.89	NHS Reference Costs 2017-18
Hypertension	[redacted]	3.7%	3.7%	3.7%	3.7%	EB04Z	£495.78	
Hypokalemia	13.0%	11.0%	11.0%	11.0%	11.0%	KC05J, KC05K, KC05L, KC05M, KC05N	£339.43	
Hyponatraemia	[redacted]	2.8%	2.8%	2.8%	2.8%	-	£339.43	Assumed equal to hypokalemia
Hypophosphatemia	[redacted]	3.7%	3.7%	3.7%	3.7%	-	£339.43	
Hypotension	[redacted]	2.8%	2.8%	2.8%	2.8%	-	£2,136.98	Pixantrone NICE submission[80]
Leucopenia	[redacted]	0.0%	0.0%	0.0%	0.0%	-	£803.28	Assumed equal to white blood cell count decreased

Grade 3+ AEs ≥5%	Gilteritinib	Azacitidine	FLAG-IDA	MEC	LDAC	NHS reference code (if available)	Unit cost (2018 GBP)	Source for unit cost
Neutropenia	[redacted]	13.8%	13.8%	13.8%	13.8%	-	£803.28	Pixantrone NICE submission[80]
Neutrophil count decreased	[redacted]	11.0%	11.0%	11.0%	11.0%	-	£803.28	Assumed equal to neutropenia
Platelet count decreased	22.0%	24.8%	24.8%	24.8%	24.8%	-	£1,875.04	Pixantrone NICE submission[80]
Pneumonia	[redacted]	4.6%	4.6%	4.6%	4.6%	DZ11K, DZ11L, DZ11M, DZ11N, DZ11P, DZ11Q	£691.48	NHS Reference Costs 2017-18[74]
Progressive acute myeloid leukaemia	[redacted]	3.7%	3.7%	3.7%	3.7%	SA25G, SA25H, SA25J, SA25K, SA25L, SA25M	£396.30	
Sepsis	[redacted]	0.0%	0.0%	0.0%	0.0%	WJ06G, WJ06H, WJ06J	£343.88	
Thrombocytopenia	22.8%	16.5%	16.5%	16.5%	16.5%	SA12G, SA12H, SA12J, SA12K	£280.28	
White blood cell count decreased	[redacted]	17.4%	17.4%	17.4%	17.4%	-	£803.28	Pixantrone NICE submission[80]
Total AE cost (2018 GBP)	[redacted]	£1,827.94	£1,827.94	£1,827.94	£1,827.94			

B.3.5.4 Miscellaneous unit costs and resource use

B.3.5.4.1 Subsequent HSCT disutility

The model assumes patients can receive subsequent HSCT after initial treatment. The cost and disutility of subsequent HSCT are added separately for the proportion of patients who received subsequent HSCT for each arm. The rates of subsequent HSCT are summarised in Table 36 and have been sourced from the ADMIRAL trial data. Patients in the BSC arm are assumed not to receive any subsequent HSCT.

HSCT costs have been considered in two parts: stem cell harvesting cost and the cost associated with the initial HSCT procedure. Both costs are based on the NHS reference costs 2017-2018.[74] These costs have been applied in a single cycle to represent the transplant event. The cost of initial transplantation has been computed via a weighted average (by frequency of health care resource group [HRG]) of all relevant adult allogeneic transplantation and stem cell harvesting procedures from the HRG costs. After the HSCT procedure, patients in the "EFS with HSCT" and "post-event with HSCT" states are assumed to incur monthly medical costs based on literature (see **B.3.5.2**).

Table 36 Subsequent HSCT costs

HSCT Procedure	Unit cost (2018 GBP)	NHS reference code	Source
Stem cell harvesting cost	£3,377.11	SA18Z, SA34Z	NHS Reference Costs 2017-18[74]
Cost associated with initial HSCT procedure	£37,397.24	SA20A, SA21A, SA22A, SA23A, SA38A, SA39A, SA40Z	
Total one-time HSCT costs	£40,774.35		

Abbreviations: GBP, Great Britain Pound; HSCT, haematopoietic stem cell transplantation; NHS, National Health Service; SCT, stem cell transplantation

B.3.5.4.2 FLT3 mutation testing costs

This is the requirement for FLT3 mutation testing to be conducted at the point a patient is diagnosed with relapsed or refractory disease. Only patients who test positive for the mutation would be eligible for gilteritinib. This would be conducted in all patients regardless of their previous FLT3 status. The infrastructure for this test exists given it is conducted at the point of initial diagnosis as a prognostic factor and more recently as a requirement for treatment with midostaurin which is used earlier in the treatment

pathway. The unit cost of the FLT3 mutation test using PCR methodology is estimated to be £154.54, based on clinical expert interviews reported in the Novartis midostaurin submission to NICE (TA523) in 2018 and adjusted for inflation. Key clinical opinion leaders confirmed that all AML patients are likely to receive prognostic FLT3 testing at the point of AML diagnosis.

B.3.5.4.3 Terminal care costs

In the model, all patients who transition to death are assumed to incur one-time terminal care costs. These have been sourced from Nuffield Trust 2014 data, which reports healthcare costs of hospital care, social care, district nursing, and general practitioner contacts in the last 90 days of life.[81] The total terminal costs per death have been converted to costs in the last 30 days to be consistent with the cycle length for the model (Table 37).

Table 37 Terminal Care Costs

Terminal care	Average cost (2018 GBP)	Source
Secondary hospital care	£6,191.22	Nuffield Trust[81]
Local authority-funded social care	£466.71	
District nursing	£618.07	
General practitioner contacts	£383.67	
Total terminal care costs in the last 90 days	£7,659.66	
Total terminal care costs in the last 30 days	£2,553.22	

Abbreviations: GBP, Great Britain Pound

B.3.5.4.4 Post-progression treatment costs

In the model, a proportion of patients in the "post-event without HSCT" state or in the "post-event with HSCT" state in the gilteritinib and salvage chemotherapy arms are assumed to receive post-progression treatments (i.e., anti-leukemic therapy) and incur post-progression treatment costs. The proportions of patients receiving post-progression treatments in the gilteritinib and the overall salvage chemotherapy arms are derived from the ADMIRAL trial data. The same inputs are considered for each individual salvage chemotherapy arm. No patients in the BSC arm are assumed to receive any subsequent anti-leukemic therapy.

For patients who receive post-progression treatment, the cost is applied at each model cycle for the proportion of patients who are newly progressed or have died. The post-progression cost estimate per cycle includes all costs related to drug, drug

administration, hospitalisation, and complication (Table 38). It has been derived from Wang et al.,[82] which sourced the average cost of all observed regimens, at second-line treatment for adult AML patients, from the UK Haematological Malignancy Research Network (HMRN). To estimate the total post-progression costs, such cost per cycle is multiplied by the average number of cycles of subsequent treatment observed in the ADMIRAL trial data.

Table 38 Post Progression Treatment Costs (Patients with and without HSCT)

Treatment	Proportion of patients with active treatment (%)	Cost per cycle (2018 GBP)	Average number of cycles	Total post-progression treatment cost (2018 GBP)
Source	ADMIRAL trial[29]	Wang et al.[82]	ADMIRAL trial[29]	
Gilteritinib	[redacted]	£5,179.09	2.60	[redacted]
Azacitidine	61.00			£8,264.47
FLAG-IDA	61.00			£8,264.47
MEC	61.00			£8,264.47
LDAC	61.00			£8,264.47
BSC	0.00			£0.00

Abbreviations: BSC, best supportive care; FLAG-IDA, the combination therapy of fludarabine, cytarabine, granulocyte colony stimulating factor, and idarubicin; GBP, Great Britain Pound; HSCT, haematopoietic stem cell transplantation; LDAC, low-dose cytarabine; MEC, the combination therapy of mitoxantrone, etoposide, and cytarabine

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

The base-case analysis inputs are summarised below, along with key assumptions made in the cost-effectiveness model.

Table 39 Key Assumptions in the Cost-Effectiveness Model

Parameter	Assumptions
Health states and utilities by health states	• All patients begin in the "treatment alone without HSCT" state following treatment initiation, and a proportion of patients transition to the "with HSCT" states after the average time to HSCT, based on the ADMIRAL trial observation, has elapsed • Patients who did not receive HSCT transition between the following three states: EFS without HSCT, post-event without HSCT, and death • Patients who received HSCT transition between the following three states: EFS with HSCT, post-event with HSCT, and death • Health state utilities are assumed to be dependent on the health states only and are independent of treatment arms
Subsequent HSCT	• Subsequent HSCTs after the initial treatment are considered to reflect the natural treatment course of R/R FLT3 mutation positive AML • The time at which patients receive subsequent HSCT is based on the average time to HSCT observed in the ADMIRAL trial • Subsequent HSCTs are assumed to directly affect patients' OS and EFS. The same efficacy benefit is considered for all patients receiving HSCT regardless of their initial treatment. This assumption is conservative, give the implied benefit from the post-HSCT OS curves for patients restarting gilteritinib (as the sample size of these patients is small, such benefit is not applied in the base case)

	• Cost and disutility of subsequent HSCTs are considered for the proportion of patients who received subsequent HSCT
Efficacy	• EFS and OS with HSCT vs. EFS and OS without HSCT are separately modelled • Efficacy inputs are estimated separately for gilteritinib, salvage chemotherapy and BSC to model EFS and OS without HSCT across different treatment arms (salvage chemotherapies are pooled into one overall chemotherapy arm) • The same efficacy inputs are used to model EFS and OS after HSCT across all treatment arms; inputs from the literature are used to model efficacy as data from the ADMIRAL trial are immature (for the HSCT patients) with limited follow-up and sample size (median follow-up post HSCT < 8 months) • Where EFS data were not available, EFS is estimated based on OS data assuming a constant cumulative HR over time • Patients receiving BSC are assumed to start in the alive and post-event state
Cure assumptions	• Patients who remain alive after year 3 are effectively cured; these patients are associated with a risk of death equivalent to a standardised mortality ratio (SMR) adjusted general population mortality, where the SMR is set to 2 based on a prior NICE submission (midostaurin) • After year 3, EFS is assumed to remain constant until it reaches OS (i.e., no further event), reflecting that patients are effectively cured and therefore not at risk of relapse and only associated with the risk of death as described above • After year 3, all patients who remain alive are assumed to incur health state costs and utilities associated with long-term AML survivors
Treatment costs	• Patients are treated based on the treatment schedule specified in the ADMIRAL trial for gilteritinib and salvage chemotherapy • Treatment duration of gilteritinib and comparators is based on observations in the ADMIRAL trial, assumed to be representative of UK practice
Medical costs	• Patients are assumed to incur different medical costs for each health state and treatment arms • The costs of pre-medications or concomitant medications (e.g., dexamethasone, ciprofloxacin) are conservatively not considered. These costs are more likely to be incurred by patients receiving comparator treatments (e.g., salvage chemotherapy). In addition, these costs are low. • All patients are assumed to have undertaken a FLT3 mutation test at initial diagnosis, however this model only considers the costs of re-testing for the FLT3 mutation at the point of relapsing or becoming refractory. Furthermore, the base case assumes that some initially FLT3 negative patients would also be tested, and therefore the costs of testing 200% of the FLT3 mutation positive population are applied. All costs are borne by the gilteritinib arm • All patients incur one-time terminal care costs before death
Post-progression treatment costs	• A proportion of patients in the "post-event without HSCT" and "post-event with HSCT" states are assumed to receive subsequent active treatment and incur post-progression treatment costs. The remaining patients are assumed to be managed with BSC and only incur medical costs associated with post-progression resource use
Adverse event	• Only grade 3/4 AE costs are considered in the model; one-time AE costs are added during the first cycle for simplification given the small impact of AE costs

Abbreviations: AE, adverse event; BSC, best supportive care; EFS, event-free survival; FLAG-IDA, the combination therapy of fludarabine, cytarabine, granulocyte colony-stimulating factor, and idarubicin; FLT3, FMS-like tyrosine kinase 3; HR, hazard ratio; HSCT, haematopoietic stem cell transplantation; LDAC, low dose cytarabine; MEC, the combination therapy of mitoxantrone, etoposide, and cytarabine; OS, overall survival

Using the base-case model inputs, the effectiveness and costs for each treatment arm are evaluated and compared. The outcomes evaluated included total costs and each cost component; including treatment costs, treatment-associated AE costs, subsequent HSCT costs, medical costs, and post-progression treatment costs, and

effectiveness measured as total LYs and QALYs. The total incremental costs, total incremental LYs, total incremental QALYs, and ICERs (presented as incremental cost per LY and incremental cost per QALY) are also calculated comparing gilteritinib vs. each comparator. For the weighted comparator, weights are assigned to the summarised cost and effectiveness estimates of all considered individual comparators to derive the effectiveness and cost.

B.3.7 Base case results

Total QALYs and total costs, estimated over a lifetime horizon, are summarised in Table 40, along with the incremental cost per QALY gained. For gilteritinib, drug and drug administration costs are the primary driver of total costs ([redacted] of total cost). For azacitidine, FLAG-IDA, MEC, LDAC, BSC and the weighted comparator, medical costs are the primary driver for the total costs ([redacted] respectively).

Considering both cost and effectiveness outcomes, the cost per QALY gained with gilteritinib relative to the weighted comparator was £47,695. The cost per QALY gained relative to the individual comparators range between £35,773 versus BSC to £52,954 versus LDAC.

Table 40 Base Case Results, Deterministic

	Gilteritinib	Azacitidine	FLAG_IDA	MEC	LDAC	BSC	Weighted Comparator
Comparator weights		[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
Costs (2018 GBP)							
Total Costs	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
Effectiveness							
Total LYs	3.033	1.749	1.749	1.749	1.749	0.330	1.749
Total QALYs	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
Incremental Changes							
Incremental Costs		[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
Incremental LYs		1.284	1.284	1.284	1.284	2.702	1.284
Incremental QALYs		[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
Incremental cost-effectiveness ratio (ICER) (2018 GBP)							
Incremental Cost per LY Gained		[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
Incremental Cost per QALY Gained		**£ 44,663**	**£ 47,235**	**£ 48,512**	**£ 52,954**	**£ 35,773**	**£ 47,695**

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

Table 41 Base Case Results, Probabilistic

	Gilteritinib	Azacitidine	FLAG_IDA	MEC	LDAC	BSC	Weighted Comparator
Comparator weights		[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
Costs (2018 GBP)							
Total Costs	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
Effectiveness							
Total LYs	3.400	2.032	2.042	2.040	2.043	0.336	2.039
Total QALYs	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
Incremental Changes							
Incremental Costs		[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
Incremental LYs		1.368	1.358	1.360	1.357	3.064	1.361
Incremental QALYs		[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
Incremental cost-effectiveness ratio (ICER) (2018 GBP)							
Incremental Cost per LY Gained		[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
Incremental Cost per QALY Gained		**£ 41,755**	**£ 44,458**	**£ 45,377**	**£ 49,936**	**£ 31,205**	**£ 44,750**

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

B.3.8 Sensitivity analyses

B.3.8.1. Deterministic sensitivity analysis

To assess the robustness of the model results, DSAs were conducted by varying one model input or assumption at a time. The list of varied model inputs is provided in Table 42. The modelled parameters have been varied according to their 95% confidence interval (CI) or their range if such information is reported in the original source. The modelled parameters have been varied by ± 25% from the base-case if such information is not provided.

Table 42 DSA Inputs - Model Parameters

Parameter	Base-case input	DSA input
Efficacy		
Cumulative HR between EFS and OS for HSCT state		
Cumulative HR between EFS and OS for HSCT state	0.89	0.67-1.00
Subsequent HSCT		
Subsequent HSCT rate - gilteritinib	25.51%	[redacted]
Subsequent HSCT rate – comparators	Azacitidine: 15.32% FLAG-IDA:15.32% MEC:15.32% LDAC:15.32% BSC: N/A	Azacitidine 95% CI: [redacted] FLAG-IDA 95% CI: [redacted] MEC 95% CI: [redacted] LDAC 95% CI: [redacted] BSC 95% CI: N/A
Utility and disutility		
Utility for EFS without HSCT	[redacted]	95% CI: [redacted]
Utility for post-event without HSCT	[redacted]	95% CI: [redacted]
Utility for EFS with HSCT	[redacted]	95% CI: [redacted]
Utility for post-event with HSCT	[redacted]	95% CI: [redacted]
Utility for AML long-term survivors	[redacted]	-25% of base-case
AE disutility – gilteritinib	[redacted]	±25% of base-case
AE disutility – comparators	Azacitidine: [redacted] FLAG-IDA: [redacted] MEC: [redacted] LDAC: [redacted] BSC: [redacted]	±25% of base-case
HSCT disutility duration	[redacted]	[redacted]
Subsequent HSCT disutility	[redacted]	95% CI: -[redacted] - [redacted]
Cost		
Treatment cost – gilteritinib	[redacted]	±25% of base-case
Treatment cost – comparators	Azacitidine: £13,698 FLAG-IDA: £3,336 MEC: £1,849 LDAC: £4,048 BSC: N/A	±25% of base-case
Medical cost before relapse/progression with HSCT	£170	±25% of base-case

Parameter	Base-case input	DSA input
Medical cost after relapse/progression with HSCT	[redacted]	±25% of base-case
Hospitalisation cost before relapse/progression without HSCT – gilteritinib	[redacted]	±25% of base-case
Hospitalisation cost before relapse/progression without HSCT – comparators	Azacitidine: £690 FLAG-IDA: £3,754 MEC:£3,754 LDAC: £690 BSC: N/A	±25% of base-case
Medical cost excluding cost of hospitalisation - EFS without HSCT	[redacted]	±25% of base-case
Medical cost including cost of hospitalisation - post-event without HSCT	[redacted]	±25% of base-case
Medical cost for AML long-term survivors	£0	Excluded
AE cost – gilteritinib	[redacted]	±25% of base-case
AE cost – comparators	Azacitidine: £1,828 FLAG-IDA: £1,828 MEC: £1,828 LDAC: £1,828 BSC: N/A	±25% of base-case
FLT3 testing cost	£154	±25% of base-case
Terminal care cost	£2,553	±25% of base-case
Subsequent HSCT cost	£40,774	±25% of base-case
Post-progression treatment cost – gilteritinib	[redacted]	±25% of base-case
Post-progression treatment cost – comparators	Azacitidine:£8,264 FLAG-IDA:£8,264 MEC: £8,264 LDAC: £8,264 BSC: N/A	±25% of base-case
SMR for long-term survivors		
SMR for long-term survivors	2.00	1.00-4.00
Discount rate		
Discount rate for cost and effectiveness	3.50%	1.50%-6.00%

Abbreviations: AE, adverse event; AML, acute myeloid leukaemia; BSC, best supportive care; EFS, event-free survival; FLAG-IDA, the combination therapy of fludarabine, cytarabine, granulocyte colony stimulating factor, and idarubicin; FLT3, fms-like tyrosine kinase 3; HR, hazard ratio; HSCT, hematopoietic stem cell transplantation; LDAC, low-dose cytarabine; MEC, the combination therapy of mitoxantrone, etoposide, and cytarabine; OS, overall survival; SMR, standard mortality ratio

In addition, the DSA also evaluates different efficacy and cost scenarios. The types of scenarios explored are detailed in Table 43. Extensive scenario analyses have been performed on possible drivers of the results to consider different parametric functions for gilteritinib, the use of different post-HSCT survival and long-term survival

assumptions, alternative input sources for core efficacy and cost inputs, and different time horizons.

Table 43 DSA Inputs - Modelling Scenarios

Scenarios	Base-case input	DSA input
Effectiveness based on alternative parametric functions for gilteritinib		
OS/EFS	Log-logistic	Exponential, Weibull, log-normal, Gompertz, generalised gamma
Effectiveness based on alternative extrapolation approaches for salvage chemotherapies		
OS - HR for comparators vs. gilteritinib	Log-logistic	HRs applied Azacitidine 95% CI: [redacted] FLAG-IDA 95% CI: [redacted] MEC 95% CI: [redacted] LDAC 95% CI: [redacted] BSC 95% CI: [redacted]
EFS - HR for comparators vs. gilteritinib	Log-logistic	HRs applied Azacitidine 95% CI: [redacted] FLAG-IDA 95% CI: [redacted] MEC 95% CI: [redacted] LDAC 95% CI: [redacted] BSC 95% CI: N/A
Time horizon		
Time Horizon	Lifetime	20 years; 30 years
Alternative cost scenarios		
Post-progression treatment	Consider post-progression treatment	Do not consider post-progression treatment
PAS discount on all comparators	Do not consider PAS discount	Consider 10% PAS discount
FLT3 re-testing rate	200% re-testing rate	100% re-testing rate
Vial sharing	Do not consider vial sharing	Consider vial sharing
Alternative utility scenarios		
Data source for health state utilities and disutilities	Health state utility input source - ADMIRAL trial	Health state utility input source - Joshi et al., 2019
Utility for long-term survivors	1	0.940 based on Joshi et al.[66]
AE disutilities for each treatment	Consider AE disutilities for each treatment	Do not consider AE disutilities for each treatment
Alternative scenarios for HSCT assumptions		
HSCT rates for gilteritinib	Based on the ADMIRAL trial[29]	
HSCT rates for patients on salvage chemotherapy arms	HSCT rates based on the ADMIRAL trial[29]	Do not consider HSCT use for salvage chemotherapy arms;
Model cycle to introduce HSCT states	Average time to HSCT - gilteritinib: [redacted] months;	At model start
HSCT rates for all arms	Consider HSCT use based on ADMIRAL trial observation[29]	Do not consider HSCT use; only consider HSCT use among patients who achieved CRc
Alternative modelling scenarios		
Long-term cost and utilities	Consider long-term cost and utilities starting from month 37	Do not consider long term cost and utilities
Cost and benefit associated with post-HSCT gilteritinib	Consider both cost and benefit	Consider neither cost nor benefit

Scenarios	Base-case input	DSA input
use as maintenance therapy		
Long-term survival assumption after cure point	Consider long-term survivors will have the same mortality as normal population	Use the same efficacy input and extrapolation approach before the cure point to model survival after cure
Post-HSCT survival assumption	Same survival for all patients	Different survival for patients with vs. without CRc
Post-HSCT OS input source	Evers et al., 2018	Ustun et al., 2017
OS extrapolation for salvage chemotherapy comparators in without HSCT state	Use the base-case log-logistic curve	Use different HRs specific to the high-dose chemotherapy and low-dose chemotherapy for respective salvage chemotherapy comparators

Abbreviations: CRc, composite complete remission; DSA, deterministic sensitivity analysis; EFS, event-free survival; FLT3, fms-like tyrosine kinase 3; HR, hazard ratio; HSCT, hematopoietic stem cell transplantation; OS, overall survival; PAS, patient access scheme

The DSA results indicate that the model is robust to most scenario changes. Across all comparator arms, the model results are most sensitive to the following factors:

- Consideration of subsequent HSCT
- Subsequent HSCT rates
- Treatment costs of gilteritinib and comparator arms
- Discount rates for cost and effectiveness
- Long-term survival assumptions after cure point
- Long-term cost and utilities assumptions

B.3.8.2. Probabilistic sensitivity analysis

A probabilistic sensitivity analysis has been conducted to estimate the probability for gilteritinib to be cost-effective compared to a weighted combination of the comparators based on a willingness-to-pay (WTP) threshold of £50,000. A Monte-Carlo simulation with 5,000 iterations has been conducted. In each iteration, key efficacy, utility and cost inputs were randomly drawn from specified distributions to inform the possible range of the inputs. The results were presented as a cost-effectiveness scatter plot and a cost-effectiveness acceptability curve (CEAC) comparing gilteritinib with each comparator. All the model parameters that have been varied in the PSA and their associated distributions are summarised in Table 44. Whenever available, the standard error (SE) of the model input is directly obtained from the same data source that informed the mean value. In the absence of data on the variability of health state

costs, utilities or HR values, the SE for each parameter is assumed to be equal to 25% of the mean estimate.

Table 44 PSA Parameters

Parameters	Descriptions
Effectiveness for patients without HSCT	In the base-case, OS and EFS are modelled using parametric functions. The uncertainty in the survival probabilities is represented through the joint variance-covariance matrix of parameter estimates using normal distributions
OS of each treatment	
EFS of each treatment	
Effectiveness for patients with HSCT	Uncertainty in the cumulative HR between OS and EFS in the base-case is modelled using lognormal distributions with SEs assumed to be 25% of the mean. The uncertainty in the survival probabilities of OS is represented through the joint variance-covariance matrix of parameter estimates using normal distributions
OS	
EFS	
Utility for health states	Utilities are modelled using Beta distributions with the mean values as specified in the base-case model. SEs were obtained from the ADMIRAL trial or literature. It is assumed that the utility for "Post-event" cannot exceed that of "EFS"
Utility for EFS without HSCT	
Utility for post-event without HSCT	
Utility for EFS with HSCT	
Utility for post-event with HSCT	
Utility for AML long-term Survivors	
Subsequent HSCT	Subsequent HSCT rates and disutilities are modelled using Beta distributions with the mean values as specified in the base-case model. SEs are obtained from the ADMIRAL trial. Subsequent HSCT cost is modelled using a Gamma distribution with mean values as specified in the base-case model. SE is assumed to be 25% of the mean
Subsequent HSCT rate	
Subsequent HSCT cost	
Subsequent HSCT disutility	
Treatment cost	Treatment duration and post-progression treatment costs are modelled using Gamma distributions with the mean values as specified in the base-case model. SEs of treatment duration for gilteritinib and chemotherapies are obtained from the ADMIRAL trial; SEs of post-progression treatment cost are assumed to be 25% of the mean
Treatment duration	
Post-progression treatment cost	
Medical cost	Medical costs are modelled using Gamma distributions with the mean values as specified in the base-case model. SEs are assumed to be 25% of the mean
Hospitalisation cost - EFS without HSCT	
Health state costs	
Testing cost	
Terminal care cost	
Safety	
AE cost of each treatment	AE costs are modelled using Gamma distributions with the mean values as specified in the base-case model. SEs are assumed to be 25% of the mean. AE disutilities are modelled using Beta distributions with mean values as specified in the base-case model. SEs are assumed to be 25% of the mean
AE disutilities of each treatment	
Patient characteristics	Age and BSA are modelled using normal distributions with mean values as specified in the base-case model. SEs are obtained from the ADMIRAL trial
Age	
Body surface area (BSA)	

Abbreviations: AE, adverse event; AML, acute myeloid leukaemia; BSA, body surface area; EFS, event-free survival; OS, overall survival; HR, hazard ratio; HSCT, haematopoietic stem cell transplantation; SE, standard error

B.3.8.3. Summary of sensitivity analyses results

Summaries of the deterministic and probabilistic sensitivity analyses are presented below.

B.3.8.3.1 Deterministic sensitivity analysis results

The Tornado diagrams below present the impact of one-way changes in model parameters for gilteritinib versus the weighted comparator and the individual comparators. Parameters were varied according to their 95% confidence intervals, if available. If confidence intervals were not available or applicable, parameters were varied by ±25% of their expected value.

Figure 21 DSA results ranked by impact on ICER values (gilteritinib vs. weighted comparator)

Figure 22 Top 20 DSA results ranked by impact on ICER values (gilteritinib vs. azacitidine)

Figure 23 Top 20 DSA results ranked by impact on ICER values (gilteritinib vs. Flag-Ida)

Figure 24 Top 20 DSA results ranked by impact on ICER values (gilteritinib vs. MEC)

Figure 25 Top 20 DSA results ranked by impact on ICER values (gilteritinib vs. MEC)

Figure 26 Top 20 DSA results ranked by impact on ICER values (gilteritinib vs. LDAC)

Figure 27 Top 20 DSA results ranked by impact on ICER values (gilteritinib vs. Best Supportive Care)

B.3.8.3.2 Probabilistic sensitivity analysis results

A probabilistic sensitivity analysis was conducted, using random draws from defined probability distributions across 5,000 iterations. The share of iterations for each comparator treatment reflected the share observed in the ADMIRAL trial. The cost-effectiveness scatterplot is shown in Figure 28 and Figure 29 shows a cost-effectiveness acceptability curve for gilteritinib versus the weighted comparator. Compared with a weighted comparator of azacitidine, FLAG-IDA, MEC, LDAC, the probability of gilteritinib being cost-effective at a WTP threshold of £50,000 per QALY gained was 80.2%.

Figure 28 Cost-Effectiveness Scatterplot, gilteritinib vs Weighted Comparators

Figure 29 Cost-Effectiveness Acceptability Curve, gilteritinib vs Weighted Comparators

A cost-effectiveness acceptability frontier for gilteritinib versus each treatment is included in the Excel model, however, since individual comparators have all used the pooled salvage chemotherapy effectiveness value, this is of limited use.

B.3.8.4. Additional scenario analyses

Several additional scenario analyses have been conducted based on uncertainties of previous appraisals.

B.3.8.4.1 Scenario 1

In the base case, post-event without HSCT has a lower utility than post-event with HSCT ( ). This may not be plausible considering the hardships of the stem cell transplant. In this scenario, both post-event utilities are set to .

Over a lifetime horizon, the total QALYs for gilteritinib, azacitidine, FLAG-IDA, MEC, LDAC, BSC and weighted comparator are estimated to be , respectively, and the model estimates the total costs to be , respectively,

during that period. For gilteritinib, drug and drug administration costs are the primary driver for the total costs (56%). For azacitidine, FLAG-IDA, MEC, LDAC, BSC and weighted comparator, medical costs are the primary driver for the total costs (42%, 61%, 63%, 52%, 100%, 55% respectively).

Considering both cost and effectiveness outcomes, deterministic ICERs of £44,674, £47,247, £48,525, £52,967, £35,787 and £47,707 per QALY gained are estimated for patients treated with azacitidine, FLAG-IDA, MEC, LDAC, BSC and weighted comparator compared to gilteritinib, respectively. The cost-effectiveness results barely change as all the treatments are influenced by the same HSCT cost and effectiveness inputs.

Table 45 Scenario 1 Results

	Gilteritinib	Azacitidine	FLAG-IDA	MEC	LDAC	BSC	Weighted comparator
Total costs	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
Treatment costs	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
Treatment associated adverse event costs	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
HSCT Costs	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
Medical Costs	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
Post-Progression treatment Costs	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
Total LYs	3.033	1.749	1.749	1.749	1.749	0.330	1.749
Total QALYs	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
Incremental costs		[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
Incremental LYs		[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
Incremental QALYs		[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
Incremental cost per LY gained		[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
Incremental cost per QALY gained		**£ 44,674**	**£ 47,247**	**£ 48,525**	**£ 52,967**	**£ 35,787**	**£ 47,707**

B.3.8.4.2 Scenario 2

In the base case, the weighted comparator has weights based on the ADMIRAL trial data assigned to each individual salvage chemotherapy arm. Specifically, the weights assigned to azacitidine, FLAG-IDA, MEC, and LDAC were [redacted], respectively, with BSC being assigned a weight of 0. In this scenario, we consider incorporating BSC into the weighted comparator with a weight of 20%. The weights of the salvage chemotherapies make up the remaining 80%, keeping their weights relative to each other the same. Specifically, the weights assigned to azacitidine, FLAG-IDA, MEC, LDAC and BSC, are [redacted]

Over a lifetime horizon, the total QALYs for the weighted comparator are estimated to be [redacted], and the model estimates the total costs as [redacted] during that period. For the weighted comparator, medical costs remain the primary driver for the total costs ([redacted] of the total costs). Considering both cost and effectiveness outcomes, a deterministic ICER of [redacted] per QALY gained is estimated for patients treated with the weighted comparator compared to gilteritinib. Other costs and ICERs for the individual salvage chemotherapy arms are identical to those of the base-case.

Table 46 Scenario 2 Results

	Gilteritinib	Azacitidine	FLAG-IDA	MEC	LDAC	BSC	Weighted comparator
Total costs	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
Treatment costs	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
Treatment associated adverse event costs	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
HSCT Costs	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
Medical Costs	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
Post-Progression treatment Costs	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
Total LYs	3.033	1.749	1.749	1.749	1.749	0.330	1.465
Total QALYs	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
Incremental costs		[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
Incremental LYs		[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]

Incremental QALYs		[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
Incremental cost per LY gained		[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
Incremental cost per QALY gained		£ 35,748	£ 38,320	£ 39,597	£ 44,039	£ 31,512	£ 36,284

B.3.8.4.3 Scenario 3

In the base case, the "cure point" is set to 3 years. As per the prior NICE submission of midostaurin, the SMR is set to two based on clinical expert opinion. Key opinion leader clinicians suggested 2 to 3 years as a plausible cure point. In this scenario, the cure point is moved to 2 years. Long-term medical costs and utility values are also applied after 2 years instead of 3 in this scenario.

Over a lifetime horizon, the total QALYs for gilteritinib, azacitidine, FLAG-IDA, MEC, LDAC, BSC and weighted comparator are estimated as [redacted], respectively. The model estimates the total costs as [redacted], respectively, during that period. For gilteritinib, drug and drug administration costs are the primary driver for the total costs ([redacted]). For azacitidine, FLAG-IDA, MEC, LDAC, BSC and weighted comparator, medical costs are the primary driver for the total costs ([redacted] respectively).

Considering both cost and effectiveness outcomes, deterministic ICERs of £36,099, £38,498, £39,553, £42,942, £27,762, £38,769 per QALY gained are estimated for patients treated with azacitidine, FLAG-IDA, MEC, LDAC, BSC and weighted comparator compared to gilteritinib, respectively.

Table 47 Scenario 3 Results

	Gilteritinib	Azacitidine	FLAG-IDA	MEC	LDAC	BSC	Weighted comparator
Total costs	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
Treatment costs	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
Treatment associated adverse event costs	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]

HSCT Costs	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
Medical Costs	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
Post-Progression treatment Costs	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
Total LYs	3.645	2.125	2.125	2.125	2.125	0.344	2.125
Total QALYs	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
Incremental costs		[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
Incremental LYs		[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
Incremental QALYs		[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
Incremental cost per LY gained		[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
Incremental cost per QALY gained		£ 36,099	£ 38,498	£ 39,553	£ 42,942	£ 27,762	£ 38,769

B.3.8.4.4 Scenario 4

In this scenario, there is a cure point at 3 years, with an SMR of 3 during the fourth year, an SMR of 2 the fifth year and an SMR of 1 in all subsequent years. This implies a gradual decrease to the background mortality rate, with patients having an equal mortality rate to the general population after 5 years. Long-term medical costs and utility values are applied after three years in this scenario.

Over a lifetime horizon, the total QALYs for gilteritinib, azacitidine, FLAG-IDA, MEC, LDAC, BSC and weighted comparator are estimated as [redacted], respectively, and the model estimates the total costs as [redacted], respectively, during that period. For gilteritinib, drug and drug administration costs are the primary driver for the total costs ([redacted]). For azacitidine, FLAG-IDA, MEC, LDAC, BSC and weighted comparator, medical costs are the primary driver for the total costs ([redacted], respectively).

Considering both cost and effectiveness outcomes, deterministic ICERs £39,100, £41,913, £43,310, £48,166, £35,064, £42,416 per QALY gained are estimated for patients treated with azacitidine, FLAG-IDA, MEC, LDAC, BSC and weighted comparator compared to gilteritinib, respectively.

Table 48 Scenario 4 Results

	Gilteritinib	Azacitidine	FLAG-IDA	MEC	LDAC	BSC	Weighted comparator
Total costs	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
Treatment costs	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
Treatment associated adverse event costs	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
HSCT Costs	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
Medical Costs	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
Post-Progression treatment Costs	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
Total LYs	2.776	1.596	1.596	1.596	1.596	0.329	1.596
Total QALYs	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
Incremental costs		[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
Incremental LYs		[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
Incremental QALYs		[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
Incremental cost per LY gained		[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
Incremental cost per QALY gained		£ 39,100	£ 41,913	£ 43,310	£ 48,166	£ 35,064	£ 42,416

B.3.8.4.5 Scenario 5

Most of the patients (77%) in the ADMIRAL trial who received HSCT achieved a composite complete remission (CRc).[29] In this final scenario, we only consider HSCT use among patients who achieved CRc in the trial observation. This reduces the HSCT

rate for gilteritinib from 25.5% to [redacted] and from 15.3% to [redacted] for salvage chemotherapy. The HSCT rate remains at 0% for BSC.

Over a lifetime horizon, the total QALYs for gilteritinib, azacitidine, FLAG-IDA, MEC, LDAC, BSC and weighted comparator are estimated as [redacted], respectively, and the model estimates the total costs as [redacted], respectively, during that period. For gilteritinib, drug and drug administration costs are the primary driver for the total costs ([redacted]). For azacitidine, FLAG-IDA, MEC, LDAC, BSC and weighted comparator, medical costs are the primary driver for the total costs ([redacted], respectively).

Considering both cost and effectiveness outcomes, deterministic ICERs of £41,761, £44,134, £45,304, £49,353, £39,222, £44,549 per QALY gained are estimated for patients treated with azacitidine, FLAG-IDA, MEC, LDAC, BSC and weighted comparator compared to gilteritinib, respectively.

Table 49 Scenario 5 Results

	Gilteritinib	Azacitidine	FLAG-IDA	MEC	LDAC	BSC	Weighted comparator
Total costs	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
Treatment costs	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
Treatment associated adverse event costs	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
HSCT Costs	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
Medical Costs	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
Post-Progression treatment Costs	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
Total LYs	2.768	1.367	1.367	1.367	1.367	0.330	1.367
Total QALYs	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
Incremental costs		[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
Incremental LYs		[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
Incremental QALYs		[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]

Incremental cost per LY gained		[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
Incremental cost per QALY gained		£ 41,761	£ 44,134	£ 45,304	£ 49,353	£ 39,222	£ 44,549

B.3.9 Subgroup analysis

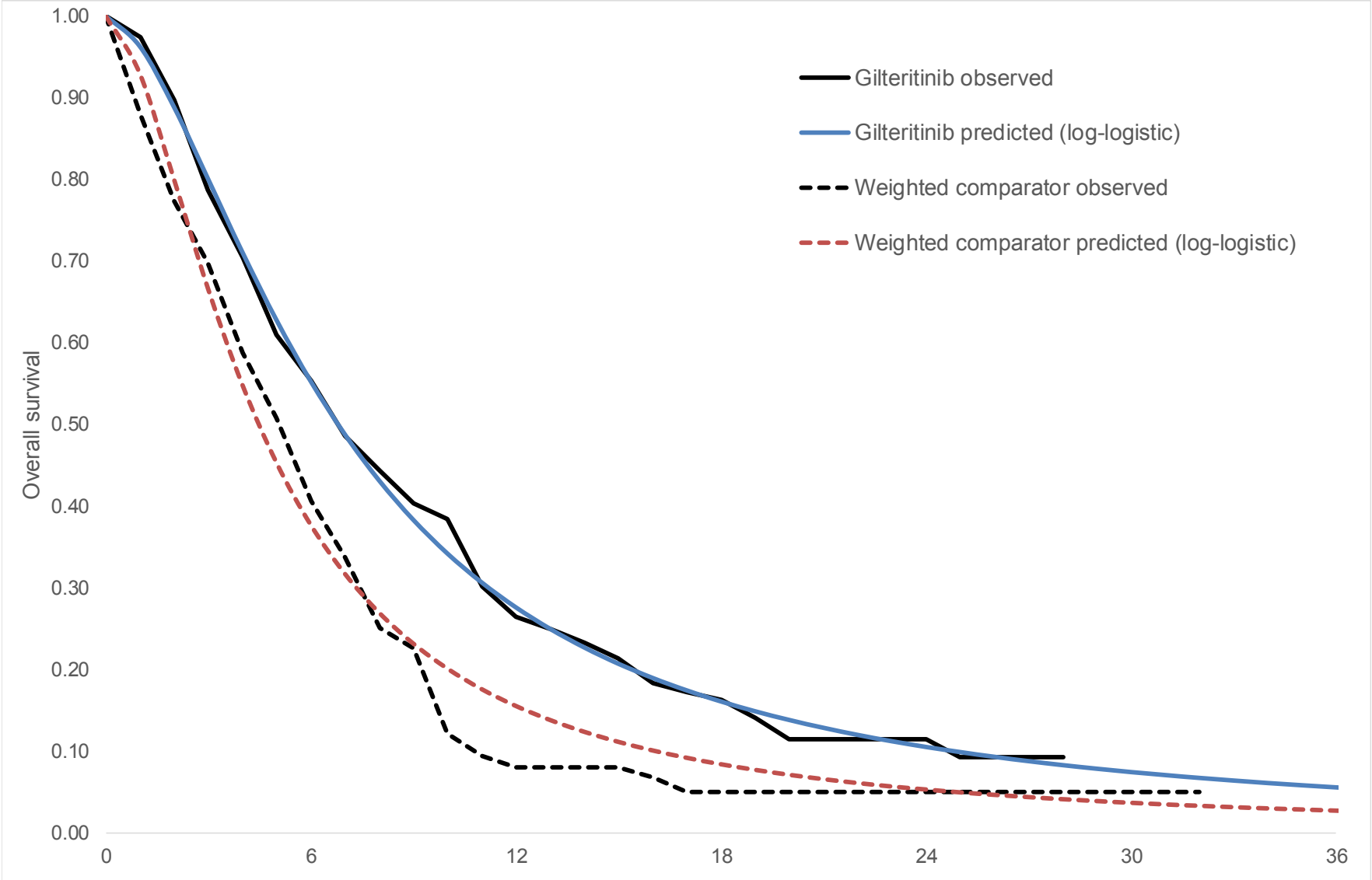
None conducted.

B.3.10 Validation of cost-effectiveness analysis

The current model uses data for gilteritinib and its comparators from the ADMIRAL phase III clinical trial to simulate the survival of patients with relapsed or refractory FLT3 mutation positive AML over a lifetime horizon. Clinical trial data are not available for the entire time horizon, so parametric extrapolation was used to reflect the progress of the disease beyond the period of the trial.

The validity of the predicted survival was assessed by comparing modelled efficacy outcomes against the original sources that informed the efficacy inputs. As the model estimates efficacy stratified by HSCT status, the primary comparison was between observed OS and the predicted OS curves without HSCT based on log-logistic distributions. The weighted comparator observed and predicted OS curves are based on the average of the salvage chemotherapy alternatives included in the ADMIRAL phase III trial weighted by the proportions observed in the trial. These curves are illustrated in Figure 30.

Figure 30 Observed vs Predicted Overall Survival, gilteritinib and Weighted Comparator



Upon visual inspection, the predicted curves fit the observed data well, although the curves provide a better fit earlier in the trial where there are more observations. Of note, the predicted salvage chemotherapy curve appears to slightly overestimate the survival of this patient group, representing a conservative model assumption.

External validation has also been performed to compare the model's predicted long-term survival with the OS data reported in the literature. It is challenging to validate the gilteritinib arm externally because FLT3 inhibitors have not yet been approved for relapsed or refractory AML indication in most countries. Therefore, there is limited external evidence to inform the long-term survival of this type of treatment. External validation of the salvage chemotherapy arm predictions was more straightforward. In general, the long-term OS predictions are consistent with the data reported in the literature.[47,83]

The predicted OS at year 5 for salvage chemotherapy was estimated as 7% in the CEA model, while the observed 5-year OS of relapsed or refractory AML patients receiving standard salvage chemotherapy ranges between 6% to 7%. Long-term OS data for patients with both relapsed or refractory AML and FLT3 mutation is limited and only one study has been identified.[64] The reported 5-year OS for patients with FLT3 mutation positive relapsed or refractory AML receiving conventional chemotherapy was 8%.[47,83]

Finally, the assumptions related to the cure point and the survival rate after the cure point have been validated by clinical experts. Given the uncertainty surrounding long-term OS extrapolation, scenario analyses have been conducted using alternative cure assumptions. The model results are robust to changing scenarios.

B.3.11 Interpretation and conclusions of economic evidence

Gilteritinib is a cost-effective option compared to salvage chemotherapy and meets the NICE definition for end of life criteria.

The model shows that gilteritinib offers marked survival benefits to patients with relapsed or refractory FLT3 mutation positive AML in terms of LYs and QALYs, in

comparison with salvage chemotherapy and best supportive care. The results demonstrate that gilteritinib is a highly effective treatment with good economic value. With an assumed per-cycle price of [redacted] (with a PAS simple discount of [redacted]), the base case cost per QALY gained with gilteritinib relative to the weighted comparator was £47,695. The cost per QALY gained relative to the individual comparators ranged between £35,773 and £52,954. Deterministic and probabilistic sensitivity analyses suggest that the cost-effectiveness results are robust across plausible ranges, and the PSA suggested an 80% likelihood of gilteritinib being cost-effective relative to the weighted comparator at a threshold willingness to pay of £50,000.

B.4 References

1. Deschler B, Lübbert M. Acute myeloid leukemia: epidemiology and etiology. *Cancer*. 2006;107(9):2099-2107. doi:10.1002/cncr.22233

2. Cancer Research UK. Acute myeloid leukaemia (AML) incidence statistics. Cancer Research UK. https://www.cancerresearchuk.org/health-professional/cancer-statistics/statistics-by-cancer-type/leukaemia-aml/incidence. Published May 15, 2015. Accessed June 16, 2019.

3. Meyers J, Yu Y, Kaye JA, Davis KL. Medicare fee-for-service enrollees with primary acute myeloid leukemia: an analysis of treatment patterns, survival, and healthcare resource utilization and costs. *Appl Health Econ Health Policy*. 2013;11(3):275-286. doi:10.1007/s40258-013-0032-2

4. Oran B, Weisdorf DJ. Survival for older patients with acute myeloid leukemia: a population-based study. *Haematologica*. 2012;97(12):1916-1924. doi:10.3324/haematol.2012.066100

5. Fey MF, Buske C, ESMO Guidelines Working Group. Acute myeloblastic leukaemias in adult patients: ESMO Clinical Practice Guidelines for diagnosis, treatment and follow-up. *Ann Oncol Off J Eur Soc Med Oncol*. 2013;24 Suppl 6:vi138-143. doi:10.1093/annonc/mdt320

6. Döhner H, Estey E, Grimwade D, et al. Diagnosis and management of AML in adults: 2017 ELN recommendations from an international expert panel. *Blood*. 2017;129(4):424-447. doi:10.1182/blood-2016-08-733196

7. London Cancer. Acute Myeloid Leukaemia Guidelines 2015-16 v1.0. December 2015. http://www.londoncancer.org/media/111744/acute-myeloid-leukaemia-london-cancer-guidelines-2015-.pdf.

8. Ramos NR, Mo CC, Karp JE, Hourigan CS. Current Approaches in the Treatment of Relapsed and Refractory Acute Myeloid Leukemia. *J Clin Med*. 2015;4(4):665-695. doi:10.3390/jcm4040665

9. Byrd JC, Mrózek K, Dodge RK, et al. Pretreatment cytogenetic abnormalities are predictive of induction success, cumulative incidence of relapse, and overall survival in adult patients with de novo acute myeloid leukemia: results from Cancer and Leukemia Group B (CALGB 8461). *Blood*. 2002;100(13):4325-4336. doi:10.1182/blood-2002-03-0772

10. Goldstone AH, Burnett AK, Wheatley K, et al. Attempts to improve treatment outcomes in acute myeloid leukemia (AML) in older patients: the results of the United Kingdom Medical Research Council AML11 trial. *Blood*. 2001;98(5):1302-1311. doi:10.1182/blood.v98.5.1302

11. Ossenkoppele GJ, Janssen JJWM, van de Loosdrecht AA. Risk factors for relapse after allogeneic transplantation in acute myeloid leukemia. *Haematologica*. 2016;101(1):20-25. doi:10.3324/haematol.2015.139105

12. National Comprehensive Cancer Network. NCCN Guidelines - Acute Myeloid Leukemia, Version 3.2018. 2018.

13. Conway O'Brien E, Prideaux S, Chevassut T. The epigenetic landscape of acute myeloid leukemia. *Adv Hematol*. 2014;2014:103175. doi:10.1155/2014/103175

14. Kayser S, Schlenk RF. Targeting the FLT3 mutation in acute myeloid leukemia. *Eur Oncol Haematol*. 2017;13(2):139-146. doi:10.17925/EOH.2017.13.02.139

15. Meshinchi S, Appelbaum FR. Structural and functional alterations of FLT3 in acute myeloid leukemia. *Clin Cancer Res Off J Am Assoc Cancer Res*. 2009;15(13):4263-4269. doi:10.1158/1078-0432.CCR-08-1123

16. Sotak ML, Marin M, Coombs J, Schiller GJ, Teitelbaum A. Burden of Illness of FLT-3 Mutated Acute Myeloid Leukemia (AML). *Blood*. 2011;118(21):4765-4765.

17. Patel JP, Gönen M, Figueroa ME, et al. Prognostic Relevance of Integrated Genetic Profiling in Acute Myeloid Leukemia. *N Engl J Med*. 2012;366(12):1079-1089. doi:10.1056/NEJMoa1112304

18. Alvarado Y, Kantarjian HM, Luthra R, et al. Treatment with FLT3 inhibitor in patients with FLT3-mutated acute myeloid leukemia is associated with development of secondary FLT3-tyrosine kinase domain mutations. *Cancer*. 2014;120(14):2142-2149. doi:10.1002/cncr.28705

19. Döhner H, Weisdorf DJ, Bloomfield CD. Acute Myeloid Leukemia. *N Engl J Med*. 2015;373(12):1136-1152. doi:10.1056/NEJMra1406184

20. Levis M. FLT3 mutations in acute myeloid leukemia: what is the best approach in 2013? *Hematol Am Soc Hematol Educ Program*. 2013;2013:220-226. doi:10.1182/asheducation-2013.1.220

21. Ilyas AM, Ahmad S, Faheem M, et al. Next Generation Sequencing of Acute Myeloid Leukemia: Influencing Prognosis. *BMC Genomics*. 2015;16(Suppl 1):S5. doi:10.1186/1471-2164-16-S1-S5

22. Gerds AT, Appelbaum FR. To Transplant or Not to Transplant for Adult Acute Myeloid Leukemia: An Ever-Evolving Decision. *Clin Adv Hematol Oncol HO*. 2012;10(10):655-662.

23. British Committee for Standards in Haematology, Milligan DW, Grimwade D, et al. Guidelines on the management of acute myeloid leukaemia in adults. *Br J Haematol*. 2006;135(4):450-474. doi:10.1111/j.1365-2141.2006.06314.x

24. Dewing DRT, Tholouli DE, Dennis DM. Guidelines for the management of Acute Myeloid Leukaemia. :27.

25. National Health Service. Guidelines for the diagnosis and management of Adult Acute Leukaemia within Anglia Cancer Network. 2013. https://www.canceralliance.co.uk/download/0fae2727-4bb9-47e7-8cab-34a031060c03.

26. National Health Service. Guidelines for the Management of Acute Myeloid Leukaemia in Adults (AML). 2011. https://www.uhb.nhs.uk/Downloads/pdf/CancerPbAcuteMyeloidLeukaemia.pdf.

27. European Medicines Agency. Rydapt. European Medicines Agency. https://www.ema.europa.eu/en/medicines/human/EPAR/rydapt. Published September 17, 2018. Accessed June 16, 2019.

28. Perl AE, Altman JK, Cortes J, et al. Selective inhibition of FLT3 by gilteritinib in relapsed or refractory acute myeloid leukaemia: a multicentre, first-in-human, open-label, phase 1-2 study. *Lancet Oncol*. 2017;18(8):1061-1075. doi:10.1016/S1470-2045(17)30416-3

29. Astellas Data on File: ADMIRAL CSR

30. European Medicines Agency. Good clinical practice. European Medicines Agency. https://www.ema.europa.eu/en/human-regulatory/research-development/compliance/good-clinical-practice. Published September 17, 2018. Accessed June 16, 2019.

31. Griffin JD, Yang H, Song Y, Kinrich D, Bui CN. Treatment Patterns and Healthcare Resource Utilization in Patients with FLT3-Mut and FLT3-Wt Acute Myeloid Leukemia: A Multi-Country Medical Chart Study. *Blood*. 2017;130(Suppl 1):2186-2186.

32. Woods B, Sideris E, Palmer S, Latimer N, Soares M. Partitioned survival analysis for decision modelling in health care: a critical review. June 2017. http://scharr.dept.shef.ac.uk/nicedsu/wp-content/uploads/sites/7/2017/06/Partitioned-Survival-Analysis-final-report.pdf.

33. Nath R, Reddy V, Kapur A, et al. Survival of Relapsed/Refractory Acute Myeloid Leukemia (R/R AML) Patients Receiving Stem Cell Transplantation (SCT). *Biol Blood Marrow Transplant*. 2019;25(3):S125. doi:10.1016/j.bbmt.2018.12.407

34. Hills RK, Burnett AK, Gale R, Linch DC, Gilkes A, Russell NH. Outcomes in Relapsed/Refractory Patients with FLT3-ITD Mutated AML Are Poor When Treated with Non-Targeted Therapy with a Potential Role for Stem Cell Transplantation: Results from the NCRI AML Trials. *Blood*. 2018;132(Suppl 1):1392-1392. doi:10.1182/blood-2018-99-116542

35. Aly A, Ray S, Shah A, Botteman M. Real-World Differences in Characteristics and Survival of Relapsed AML Patients with and without Transplant. *Blood*. 2018;132(Suppl 1):2297-2297. doi:10.1182/blood-2018-99-118621

36. Schiller GJ, Tuttle P, Desai P. Allogeneic Hematopoietic Stem Cell Transplantation in FLT3-ITD-Positive Acute Myelogenous Leukemia: The Role for FLT3 Tyrosine Kinase Inhibitors Post-Transplantation. *Biol Blood Marrow Transplant J Am Soc Blood Marrow Transplant*. 2016;22(6):982-990. doi:10.1016/j.bbmt.2016.01.013

37. Lin P-H, Lin C-C, Yang H-I, et al. Prognostic impact of allogeneic hematopoietic stem cell transplantation for acute myeloid leukemia patients with internal tandem duplication of FLT3. *Leuk Res*. 2013;37(3):287-292. doi:10.1016/j.leukres.2012.10.005

38. National Institute for Health and Care Excellence (NICE). Brentuximab vedotin for treating relapsed or refractory systemic anaplastic large cell lymphoma. Technology appraisal guidance [ID512]. August 2017. https://www.nice.org.uk/guidance/ta478/documents/committee-papers-2.

39. Hettle R, Corbett M, Hinde S, et al. The assessment and appraisal of regenerative medicines and cell therapy products: an exploration of methods for review, economic evaluation and appraisal. *Health Technol Assess Winch Engl*. 2017;21(7):1-204. doi:10.3310/hta21070

40. National Institute for Health and Care Excellence (NICE). Axicabtagene ciloleucel for treating diffuse large B-cell lymphoma and primary mediastinal large B-cell lymphoma after 2 or more systemic therapies. Technology appraisal guidance [TA559]. NICE. https://www.nice.org.uk/guidance/TA559. Published January 23, 2019.

41. National Institute for Health and Care Excellence (NICE). Midostaurin for untreated acute myeloid leukaemia. Technology appraisal guidance [TA523]. NICE. https://www.nice.org.uk/guidance/ta523/history. Published June 13, 2018.

42. National Institute for Health and Care Excellence (NICE). Azacitidine for the treatment of myelodysplastic syndromes, chronic myelomonocytic leukaemia and acute myeloid leukaemia. Technology appraisal guidance [TA218]. March 2011. https://www.nice.org.uk/guidance/ta218/resources/azacitidine-for-the-treatment-of-myelodysplastic-syndromes-chronic-myelomonocytic-leukaemia-and-acute-myeloid-leukaemia-pdf-82600256379589.

43. Evers G, Beelen DW, Braess J, et al. Outcome of Patients with Acute Myeloid Leukemia (AML) Undergoing Allogeneic Hematopoietic Stem Cell Transplantation (HSCT) Beyond First Complete Remission (CR1). *Blood*. 2018;132(Suppl 1):4649-4649. doi:10.1182/blood-2018-99-116964

44. National Institute for Health and Care Excellence (NICE). Tisagenlecleucel for treating relapsed or refractory B-cell acute lymphoblastic leukaemia in people aged up to 25 years. Technology appraisal guidance [TA554]. NICE. https://www.nice.org.uk/guidance/ta554. Published December 21, 2018.

45. Stahl M, DeVeaux M, Montesinos P, et al. Performance of the Medical Research Council (MRC) and the Leukemia Research Foundation (LRF) score in predicting survival benefit with

hypomethylating agent use in patients with relapsed or refractory acute myeloid leukemia. *Leuk Lymphoma*. 2019;60(1):246-249. doi:10.1080/10428194.2018.1468893

46. Latimer NR. *NICE DSU Technical Support Document 14: Survival Analysis For Economic Evaluations Alongside Clinical Trials - Extrapolation with Patient-Level Data*. London: National Institute for Health and Care Excellence (NICE); 2013. http://nicedsu.org.uk/wp-content/uploads/2016/03/NICE-DSU-TSD-Survival-analysis.updated-March-2013.v2.pdf.

47. Sarkozy C, Gardin C, Gachard N, et al. Outcome of older patients with acute myeloid leukemia in first relapse. *Am J Hematol*. 2013;88(9):758-764. doi:10.1002/ajh.23498

48. Ustun C, Ruggeri A, Weisdorf DJ. UCB HCT in FLT3+ AML. *Oncotarget*. 2017;8(47):81733-81734. doi:10.18632/oncotarget.21048

49. Astellas Data on File: UK HSCT Rates in 2017

50. Deol A, Sengsayadeth S, Ahn KW, et al. Does FLT3 mutation impact survival after hematopoietic stem cell transplantation for acute myeloid leukemia? A Center for International Blood and Marrow Transplant Research (CIBMTR) analysis. *Cancer*. 2016;122(19):3005-3014. doi:10.1002/cncr.30140

51. Guyot P, Ades A, Ouwens MJ, Welton NJ. Enhanced secondary analysis of survival data: reconstructing the data from published Kaplan-Meier survival curves. *BMC Med Res Methodol*. 2012;12(1):9. doi:10.1186/1471-2288-12-9

52. Steckel NK, Groth C, Mikesch J-H, et al. High-dose melphalan-based sequential conditioning chemotherapy followed by allogeneic haematopoietic stem cell transplantation in adult patients with relapsed or refractory acute myeloid leukaemia. *Br J Haematol*. 2018;180(6):840-853. doi:10.1111/bjh.15137

53. Fong CY, Grigoriadis G, Hocking J, et al. Fludarabine, cytarabine, granulocyte-colony stimulating factor and amsacrine: an effective salvage therapy option for acute myeloid leukemia at first relapse. *Leuk Lymphoma*. 2013;54(2):336-341. doi:10.3109/10428194.2012.713479

54. Frazer J, Couban S, Doucette S, Shivakumar S. Characteristics predicting outcomes of allogeneic stem-cell transplantation in relapsed acute myelogenous leukemia. *Curr Oncol*. 2017;24(2):e123-e130. doi:10.3747/co.24.3485

55. Schmid C, Schleuning M, Schwerdtfeger R, et al. Long-term survival in refractory acute myeloid leukemia after sequential treatment with chemotherapy and reduced-intensity conditioning for allogeneic stem cell transplantation. *Blood*. 2006;108(3):1092-1099. doi:10.1182/blood-2005-10-4165

56. Oran B, Cortes J, Beitinjaneh A, et al. Allogeneic Transplantation in First Remission Improves Outcomes Irrespective of FLT3-ITD Allelic Ratio in FLT3-ITD-Positive Acute Myelogenous Leukemia. *Biol Blood Marrow Transplant J Am Soc Blood Marrow Transplant*. 2016;22(7):1218-1226. doi:10.1016/j.bbmt.2016.03.027

57. Schlenk RF, Kayser S, Bullinger L, et al. Differential impact of allelic ratio and insertion site in FLT3-ITD-positive AML with respect to allogeneic transplantation. *Blood*. 2014;124(23):3441-3449. doi:10.1182/blood-2014-05-578070

58. Song Y, Magenau J, Li Y, et al. FLT3 mutational status is an independent risk factor for adverse outcomes after allogeneic transplantation in AML. *Bone Marrow Transplant*. 2016;51(4):511-520. doi:10.1038/bmt.2015.170

59. Ustun C, Giannotti F, Zhang M-J, et al. Outcomes of UCB transplantation are comparable in FLT3+ AML: results of CIBMTR, EUROCORD and EBMT collaborative analysis. *Leukemia*. 2017;31(6):1408-1414. doi:10.1038/leu.2017.42

60. Schlenk RF, Döhner H, Döhner K, et al. Event-Free Survival Is a Surrogate for Overall Survival in Patients Treated for Acute Myeloid Leukemia. *Blood*. 2015;126(23):3744-3744.

61. Shimoni A, Labopin M, Savani B, et al. Long-term survival and late events after allogeneic stem cell transplantation from HLA-matched siblings for acute myeloid leukemia with myeloablative compared to reduced-intensity conditioning: a report on behalf of the acute leukemia working party of European group for blood and marrow transplantation. *J Hematol OncolJ Hematol Oncol*. 2016;9(1). doi:10.1186/s13045-016-0347-1

62. Bejanyan N, Weisdorf DJ, Logan BR, et al. Survival of patients with acute myeloid leukemia relapsing after allogeneic hematopoietic cell transplantation: a center for international blood and marrow transplant research study. *Biol Blood Marrow Transplant J Am Soc Blood Marrow Transplant*. 2015;21(3):454-459. doi:10.1016/j.bbmt.2014.11.007

63. Othus M, Garcia-Manero G, Godwin J, et al. Associations between Complete Remissions (CRs) with 7+3 Induction Chemotherapy for Acute Myeloid Leukemia and 2-3 Year Survival ("Potential Cure") over the Past Four Decades: Analysis of SWOG Trial Data. *Blood*. 2017;130(Suppl 1):1301-1301.

64. Takahashi K, Kantarjian H, Pemmaraju N, et al. Salvage therapy using FLT3 inhibitors may improve long-term outcome of relapsed or refractory AML in patients with FLT3-ITD. *Br J Haematol*. 2013;161(5):659-666. doi:10.1111/bjh.12299

65. van Hout B, Janssen MF, Feng Y-S, et al. Interim scoring for the EQ-5D-5L: mapping the EQ-5D-5L to EQ-5D-3L value sets. *Value Health J Int Soc Pharmacoeconomics Outcomes Res*. 2012;15(5):708-715. doi:10.1016/j.jval.2012.02.008

66. 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. doi:10.1007/s40273-018-0704-8

67. Janssen B, Szende A. Population Norms for the EQ-5D. In: Szende A, Janssen B, Cabases J, eds. *Self-Reported Population Health: An International Perspective Based on EQ-5D*. Dordrecht: Springer Netherlands; 2014:19-30. doi:10.1007/978-94-007-7596-1_3

68. Swinburn P, Lloyd A, Nathan P, Choueiri TK, Cella D, Neary MP. Elicitation of health state utilities in metastatic renal cell carcinoma. *Curr Med Res Opin*. 2010;26(5):1091-1096. doi:10.1185/03007991003712258

69. Doyle S, Lloyd A, Walker M. Health state utility scores in advanced non-small cell lung cancer. *Lung Cancer Amst Neth*. 2008;62(3):374-380. doi:10.1016/j.lungcan.2008.03.019

70. Lloyd A, Nafees B, Narewska J, Dewilde S, Watkins J. Health state utilities for metastatic breast cancer. *Br J Cancer*. 2006;95(6):683-690. doi:10.1038/sj.bjc.6603326

71. Nafees B, Stafford M, Gavriel S, Bhalla S, Watkins J. Health state utilities for non small cell lung cancer. *Health Qual Life Outcomes*. 2008;6:84. doi:10.1186/1477-7525-6-84

72. Monthly Index of Medical Specialities (MIMS). Drug listings by brand and generic drugs. MIMS. https://www.mims.co.uk/drugs.

73. UK Department of Health and Social Care. Drugs and pharmaceutical electronic market information tool (eMIT). GOV.UK. https://www.gov.uk/government/publications/drugs-and-pharmaceutical-electronic-market-information-emit. Published April 8, 2019.

74. UK Department of Health National Health Services (NHS). NHS Reference Costs 2017-2018. National Schedule of Reference Costs: The Main Schedule. NHS. https://improvement.nhs.uk/resources/reference-costs/.

75. NHS Digital. TRUD - NHS Classifications OPCS-4. https://hscic.kahootz.com/gf2.ti/f/762498/27838501.1/PDF/-/ChemRegClinCodingStandGuidApl2017.pdf. Accessed June 16, 2019.

76. Astellas Data on File: Adelphi AML Treatment Pattern Study

77. Curtis LA, Burns A. Unit Costs of Health and Social Care 2018. doi:Curtis, Lesley A. and Burns, Amanda (2018) Unit Costs of Health and Social Care 2018. Project report. University of Kent 10.22024/UniKent/01.02.70995 <https://doi.org/10.22024/UniKent%2F01.02.70995>. (doi:10.22024/UniKent/01.02.70995 <https://doi.org/10.22024/UniKent%2F01.02.70995>)

78. UK Department of Health National Health Services (NHS). NHS Blood and Transplant Price List 2017-18. 2017.

79. 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 CE*. 2018;16:33. doi:10.1186/s12962-018-0153-4

80. NICE. Pixantrone monotherapy for treating multiply relapsed or refractory aggressive non-Hodgkin's B−cell lymphoma | Guidance. https://www.nice.org.uk/guidance/ta306. Published February 26, 2014. Accessed June 16, 2019.

81. Theo Georghiou, Martin Bardsley. *Exploring the Cost of Care at the End of Life*. The Nuffield Trust; 2014. https://www.nuffieldtrust.org.uk/files/2017-01/end-of-life-care-web-final.pdf. Accessed June 16, 2019.

82. Han-I W. Cancer treatment pathways: AML a detailed costing study example. 2011.

83. Rowe JM, Li X, Cassileth PA, 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):546-546.

B.5 Appendices

NATIONAL INSTITUTE FOR HEALTH AND CARE EXCELLENCE

Single technology appraisal

Gilteritinib for treating relapsed or refractory acute myeloid leukaemia [ID1484]

Clarification questions

July 2019

File name	Version	Contains confidential information	Date
ID1484 Gilteritinib AML Clarification Questions ACIC	Redacted	Yes	25th July 2019

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

Literature searching

A1. The systematic literature review (SLR) search was undertaken up until October 2018 (over 6 months ago). Please confirm that no new or completed studies have been published since then. Alternatively, please update the SLR searches and confirm that no further studies are relevant to the decision problem.

The SLR searches for clinical evidence in EMBASE, Medline and Cochrane have been updated. The searches in Cochrane were limited to the period between 01 October 2018 and 15 July 2019. For EMBASE and Medline the timeframe was restricted to between 01 January 2018 and the 17 July 2019. The results of these two searches are given in Table 1 and Table 2, respectively.

Overall, seven relevant publications are new. A list of these publications and their abstracts are given in Table 3. Of these, four provide additional data on studies already captured in the initial SLR: two publications relate to QuANTUM-R (Cortes 2019a[1]/2019b[2]), one to ADMIRAL (Wang 2018[3]), and one provides safety data for quizartinib (Kang 2018[4]). In addition, Foran 2018[5] provides data for the North American Intergroup E2906 phase III trial in patients aged ≥60 years, Hills 2018[6] provides data for the NCRI AML trials, and Marconi 2018[7] provides data for patients treated with tyrosine kinase inhibitors without specifying which drugs patients received.

Table 1 Searches in Cochrane

Concept	ID	Search String	Hits
Disease	#1	MeSH descriptor: "leukemia, myeloid, acute" explode all trees	4,224
	#2	“acute granulocytic leukemia”:ti,ab,kw or “acute granulocytic leukaemia”:ti,ab,kw or “acute nonlymphocytic leukemia”:ti,ab,kw or “acute nonlymphocytic leukaemia”:ti,ab,kw or “acute non-lymphocytic leukemia”:ti,ab,kw or “acute non-lymphocytic leukaemia”:ti,ab,kw or “acute myelogenous leukemia”:ti,ab,kw or “acute myelogenous leukaemia”:ti,ab,kw	889
	#3	“acute myeloid leukemia”:ti,ab,kw or “acute myeloid leukaemia”:ti,ab,kw or “acute myelocytic leukemia”:ti,ab,kw or “acute myelocytic leukaemia”:ti,ab,kw or “acute myeloblastic leukemia”:ti,ab,kw or “acute myeloblastic leukaemia”:ti,ab,kw or “acute non-lymphoblastic leukemia”:ti,ab,kw or “acute non-lymphoblastic leukaemia”:ti,ab,kw or “acute nonlymphoblastic leukemia”:ti,ab,kw or “acute nonlymphoblastic leukemia”:ti,ab,kw	3,540
	#4	#1 OR #2 OR #3	4,725
	#5	"refractory":ti,ab,kw or "relapsed":ti,ab,kw or "relapse":ti,ab,kw	44,406
	#6	MeSH descriptor: "fms like tyrosine kinase 3" explode all trees	35
	#7	"cd135":ti,ab,kw or "flt3":ti,ab,kw or "flt 3":ti,ab,kw or "fms like tyrosine kinase 3":ti,ab,kw	478
	#8	"poor cytogenetics":ti,ab,kw or "high risk cytogenetics":ti,ab,kw or "high risk cytogenetic":ti,ab,kw or "poor risk cytogenetics":ti,ab,kw or "poor risk cytogenetic":ti,ab,kw	230
	#9	#6 OR #7 OR #8	691
	#10	#4 AND #5 AND #9	232
Therapies	#11	MeSH descriptor: "cytarabine" explode all trees	1,194
	#12	MeSH descriptor: "azacitidine" explode all trees	242
	#13	MeSH descriptor: "mitoxantrone" explode all trees	478
	#14	MeSH descriptor: "etoposide" explode all trees	1,624
	#15	MeSH descriptor: "granulocyte colony stimulating factor" explode all trees	0
	#16	MeSH descriptor: "idarubicin" explode all trees	266
	#17	"gilteritinib":ti,ab,kw or "asp2215":ti,ab,kw	34
	#18	"hypomethylating":ti,ab,kw or "5 azacytidine":ti,ab,kw or "decitabine":ti,ab,kw or “Dacogen“:ti,ab,kw	556
	#19	"sorafenib":ti,ab,kw or "nexavar":ti,ab,kw or "bay 43 9006":ti,ab,kw or "bay 54 9085":ti,ab,kw	1,634
	#20	"quizartinib":ti,ab,kw or "ac220":ti,ab,kw or "ibrutinib":ti,ab,kw or "Imbruvica":ti,ab,kw or "PCI-32765":ti,ab,kw or "CRA-032765":ti,ab,kw or "JNJ-54179060":ti,ab,kw	480
	#21	"crenolanib":ti,ab,kw or "cp 868,596":ti,ab,kw	24
	#22	"cytarabine":ti,ab,kw or "Depocyt":ti,ab,kw or "AraC":ti,ab,kw or "cytosine arabinoside":ti,ab,kw or "lodac":ti,ab,kw	3,197
	#23	“Azacitidine“:ti,ab,kw or “vidaza“:ti,ab,kw or “mitoxantrone“:ti,ab,kw or “Novantrone“:ti,ab,kw or “etoposide“:ti,ab,kw or "hidac":ti,ab,kw or “VePesid“:ti,ab,kw or “Etopophos“:ti,ab,kw or "mec":ti,ab,kw	6,149
	#24	“G-CSF“:ti,ab,kw or “granulocyte colony stimulating factor“:ti,ab,kw or “Neupogen“:ti,ab,kw or “fludarabine“:ti,ab,kw or “idarubicin“:ti,ab,kw or “Idamycin pfs“:ti,ab,kw or "flag ida":ti,ab,kw	6,058
	#25	#11 OR #12 OR #13 OR #14 OR #15 OR #16 OR #17 OR #18 OR #19 OR #20 OR #21 OR #22 OR #23 OR #24	15,095

Concept	ID	Search String	Hits
Study design	#26	MeSH descriptor: “Double Blind Method” explode all trees	131,768
	#27	MeSH descriptor: “Single Blind Method” explode all trees	19,354
	#28	MeSH descriptor: "cohort studies" explode all trees	142,437
	#29	MeSH descriptor: “clinical trial” explode all trees	146
	#30	MeSH descriptor: "clinical trials as topic" explode all trees	47,250
	#31	MeSH descriptor: "placebos" explode all trees	23,558
	#32	MeSH descriptor: "randomized controlled trial" explode all trees	125
	#33	MeSH descriptor: "random allocation" explode all trees	20,586
	#34	MeSH descriptor: "randomized controlled trials as topic" explode all trees	13,878
	#35	"controlled clinical trial":pt or “randomized controlled trial”: pt or "multicenter study":pt or "clinical trial":pt	560,412
	#36	“double blind”:ti,ab,kw or “double blinded”:ti,ab,kw or RCT:ti,ab,kw or Randomisation:ti,ab,kw or Randomization:ti,ab,kw or controlled:ti,ab,kw or controled:ti,ab,kw or control:ti,ab,kw or Placebo:ti,ab,kw or Trial:ti,ab,kw or “randomly allocated”:ti,ab,kw or “prospective study”:ti,ab,kw or “prospective studies”:ti,ab,kw or “prospective trial”:ti,ab,kw or “prospective trials”:ti,ab,kw or “clinical trial”:ti,ab,kw or “clinical trials”:ti,ab,kw	1,128,009
	#37	(Study:ti,ab,kw or studies:ti,ab,kw) and (open:ti,ab,kw or "open-label":ti,ab,kw or "non-randomised":ti,ab,kw or "non-randomized":ti,ab,kw or "cohort":ti,ab,kw or "single-arm":ti,ab,kw)	131,179
	#38	#26 OR #27 OR #28 OR #29 OR #30 OR #31 OR #32 OR #33 OR #34 OR #35 OR #36 OR #37	1,242,603
	#39	MeSH descriptor: [Observational Study] explode all trees	3
	#40	MeSH descriptor: [Observational Studies as Topic] explode all trees	48
	#41	#39 OR #40	51
	#42	#38 OR #41	1,242,603
All	#43	#10 AND #25 AND #38	172
	#44	#10 AND #25 AND #42	172
	#45	**Since Oct 01, 2018**	**48**

Table 2 Searches in EMBASE

#	Searches	Results
1	exp acute myeloid leukemia/ or exp leukemia, myeloid, acute/	87,041
2	("acute granulocytic leukemia" or "acute granulocytic leukaemia" or "acute nonlymphocytic leukemia" or "acute nonlymphocytic leukaemia" or "acute non-lymphocytic leukemia" or "acute non-lymphocytic leukaemia" or "acute myelogenous leukemia" or "acute myelogenous leukaemia").ti,ab.	18,870

#	Searches	Results
3	("acute myeloid leukemia" or "acute myeloid leukaemia" or "acute myelocytic leukemia" or "acute myelocytic leukaemia" or "acute myeloblastic leukemia" or "acute myeloblastic leukaemia" or "acute non-lymphoblastic leukemia" or "acute non-lymphoblastic leukaemia" or "acute nonlymphoblastic leukemia" or "acute nonlymphoblastic leukemia").ti,ab.	91,207
4	1 or 2 or 3	157,551
5	(refractory or relapsed or relapse).ti,ab.	626,312
6	exp CD135 antigen/ or exp fms like tyrosine kinase 3/	8,164
7	(cd135 or flt3 or flt 3 or fms like tyrosine kinase 3).ti,ab.	17,342
8	("poor cytogenetics" or "high risk cytogenetics" or "high risk cytogenetic" or "poor risk cytogenetics" or "poor risk cytogenetic").ti,ab.	2,347
9	6 or 7 or 8	21,267
10	exp cytarabine/ or exp azacitidine/ or exp mitoxantrone/ or exp etoposide/ or exp granulocyte colony stimulating factor/ or exp idarubicin/	236,279
11	("gilteritinib" or "asp2215").ti,ab.	138
12	(hypomethylating or azacytidine or decitabine or Dacogen).ti,ab.	14,754
13	("sorafenib" or "nexavar" or "bay 43 9006" or "bay 54 9085").ti,ab.	21,448
14	("quizartinib" or "ac220" or "ibrutinib" or "Imbruvica" or "PCI-32765" or "CRA-032765" or "JNJ-54179060").ti,ab.	5,907
15	("crenolanib" or "cp 868,596").ti,ab.	221
16	(cytarabine or Depocyt or AraC or cytosine arabinoside or lodac).ti,ab.	30,075
17	(Azacitidine or vidaza or mitoxantrone or Novantrone or etoposide or hidac or VePesid or Etopophos or mec).ti,ab.	74,989
18	(G-CSF or granulocyte colony stimulating factor or Neupogen or fludarabine or idarubicin or Idamycin pfs or flag ida).ti,ab.	69,468
19	exp Double Blind Method/ or exp Single Blind Method/ or exp Single Blind Method/ or exp cohort studies/ or exp clinical trial/ or exp clinical trials as topic/ or exp placebos/ or exp double blind procedure/ or exp single blind procedure/ or exp cohort analysis/ or exp clinical trial/ or exp placebo/ or exp "clinical trial (topic)"/	5,008,007
20	exp controlled clinical trial/ or (controlled clinical trial or randomized controlled trial or multicenter study or clinical trial).mp.	2,907,995
21	((Study or studies) and (open or "open-label" or "non-randomised" or "non-randomized" or "cohort" or "single-arm")).ti,ab.	1,563,777
22	exp randomization/ or exp randomized controlled trial/ or exp "randomized controlled trial (topic)"/ or (double blind or double blinded or RCT or Randomi* or controlled or controled or control or Placebo or Trial or randomly allocated or prospective stud* or prospective trial* or clinical trial*).ti,ab.	8,963,438
23	4 and 5 and 9	3,775
24	10 or 11 or 12 or 13 or 14 or 15 or 16 or 17 or 18	328,449
25	19 or 20 or 21 or 22	12,248,169
26	23 and 24 and 25	1,071
27	remove duplicates from 26	937

#	Searches	Results
28	limit 27 to yr="2018 -Current"	132

Table 3 New Studies Identified in the Searches in Cochrane, EMBASE and Medline conducted from 2018 onwards

Author	Full reference	Abstract
Cortes et al 2019	Cortes JE et al. Efficacy and safety of single-agent quizartinib (Q), a potent and selective FLT3 inhibitor (FLT3i), in patients (pts) with FLT3-internal tandem duplication (FLT3- ITD)-mutated relapsed/refractory (R/R) acute myeloid leukemia (AML) enrolled in the global, phase 3, randomized controlled QuANTUM-R trial. British Journal of Haematology. 2019; 185 (Supplement 1):7	FLT3-ITD mutations occur in = 25% of pts with AML and are associated with high leukemic burden, high risk of relapse, decreased response to salvage therapy, and shorter overall survival (OS). Pts with R/R FLT3-ITD AML have worse prognosis and high unmet medical need. Q is a once-daily, oral, highly potent and selective FLT3i with promising single-agent activity and a manageable safety profile. QuANTUM-R was the first global, phase 3, randomized controlled trial (NCT02039726) to show that a FLT3i prolonged OS vs chemotherapy (SC) in pts with R/R FLT3-ITD AML. Pts aged >= 18 years with FLT3-ITD AML refractory to or relapsed (duration of first remission <= 6 mo) after standard AML therapy, w/wo hematopoietic stem cell transplant (HSCT) were randomized 2:1 to receive Q (53.0 mg [26.5-mg lead-in]) or 1 of 3 preselected investigator's choice (IC) SC: low-dose cytarabine (LoDAC); mitoxantrone, etoposide, and intermediate-dose cytarabine (MEC); or fludarabine, cytarabine, and granulocyte colony stimulating factor with idarubicin (FLAG-IDA). Pts receiving HSCT in the Q arm could resume Q after HSCT. Up to 2 cycles of MEC or FLAG-IDA were permitted; Q or LoDAC were given until lack of benefit, unacceptable toxicity, or HSCT. Prior therapy with midostaurin was allowed. Primary and secondary endpoints were OS and event-free survival (EFS), respectively. Sensitivity analyses for OS and EFS were conducted. Predefined subgroup analyses of OS were also performed. Exploratory endpoints included response rates, duration of CRc, and transplant rate. Treatment-emergent adverse events (TEAEs) included AEs <= 30 days after last dose and treatment-related AEs > 30 days. 367 pts were randomized; 245 to Q and 122 to IC SC (LoDAC, n = 29; MEC, n = 40; FLAG-IDA, n = 53). Four pts randomized to Q and 28 pts randomized to SC did not receive therapy. Median follow- up was 23.5 mo. Six pts were still on initial Q treatment at data cutoff vs 0 for SC. Treatment groups were balanced for baseline characteristics, such as age, response to prior therapy, transplant history, and FLT3-ITD allelic burden. OS hazard ratio (HR) of Q relative to SC was 0.76 (95% CI, 0.58-0.98; stratified log-rank test, 1- sided P = .0177). Median OS was 6.2 (95% CI, 5.3-7.2) vs 4.7 (95% CI, 4.0-5.5) mo, with an estimated 12-mo OS probability of 27% vs 20% in Q and SC arms, respectively. EFS HR was 0.90 (95% CI, 0.70-1.16; stratified log-rank test, 1-sided P = .1071); median EFS was 1.4 (95% CI, 0.0-1.9) vs 0.9 (95% CI, 0.4-1.3) mo, respectively. Sensitivity analyses of OS (censoring for SCT, for use of subsequent FLT3 inhibitors, and in per-protocol population) and EFS (in perprotocol population) supported benefit of Q vs SC, as did OS subgroup analyses. CRc was 48% and 27% in Q and SC arms, respectively. Duration of CRc was 12.1 (95% CI, 10.4-27.1) vs 5.0 (95% CI, 3.3-12.6) wk. Transplant rate was 32% (Q arm) and 12% (SC arm). TEAE rates were comparable, despite longer treatment duration in Q vs SC arms. Most common grade >= 3 TEAEs in both arms were infections and those associated with cytopenia. Only 2 pts discontinued Q due to QT prolongation (both grade 2). QTcF > 500 ms (grade 3) by central laboratory was 3% in the Q arm; no grade 4 QTcF occurred. This report confirms the OS benefit observed with single-agent

Author	Full reference	Abstract
		Q vs SC in pts with R/R FLT3-ITD AML and the favorable Q safety profile, providing evidence of meaningful clinical benefit in pts who have few options.
Cortes et al 2019	Cortes JE et al. Quizartinib versus salvage chemotherapy in relapsed or refractory FLT3-ITD acute myeloid leukaemia (QuANTUM-R): a multicentre, randomised, controlled, open-label, phase 3 trial. The Lancet Oncology 2019; 20(7):984-997	Background: Patients with relapsed or refractory FLT3 internal tandem duplication (FLT3-ITD)-positive acute myeloid leukaemia have a poor prognosis, including high frequency of relapse, poorer response to salvage therapy, and shorter overall survival than those with FLT3 wild-type disease. We aimed to assess whether single-agent quizartinib, an oral, highly potent and selective type II FLT3 inhibitor, improves overall survival versus salvage chemotherapy.; Method(s): QuANTUM-R is a randomised, controlled, phase 3 trial done at 152 hospitals and cancer centres in 19 countries. Eligible patients aged 18 years or older with ECOG performance status 0-2 with relapsed or refractory (duration of first composite complete remission <=6 months) FLT3-ITD acute myeloid leukaemia after standard therapy with or without allogeneic haemopoietic stem-cell transplantation were randomly assigned (2:1; permuted block size of 6; stratified by response to previous therapy and choice of chemotherapy via a phone-based and web-based interactive response system) to quizartinib (60 mg [30 mg lead-in] orally once daily) or investigator's choice of preselected chemotherapy: subcutaneous low-dose cytarabine (subcutaneous injection of cytarabine 20 mg twice daily on days 1-10 of 28-day cycles); intravenous infusions of mitoxantrone (8 mg/m² per day), etoposide (100 mg/m² per day), and cytarabine (1000 mg/m² per day on days 1-5 of up to two 28-day cycles); or intravenous granulocyte colony-stimulating factor (300 mug/m² per day or 5 mug/kg per day subcutaneously on days 1-5), fludarabine (intravenous infusion 30 mg/m² per day on days 2-6), cytarabine (intravenous infusion 2000 mg/m² per day on days 2-6), and idarubicin (intravenous infusion 10 mg/m² per day on days 2-4 in up to two 28-day cycles). Patients proceeding to haemopoietic stem-cell transplantation after quizartinib could resume quizartinib after haemopoietic stem-cell transplantation. The primary endpoint was overall survival in the intention-to-treat population. This trial is registered with ClinicalTrials.gov, number NCT02039726, and follow-up is ongoing.; Finding(s): Between May 7, 2014, and Sept 13, 2017, 367 patients were enrolled, of whom 245 were randomly allocated to quizartinib and 122 to chemotherapy. Four patients in the quizartinib group and 28 in the chemotherapy group were not treated. Median follow-up was 23.5 months (IQR 15.4-32.3). Overall survival was longer for quizartinib than for chemotherapy (hazard ratio 0.76 [95% CI 0.58-0.98; p=0.02]). Median overall survival was 6.2 months (5.3-7.2) in the quizartinib group and 4.7 months (4.0-5.5) in the chemotherapy group. The most common non-haematological grade 3-5 treatment-emergent adverse events (within <=30 days of last dose or >30 days if suspected to be a treatment-related event) for quizartinib (241 patients) and chemotherapy (94 patients) were sepsis or septic shock (46 patients [19%] for quizartinib vs 18 [19%] for chemotherapy), pneumonia (29 [12%] vs eight [9%]), and hypokalaemia (28 [12%] vs eight [9%]). The most frequent treatment-related serious adverse events were febrile neutropenia (18 patients [7%]), sepsis or septic shock (11 [5%]), QT prolongation (five [2%]), and nausea (five [2%]) in the quizartinib group, and febrile neutropenia (five [5%]), sepsis or septic shock (four [4%]), pneumonia (two [2%]), and pyrexia (two [2%]) in the chemotherapy group. Grade 3 QT prolongation in the quizartinib group was uncommon (eight [3%] by central reading, ten [4%] by investigator report); no grade 4 events occurred. There were 80 (33%) treatment-emergent deaths in the quizartinib

Author	Full reference	Abstract
		group (31 [13%] of which were due to adverse events) and 16 (17%) in the chemotherapy group (nine [10%] of which were due to adverse events).; Interpretation(s): Treatment with quizartinib had a survival benefit versus salvage chemotherapy and had a manageable safety profile in patients with rapidly proliferative disease and very poor prognosis. Quizartinib could be considered a new standard of care. Given that there are only a few available treatment options, this study highlights the value of targeting the FLT3-ITD driver mutation with a highly potent and selective FLT3 inhibitor.; Funding(s): Daiichi Sankyo.; Copyright © 2019 Elsevier Ltd
Foran et al 2018	Foran JM et al. FLT3-ITD mutations are prevalent and significantly impact outcome after intensive therapy in elderly adults with acute myeloid leukemia (AML): Analysis of the north american intergroup E2906 phase III trial in patients age >=60 years. Blood 2018;132 (Suppl. 1)	Background Activating mutations in FLT3-and in particular FLT3-ITD (internal tandem duplication)-are common in younger patients with AML, and are associated with relapse and with inferior overall survival (OS), however their prevalence and impact in older adults remains uncertain. We performed an analysis of FLT3-ITD mutations in patients age >=60 yrs in the large prospective multicenter ECOG-ACRIN (E-A) E2906 Phase III trial. Methods Eligible patients (PTS) age >=60 yrs (n = 727) were randomized to 'Standard' therapy with 7&3 (Dauno 60mg/m2) induction, and 2 cycles of intermediate dose Ara-C (1.5g/m2 x 1 2 doses; 6 doses if age >=70yrs) consolidation (Arm A); or single agent clofarabine (CLO) induction and consolidation (2 cycles) (Arm B). As previously reported (Foran et al, ASH #21 7a, 201 5), there was superior overall survival (OS) with standard therapy. AML diagnostic samples collected prospectively in the central E-A Leukemia Translational Research Lab (LTL) were used to detect mutations in the FLT3 gene by PCR using cDNA from total RNA. LTL investigators were blinded to treatment assignment. Patients with no RNA available or blast count <1 0% (threshold of sensitivity for the PCR assay) on the submitted sample were excluded. Statistical analysis was performed using X² (categories) and Wilcoxon rank sum (continuous) tests to compare baseline patient and disease characteristics. Log-rank tests and multivariate Cox models stratified by treatment arm and adjusted for patient and disease variable (including WBC, cytogenetics, sex, performance status, secondary AML) were used to examine FLT3-ITD effect on OS and disease-free survival $DFS; relapse or death after complete remission (CR)/CRi (CR with incomplete CBC recovery). Results In the first 231 pts tested, FLT3-ITD mutations were identified in 43 (18.6%) pts, and the remainder were FLT3-ITD-negative (i.e. nonITD). The proportion with FLT3-ITD was the same for patients age 60-69 yrs (1 8.8%) vs. >=70 yrs (1 8.3%). In comparison to non-ITD, FLT3-ITD+ pts had significantly higher WBC (median 1 8.1 vs. 6.3, p=0.002) and BM blasts (80% vs. 51.5%, p=0.0004) at AML diagnosis, and were more likely to have intermediate risk (79.1% vs. 59%) vs. unfavorable risk (7.0 % vs. 31.4%) cytogenetics (p=0.002). There was no difference in CR/CRi rate overall (p=0.40), however standard (Arm A) pts with FLT3-ITD had a significantly higher CR/CRi rate (78.9% vs. non-ITD 51.5%, p=0.04). With median follow-up 53.5 months, FLT3-ITD patients tended to have worse OS (HR 1.24, 95%CI 0.85-1.81) but this was not statistically significant (Figure 1, p=0.26), and results were similar for both Arms A/B. DFS similarly tended to be worse overall for FLT3-ITD (HR 1.44, 95%CI 0.65-2.43) (p=0 1 7) and DFS was significantly worse for FLT3-ITD+ Arm A/Standard pts (Figure 2, p=0.033), while Arm B/CLO pts had a worse DFS regardless of mutation status (Arm B, FLT3-ITD vs. non-ITD, p=0.93). More patients with FLT3-ITD underwent allogeneic transplantation (25.6% vs. non-19.7%), although this was not significant (p=0.41). Conclusions Jump to FLT3-ITD mutations are prevalent in older and elderly (age >=70 yrs) patients with

Author	Full reference	Abstract
		AML, and while they occur at somewhat lower rates than reported for younger adults, FLT3-ITD+ AML in this population has a leukemia phenotype similar to that reported in younger patients. Despite significantly higher CR/CRi rates with Standard therapy, older pts with FLT3-ITD also have significantly worse DFS following intensified Ara-C consolidation therapy than non-ITD pts. These results support routine assessment of FLT3-ITD status in older AML patients, and the incorporation of novel post-remission treatment strategies to improve outcome in FLT3-ITD+ pts >=60 yrs.
Hills et al 2018	Hills RK et al. Outcomes in relapsed/refractory patients with FLT3-ITD mutated AML are poor when treated with non-targeted therapy with a potential role for stem cell transplantation: Results from the ncri AML trials. Blood 2018; 132 (Suppl. 1)	Introduction: Patients with Acute Myeloid Leukaemia (AML) who harbour a FLT3-ITD mutation have a worse prognosis characterised by increased early relapse. Outcomes following relapse in such patients are typically poor, not only because of earlier relapse, but also because of worse performance post-relapse than those who have the same remission duration but are FLT3-ITD WT. In a single centre retrospective study among relapsed patients, Ravandi et al $1@ found a remission rate of 24% and median survival of only 1 3 weeks. There are therefore twin challenges in this population: first to reduce the early relapse rate, and also to develop more effective treatments post relapse. In the recent QUANTUM-R trial $2@ for patients with relapsed or refractory disease, single agent quizartinib (AC220) was found to significantly improve median survival from 20.4 weeks to 27 weeks when compared to "doctor's choice" treatment (low-dose ara-C, MEC or FLAG-Ida). There is however, no directly comparable data for the population in the QUANTUM-R trial. To contextualise these results, especially given the potentially different outcomes by control treatment, we looked at outcomes in the UK NCRI AML1 5,16,17 trials in patients satisfying the eligibility criteria of QUANTUM-R.; Method(s): Patients aged 1 8+ in the UK NCRI AML1 5,16,17 trials were identified who harboured a FLT3 ITD mutation, were treated with intensive chemotherapy, and were either refractory to two courses of induction therapy, or relapsed within six months of transplant, or did not receive a prior transplant and had a remission duration of 6 months or less. Patients were grouped hierarchically as refractory, relapsed post-transplant, or relapsed without prior transplant. Eligibility was established at the point a patient first became eligible for analysis. The primary outcome was overall survival (OS), measured from point of eligibility, with subsequent remission with or without count recovery as secondary outcome. A sensitivity analysis was performed excluding those who died within 21 days of eligibility to eliminate patients who might be thought of as too unwell to enter a post-relapse trial. Cox regression was used to identify prognostic factors for survival.; Result(s): A total of 264 patients were identified (refractory N = 58, relapsed post SCT N = 49, relapsed without SCT N = 1 57). The median age was 51 (range 1 8-84); 2 5% of patients were aged 60 or older; 44% were male, 95% had intermediate cytogenetics; 11% had secondary disease. Split by age, among those under 60 45 were refractory, 44 relapsed post SCT and 11 0 relapsed without SCT; for ages 60+ the figures were 1 3 vs 5 vs 47. Overall 1 7% of patients experienced a subsequent remission; with median survival of 86 days and 1 year OS of 1 3%. If deaths within 21 days were excluded, the remission rate improved to 21%; with a median survival of 1 33 days and 1 year OS of 16%. In multivariable Cox regression, age group HR for age>60 1.81 (1.33-2.47) p=0.0002) and route to eligibility (HR refractory vs relapsed no SCT 0.77 (0.55-1.07); relapsed post SCT vs no SCT 1.58 (1.11-2.25) p=0.003) were the only factors affecting survival-in particular sex, secondary disease, and ITD allelic burden were not significant. In the sensitivity analyses, only age was significant (HR 1.77 (1.24-2.53) p=0.001); with

Author	Full reference	Abstract
		route to eligibility not significant (p=0.1 4). Among patients with post-relapse treatment information, 65% were treated intensively, 8% non-intensively, and 20% with palliation-other received experimental therapies. When restricting attention to those treated intensively, median survival was 1 30 days with 1 7% 1 year OS. F were not materially changed if early death was excluded. Jump to Of 215 patients who had not relapsed post transplant, 53 (25%) received a transplant post-eligibility. In these 56 patients, median survival was 301 days with 42% alive at one year.; Conclusion(s): In relapsed/refractory AML, outcomes for FLT3-ITD mutated patients are generally poor and worse for older patients. Applying the eligibility criteria of QUANTUM-R and excluding early deaths gives outcomes comparable to the control group of the QUANTUM-R study. In the 25% of patients who proceeded to transplant survival was extended indicating that a treatment which can deliver patients to transplant has the potential to improve patient outcomes.
Kang et al 2018	Kang D et al. Concentration-QT analysis of quizartinib in patients with relapsed/refractory AML. Journal of Pharmacokinetics and Pharmacodynamics 2018; 45 (Supplement 1): S23-S24	Objectives: Quizartinib is a highly potent and selective FLT3 inhibitor, and has shown high clinical activity in patients with relapsed/refractory acute myeloid leukemia (AML) with FLT3-ITD mutations. In this analysis, we evaluated the relationship between pharmacokinetic exposures of quizartinib and active metabolite AC886 and QTc interval.; Method(s): Data were obtained from a Phase 2 Study (2689-CL-2004; NCT #01565668) evaluating the safety and efficacy of quizartinib with planned doses of 30 mg/d and 60 mg/d in relapsed/refractory AML patients with FLT3-ITD mutations. Serial triplicate centrally reviewed electrocardiograms, together with time-matched PK samples, were collected over 24 h following a single dose on cycle 1 day1 and at steady state on cycle 1 day 15. Different base structural models, correction terms for QTc, potential hysteresis, circadian rhythm correction and model parameter distribution, were thoroughly evaluated. Covariates evaluated include baseline QTcF, patient demographics (sex, age, body weight, race), low electrolyte (Ca, K, Mg) levels, and concomitant use of QT prolonging agents.; Result(s): Analysis included 868 time-matched mean QTc and concentration measurements from 73 patients. QTcF increases linearly with respect to concentrations of quizartinib and AC886 (Figure 1), with 15-fold higher slope for quizartinib than AC886. Race was identified as a statistically significant covariate on baseline QTcF, with baseline QTcF being approximately 4% higher in the white race as compared to others. Model predicted mean QTcF increase from baseline was 7.36 and 19.3 ms (upper bound of two-sided 90% CI: 8.90 and 23.3 ms) respectively, for quizartinib 30 mg/d and 60 mg/d. An alternative model with quizartinib concentration alone as a predictor provided similar results.; Conclusion(s): Analysis suggested concentration-dependent QTc prolongation of quizartinib. Results support clinical recommendation of dose reduction in patients receiving strong CYP3A inhibitors, where quizartinib exposure is increased 2-fold in the presence of such agents.
Marconi et al 2018	Marconi G et al. Tyrosine kinase inhibitors (TKI) in relapsed/refractory (RR) patients with FLT3-ITD positive acute myeloid leukemia (AML) confer better survival than	Background: Approximately 20-30% of AML patients harbor internal tandem duplication (ITD) of FLT3 gene. FLT-ITD mutations are associated with a poor prognosis, due to a high relapse rate. Several drugs have been developed to inhibit FLT3. However, R/R FLT3-ITD AML patients still represent an unmet clinical need.; Aim(s): Since no prospective randomized studies comparing the role of chemotherapy and TKIs in R/R FLT3 ITD AML patients have been conducted, our aim is to assess outcome, safety and duration of hospitalizations in two retrospective groups of patients, referred to or diagnosed at our Institution, and treated with TKIs or chemotherapy, respectively.; Method(s): We retrospectively collected and analyzed

Author	Full reference	Abstract
	chemotherapy, due to a better safety profile. HemaSphere 2018; 2 (Supplement 2): 450-451	clinical and biological data of 58 consecutive FLT3-ITD AML patients, treated at our Institution from 2004 to 2017 with chemotherapy (3+7 like regimens; 3+7 like regimens with the addition of a third agent; fludarabine based regimens) and/or single agent TKIs (Sorafenib, Ponatinib, Quizartinib, Gilteritinib, Midostaurin). All the patients underwent any kind of therapy after informed consent was signed.; Result(s): We compared patients who received at least once in their life, as salvage treatment, a TKI inhibitor ("TKI" group; N=36) with patients that were treated exclusively with conventional cytotoxic ("conventional group"; N=22). The median age of the entire population was 59 years (range 17- 74); there were no significant differences in patient age, white blood cells count, platelet count and ELN risk at diagnosis between the two groups. Fifty-one out of 58 patients (86%) relapsed after (N=22; 36%) or were refractory (N=29; 50%) to the first course of induction chemotherapy. Second- line therapy included salvage chemotherapy (N=32/51, 63%), a TKI as single agent (N=12/51, 23%) or best supportive therapy (N=14%, 7/51). Forty-one patients experienced a 2nd relapse, or were persistently refractory, and of these 18 received a single agent TKI as salvage treatment. Six patients received a TKI in 3rd or further relapse. Standard chemotherapy compared with TKIs did not show an increased efficacy in terms of CR (25% vs 16.7%), and it was not a better bridge-to-transplant option. However, among R/R patients, we observed an advantage in terms of OS for patients of the "TKI" group compared with "conventional" group (median OS from R/R of 10 months [95% CI, 5.89-14.12] and 4 months [95% CI, 3.12- 4.90], respectively; p= .017). Finally, as far as toxicity is concerned, patients in "TKI" group experienced a lower number of AEs during treatment with TKIs (1.63 mean AEs in each TKI line vs 3.03 mean AEs in each chemotherapy line, excluding stem cell transplant; p< .001; grade III-IV 6/23 and grade V 2/23 with TKI; grade III-IV 24/66 and grade V 12/66 with chemotherapy). Furthermore, AEs during TKI therapy were less severe if compared with AEs during chemotherapy (Figure 1, p= .049). We also noted a trend toward less day spent in hospital per month by patients during TKI treatment, compared to patients treated with standard chemotherapy (including post chemotherapy remission period): 10.5 days and 16.7 days in the two groups, respectively. Summary/Conclusion: Our study, even if in a retrospective set, reports a survival advantage of TKI in R/R FLT3 ITD AML patients, compared with conventional approaches. Such an advantage is due to the lower number and grade of AEs of "TKI" group. For their safety profile, TKIs are probably a better option to bridge patients to transplant, thanks to a lower risk of toxicity (Figure Presented).
Wang et al 2018	Wang J et al. A phase III randomized study of gilteritinib versus salvage chemotherapy in FLT3 mutation-positive subjects with relapsed or refractory acute myeloid leukemia. Annals of Oncology. 2018. 29 (Supplement 9)	Background: The highly potent, selective fms-like tyrosine kinase 3 (FLT3)/AXL inhibitor, gilteritinib (ASP2215), showed strong antileukemic activity at doses >=80 mg/day in patients with FLT3 mutation-positive (FLT3) relapsed/refractory (R/R) acute myeloid leukemia (AML). This phase 3 trial was designed to compare the efficacy and safety of gilteritinib versus salvage chemotherapy in FLT3 subjects with R/R AML. Trial design: This phase 3, open-label randomized multicenter trial (NCT03182244) will enroll approximately 320 adult subjects (aged >=18 years; Eastern Cooperative Oncology Group [ECOG] performance status <=2) with FLT3 R/R AML from mu50 centers across China, Russia, Singapore, Thailand, and Malaysia. Subjects will be randomized (1:1) to receive 28- day cycles of once-daily oral gilteritinib (120 mg) or salvage chemotherapy. The salvage chemotherapy regimen will be selected by the investigator from the following predetermined options: LoDAC (intravenous [IV]/subcutaneous [SC] cytarabine 20 mg BID for 10 days),

Author	Full reference	Abstract
		MEC (IV mitoxantrone 6 mg/m /d plus IV etoposide 100 mg/m /d plus IV cytarabine 1000 mg/m /d, Days 1-5), or FLAG (granulocyte colony-stimulating factor SC/IV 300 mug/m /d, Days 1-5; IV fludarabine 30 mg/m /d, Days 2-6; IV cytarabine 2000 mg/m /d, Days 2-6). Subjects receiving gilteritinib or LoDAC will continue treatment until a discontinuation criterion is met; those receiving MEC or FLAG will be assessed for response on or after Day 15 of Cycle 1 and will receive a second cycle of MEC/FLAG chemotherapy if bone marrow (BM) cellularity is >= 20% with >=50% reduction in BM blasts. If BM cellularity is > 5% to < 20%, the decision to administer a second cycle of MEC/FLAG chemotherapy will be made by the investigator. The primary endpoint is overall survival; key secondary endpoints are eventfree survival and complete remission rate. Safety endpoints include the incidence of adverse events, results from laboratory investigations and vital sign examinations, findings from electrocardiograms, and changes in ECOG performance status. A formal interim analysis is planned when approximately 50% of deaths have occurred.

A2. CS, Appendix D: Identification, selection and synthesis of clinical evidence, Section D.1.1, Study selection Table 1 Clinical review scope - inclusion and exclusion criteria. The study design of interest includes observational studies. However, the RCT search filter applied to the clinical search strategy for both Medline and Embase (Appendix I: Search strategies 8.1.2.) would exclude observational studies. Please provide database search strategies for identifying observational studies for the clinical review.

The searches included also a filter for single-arm and cohort studies which would already have identified many of the observational studies. However, Astellas agrees that no specific term for observational studies was included.

The searches in EMBASE, Medline and Cochrane have been re-run with and without terms specific for observational studies (Table 4). Adding these terms did not yield any additional studies in Cochrane (Table 4) and only two in EMBASE/Medline. Neither of the two additional studies were relevant.

Table 4 EMBASE/Medline Searches for Clinical Evidence With and Without Terms Specific for Observational Studies

#	Searches	Results
1	exp acute myeloid leukemia/ or exp leukemia, myeloid, acute/	87,041
2	("acute granulocytic leukemia" or "acute granulocytic leukaemia" or "acute nonlymphocytic leukemia" or "acute nonlymphocytic leukaemia" or "acute non-lymphocytic leukemia" or "acute non-lymphocytic leukaemia" or "acute myelogenous leukemia" or "acute myelogenous leukaemia").ti,ab.	18,870
3	("acute myeloid leukemia" or "acute myeloid leukaemia" or "acute myelocytic leukemia" or "acute myelocytic leukaemia" or "acute myeloblastic leukemia" or "acute myeloblastic leukaemia" or "acute non-lymphoblastic leukemia" or "acute non-lymphoblastic leukaemia" or "acute nonlymphoblastic leukemia" or "acute nonlymphoblastic leukemia").ti,ab.	91,207
4	1 or 2 or 3	157,551
5	(refractory or relapsed or relapse).ti,ab.	626,312
6	exp CD135 antigen/ or exp fms like tyrosine kinase 3/	8,164
7	(cd135 or flt3 or flt 3 or fms like tyrosine kinase 3).ti,ab.	17,342
8	("poor cytogenetics" or "high risk cytogenetics" or "high risk cytogenetic" or "poor risk cytogenetics" or "poor risk cytogenetic").ti,ab.	2,347
9	6 or 7 or 8	21,267
10	exp cytarabine/ or exp azacitidine/ or exp mitoxantrone/ or exp etoposide/ or exp granulocyte colony stimulating factor/ or exp idarubicin/	236,279
11	("gilteritinib" or "asp2215").ti,ab.	138
12	(hypomethylating or azacytidine or decitabine or Dacogen).ti,ab.	14,754
13	("sorafenib" or "nexavar" or "bay 43 9006" or "bay 54 9085").ti,ab.	21,448
14	("quizartinib" or "ac220" or "ibrutinib" or "Imbruvica" or "PCI-32765" or "CRA-032765" or "JNJ-54179060").ti,ab.	5,907
15	("crenolanib" or "cp 868,596").ti,ab.	221
16	(cytarabine or Depocyt or AraC or cytosine arabinoside or lodac).ti,ab.	30,075

#	Searches	Results
17	(Azacitidine or vidaza or mitoxantrone or Novantrone or etoposide or hidac or VePesid or Etopophos or mec).ti,ab.	74,989
18	(G-CSF or granulocyte colony stimulating factor or Neupogen or fludarabine or idarubicin or Idamycin pfs or flag ida).ti,ab.	69,468
19	exp Double Blind Method/ or exp Single Blind Method/ or exp Single Blind Method/ or exp cohort studies/ or exp clinical trial/ or exp clinical trials as topic/ or exp placebos/ or exp double blind procedure/ or exp single blind procedure/ or exp cohort analysis/ or exp clinical trial/ or exp placebo/ or exp "clinical trial (topic)"/	5,008,007
20	exp controlled clinical trial/ or (controlled clinical trial or randomized controlled trial or multicenter study or clinical trial).mp.	2,907,995
21	((Study or studies) and (open or "open-label" or "non-randomised" or "non-randomized" or "cohort" or "single-arm")).mp. or observational*.ti,ab. or exp observational study/ or exp observational studies as topic/ [mp=ti, ab, hw, tn, ot, dm, mf, dv, kw, fx, dq, nm, kf, ox, px, rx, ui, sy]	2,445,171
22	exp randomization/ or exp randomized controlled trial/ or exp "randomized controlled trial (topic)"/ or (double blind or double blinded or RCT or Randomi* or controlled or controled or control or Placebo or Trial or randomly allocated or prospective stud* or prospective trial* or clinical trial*).ti,ab.	8,963,438
23	4 and 5 and 9	3,775
24	10 or 11 or 12 or 13 or 14 or 15 or 16 or 17 or 18	328,449
25	19 or 20 or 21 or 22	12,596,321
26	23 and 24 and 25	1,085
27	remove duplicates from 26	948
28	((Study or studies) and (open or "open-label" or "non-randomised" or "non-randomized" or "cohort" or "single-arm")).mp. [mp=ti, ab, hw, tn, ot, dm, mf, dv, kw, fx, dq, nm, kf, ox, px, rx, ui, sy]	2,084,037
29	19 or 20 or 28 or 22	12,403,649
30	23 and 24 and 29	1,083
31	remove duplicates from 30	946
32	27 not 31	2

ADMIRAL study design and statistical analyses

A3. CS, Section B.2.3, Table 6, pages 23 to 26. Please clarify the following exclusion criteria applied in the ADMIRAL trial:

(a) "Breakpoint cluster region-Abelson murine leukaemia" - Should this read "BCR-ABL leukaemia"?

(b) "Clinically significant abnormality of coagulation profile" - How was this defined?

(c) "…unless a screening ECG done within 3 months before study entry resulted in a LVEF of ≥45%." – Should this refer to echocardiogram rather than ECG?

(d) "Hypokalaemia or hypomagnesaemia" – How was this defined? If it was corrected, were these patients eligible for inclusion in ADMIRAL?

(e) "other uncontrolled hepatic disorder" – What did these include? Was severity of the disorder defined?

a) Either terminology is correct.

b) This was based on Investigator discretion.

c) Yes, and the period should have said within 1 month of study entry.

d) Defined as lower than lower limit of normal (LLN); patients were excluded.

e) This was based on Investigator discretion and severity was not defined.

A4. CS, Section B.2.4, page 28. Please provide details of the sample size calculation for ADMIRAL, including all assumptions made.

This is a group sequential design based on co-primary endpoint of OS using the O'Brien-Fleming boundaries (non-binding) as implemented by Lan-DeMets alpha/beta spending method (East)[8]. The overall 0.025 one-sided type I error rate is allocated by 0.0005 and 0.0245 (0.001 and 0.049 for two-sided type I error rate) for the two co-primary efficacy endpoints of CR/CRh and OS, respectively. The type I error (alpha) in the first interim analysis was not be recycled in the second interim analysis and final analysis. The first interim analysis was planned when approximately 141 subjects were randomised into gilteritinib arm and at least 112 days (4 treatment cycles) post first dose or randomisation (for subjects who received no study drug). The second interim analysis was planned when approximately 129 death events had occurred and the final analysis was planned when approximately 258 death events had occurred.

OS: Approximately 369 subjects (the planned sample size with 10% dropout rate) were to be randomised in a 2:1 ratio to receive gilteritinib or salvage chemotherapy (246 subjects in the gilteritinib treatment arm and 123 subjects in the salvage chemotherapy arm). The planned 258 death events was to provide 90% power to detect a difference in OS between the gilteritinib arm with 7.7 months median survival time and salvage chemotherapy arm with 5 months median survival time (hazard ratio = 0.65) at the overall 1-sided 0.0245 significance level.

CR/CRh rate: The first interim analysis was to be conducted only to evaluate the co-primary endpoint of CR/CRh. One hundred and forty-one subjects randomised to gilteritinib arm (211 subjects in total: 141 in the gilteritinib arm and 70 in the salvage chemotherapy arm) with a minimum follow-up of 4 treatment cycles were considered to achieve a maximum width of 15.78% for the two-sided 95% exact confidence interval (CI) when the CR/CRh was

expected to be in the 5% to 30% range. A sample size of 141 subjects provided 80% power to exclude a CR/CRh rate of 12% using the two-sided 95% exact CI when the CR/CRh rate of gilteritinib was assumed to be 21%.

A5. CS, Section B.2.4, page 28. Please provide details of how the randomisation was conducted within ADMIRAL, including block size.

ADMIRAL used stratified randomisation rather than a block randomisation[9]. Randomisation and study drug assignment were performed via Interactive Response Technology (IRT). Prior to the initiation of the study treatment, the site staff contacted the IRT in order to determine the randomly assigned treatment.

Subjects were randomised in a 2:1 ratio to receive gilteritinib or salvage chemotherapy. Randomisation was stratified by response to first-line AML therapy and pre-selected salvage chemotherapy:

Response to first-line therapy:

- Relapse within 6 months after allogeneic HSCT
- Relapse after 6 months after allogeneic HSCT
- Primary refractory without HSCT
- Relapse within 6 months after CRc and no HSCT
- Relapse after 6 months after CRc and no HSCT

Preselected chemotherapy:

- High intensity chemotherapy (FLAG-IDA, MEC)
- Low intensity chemotherapy (LoDAC or azacitidine).

A6. CS, Section B.2.4, page 28. Please clarify whether there are any known covariates that are predictive of outcome in addition to response to first-line AML and pre-selected salvage chemotherapy (e.g. prior use of a FLT3 inhibitor). Please provide an analysis of the co-primary outcomes adjusting for these and the stratification variables.

Astellas has not identified any covariates that are predictive of outcomes. We have analysed a number of sub-groups per the Forest Plot presented in the Company evidence submission, page 39, Figure 7. Gilteritinib appears to be effective across populations.

A7. CS, Section B.2.4, page 28. Please clarify why the CR/CRh outcome was not stratified for response to first-line AML and pre-selected salvage chemotherapy. Also, it is not generally true that

non-overlapping arm-specific confidence intervals for a parameter implies that there is no difference between treatments. Does this fact affect any inferences based on arm-specific confidence intervals in the ADMIRAL trial?

The ADMIRAL trial was stratified based on first line therapy (see Question A5), with the primary endpoint of OS. Following feedback from the FDA CR/CRh was added as a co-primary endpoint[9] but the stratification remained per the initial approach.

Astellas accepts that confidence intervals are not ideal predictors when numbers are small. However this does not affect any inferences .

A8. PRIORITY. CS, Section B.2.6, page 38. Please clarify why there is variable patient follow-up in the first month in ADMIRAL. In addition, please provide a statistical justification regarding why it is necessary to treat EFS events occurring during the first month as occurring on Day 0.

In the first 30 days from randomisation, patients were seen on Day 1, 4, 8, 15 and 30 as per the ADMIRAL protocol. Overall, 89.2% (331/371) of patients reached the 30-day follow-up evaluation. The most frequent 30-day follow-up evaluation status was completed (50.9% [189/371]) and death (28.8% [107/371]). The long-term follow-up evaluation was reached by 76.8% (285/371) of patients. Any patients who discontinued or failed treatment within first 30 days were treated as reaching EFS on Day 0 i.e., the randomisation day. This was not based on statistics, but was based on discussion and agreement with the FDA[9].

Per the SAP: “EFS is defined as the time from the date of randomization until the date of documented relapse (excluding relapse after PR), treatment failure or death from any cause within 30 days after the last dose of study drug, whichever occurs first [earliest of (relapse date, treatment failure date, death date) – randomization date + 1]. If a subject experiences relapse or death within 30 days after the last dose of study drug, the subject is defined as having EFS event related to either “relapse” or “death”, and the event date is the date of relapse or death. If a subject discontinues the treatment and fails to achieve any of the response of CR, CRp or CRi during the treatment period (subject with best response of [partial remission] and [non-response]), the subject is defined as having EFS event related to treatment failure, and the event date is the randomization date. Subjects that discontinue the treatment with post-treatment disease assessment and best response of NE will be censored.”

A9. PRIORITY. CS, Section B.2.3, Table 7, page 27. Please provide protocol definitions of all clinical endpoints in CHRYSALIS and ADMIRAL. Please ensure that this includes the protocol definition for EFS, as well as definitions of the events of relapse and progression.

Table 5 below provides all definitions for ADMIRAL and CHRYSALIS.

Table 5 Protocol Definitions

Definition	Description
Complete Remission (CR)	For subjects to be classified as being in CR at a post-baseline visit, they must have bone marrow regenerating normal hematopoietic cells and achieve a morphologic leukaemia-free state and must have an ANC ≥ 1 x 10^9/L and platelet count ≥ 100 x 10^9/L and normal marrow differential with < 5% blasts, and they will be RBC and platelet transfusion independent (defined as 1 week without RBC transfusion and 1 week without platelet transfusion). There should be no evidence of extramedullary leukemia
Complete Remission with Partial Hematologic Recovery (CRh)	At a post baseline visit, subjects will be classified as CRh if they have marrow blasts < 5%, partial haematologic recovery ANC ≥ 0.5 x 10^9/L and platelets ≥ 50 x 10^9/L, no evidence of extramedullary leukemia and cannot be classified as CR
Complete Remission with Incomplete Platelet Recovery (CRp)	For subjects to be classified as being in CRp at a post-baseline visit, they must achieve CR except for incomplete platelet recovery (< 100 x 10^9/L)
Complete Remission with Incomplete Haematologic Recovery (CRi)	For subjects to be classified as being in CRi at a post-baseline visit, they must fulfill all the criteria for CR except for incomplete haematological recovery with residual neutropenia < 1 x 10^9/L with or without complete platelet recovery. RBC and platelet transfusion independence is not required
Composite Complete Remission (CRc)	For subjects to be classified as being in CRc at a post-baseline visit, they must either achieve CR, CRp or CRi at the visit
Partial Remission (PR)	For subjects to be classified as being in PR at a post-baseline visit, they must have bone marrow regenerating normal haematopoietic cells with evidence of peripheral recovery with no (or only a few regenerating) circulating blasts and with a decrease of at least 50% in the percentage of blasts in the bone marrow aspirate with the total marrow blasts between 5% and 25%. A value of less or equal than 5% blasts is also considered a PR if Auer rods are present
Best response	Best response was defined as the best measured response to treatment for all post baseline visits (in the order of CR, CRp, CRi, PR, NR and not evaluable). Patients with best responses of CR, CRp, CRi or PR were considered responders. Patients who did not achieve at least a best response of PR were considered non-responders
Duration of remission (DOR)	Time from achieving remission to relapse. Duration of remission included duration of CRc, duration of CR/CRh, duration of CRh, duration of CR and duration of response (CRc + PR)
Overall survival (OS)	OS was defined as the time from the date of randomisation until the date of death from any cause (death date – randomisation date + 1). For a patient who was not known to have died by the end of study follow-up, OS was censored at the date of last contact (date of last contact – randomized date + 1). The date of last contact was the latest date that the patient was known to be alive by the cut-off date. The last contact date was derived for patients alive at the analysis cut-off date. Patients with a last contact date beyond the analysis cut-off date were censored at the analysis cut-off date
Leukaemia-free survival (LFS)	LFS was defined as the time from the date of first CRc until the date of documented relapse (excluding relapse from PR) or death for patients who achieved CRc (relapse date or death date – first CRc disease assessment date + 1). For a patient who was not known to have relapsed or died, LFS was censored

Definition	Description
	on the date of last relapse-free disease assessment date (last relapse-free disease assessment date – first CRc disease assessment date + 1)
Event-free survival (EFS)	EFS was defined as the time from the date of randomisation until the date of documented relapse (excluding relapse after PR), treatment failure or death from any cause within 30 days after the last dose of study drug, whichever occurred first (earliest of [relapse date, treatment failure date, death date] – randomisation date + 1). If a patient experienced relapse or death within 30 days after the last dose of study drug, the patient was defined as having an EFS event related to either “relapse” or “death”, and the event date was the date of relapse or death
Transfusion conversion rate; transfusion maintenance rate	Transfusion conversion rate and transfusion maintenance rate were only defined for the patients in the gilteritinib arm. For the purpose of defining transfusion conversion rate and transfusion maintenance rate, transfusion status (independent vs. dependent) during the baseline period and during the post-baseline period was defined as follows for patients who took at least 1 dose of study drug: Baseline transfusion status: ● The baseline period was defined as the period from 28 days prior to the first dose to 28 days after the first dose. For patients who were on treatment < 28 days, the baseline period was from 28 days prior to the first dose until the end of treatment. ● Patients were classified as baseline transfusion independent if there were no RBC or platelet transfusions within the baseline period; otherwise, the patient was baseline transfusion dependent. Post-baseline transfusion status: ● The post-baseline period was defined as the period from 29 days after the first dose until the last dose. ● For patients who were on treatment ≥ 84 days, they were classified as post baseline transfusion independent if there was 1 consecutive period of 56 days without any RBC or platelet transfusion within the post baseline period. ● For patients who were on treatment > 28 days but < 84 days, if there was no RBC or platelet transfusion within the post-baseline period, then post-baseline transfusion status was not evaluable. ● For patients who were on treatment ≤ 28 days, post-baseline transfusion status was not evaluable. ● Otherwise, the patient was considered post-baseline transfusion dependent. Both transfusion conversion rate and maintenance rate were defined for patients who had evaluable post-baseline transfusion status. Transfusion conversion rate was defined as the number of patients who were transfusion dependent during the baseline period but become transfusion independent during the post-baseline period divided by the total number of patients who were transfusion dependent during the baseline period. Transfusion maintenance rate was defined as the number of patients who were transfusion independent during the baseline period and still maintained transfusion independence during the post-baseline period divided by the total number of patients who were transfusion independent during the baseline period
Pharmacodynamics (Chrysalis)	These were different response endpoints as defined above (CR, CRc, CRh, CRi, CRp, PR, BR)
Transplantation rate	The transplantation rate was defined as the percentage of patients who underwent HSCT during the study period

Definition	Description
Brief Fatigue Inventory (BFI) patient-reported fatigue.	The BFI [10] was developed to assess the severity of fatigue and the impact of fatigue on daily functioning in patients with fatigue due to cancer and cancer treatment. The BFI short form has 9 items and a 24-hour recall. A global fatigue score is computed by averaging the 9 items. The BFI was administered at site visits directly to the patients via an electronic PRO device. A higher BFI fatigue score indicates a more unfavorable outcome

A10. PRIORITY. CS, Section B.2.6, pages 32 to 42. Do the results for CHRYSALIS and ADMIRAL presented in the CS reflect the final analyses? If not, please provide further information regarding when the final analyses will be available.

Both datasets are final.

A11. CS, Section B.2.3, Table 6, page 23. How many UK centres were involved in ADMIRAL?

Four (London, Manchester, Bournemouth, Plymouth).

A12. CS, Section B.2.5, page 29. The maximum tolerated dose in CHRYSALIS was estimated to be 300mg/day. Please explain why a daily dose of 120mg gilteritinib was selected for evaluation in ADMIRAL.

A total of 31 patients in the safety analysis set experienced a Drug Limited Toxicity (DLT) during the study, with 9.5%, 10.8%,16.7%, 18.8% and 66.7% of patients at dose levels of 80, 120, 200, 300 and 450 mg gilteritinib, respectively, experiencing a DLT. The maximum tolerated dose (MTD) for the study was determined to be 300mg. Based on exposure, response and safety data, a starting dose of 120mg gilteritinib was expected to result in adequate drug exposure for clinical efficacy for phase 3 studies in patients with FLT3 mutation positive relapsed/refractory AML, while providing an acceptable safety profile without the need for dose adjustment in patients receiving concomitant treatment with strong or moderate CYP3A4 inhibitors.

A13. CS, Section B.2.3, Table 6, page 23. Given that ADMIRAL involved data collection across 14 countries, please clarify what is meant by the statement *"Data was collected from centres involved in the trial and reflected expected UK clinical practice".*

Astellas conducted in-depth consultations with clinical experts which confirmed the approach taken in the ADMIRAL trial reflected UK caseload.

Clinical study results

A14. CS, Section B.2.6, pages 32 to 41. The clinical section of the CS does not include any results from ADMIRAL for the EQ-5D.

(a) Please clarify why the EQ-5D is not listed as an outcome in CS Table 6.

(b) Was there a difference in change from baseline between the randomised treatment groups in ADMIRAL?

a) EQ-5D was described in the ADMIRAL protocol as an exploratory endpoint and therefore not included in the summary Table in question. Please see Data on File ADMIRAL CSR 8.1.3.3 for a summary, and tables 12.3.12 to 12.3.13.5 for specific results[11].

b) Further to the above, additional post hoc analyses of data from the EQ-5D VAS instrument in ADMIRAL was conducted using patients in the gilteritinib arm only. Longitudinal change from baseline in PRO scores was analyzed using a restricted maximum likelihood (REML) based Mixed Model Repeated Measures (MMRM) approach. The MMRM analyses used all available data and assumed that missing observations were missing at random. The model included the analysis visit and the stratification factors (pre-selected salvage chemotherapy and response to first-line acute myeloid leukaemia therapy) as discrete parameters, the baseline PRO score as a covariate along with the baseline PRO score by visit interaction and patient as a random effect. Compound symmetry variance-covariance matrix was used to model the covariance structure among each patient's repeated measures. Due to substantial dropout from cycle 28 onwards (the number of patients expected to complete a PRO assessment was <5); the post hoc analyses included all time points until cycle 27 (inclusive).

As with all other PRO instruments assessed in the trial, overall change from baseline on the EQ-5D-5L VAS was small and non-meaningful (LS mean: 7.793, 95% CI: 4.876, 10.709). Based on the available literature, 12 points was considered a clinically meaningful change (using an upper bound of the range presented in the literature[12]. However, larger, clinically meaningful changes from baseline were observed at cycle 24 (LS mean: 15.022, 95% CI: 3.231, 26.814) and cycle 27 (LS mean: 13.360, 95% CI: -1.054, 27.773). Overall, patients reported quality of life was maintained throughout the study period.

In summary, analyses of changes on the EQ-5D VAS from baseline indicate that patients maintained their initially reported scores through cycle 27. Though some changes were observed at various time points, overall scores remained stable and patient reported outcomes were maintained.

A15. CS, Appendix D, pages 21 to 117. The CHRYSALIS study is sometimes referred to as a randomised study, and sometimes as a non-comparative study.

(a) Please clarify if patients were randomised within the dose expansion phase of the study.

(b) Please give the method of random sequence generation and allocation.

(c) Please clarify the method used to allocate patients to randomised groups (e.g. was it Interactive Response Technology like the ADMIRAL trial) - was this centrally or per centre?

a) The full description of the study is: open-label, dose escalation, first-in-human study in patients with relapsed or refractory AML, with concomitant expansion cohort for multiple doses. Patients were enrolled on the lowest dose (20mg) then escalated. The second phase of the study was a dose expansion cohort. There was a randomisation in to the dose expansion phase of the study.

b) As a dose level was decided to be expanded, up to 17 patients could be enrolled for the dose level in the dose expansion phase (to have a total of 20 patients enrolled at a dose level including the patients from dose escalation cohort). When more than 1 dose level was expanded in the dose expansion phase (Cohort 2), the newly enrolled patients were to be randomised to one of the open expanded dose levels, based on the relative chance of (20 - n) in each dose level, where n is the number of patients already enrolled in the dose level, including both the dose escalation and expansion phases.

c) Randomisation was performed via IRT during the dose expansion phase (Cohort 2).

Systematic review

A16. CS, Appendix D, page 29. The text states *"Among the excluded studies, the following merit special mention. Seventeen publications were retrieved in the searches but excluded from abstraction as these concerned reporting of meta-analyses or systematic reviews of clinical studies already included in the review."* Please provide the PDFs for these studies.

References were provided for the Appendix SLR section in file SLR Refs.zip

A17. CS, Appendix D, Figure 1, page 29. Please provide a table of all 68 full text exclusions from the PRISMA diagram with reasons for exclusion.

Reasons for exclusion of the sixty-eight studies are provided in Table 6 below.

Table 6 Reasons for Exclusion for the 68 Publications

Publication	Reason for exclusion
Alattar ML et al. Response rates in patients with relapsed/refractory acute myeloid leukemia with FLT3-ITDmutation using 5-azacitadine plus sorafenib. Journal of Clinical Oncology. 2012: 30 (15 SUPPL. 1) (no pagination)	More results subsequently published in a peer-reviewed journal
Al-Kali A et al. Patterns of molecular response to and relapse after combination of sorafenib, idarubicin, and cytarabine in patients with FLT3 mutant acute myeloid leukemia. Clinical Lymphoma, Myeloma and Leukemia. 11 (4) (pp 361-366), 2011	No relevant endpoints
Altman JK, et al. Antileukemic activity and tolerability of ASP2215 80mg and greater in FLT3 mutation-positive subjects with relapsed or refractory acute myeloid leukemia: Results from a phase 1/2, open-label, dose-escalation/dose-response study. Blood. 2015; 126 (23):321	More results subsequently published in a peer-reviewed journal
Altman JK, et al. Deep molecular response to gilteritinib to improve survival in FLT3 mutation-positive relapsed/refractory acute myeloid leukemia. Journal of Clinical Oncology. Conference: 2017 Annual Meeting of the American Society of Clinical Oncology, ASCO. United States. 35 (15 Supplement 1) (no pagination), 2017. Date of Publication: 20 Jun 2017.	More results subsequently published in a peer-reviewed journal
Anonymous. Correction to Selective inhibition of FLT3 by gilteritinib in relapsed or refractory acute myeloid leukaemia: a multicentre, first-in-human, open-label, phase 1-2 study (Lancet Oncol (2017) 18 (1061-75)(S1470204517304163)(10.1016/S1470-2045(17)30416-3)). The Lancet Oncology. 19 (7) (pp e335), 2018. Date of Publication: July 2018.	The correction had already been applied to the abstracted study
Anonymous. Correction: Selective inhibition of FLT3 by gilteritinib in relapsed or refractory acute myeloid leukaemia: a multicentre, first-in-human, open-label, phase 1-2 study (The Lancet Oncology (2017) 18(8) (1061-1075) (S1470204517304163)(10.1016/S1470-2045(17)30416-3)). The Lancet Oncology. 18 (12) (pp e711), 2017. Date of Publication: December 2017.	The correction had already been applied to the abstracted study
Antar A et al. Inhibition of FLT3 in AML: A focus on sorafenib. Bone Marrow Transplantation. 52 (3) (pp 344-351), 2017. Date of Publication: 01 Mar 2017.	No results are provided
Best-Aguilera C et al. Treatment of Acute Myeloid Leukemia with the FLT3 Gene Mutation. Current Oncology Reports. 19 (3) (no pagination), 2017. Article Number: 21. Date of Publication: 01 Mar 2017.	No results are provided
Boddu P et al. Outcomes by treatment setting and genomic profile in patients with AML on cladribine, idarubicin, and cytarabine. Blood. Conference: 59th Annual Meeting of the American Society of Hematology, ASH 2017. United States. 130 (Supplement 1) (no pagination), 2017. Date of Publication: December 2017.	No separate results for FLT3 in abstract; poster could not be found
Borthakur G et al. Phase I study of sorafenib in patients with refractory or relapsed acute leukemias. Haematologica. 96 (1) (pp 62-68), 2011. Date of Publication: January 2011.	No relevant endpoints
Burgues J. Treatment with flag-IDA or flago-IDA regimen in adult patients with relapsed/refractory acute myeloid leukemia. Retrospective analysis of the pethema AML registry. Haematologica. Conference: 18th Congress of the European Hematology Association. Stockholm Sweden. Conference Publication: (var.pagings). 98 (SUPPL. 1) (pp 25), 2013. Date of Publication: 01 Jun 2013.	No separate results for FLT3 in abstract; poster could not be found

Publication	Reason for exclusion
Canaani J et al. Use of FLT3 inhibitors to bridge relapsed/refractory AML patients to an allogeneic stem cell transplant. Biology of Blood and Marrow Transplantation. Conference: 2016 BMT Tandem Meetings. Honolulu, HI United States. Conference Publication: (var.pagings). 22 (3 SUPPL. 1) (pp S199-S200), 2016. Date of Publication: March 2016.	The abstract did not provide information on exact therapy; poster could not be found
Cloe A, Larson RA, Cheng JX. FLT3 inhibitors for the treatment of acute myeloid leukemia: An evaluation of efficacy of target inhibition and relationship to disease progression. Blood. Conference: 57th Annual Meeting of the American Society of Hematology, ASH 2015. San Diego, CA United States. Conference Publication: (var.pagings). 126 (23) (pp 4940), 2015. Date of Publication: 03 Dec 2015.	No results are provided
Cooper TM, et al. A Phase I Study of Quizartinib Combined with Chemotherapy in Relapsed Childhood Leukemia: A Therapeutic Advances in Childhood Leukemia & Lymphoma (TACL) Study. Clinical Cancer Research. 22(16):4014-22, 2016 Aug 15.	No separate results for FLT3 - children only
Daver NG, et al. First-in-human study of FLX925, an orally administered FLT3/CDK4/CDK6 inhibitor, in subjects with relapsed or refractory acute myeloid leukemia (AML). Journal of Clinical Oncology. Conference: 2015 Annual Meeting of the American Society of Clinical Oncology, ASCO. Chicago, IL United States. Conference Publication: (var.pagings). 33 (15 SUPPL. 1) (no pagination), 2015. Date of Publication: 20 May 2015.	Clinical research was stopped
Dohner H, et al. Efficacy and safety of quizartinib (AC220) in patients age >=60 years with FLT3-ITD-positive relapsed/refractory acute myeloid leukemia (AML). Haematologica. Conference: 18th Congress of the European Hematology Association. Stockholm Sweden. Conference Publication: (var.pagings). 98 (SUPPL. 1) (pp 233), 2013. Date of Publication: 01 Jun 2013.	More results subsequently published in a peer-reviewed journal
Fathi AT, Chen YB. Treatment of FLT3-ITD acute myeloid leukemia. American Journal of Blood Research. 1(2):175-89, 2011.	No results are provided
Gill H, Leung AY, Kwong Y-L. Molecularly targeted therapy in acute myeloid leukemia. Future Oncology. 12 (6) (pp 827-838), 2016. Date of Publication: March 2016.	No results are provided
Giri S et al. Sorafenib in Relapsed AML With FMS-Like Receptor Tyrosine Kinase-3 Internal Tandem Duplication Mutation. Journal of the National Comprehensive Cancer Network : JNCCN. 13 (5) (pp 508-514), 2015. Date of Publication: 01 May 2015.	Case report only
Halpern AB et al. Single center experience treating adults with FLT3-mutated acute myeloid leukemia (AML). Journal of Clinical Oncology. Conference: 2012 Annual Meeting of the American Society of Clinical Oncology, ASCO. Chicago, IL United States. Conference Publication: (var.pagings). 30 (15 SUPPL. 1) (no pagination), 2012. Date of Publication: 20 May 2012.	No separate results for FLT3 in abstract; poster could not be found
Hassanein M et al. FLT3 Inhibitors for Treating Acute Myeloid Leukemia. Clinical Lymphoma, Myeloma and Leukemia. 16 (10) (pp 543-549), 2016. Date of Publication: 01 Oct 2016.	No results are provided
Illmer T, Ehninger G. FLT3 kinase inhibitors in the management of acute myeloid leukemia. Clinical Lymphoma and Myeloma. 8 (SUPPL. 1) (pp S24-S34), 2008. Date of Publication: 2008.	No results are provided
Itzykson R et al. Azacitidine for the treatment of relapsed and refractory AML in older patients. Leukemia Research. 39(2):124-30, 2015 Feb.	No FLT3-related data

Publication	Reason for exclusion
Ivanoff S et al. 5-Azacytidine treatment for relapsed or refractory acute myeloid leukemia after intensive chemotherapy. American Journal of Hematology. 88(7):601-5, 2013 Jul.	No FLT3-related data
Jabbour E et al. Phase 2 study of low-dose clofarabine plus cytarabine for patients with higher-risk myelodysplastic syndrome who have relapsed or are refractory to hypomethylating agents. Cancer. 123 (4) (pp 629-637), 2017. Date of Publication: 15 Feb 2017.	No separate results for FLT3
Kalaycio M et al. Chemotherapy for acute myelogenous leukemia in the elderly with cytarabine, mitoxantrone, and granulocyte-macrophage colony-stimulating factor. American Journal of Clinical Oncology: Cancer Clinical Trials. 24 (1) (pp 58-63), 2001. Date of Publication: 2001.	No FLT3-related data
Kavanagh S et al. AML refractory to primary induction with Ida-FLAG has a poor clinical outcome. Leukemia Research. 68 (pp 22-28), 2018. Date of Publication: May 2018.	No FLT3-related data
Kaya AH et al. Efficacy of CLARA in recurrent/refractory acute myeloid leukaemia patients unresponsive to FLAG chemotherapy. Journal of Chemotherapy. 30 (1) (pp 44-48), 2018. Date of Publication: 02 Jan 2018.	No FLT3-related data
Khaled S, et al. Concordance between bone marrow and peripheral blood samples for assessment of FLT3 internal tandem duplication (ITD) mutations: Data from patients screened for participation in quantum-r, a global, randomized, open-label, phase 3 study examining the effect of quizartinib monotherapy vs salvage chemotherapy on overall survival in patients with FLT3-itd-mutated AML who are refractory to or have relapsed after first-line therapy. Blood. Conference: 59th Annual Meeting of the American Society of Hematology, ASH 2017. United States. 130 (Supplement 1) (no pagination), 2017. Date of Publication: December 2017.	No relevant endpoints
Lange A et al. The sorafenib anti-relapse effect after alloHSCT is associated with heightened alloreactivity and accumulation of CD8+PD-1+ (CD279+) lymphocytes in marrow. PLoS ONE. 13 (1) (no pagination), 2018. Article Number: e0190525. Date of Publication: January 2018.	No relevant endpoints
Larrosa-Garcia M, Baer MR. FLT3 Inhibitors in acute myeloid leukemia: Current status & future directions. Molecular Cancer Therapeutics. 16 (6) (pp 991-1001), 2017. Date of Publication: June 2017.	No results are provided
Levis M, et al. Evaluation of the impact of signal ratio on overall survival in FLT3-mutation-positive relapsed/refractory acute myeloid leukemia following once-daily treatment with gilteritinib. Haematologica. Conference: 22th Congress of the European Hematology Association. Spain. 102 (Supplement 2) (pp 216-217), 2017. Date of Publication: June 2017	More results subsequently published in a peer-reviewed journal
Levis M. Novel FLT3 inhibitors and targeted therapies in AML. Annals of Hematology. Conference: Acute Leukemias XV: Biology and Treatment Strategies. Munich Germany. Conference Publication: (var.pagings). 94 (1 SUPPL. 1) (pp S35-S37), 2015. Date of Publication: February 2015.	No results are provided
Levis MJ et al. Final results of a phase 2 open-label, monotherapy efficacy and safety study of quizartinib (AC220) in patients with FLT3-ITD positive or negative relapsed/refractory acute myeloid leukemia after second-line chemotherapy or hematopoietic stem cell transplantation. Blood. Conference: 54th Annual Meeting of the American Society of Hematology, ASH 2012. Atlanta, GA United States. Conference Publication: (var.pagings). 120 (21) (no pagination), 2012. Date of Publication: 16 Nov 2012.	More results subsequently published in a peer-reviewed journal

Publication	Reason for exclusion
Levis MJ et al. Results of a first-in-human, phase I/II trial of ASP2215, a selective, potent inhibitor of FLT3/Axl in patients with relapsed or refractory (R/R) acute myeloid leukemia (AML). Journal of Clinical Oncology. Conference: 2015 Annual Meeting of the American Society of Clinical Oncology, ASCO. Chicago, IL United States. Conference Publication: (var.pagings). 33 (15 SUPPL. 1) (no pagination), 2015. Date of Publication: 20 May 2015.	More results subsequently published in a peer-reviewed journal
Levis MJ, et al. Evaluation of the impact of minimal residual disease, FLT3 allelic ratio, and FLT3 mutation status on overall survival in FLT3 mutation-positive patients with relapsed/refractory (R/R) acute myeloid leukemia (AML) in the chrysalis phase 1/2 study. Blood. Conference: 59th Annual Meeting of the American Society of Hematology, ASH 2017. United States. 130 (Supplement 1) (no pagination), 2017. Date of Publication: December 2017.	More results subsequently published in a peer-reviewed journal
Liegel J, Courville E, Sachs Z, Ustun C. Use of sorafenib for post-transplant relapse in FLT3/ITD-positive acute myelogenous leukemia: Maturation induction and cytotoxic effect. Haematologica. 99 (11) (pp e222-e224), 2014. Date of Publication: 01 Nov 2014.	Case study
McMahon CM et al. Mechanisms of acquired resistance to gilteritinib therapy in relapsed and refractory FLT3-mutated acute myeloid leukemia. Blood. Conference: 59th Annual Meeting of the American Society of Hematology, ASH 2017. United States. 130 (Supplement 1) (no pagination), 2017. Date of Publication: December 2017.	No relevant endpoints
Nazha A, et al. A phase I/II Trial of combination of midostaurin (PKC412) and 5-azacytidine (5-AZA) for the treatment of patients with refractory or relapsed (R/R) acute myeloid leukemia (AML) and myelodysplastic syndrome (MDS). Blood. 2012; 120 (21) (no pagination)	The abstract did not provide separate information by therapy; poster could not be found
Nazha A, et al. A phase I/II trial of combination of PKC412 and 5-azacytidine (AZA) for the treatment of patients with refractory or relapsed (R/R) acute myeloid leukemia (AML) and myelodysplastic syndrome (MDS). Journal of Clinical Oncology. 2012 30 (15 SUPPL. 1) (no pagination)	The abstract did not provide separate information by therapy; poster could not be found
Ohanian M et al. Final report of phase II study of sorafenib and 5-azacytidine in patients with relapsed or untreated acute myeloid leukemia and FLT3-ITD mutation. Blood. Conference: 55th Annual Meeting of the American Society of Hematology, ASH 2013. New Orleans, LA United States. Conference Publication: (var.pagings). 122 (21) (no pagination), 2013. Date of Publication: 21 Oct 2013.	More results subsequently published in a peer-reviewed journal
Perl A et al. Results of a first-in-human, phase 1/2 trial of ASP2215, a selective, potent oral inhibitor of FLT3/AXL, in patients with relapsed or refractory acute myeloid Leukemia. Haematologica. Conference: 20th Congress of the European Hematology Association. Vienna Austria. Conference Publication: (var.pagings). 100 (SUPPL. 1) (pp 317-318), 2015. Date of Publication: 22 Jun 2015.	More recent data abstracted
Perl AE, et al. A phase 3, open-label, randomized study of the FLT3 inhibitor gilteritinib versus salvage chemotherapy in adults with first relapse or primary refractory FLT3 mutation-positive acute myeloid leukemia. Journal of Clinical Oncology. 2016 34 (Supplement 15) (no pagination)	No results are provided
Prebet T et al. Improved outcome of patients with low- and intermediate-risk cytogenetics acute myeloid leukemia (AML) in first relapse with gemtuzumab and cytarabine versus cytarabine: Results of a retrospective comparative study. Cancer. 117 (5) (pp 974-981), 2011. Date of Publication: 01 Mar 2011.	No separate results for FLT3
Ravandi F et al. Final report of phase ii trial of combination of sorafenib and 5-azacytidine in patients with FLT3-ITD positive acute myeloid leukemia. Haematologica. Conference: 18th Congress of the European Hematology	More results subsequently published in a peer-reviewed journal

Publication	Reason for exclusion
Association. Stockholm Sweden. Conference Publication: (var.pagings). 98 (SUPPL. 1) (pp 248), 2013. Date of Publication: 01 Jun 2013.	
Roboz GJ. Alliance trials for AML. Annals of Hematology. Conference: Acute Leukemias XV: Biology and Treatment Strategies. Munich Germany. Conference Publication: (var.pagings). 94 (1 SUPPL. 1) (pp S43-S46), 2015. Date of Publication: February 2015.	No FLT3-related data
Saygin C, Carraway HE. Emerging therapies for acute myeloid leukemia. Journal of Hematology and Oncology. 10 (1) (no pagination), 2017. Article Number: 93. Date of Publication: 18 Apr 2017.	No results are provided
Schiller GJ, et al. Final results of a randomized phase 2 study showing the clinical benefit of quizartinib (AC220) in patients with FLT3-ITD positive relapsed or refractory acute myeloid leukemia. Journal of Clinical Oncology. Conference: 2014 Annual Meeting of the American Society of Clinical Oncology, ASCO. Chicago, IL United States. Conference Publication: (var.pagings). 32 (15 SUPPL. 1) (no pagination), 2014. Date of Publication: 20 May 2014.	More results subsequently published in a peer-reviewed journal
Schmalbrock LK, et al. Characterization of FLT3 mutations at diagnosis, refractory disease or relapse in aml patients treated with midostaurin within the CALGB 10603 (ratify) and AMLSG 16-10 trials. Haematologica. Conference: 22th Congress of the European Hematology Association. Spain. 102 (Supplement 2) (pp 358-359), 2017. Date of Publication: June 2017.	In vitro study
Schroeder T et al. Azacitidine and donor lymphocyte infusions as first salvage therapy for relapse of AML or MDS after allogeneic stem cell transplantation. Leukemia. 27 (6) (pp 1229-1235), 2013. Date of Publication: June 2013.	No FLT3-related data
Sibon D et al. Use of clofarabine in the treatment of relapsed or refractory acute myeloid leukemia in adults: The french experience. Blood. Conference: 53rd Annual Meeting of the American Society of Hematology, ASH 2011. San Diego, CA United States. Conference Publication: (var.pagings). 118 (21) (no pagination), 2011. Date of Publication: 18 Nov 2011.	No separate results for FLT3 in abstract; poster could not be found
Sid S et al. Treatment of Post-transplant Relapse of FLT3-ITD Mutated AML Using 5-Azacytidine and Sorafenib Bitherapy. Clinical Lymphoma, Myeloma and Leukemia. 17 (4) (pp 241-242), 2017. Date of Publication: 01 Apr 2017.	Case series
Smith CC et al. Pharmacokinetic profile and pharmacodynamic effects of ASP2215, a selective, potent inhibitor of FLT3/AXL, in patients with relapsed or refractory acute myeloid leukemia: Results from a first-in-human phase 1/2 study. Blood. Conference: 57th Annual Meeting of the American Society of Hematology, ASH 2015. San Diego, CA United States. Conference Publication: (var.pagings). 126 (23) (pp 4836), 2015. Date of Publication: 03 Dec 2015.	No relevant endpoints
Smith CC et al. Pharmacokinetics and pharmacodynamics of gilteritinib in patients with relapsed or refractory acute myeloid leukemia. Journal of Clinical Oncology. Conference: 2016 Annual Meeting of the American Society of Clinical Oncology, ASCO 2016. United States. 34 (Supplement 15) (no pagination), 2016. Date of Publication: May 2016.	No relevant endpoints
Smith CC, et al. Comparative assessment of FLT3 variant allele frequency by capillary electrophoresis and next-generation sequencing in FLT3^{mut+} patients with relapsed/refractory (R/R) acute myeloid leukemia (AML) who received gilteritinib therapy. Blood. Conference: 59th Annual Meeting of the American Society of Hematology, ASH 2017. United States. 130 (Supplement 1) (no pagination), 2017. Date of Publication: December 2017.	No relevant endpoints
Song Y et al. Navigating care in a challenging case of t(6;9) and flt3 positive aml. Journal of Investigative Medicine. Conference: 2015 Combined Annual Meeting of the Central Society for Clinical and Translational Research and the	No results are available

Publication	Reason for exclusion
Midwestern Section American Federation for Medical Research. Chicago, IL United States. Conference Publication: (var.pagings). 63 (4) (pp 689-690), 2015. Date of Publication: April 2015.	
Stelljes M et al. Allogeneic transplantation versus chemotherapy as postremission therapy for acute myeloid leukemia: A prospective matched pairs analysis. Journal of Clinical Oncology. 32 (4) (pp 288-296), 2014. Date of Publication: 01 Feb 2014.	No FLT3-related data
Sternberg DW, Licht JD. Therapeutic intervention in leukemias that express the activated fms-like tyrosine kinase 3 (FLT3): Opportunities and challenges. Current Opinion in Hematology. 12 (1) (pp 7-13), 2005. Date of Publication: January 2005.	No results are provided
Stone RM. FLT3 Inhibitors in Acute Myeloid Leukemia: An Update. Annals of Hematology. Conference: Acute Leukemias XIII: Biology and Treatment Strategies. Munich Germany. Conference Publication: (var.pagings). 90 (SUPPL. 1) (pp S70-S72), 2011. Date of Publication: February 2011.	No results are provided
Swaminathan M et al. The combination of quizartinib with azacitidine or low dose cytarabine is highly active in patients (PTS) with FLT3-ITD mutated myeloid leukemias: Interim report of a phase I/II trial. Blood. Conference: 59th Annual Meeting of the American Society of Hematology, ASH 2017. United States. 130 (Supplement 1) (no pagination), 2017. Date of Publication: December 2017.	No separate results for FLT3 in abstract; poster could not be found
Swaminathan M et al. The combination of quizartinib with azacitidine or low dose cytarabine is highly active in patients (pts) with FLT3-ITD mutated myeloid leukemias: Interim report of a phase I/II trial. Clinical Lymphoma, Myeloma and Leukemia. Conference: 28th International Association for Comparative Research on Leukemia and Related Diseases Symposium, IACRLRD 2017. United States. 17 (10 Supplement 1) (pp S3), 2017. Date of Publication: October 2017.	No separate results for FLT3 in abstract; poster could not be retrieved
Swords R, Freeman C, Giles F. Targeting the FMS-like tyrosine kinase 3 in acute myeloid leukemia. Leukemia. 26 (10) (pp 2176-2185), 2012. Date of Publication: October 2012.	No results are provided
Thomas CM, Campbell P. FLT3 inhibitors in acute myeloid leukemia: Current and future. Journal of Oncology Pharmacy Practice. :1078155218802620, 2018 Sep 30	No results are provided
Tickenbrock L, et al. Emerging Flt3 kinase inhibitors in the treatment of leukaemia. Expert Opinion on Emerging Drugs. 11 (1) (pp 153-165), 2006. Date of Publication: March 2006.	No results are provided
Usuki K et al. Clinical profile of gilteritinib in Japanese patients with relapsed/refractory acute myeloid leukemia: An open-label phase 1 study. Cancer Science. 109(10):3235-3244, 2018 Oct.	No FLT3-related data
Uy GL et al. Addition of sorafenib to chemotherapy improves the overall survival of older adults with FLT3-ITD mutated acute myeloid leukemia (AML) (alliance C11001). Blood. Conference: 57th Annual Meeting of the American Society of Hematology, ASH 2015. San Diego, CA United States. Conference Publication: (var.pagings). 126 (23) (pp 319), 2015. Date of Publication: 03 Dec 2015.	No data for relapsed/refractory
Zappone E et al. FLT3 inhibitors in the management of acute myeloid leukemia. Anti-Cancer Agents in Medicinal Chemistry. 17 (8) (pp 1028-1032), 2017. Date of Publication: 01 Jul 2017.	No results are provided
Zhang W, et al. Combinatorial targeting of XPO1 and FLT3 exerts synergistic anti-leukemia effects through induction of differentiation and apoptosis in FLT3-mutated acute myeloid leukemias: from concept to clinical trial. Haematologica. 103(10):1642-1653, 2018 Oct.	No relevant endpoints

Subgroup analyses

A18. CS, Section B.2.7, page 39. Please provide subgroup analyses of OS with the stratification variables included in models irrespective of their statistical significance and with additional baseline characteristics (particularly sex, race, baseline ECOG and region) included in full and reduced models assessed using likelihood ratio tests. Also, unless there is a justification for the risk of death changing when patients become 65 years of age, please include age as a continuous variable, ideally allowing for non-linear relationships using splines.

Astellas is still exploring the provision of this data and will provide an update in due course.

Adverse events

A19. CS, Section B.2.10, Page 40. With reference to the safety profile of gilteritinib in ADMIRAL, the text on page 40 states that *"Gilteritinib was generally well tolerated."* The next sentence states *"Almost all patients experienced at least one TEAE."* These two statements appear to conflict with one another. Please clarify.

Overall it is considered that gilteritinib was generally well tolerated. In discussing the AEs it should be borne in mind that the duration of treatment with gilteritinib was longer compared to the comparator salvage chemotherapy arm (median 126 days vs. 28 days[11]) and therefore patients accumulated more AEs. When adjusted by patient-year of exposure, the incidence of all types of TEAE was lower in the gilteritinib arm than the salvage chemotherapy arm.

In addition, many AEs were considered to be tolerated and manageable (e.g., diarrhoea, constipation, nausea, cough). Therefore, compared to the comparator salvage chemotherapy, gilteritinib was generally well tolerated.

A20. CS, Section B.2.10, page 40. This CS presents an analysis of AEs by patient-years of exposure. Given that salvage chemotherapy was given for a single cycle in most cases (typically 5-7 days), please clarify how patient-years of exposure have been calculated

The 'patient-years of exposure' per treatment group was calculated as the sum of duration of exposure of all patients per treatment group.

In the gilteritinib group patient-year = number of patients*mean [duration of exposure]; i.e. [redacted] (patient-days), which then converts to patient-years [redacted]

In the salvage chemotherapy group, patient-year = number of patients*mean [duration of exposure] i.e. [redacted] (patient-days), which then converts to patient-years [redacted].

Section B: Clarification on cost-effectiveness data

Comparators

B1. PRIORITY. CS, Section B.3, Table 16, page 44. Please clarify why indirect comparisons were not performed between the treatments included in NICE scope.

The ADMIRAL RCT directly compares gilteritinib with the relevant comparators defined in the Scope. This is the best option for comparative data in the specific target population is FLT3 mutation positive AML.

Table 7 highlights limitations of NMA in this target population and provides the reasons for not including them in the evidence network and include:

- Relevant studies being ongoing with no data yet available
- Relevant studies being single arm studies or not informing the evidence network because they were dose-finding studies.

B2. PRIORITY. CS, Appendix D, page 30. Please provide a table summarising reasons why each of the studies of comparators included in the systematic review could not contribute to an NMA / indirect comparisons.

The reasons for exclusion from the NMA are provided in Table 7.

Table 7 Reasons for Exclusion from the NMA

Study acronym - NCT ID	Reference	Treatment arms	Reason for exclusion
CHRYSALIS NCT02014558	Perl 2016; Perl 2017B	Seven doses of gilteritinib	Dose finding study. It could not enrich the network evidence
2215-CL-0303 NCT03182244	CT.Gov	Gilteritinib vs salvage chemo (LoDAC, MEC, G-CSF, FLAG)	No results
2215-CL-1101 NCT02421939	CT.Gov	Gilteritinib + atezolizumab	No results
2215-CL-9100 NCT03070093	CT.Gov	Gilteritinib	No results
2215-CL-9200 NCT03409081	CT.Gov	Gilteritinib	No results
M16-802 NCT03625505	CT.Gov	Venetoclax + Gilteritinib	Single arm
2018-0608 NCT03735875	CT.Gov	Venetoclax + quizatinib 30 mg/d	Single arm
2689-CL-2004 NCT01565668	Cortes 2013A; Cortes 2018A; Martinelli 2014; Russell 2014	Quizartinib 30 mg/d vs 0 mg/d	Dose finding study. It could not enrich the network evidence

Study acronym - NCT ID	Reference	Treatment arms	Reason for exclusion
ACE NCT00989261	Cortes 2013B; Cortes 2018B; Hills 2015; Levis 2013; Martinelli 2013; Martinelli 2014	Quizartinib 200mg	Single arm
AC220-A-J201 NCT02984995	CT.Gov	Quizartinib	Single arm; no results
AC220-A-U203 NCT03746912	CT.Gov	Quizartinib	Single arm; no results
ARO-004 ARO-005 NCT01522469 NCT01657682	Cortes 2016A	Crenolanib	Single arm
ARO-007 NCT02298166	CT.Gov	Crenolanib + chemotherapy vs placebo + chemotherapy	No results
ARO-013 NCT03250338	CT.Gov	Crenolanib + chemotherapy (HAM+FLAG-IDA) vs chemotherapy	No results
N/A	Iyer 2016	HAM followed by crenolanib	Single arm
N/A	Randhawa 2014	Crenolanib 200 mg/m/day TID	Single arm
2010-0511 NCT01254890	Ravandi 2013	Sorafenib (400 mg orally BID) + azacitidine	Single arm
AML004 NCT03622541	CT.Gov	Sorafenib	Single arm
KCP-330-001 NCT01607892	Daver 2017	Sorafenib (400mg BID) + selinexor	Single arm
SIRA NCT02867891	CT.Gov; Lohmeyer 2018	Sorafenib	No results
N/A	Fleischmann 2016	Sorafenib	Single arm
N/A	Fleischmann 2017	Sorafenib	Single arm
N/A	Freitas 2016	Sorafenib 400mg BID	Single arm
N/A	Metzelder 2009	Sorafenib 400 mg BID	Single arm
N/A	Metzelder 2010	Sorafenib 400 mg BID	Single arm
N/A	Metzelder 2012	Sorafenib 400 mg BID	Single arm
N/A	Metzelder 2017	Sorafenib	Single arm
N/A	Rautenberg 2017	Sorafenib (400 mg BID) + azacitidine (75 mg/m^2 for 7 d every 28 d)	Single arm
N/A	Schroeder 2009	Sorafenib 800mg QD	Single arm
N/A	Sharma 2011	Sorafenib 400mg BID or 600mg BID +/- chemotherapy	Single arm
N/A	Sid 2017	Sorafenib (400 mg BID) + azacitidine (75 mg/m^2 for 7 d every 28 d)	Single arm
N/A	Xuan 2018	Sorafenib (400mg BID) + chemo + donor lymphocyte infusions	Single arm
NCT03642236	CT.Gov	Ibrutinib 420mg + sorafenib 0.4mg BID	No results

Study acronym - NCT ID	Reference	Treatment arms	Reason for exclusion
N/A	Chevallier 2010	Gemtuzumab ozogamicin (9 mg/m2 at day 4) + cytarabine (1 g/m2 BID for days 1–5) + mitoxantrone (12 mg/m2/d for days 1–3)	Single arm

B3. CS, Section B.3.7, Table 40, page 101. The base case results are presented for gilteritinib versus a weighted comparator of salvage chemotherapies or against specific salvage chemotherapy regimens.

(a) Please explain why the comparisons against individual regimens use pooled data for the overall trial comparator group, rather than regimen-specific data.

(b) Which of these comparisons represents the company's base case – the weighted comparator or the individual comparisons?

a) ADMIRAL was designed to compare gilteritinib vs. salvage chemotherapy. Performing comparisons vs. individual comparators would not be appropriate given the small number of patients receiving individual regimes. Analyses of treatment effect with regard to survival would not be powered to evaluate these outcomes and would differ in terms of costs.

b) The weighted comparator is the base case, which aligns with the NICE decision problem and reflects UK clinical practice.

Model structure / approach

B4. PRIORITY. CS, Section B.3.2.3.1, page 49. Please clarify why it was necessary to stratify the partitioned survival model according to receipt/non-receipt of HSCT? Please also provide further details regarding why a state transition approach was not used.

The model structure stratified the partitioned survival model according to the receipt of HSCT based on the following considerations:

- HSCT is a key clinical event that is prognostic of patient outcomes in AML. Several studies have investigated the relationship between HSCT status and the long-term survival outcomes among patients with relapsed or refractory AML. These studies consistently found that HSCT appeared to be an important prognostic determinant of overall survival. Patients with HSCT had substantially longer median survival and significantly decreased mortality risk compared to those without HSCT[13,14,15]. Similar evidence were seen among patients with FLT3 mutation positive AML[16,17].

- The selected model structure was developed incorporating suggestions from PRIMA advice. In the model submitted for PRIMA review, a three-state partition survival structure was used without the stratification by HSCT. Feedback from PRIMA noted that although the three-state model is most commonly used for advanced cancer therapies, the heterogeneity within the event-free health state, particularly in regards to HSCT status, cannot be fully modelled by one parametric curve or represented by same utility and cost parameters. To address this comment, the three-state partition survival model was further stratified by the HSCT status.
- In addition, the selected model structure was shared with and endorsed by HTA experts from England, France, Spain, Italy and Canada at an advisory board meeting. It is also aligned with selected prior HTA submission in advanced or relapsed/refractory haematological cancers[18,19,20].

The state transition approach was not used due to its inherent constraints and limitations versus the current model structure.

- Fundamentally, both the state transition model and the submitted stratified partitioned survival model, distributed patients to various health states to estimate costs and effectiveness. For the current decision problem, extrapolation would be needed under both model types. A state transition approach, however, can be challenging when incorporating time varying transition probabilities across various health states, which is required for the current decision problem.
- The state transition approach is not an ideal modelling approach to incorporate time varying transition probabilities across various health states. State transition approach incorporates an explicit link across different health states, and therefore requires estimation of transition probabilities for each possible transition across health states. To estimate all required transition probabilities, substantial data and assumptions would be needed. For example, transitions between EFS with HSCT state and alive and post-event with HSCT state would need to be explicitly estimated. However, given the rate of transitions between these two states are not constant (EFS and OS curves would converge over time for long-term survivors), time varying transition probabilities would be required. In a state transition model, it could be challenging to incorporate time-varying transition probabilities for all considered health states, especially for the interim health states that patients may move into at different time points as they pass through the model. Consideration of time-varying transition probabilities is resolved more clearly using the submitted modelling approach.
- Astellas believes a state transition model would not have provided additional value compared to the current model structure. The current structure allowed the use of trial

data directly to estimate the distributions of patients across health states over time, which allows a satisfactory fit to the trial data. The only major assumption in the current model structure is around the timing of HSCT. The current model structure assumes HSCT happens during a decision-tree period. The length of the decision tree period was informed by the average time from randomization to HSCT as observed in the ADMIRAL trial. The state transition model can potentially allow transition to HSCT over time, however, it would be challenging to consider time varying transition between EFS with HSCT and post-event with HSCT after patient received HSCT as these are interim health states that patients move into at different time points. Based on the existing evidence, the rate of transitions between these two states are not constant (EFS and OS curves would converge over time for long-term survivors). Therefore, the current model structure is more appropriate as it could more accurately track the clinical pathway of the target population.

B5. CS, Section B.3.2.3.1, Table 18, page 52. Table 18 of the CS lists concerns raised in the midostaurin appraisal. However, this information relates to a different technology used at a different point in the AML pathway. Please clarify which of the issues described in Table 18 are relevant to the present gilteritinib appraisal.

Concerns described here reflect those proposed and discussed during the Decision Problem meeting, and while Astellas agrees that these describe a different indication, Astellas hoped that acknowledging the preferences conveyed by NICE regarding midostaurin appraisal would be a pragmatic foundation for key assumptions in our own submission.

Assumptions of cure

B6. PRIORITY. CS, Section B.3.2, page 55. Please provide the evidence used to support the assumption of a cure in: (a) patients who undergo HSCT and (b) patients who do not undergo HSCT. Please also comment on the evidence to support the assumed 3-year timepoint for cure.

The cure assumption and the time point to introduce cure was based on the clinical inputs and Committee comments provided for the midostaurin submission (TA 523)[21]. In the midostaurin submission for FLT3 mutation positive AML patients, the Committee concluded that surviving patients (those with and without HSCT) after 3 years should enter a cured state and there should be no health state costs for these patients. Specifically, the Committee suggested that patients' disease may be cured by chemotherapy alone, and indicated that the mortality risk might be lower after the cure point for FLT3 mutation positive AML patients managed by chemotherapy alone.

The 3-year period is consistently cited in existing AML submissions and literature [21,22,23,24], representing a clinically important time point for patients to reach given the limited risk of relapses. This assumption was further validated based on clinical inputs and the observed plateauing after year 3 in the reported OS curves of relapsed or refractory AML patients in the literature[23,25].

Additionally, the cure assumption was also introduced to reduce the uncertainty associated with the survival extrapolation. In the ADMIRAL trial, data was available for up to 33 months after randomization. However, the survival curves were based on very small number of patients after year 2, and may not be stable to inform long-term extrapolation. With the current assumption, all patients who remained alive, regardless of the prior treatment, were assumed to have the same mortality risk after the cure point (i.e., year 3), which helps reduce uncertainty from the extrapolation and can also be considered as conservative as no additional benefit from gilteritinib vs. comparators was assumed after year 3.

In the submission model, a majority of the patients who remained survive at year 3 were those with HSCT. Specifically, 90% of the survivors at year 3 received HSCT in the salvage chemotherapy arm, and 77% of the survivors at year 3 received HSCT in the gilteritinib arm.

B7. PRIORITY. CS, Section B.1.3, page 14. The text states *"The goal of treatment in AML is to induce remission and prevent relapse, with haematopoietic stem cell transplant (HSCT) being the only treatment with curative intent."* In light of this, please explain why a structural assumption of cure has been applied to patients who do not undergo HSCT.

The referenced statement was a general statement based on the current treatment landscape. HSCT was considered the only plausible cure for AML patients as no effective therapeutic options are currently licensed for this patient population. However, the treatment landscape is expected to evolve with the introduction of new innovative treatment such as gilteritinib and midostaurin.

The cure assumption and the time point to introduce cure was based on the clinical inputs and Committee comments provided for the midostaurin submission (TA 523)[21]. In the midostaurin submission for FLT3 mutation positive AML patients, the ERG and Committee introduced a cured state after 3 years for all patients who remained alive, including both patients with and without HSCT. As described above, the Committee suggested that patients' disease may be cured by chemotherapy alone, and indicated that the mortality risk might be lower after the cure point for FLT3 mutation positive AML patients managed by chemotherapy alone.

As also described above, the 3-year period is consistently cited in existing AML submissions and literature,[21,22,23,24] representing a clinically important time point for patients to reach given

the limited risk of relapses. Again, the assumption was further validated based on clinical inputs and the observed plateauing after year 3 in the reported OS curves of relapsed or refractory AML patients in the literature[23,25].

Effectiveness parameters

B8. PRIORITY. CS, Section B.3.3, page 56. Please clarify what EFS and OS data are available for patients post-HSCT from ADMIRAL and explain why these data have not been used to inform outcomes for the "with HSCT" states in the model.

In the ADMIRAL trial, only survival information was collected after patients proceeded with HSCT. However, the sample size was limited and the data was not very mature.

- Total number of patients proceeded with HSCT was 82, including 63 patients randomized to gilteritinib arm and 19 patients randomized to salvage chemotherapy arm
- The median follow-up time after HSCT was 7.5 months; only 28 (34%) patients have survival data beyond year 1, and 2 (2.5%) patients have survival data beyond year 2 due to the loss of follow-up and censoring in the clinical trial.

Because of the limited sample size and the limited follow-up data time, the survival curve from the clinical trial has considerable uncertainty. The shape of the tail was heavily impacted by the censored patients and did not reflect the actual clinical course of the target patients. This would have major impact on the parametric fitting of the survival curves, leading to clinically implausible results.

Based on these considerations, data from the ADMIRAL trial was not used. Instead, a targeted literature review was conducted to identify relevant input to inform EFS and OS post-HSCT. In addition, to confirm that external data is comparable to the trial data, a comparison between the observed post-HSCT OS from the ADMIRAL trial and the external data of HSCT (i.e. Evers et al. 2018[26]) was conducted and no statistically significant difference was found (p = 0.9428). The comparison of post-HSCT OS estimates from ADMIRAL and Evers et al. 2018[26] for the first 12 months (where there are more observations available in the ADMIRAL trial) are presented in Table 8 below. The estimates are largely similar between the two sources.

Table 8 Comparison of Post-HSCT OS Estimates from ADMIRAL and Evers

Months	ADMIRAL	Evers et al. 2018[26]
1	[redacted]	94.5%
2	[redacted]	87.5%
3	[redacted]	80.5%
4	[redacted]	68.8%
5	[redacted]	68.8%

Months	ADMIRAL	Evers et al. 2018[26]
6	[redacted]	66.4%
7	[redacted]	61.7%
8	[redacted]	61.7%
9	[redacted]	59.4%
10	[redacted]	57.0%
11	[redacted]	57.0%
12	[redacted]	57.0%

B9. PRIORITY. CS, Section B.3.3.5, page 74. The text states that *"Evers et al.[43] was selected as the most relevant data source because it included the largest sample size and had the longest follow-up time."* Neither of these criteria included in the quote appear to concern relevance. Given that you have used external evidence to inform this part of the model, please clarify why you have selected a source in which patients do not have FLT3 mutation-positive disease.

Evers 2018 was identified as the most relevant data source from the list of articles identified from a pragmatic literature review[26]. Specifically, the following criteria was considered to select relevant publications: 1) comparable patient population to the model target population; 2) relevant OS and EFS measures reported in the form of K-M curves; 3) sufficient sample size and mature follow-up to reduce uncertainty with the survival extrapolation. All considered articles have to meet the first two criteria to be considered eligible. The third criteria was mainly used to identify the best evidence among all studies that fulfilled the first two criteria.

From the search, Astellas did not identify any publications that reported OS or EFS for relapsed or refractory FLT3 mutation positive AML patients specifically. In the absence of data, we evaluated the following two populations as proxy to the target population: 1) patients with relapsed or refractory AML (including patients with and without FLT3 mutation); 2) patients with FLT3 mutation positive AML (including both newly diagnosed and relapsed or refractory AML patients). Because there is evidence that demonstrates that FLT3 mutation status does not impact survival outcomes post-HSCT,[27] publications that evaluated survival for relapsed or refractory AML patients were considered more appropriate as a proxy to inform post-HSCT survival for the target population.

A number of publications were identified that reported survival outcomes for relapsed or refractory AML patients (see Table 9 below). Among those, Evers 2018 was selected in the base-case analysis due to the following considerations:

- Evers 2018 evaluated relapsed or refractory AML patients after 1st line treatment and included all relapsed or refractory patients in the evaluation[26]. This population matches closely with the ADMIRAL trial population (patients who are relapsed after or refractory

to 1st line treatment) compared to other publications that evaluated relapsed or refractory AML patients after more than one line of chemotherapy i.e. , Steckel 2018 and Schmid 2016[28,29]. Compared to Fong 2013 and Frazer 2017, Evers has a larger sample size[30,31].

- Based on the evaluation, Evers 2018 is a recent publication with the largest sample size and longest follow-up duration and presents a population similar to the ADMIRAL population[26]. Therefore, Evers 2018 would more likely to reflect recent evidence, and more stable prediction.

Table 9 Summary of Survival Outcomes for Relapsed or Refractory AML patients

Source	Patient Population	Sample Size	Follow-up	OS at year 3	OS Definition
Evers et al., 2018[26]	R/R AML patients after 1st line treatment	Overall: 498 CR2: 128	Maximum: over 11 years Median: 6.5 years	46%	From HSCT
Steckel et al., 2018[28]	R/R AML patients after 1-2 lines of chemotherapy	Overall: 292 Relapsed after 1st line: 51	Maximum: 5 years Median: 4.8 years	41%	From HSCT
Fong et al., 2013[30]	Relapsed AML patients	58	Maximum: 59 months Median: 6.7 months	55%	From treatment start date
Frazer et al., 2017[31]	AML patients in CR2 after relapse	55	NR	46%	From HSCT
Schmid et al., 2006[29]	R/R AML patients after one or multiple lines of chemotherapy	103	Maximum: 5.6 years Median: 2.1 years	32%	From HSCT

Because the selected publication to inform HSCT survival in the base-case does not evaluate FLT3 mutation positive patients specifically, Astellas also evaluated alternative publications that specifically assessed post-HSCT survival among FLT3 mutation positive patients (including both newly diagnosed and relapsed or refractory AML patients). Ustun 2017[32] was selected as the most appropriate publication in the scenario analysis as it is the most recent publication with the largest sample size. The model results were similar using this alternative input source.

B10. CS, Section B.3.3, page 57. What evidence is there to assert that there is a proportional relationship between the EFS cumulative hazard and the OS cumulative hazard for each treatment?

A single hazard ratio is only applied to the subgroup of patients that have experienced HSCT. There are no EFS data after HSCT from the ADMIRAL trial, which, in addition, is not powered to demonstrate differences in post-HSCT survival based on prior treatment. As the efficacy inputs of HSCT patients have been sourced from the literature and as all treatment arms share the same inputs, the question of relevance here is whether a PH relationship holds overall between OS with HSCT and EFS with HSCT in our sources. To assess the PH assumption,

we inspect log-cumulative hazard plots, indicating the log of cumulative hazard versus survival time.

Figure 1 presents the log-cumulative hazard plot of EFS versus OS for the unrelated donor (URD) group in Ustun[32], used to model the ratio between post-HSCT EFS and OS. Recall that the hazard ratio derived from Ustun is applied to the OS CR2 curve from Evers[26], included in Figure 1 for completeness, to inform EFS for patients with HSCT. This approach is necessary because there are no EFS data for post-HSCT survival in Evers. To produce the figure, individual-level data were simulated from the digitised K-M curves using the algorithm outlined in Guyot[33], and information on "numbers at risk" and "number of events".

Figure 1 does not reveal a distinct pattern of non-parallelism between EFS and OS in Ustun with the curves being roughly parallel and suggesting proportional hazards. The hazards are reasonably proportional between EFS and OS, particularly throughout the beginning of follow-up (recall that, in the model, a proportional relationship is only assumed to continue up to year 3; later, EFS converges to OS as reflected by the data). The gradient of the Evers OS curve is also reasonably constant initially, suggesting that the curve is compatible with the application of a single HR.

Finally, it is worth noting that similar approaches have been considered in prior NICE submissions and have been considered appropriate by the Committee members, e.g. TA554[34], TA567[35] and TA559[36].

Figure 1 Log-cumulative Hazard Plot for HSCT Efficacy Sources

B11. PRIORITY. CS, Section B.3.3.5, pages 73 to 75. Please provide details and justification for the unanchored indirect calculation of the OS hazard ratio for patients receiving gilteritinib post-HSCT in ADMIRAL and those in CR2 patients in Evers *et al*, including evidence to support the assumption that the hazard ratio is constant over time.

Such hazard ratio (0.686) has been computed by fitting a Cox proportional hazards regression model to OS data of patients in ADMIRAL receiving gilteritinib maintenance post-HSCT vs. OS data of CR2 patients simulated from Evers [26] The patient population of Evers is comparable to that of ADMIRAL, also consisting of relapsed or refractory AML patients after first line treatment.

The indirect calculation is unanchored because of the absence of a connected network of randomised studies. There is no common comparator arm we can anchor against: Evers compares outcomes and provides survival curves after allogeneic HSCT in AML patients transplanted with either subsequent second complete remission, refractory disease or with persisting cytopenia (after salvage therapy or after induction). The HR estimate is based on a naïve comparison because patient baseline characteristics have not been reported in Evers.

Astellas acknowledges that there is little evidence to assume that the hazard ratio is constant over time and that this is a limitation to this component of the analysis. This approach has been followed because no literature is available to inform the potential benefit

associated with gilteritinib maintenance. A sensitivity analysis that does not consider costs or benefits of post-HSCT gilteritinib maintenance therapy is evaluated in the CS DSA. There is no major impact on the ICER (£49,695) with respect to the base case (£47,695) when no benefits/costs are considered.

For completeness, the Schoenfeld residuals test (Figure 3) indicated that the proportional hazard was violated (P<0.05), mainly driven by the cross-over of the survival plots from ADMIRAL vs. Evers 2018 (Shown in Figure 2). However, because of the small sample size from the ADMIRAL trial, the tail end was not very stable. At the time of the crossover, there is only three patients at risk for gilteritinib comparing to forty at the beginning of the plot. Therefore, the proportional hazard test result should be interpreted with caution and may not be informative of the expected hazard pattern between the two curves evaluated.

Figure 2 Product Limit Survival Estimates from ADMIRAL and Evers

Figure 3 Schoenfeld Residuals Test of ADMIRAL and Evers

B12. CS, Section B.3.3, pages 55 to 80.

(a) Please clarify why, with the exception of the Gompertz distribution, consideration was given in the CS only to survival models that are members of the Generalised F distribution.

(b) Do these models cover the expected underlying hazard of events?

(c) Were more flexible models explored?

For patients without HSCT, individual parametric survival models (Generalised F members and Gompertz) can follow the Kaplan-Meier curves closely (upon visual inspection of the figures in the company submission) and can capture the underlying hazard of events appropriately. Consider the three models with the best relative fit in terms of AIC: these are the log-logistic, the log-normal and the generalised gamma for both OS and EFS endpoints in both gilteritinib and salvage chemotherapy. Log-cumulative hazard plots (log(-log of the survivor function) against log(time)) for OS are produced in Figure 4 and Figure 5 for gilteritinib

and salvage chemotherapy, respectively, to illustrate the cumulative hazard observed in ADMIRAL and predicted by the parametric models. Their EFS counterparts are presented in Figure 6 and Figure 7.

Figure 4 Log-cumulative Hazard Plot for gilteritinib OS

Figure 5 Log Cumulative Hazard Plot for Salvage Chemotherapy OS

Figure 6 Log-cumulative Hazard Plot for gilteritinib EFS

Figure 7 Log-cumulative Hazard Plot for Salvage Chemotherapy EFS

Figure 4 and Figure 5 show that the three Generalised F distributions follow the observed log-cumulative hazard closely and have plausible tails for OS, particularly the log-normal and log-logistic (selected) distributions. For both interventions, the observed hazard generally decreases with time, as the gradient of the log-cumulative hazard appears to decrease. A mild initial increase in such gradient for salvage chemotherapy suggests non-monotonic behaviour

in the observed hazard (although this behaviour occurs in a region with a low density of data points). The shape parameter of the fitted log-logistic models is more than one for both treatments and this suggests that the observed hazard is non-monotonic.

We find that the log-logistic and log-normal models are sufficiently flexible to appropriately capture the hazard, which decreases throughout most of follow-up and is potentially non-monotonic. In addition, the reducing hazards of the OS data induce long tails in the observed survivor function. These confirm the suitability of log-logistic and log-normal models, whose functional forms result in long tails in the survivor function.

Similarly, Figure 6 indicates that log-logistic, log-normal and generalised gamma models follow the observed log-cumulative hazard closely and have plausible tails for gilteritinib EFS. Again, log-logistic and log-normal models are sufficiently flexible to capture a non-monotonic hazard that decreases later in the follow-up, inducing a long tail in the survivor function. It is difficult to evaluate visually (Figure 7) whether the specified parametric models give a particularly good fit to the underlying hazard of events for salvage chemotherapy EFS. However, this is not attributable to the model fitting/selection procedure and is primarily due to a low sample size (41.9% of the 124 patients in salvage chemotherapy experience early treatment failure and are dropped), heavy intermittent censoring and events being clustered at certain time points, early in follow-up, along the Kaplan-Meier curve.

It is worth noting that the data for the aforementioned endpoints are very mature (e.g. 75% to 85% or patients did not receive HSCT and mortality in this cohort was 90-95% by data cut-off). As extrapolation is only applied to a minority of survivors and up to the three-year cure point, it contributes little to the mean overall area under the curves. It is reasonable to assume that if the parametric models fit the observed hazard of events well, they will also extrapolate the expected underlying hazard well. Hence, the best-fitting members of the Generalised F distribution will cover well the hazard in the extrapolation period, as suggested by their internal validity (log-cumulative hazard plots and AIC/BIC values).

For patients with HSCT, individual parametric survival models can follow the OS Kaplan-Meier curve closely (upon visual inspection of the figures in the company submission) and can capture the underlying hazard of events appropriately. Consider the two models with the best relative fit in terms of AIC; the generalised gamma and the log-normal, and also the Gompertz model, which was eventually selected (for reasons highlighted in Question B16). The observed log-cumulative hazard and that predicted by the parametric models is presented in Figure 8. The generalised gamma distribution follows the observed log-cumulative hazard very closely. The observed hazard consistently decreases with time, as the gradient of the log-

cumulative hazard appears to decrease. This induces long tails consistent with the hazard profile of the three models.

Figure 8 Log-cumulative Hazard Plot for HSCT OS

The use of more flexible models, such as spline-based models, was explored but eventually discarded. This was due to a number of reasons. Firstly, for patients without HSCT, the evidence favouring greater complexity is very weak, and more parsimonious models should be preferred (Occam's razor). In the case of EFS salvage chemotherapy, more weakly structured models would actually overfit, following the small-scale fluctuations in the data.

Secondly, while plateaus (gradual decelerations to a steady hazard rate) of long-term survival are observed in some cases, e.g. for OS with HSCT, there are no seemingly important turning points, with significant changes in the observed hazard, throughout follow-up. The complex hazard profile of HSCT patients, with evidence of a delayed and prolonged treatment effect, may warrant additional flexibility. However, it is important to bear in mind is that flexible parametric models extrapolate beyond the trial follow-up using only the final segment of the curve. For HSCT patients, not many events take place during the tail end of the OS K-M curve. Gains in internal validity for this subgroup through the use more flexible models would come at the expense of less accurate extrapolations. These would be overly influenced by the tail end of the K-M curve and add more uncertainty than benefit.

Finally, for patients undergoing HSCT, there were no available patient-level data and we relied on published statistics. More flexible models did not easily lend themselves to the proportional

hazards modelling approach using hazard ratios. In addition, we observed that company submissions in this disease area typically fitted systematically the more standard parametric models (generalised F and Gompertz) and that more flexible spline-based models only appeared in immunotherapy submissions.

B13. CS, Section B.3.3.3, pages 65 to 67. Please confirm that the parameters of the generalised gamma distribution for EFS without HSCT were identifiable in the case of salvage chemotherapy. The BICs suggest that the log logistic distribution is the best fitting model, although there is little to choose between them based on the sample data. Please provide a clinical justification for the predicted survival function beyond the study data.

The parameters of the generalised gamma distribution are indeed identifiable in the case of EFS without HSCT for salvage chemotherapy patients. The generalised gamma has been parametrised under the probability density function:

$$f(t \mid \mu, \sigma, Q) = \frac{|Q|(Q^{-2})^{Q^{-2}}}{\sigma t \Gamma(Q^{-2})} \exp(Q^{-2}(Qw - \exp(Qw))),$$

Where $w = \frac{\ln x - \mu}{\sigma}$. The fitted model has location parameter $\mu = -0.091$, scale parameter $\sigma = 1.254$, and shape parameter $Q = -1.618$.

The choice of the log-logistic distribution to model EFS without HSCT for salvage chemotherapy has been validated by clinical expert opinion, which suggests that a non-monotonic hazard function with respect to time and the choice of an accelerated failure time model is valid. In particular, an initially increasing hazard followed by a decreasing hazard (the hazard having a single mode) is considered clinically plausible. The fitted log-logistic model has a shape parameter of more than one, which indicates a non-monotonic hazard.

Other reasons for selecting the log-logistic, based on the sample data, are that it is clearly the best-fitting model for EFS without HSCT for gilteritinib. For salvage chemotherapy, it is not totally clear which of the models gives the best fit based on internal validity (visual inspection, AIC/BIC and log-cumulative hazard plots). However, the log-logistic is clearly the best-fitting model for gilteritinib based on internal validity (lowest AIC, lowest BIC and best visual fit to the observed K-M curve and log-cumulative hazard) and external/clinical validity. NICE guidance suggests that the same distribution should be used to model each treatment for a given endpoint, as different distributions allow for different shapes and assumptions. This is a clear rationale for selecting the log-logistic, as well as parsimony across endpoints (the log-logistic is also the best-fitting model for OS without HSCT based on internal and external validity).

B14. CS, Section B.3.3, Table 25, page 72. Please provide information regarding the proportion of patients reaching transplant for each individual salvage chemotherapy regimen.

A summary of HSCT rates by treatment in the Base-Case Model is provided in Table 10

Table 10 Summary of HSCT Rates

	Proportion of ADMIRAL patients receiving HSCT (%)	Source	Proportion of individual salvage chemotherapy patients receiving HSCT (%)
Gilteritinib	[redacted]	ADMIRAL trial	N/A
Azacitidine	[redacted]	ADMIRAL trial pooled salvage chemotherapy	[redacted]
FLAG-IDA	[redacted]	ADMIRAL trial pooled salvage chemotherapy	[redacted]
MEC	[redacted]	ADMIRAL trial pooled salvage chemotherapy	[redacted]
LDAC	[redacted]	ADMIRAL trial pooled salvage chemotherapy	[redacted]
BSC	[redacted]	Assumption of no HSCT in BSC arm	N/A

B15. CS, Section B.3.3.5, page 75. The hazard function of a Gompertz distribution is monotonically increasing, monotonically decreasing (in which case some patients never experience the event) or constant (as a special case). However, it is stated that *"the log cumulative hazard plot indicates a non-monotonic hazard pattern over time, which is consistent with the underlying hazard assumptions of the Gompertz model."*

(a) Please clarify how the hazard for the Gompertz model can be non-monotonic unless it is the special case of an exponential.

(b) Please clarify whether it is clinically plausible for the hazards for OS and EFS for patients without HSCT to be decreasing monotonically or increasing then decreasing rather than increasing.

Indeed, the Gompertz model can only have a hazard function that is monotonic or constant. The log-cumulative hazard plot of OS with HSCT (Figure 8) shows the observed hazard consistently decreasing with time (the gradient of the log-cumulative hazard decreases). This is actually compatible with the fitted Gompertz model, which has a shape parameter of less than zero indicating a hazard that decreases monotonically with time.

It is clinically plausible for the OS and EFS hazards of patients without HSCT to be increasing then decreasing. The hazard of the fitted log-logistic models for these endpoints have a single mode, with an initial increasing hazard followed by a decreasing hazard.

B16. CS, Section B.3.3.5, Figure 18, page 77. This plot suggests that a Gompertz distribution provides a poor representation of the data. Please comment on the appropriateness of this model selection and its implications for the cost-effectiveness of gilteritinib.

For OS with HSCT, the Gompertz distribution is the fourth best-fitting model overall in terms of AIC/BIC. However, as Evers does not report any EFS measures, and as a single HR is used to derive the EFS inputs for HSCT patients, the candidate distributions to model OS with HSCT should only be PH-compatible models (exponential, Weibull and Gompertz). The Gompertz distribution has the best internal validity of the three, with the lowest AIC/BIC and the best fits to the K-M curve and the observed log-cumulative hazard upon visual inspection. A proportional hazards modelling approach using hazard ratios is adopted due to our reliance in published statistics for patients with HSCT.

The Gompertz model may not demonstrate a great fit to the observed K-M curve visually. However, while it does not follow the curve closely at some segments (e.g. the beginning of follow-up), that does not necessarily mean that the model is inappropriate. As substantial extrapolation is required for patients with HSCT, the plausibility of the extrapolated portion of the curve is of greater importance than the fit to the observed data (described by the tests in the previous paragraph). We draw from external data and clinical expert opinion to determine that the Gompertz model performs a plausible extrapolation.

Firstly, a monotonically decreasing hazard is clinically plausible as hazard for transplant is known to decrease with time. There are major hazards of early mortality following transplantation, e.g. due to infection and rejection, but the hazard decreases over time as the transplant confers a long-term survival benefit. Secondly, the long tail of the Gompertz distribution is consistent with HSCT's plausible curative potential. Thirdly, external data (the other publications identified in our targeted literature review; Table 26 in the CS) show the hazard profile and long tails observed in Evers. Finally, the fit of the Gompertz model improves at the "plateau" phase of the K-M.

Choosing the Gompertz distribution to model survival for HSCT patients has the implication of making gilteritinib less cost-effective as it reduces the incremental QALYs/LYs gained of gilteritinib vs. weighted comparator with respect to other distributions. The impact of using different distributions on the cost-effectiveness analysis is summarised in

Table **11** (note that the application of single HRs to the log-logistic, log-normal and generalised gamma distributions is methodologically incorrect).

Table 11: Impact of different HSCT survival distributions on cost-effectiveness

Distribution	Incremental Costs	Incremental QALYs	ICER
Gompertz	**£55,510**	**1.164**	**£47,695**
Exponential	£55,424	1.411	£39,274
Weibull	£55,485	1.271	£43,642
Log-logistic	£55,500	1.222	£45,408
Log-normal	£55,494	1.238	£44,809
Generalized gamma	£55,512	1.196	£46,423

B17. PRIORITY. Model. In the ERG report for midostaurin [TA 523], the ERG states (page 81 of the ERG report) that *"stem-cell transplant is associated with a range of complications, the most serious of these is Graft Versus Host Disease (GVHD), a life-threatening adverse event, which affects approximately 40% of SCT recipients"*. Please clarify if the gilteritinib model accounts any of these complications/AEs associated with HSCT.

AEs/complications related to HSCT, were not considered in the model because they were not commonly reported in the literature which evaluated HSCT outcomes among relapsed or refractory AML patients. In addition, some of the GVHD costs could be included in the HSCT procedure costs and inclusion of a separate cost for GVHD would lead to double counting. Based on literature, GVHD tends to occur relatively quickly after the HSCT procedure, with a median time to onset of 19 days[29].

To address Committee's comments, a targeted literature review of R/R AML patients was conducted and identified five publications that reported GVHD information after HSCT procedure.[15,28,29,31,37] The reported incidences for grade III-IV GVHD varied from 11% to 23% (see Table 12). Additionally, a scenario analysis was conducted using the median rate (i.e. 12.5%) from the literature search and event cost input of £55,145 based on prior midostaurin submission to evaluate the potential impact of considering additional cost associated with GVHD. The impact on results was limited (<£1,000 for the weighted comparator).The base-case ICER was £47,695 for the weighted comparator arm. When GVDH cost was considered, the ICER increased to £48,298.

Table 12 Summary of published GVHD rates

Source	Patient Population	Sample Size	Grade III-IV GVHD (%)	All grades GVHD (%)
Steckel et al., 2018[28]	R/R AML patients after 1-2 lines of chemotherapy	Overall: 292 Relapsed after 1st line: 51	14%	56%
Frazer et al., 2017[31]	AML patients in CR2 after relapse	55	NR	50%
Schmid et al., 2006[29]	R/R AML patients after one or multiple lines of chemotherapy	103	15%	63%
Jabbour et al., 2014[15]	Primary refractory AML patients after high-dose cytarabine-based induction therapy	28	11%	NR
Duval et al., 2010[37]	R/R AML patients after myeloablative conditioning regimen	1,652	23%	NR

Health-related quality of life

B18. CS, Section 3.4.2.3, page 84. Please clarify why the book chapter by Janssen *et al* has been used as the source of population norms for EQ-5D. Why was a more granular source not used e.g. the age-specific regression model reported by Ara and Brazier (Value in Health, 2010, vol 13, issue 5)?

In response to the Committee's comments, we explored the impact on the result if the age adjusted disutility was based on Ara and Brazier 2010[38]. The impact on results was limited (<£1,500 for the weighted comparator). The base-case ICER was £47,695 for patients treated with the weighted comparator. With the alternative population norms, the ICER increased to £48,825 per QALY gained.

B19. CS, Section B.3.4.2, page 82 and Model, worksheet "utility", cells F18 and H18. The CS states that *"The utility associated with long-term survivorship is assumed to equal that of patients in the "EFS with HSCT" state", based on similar assumptions in prior NICE submissions of midostaurin."*

(a) This text does not seem to be accurate, as the model includes an assumption of perfect health for all AML long-term survivors, and a lower value is assumed for EFS with HSCT prior to the assumed cure point. Please clarify.

(b) Please justify the assumption of perfect health (excluding age-adjustment) for long-term survivors.

In the prior NICE submissions of midostaurin (TA 523)[21], the ERG explored a scenario analysis in which a new cured health state was added after patients become long-term survivors (i.e., Year 3). The overall approach was further validated by the Committee, which claims that "surviving patients with relapsed disease entering a cured health state after 3 years was the most appropriate to overcome the model's restriction on people in the relapsed state and to better reflect clinical practice in England".

Based on the prior NICE comments, the current model used similar assumptions and considered distinct utility and cost inputs for patients who become long-term survivors. Because there is literature that directly inform utility and cost inputs for AML long-term survivors, patients were assumed to have perfect health (utility = 1) and 0 health state costs. Additional age-adjusted disutilities were also considered throughout the time horizon to incorporate potential decline in the health utility associated with aging, as such the age-adjusted dis-utilities are applied to give a net-effect of normalising for utilities for age i.e 0.818 not 1.

Costs

B20. PRIORITY. CS, Section B.3.5.1. Table 30, page 88. How is wastage being dealt with in the model? Given the poor prognosis of the patient population, would patients be prescribed a full 28 days' supply of gilteritinib? Please include wastage in the model.

In the model, wastage for unused vials for daily administration was considered. The estimated number of vials or tablets per administration was rounded up to the whole number and fraction of the vials/tablets was not considered. For gilteritinib, because the intended daily dose was 120mg and the strength for each tablet was 40mg. Three full doses were considered for each day of administration.

For patients receiving gilteritinib, the number of treatment cycles that patients would receive was estimated based on the observed average exposure time in the ADMIRAL trial. The model used the assumption that patients would only pay for gilteritinib for the days where they consumed the drug. The wastage of unused supply from prescribed tablets was not considered.

It is expected that gilteritinib will be prescribed 3 x 28 days treatment, and dispensed as three separate 28 day packs as the patient returns on a 28 day cycle. Although prognosis is a predictor of outcomes, it is not used to determine the exact duration of therapy dispensed.

In response to the Committee's comments, we added a sensitivity analysis where we extend the estimated dose cycle by 7 days to account for potential wastage due to over-prescribed

tablets. The impact on results was limited (<£2,000). The base-case ICER was £47,695 for patients treated with the weighted comparator. With the alternative wastage scenario, the ICER increased to £49,552 per QALY gained.

B21. PRIORITY. CS, Section B.3.5.1. Table 30, page 88. The net per-patient drug acquisition and administration costs are applied in the first model cycle.

(a) Why was this approach taken?

(b) How were censored observations dealt with? E.g. does the [redacted] dosing cycles reported in Table 30 account for the fact that some patients are still receiving gilteritinib?

(c) If possible, please provide a more conventional analysis in which patients discontinue treatment over time and which accounts for censoring.

The net per-patient drug acquisition and administration costs are applied in the first model cycle for the following considerations

- The use of trial observed treatment duration was a common approach used in economic models and could accurately capture duration of treatment use in accordant with the clinical trial observation, and in accordant with the efficacy data. In the ADMIRAL trial, patients pre-selected to high intensity regimens (i.e., FLAG-IDA, MEC) would only receive up to 2 cycles of the treatment depending on the patients' response status and the investigator's discretion. Therefore, all treatment use related to the high intensity regimens would be fully captured in the ADMIRAL trial.
- For gilteritinib and low intensity regimens (i.e., azacitadine, LDAC), patients would continuously receive treatment until a lack of clinical benefit, intolerance, or a protocol-defined discontinuation criterion was met. However, at the data cut-off date of September 17, 2018, majority of treatment use was already captured in the ADMIRAL trial data. In the trial, all patients receiving low intensity regimens discontinued the treatment and only 7% of patients receiving gilteritinib remained on the initial treatment.
- To further address Committee's comments, a scenario analysis was explored where the observed median treatment cycles for gilteritinib (i.e., 4.5 cycles) was extrapolated exponentially through the entire modelled horizon. Using this approach, a mean treatment cycle of 7.00 was estimated and applied in the model. To further address Committee's comments, a scenario analysis was explored where the observed median treatment cycles for gilteritinib (i.e., 4.5 cycles) was extrapolated

exponentially through the entire modelled horizon. Using this approach, a mean treatment cycle of 6.49 was estimated and applied in the model. The impact on results was limited (less than £300). The base-case ICER was £47,695 for patients treated with the weighted comparator. With the alternative treatment duration scenario, the ICER increased to £47,981 per QALY gained.

B22. CS, Section B.3.4.2. Table 29, pages 84 to 85, and model worksheet "Unit costs" cell O35 and model worksheet "Safety" P35. The model includes the costs and disutility of disease progression as an AE. Given that the model includes different costs and utilities according to EFS/no-event by virtue of its structure, please comment on whether this is double-counting costs and disutilities?

The "progressive acute myeloid leukaemia" event was one of the AEs that was captured in the CSR table. This category was included to be comprehensive and consistent with the CSR.

In response to the Committee's comments, we added a sensitivity analysis where we removed this AE category from the safety table. The impact on results was limited (<£100 for the weighted comparator). The base-case ICER was £47,695 for patients treated with the weighted comparator. With the removal of this AE category, the ICERs decreased to £47,627 per QALY gained.

B23. CS, Section B.3.5.1. Table 30, page 88 and model, worksheet "Drug cost". Please clarify the apparent discrepancy between the gilteritinib drug and administration costs reported in CS Table 30 and those estimated in model (worksheet "Drug cost" cells G24 and K24).

The drug costs per 28-day dosing cycle and the total drug and admin costs for the salvage chemotherapy options in CS Table 30 account for relative dose intensities. As the cells in the "Drug cost" worksheet (columns G and K) account for these too, there are no discrepancies between the drug and administration costs corresponding to salvage chemotherapy. For gilteritinib, on the other hand, while the relative dose intensity is accounted for in the drug costs per 28-day cycle in the worksheet, it is not accounted for in CS Table 30 (there is also a rounding error in the total costs). The drug cost per dosing cycle for gilteritinib is [redacted] x [redacted]% = [redacted] (drug cost in CS Table 30 times relative dose intensity). The total drug and admin costs for gilteritinib are [redacted] (drug and admin costs per dosing cycle times number of dosing cycles). The corrected drug and drug administration costs are presented in Table 13.

Table 13 Drug and Drug Administration Costs

Treatment	Per 28-day dosing cycle		Dosing cycles (N)	Total drug and admin costs (2018 GBP)
	Drug cost (2018 GBP)	Admin cost (2018 GBP)		
Gilteritinib	[redacted]	[redacted]	[redacted]	[redacted]
Azacitidine	£4,537.50	£1,573.79	2.24	£13,698.02
FLAG-IDA	£1,849.50	£1,418.51	1.02	£3,335.71
MEC	£456.18	£1,185.28	1.13	£1,848.99
LDAC	£77.32	£2,327.64	1.68	£4,048.06

B24. CS, Section B.1.3, page 12. The text states *"Activating mutations in FLT3, are one of the most common class of recurring mutations in patients with AML and occur in around 30% of patients with AML."* Given this estimate, please explain why only 200% of the FLT3 mutation costs are applied in the economic model.

As the target population of the model was patients with relapsed or refractory FLT3 mutation positive AML, it was assumed that most patients already had FLT3 mutation tested at the time of the initial diagnosis. Therefore, costs of initial FLT3 mutation testing were not considered in the base-case analysis. The FLT3 mutation status, however, could change throughout patients' disease course. Additionally, FLT3 testing is also considered routine monitoring procedure to understand the prognosis of the patient. Existing evidence suggested that patients might lose or regain FLT3 mutation during disease progression. [39] For patients treated with FLT3 inhibitors, it is required that patients have a confirmed FLT3 mutation. Thus, in the base-case analysis, two additional tests of FLT3 mutation were considered at the relapsed or refractory stage for gilteritinib based on clinical inputs.

Changing the proportion of patients receiving testing from 200% to 333% has minimal impact on the results (<£200). The ICERs changed to £47,872 from £47,695 for patients treated with the weighted comparator.

End of Life criteria

B25. PRIORITY. Model, worksheet 'Base Case'. Given that mean undiscounted survival in the comparator group is greater than 2 years, please clarify why you believe that NICE's End of Life criteria are met for gilteritinib.

NICE end of life criteria are:

1) The treatment is indicated for patients with a short life expectancy, normally less than 24 months and;

2) There is sufficient evidence to indicate that the treatment offers an extension to life, normally of at least an additional 3 months, compared to current NHS treatment, and;

3) The treatment is licensed or otherwise indicated, for small patient populations.

In interpreting the first criterion, to the clarification requested, the median survival is reported to be two months or less in patients diagnosed with AML receiving supportive care alone. The pivotal ADMIRAL phase 3 trial, of FLT3 mutation positive relapsed or refractory AML, showed the median overall survival in the comparator salvage chemotherapy arm was 5.6 months, when censored for HSCT this reduced to 5.3 months.

On looking at the proportion of patients predicted to be alive at indicative time points, using the model traces, it can be seen that the majority of patients have a short life expectancy.

Table 14 Proportion of Patients Alive (With and Without HSCT)

	Gilteritinib		Comparator	
	HSCT	No HSCT	HSCT	No HSCT
% alive at Year 1	[redacted]	[redacted]	[redacted]	[redacted]
% alive at Year 2	[redacted]	[redacted]	[redacted]	[redacted]
% alive at Year 5	[redacted]	[redacted]	[redacted]	[redacted]

This, with a median OS considerably below 2 years, and the majority of patients having died at this point (even at the most optimistic estimate of HSCT patients) supports meeting the end of life criteria.

Even if mean life years gained is used to measure end of life eligibility, Astellas believes that a proportion of patients will be treated with BSC, therefore the most plausible mean survival using current practice is likely to be less than the two years.

Astellas believes that life expectancy is “normally less than 24 months”.

Uncertainty

B26. Model. Please clarify whether the probability of undergoing HSCT in each group is characterised in the model as an uncertain parameter and whether this is included in the PSA. If not, please amend the model.

The probability of undergoing HSCT in each group is included as an uncertain parameter in the PSA. HSCT rates are drawn from Beta distributions (see Model, worksheet “PSA Setup”, rows 73 to 79). For each treatment, the Beta distribution models the probability of HSCT, π, as $p(\pi) = \text{Beta}(\alpha, \beta)$, where α and β are shape parameters which have been determined through the method of moments:

$$\alpha = \bar{x}\left(\frac{\bar{x}(1-\bar{x})}{s^2} - 1\right),$$

$$\beta = (1-\bar{x})\left(\frac{\bar{x}(1-\bar{x})}{s^2} - 1\right).$$

Above, $\bar{x}$ is the proportion of patients undergoing HSCT in the ADMIRAL trial (0.255 for gilteritinib and 0.153 for salvage chemotherapy) and s is the standard error obtained from ADMIRAL data (0.028 for gilteritinib and 0.032 for salvage chemotherapy). For gilteritinib, $p(\pi) = \text{Beta}(\alpha = 62.7,\ \beta = 183.3)$; for salvage chemotherapy, $p(\pi) = \text{Beta}(\alpha = 18.8,\ \beta = 104.2)$.

B27. PRIORITY. CS, Section B.3.7, Tables 40 and 41, pages 101 and 102. Please explain why there is a marked difference in estimated LYGs and QALYs between the deterministic and probabilistic results for gilteritinib and the salvage chemotherapy options?

The marked difference in estimated LYGs and QALYs between the deterministic and probabilistic results is induced by the inclusion of the rate parameter of the OS with HSCT curve in the PSA. The sampling mechanism generates some Gompertz survival curves with long tails which skew the mean LYGs and QALYs upwards for both gilteritinib and salvage chemotherapy. The differences seem to affect both treatments equally and do not considerably alter the incremental LYGs/QALYs. The probabilistic results for gilteritinib and the salvage chemotherapy options, without the inclusion of the OS with HSCT rate parameter (500 iterations only due to time constraints), are reported in Table 15. There is no longer a marked difference in estimated LYGs and QALYs between the probabilistic results and the deterministic results reported in CS Table 40.

Table 15: Probabilistic Cost-effectiveness Results (OS with HSCT rate parameter not included in PSA)

	Gilteritinib	Azacitidine	FLAG-IDA	MEC	LDAC
Comparator weights					
Total costs					
Total LYs					
Total QALYs					
Incremental costs					
Incremental LYs					
Incremental QALYs					
Incremental cost per LY gained					
Incremental cost per QALY gained		£43,556	£46,188	£47,644	£52,335

B28. CS, Section 3.8.3.2, Figure 28, page 116. Is this plane showing 5,000 samples of incremental costs and QALYs for gilteritinib versus each comparator? If not, what is the plot showing?

The figure in question presents a graph of the cost-effectiveness plane for gilteritinib versus each individual salvage chemotherapy comparator (azacitidine, FLAG-IDA, MEC and LDAC). The dots represent 5,000 simulations of incremental QALYs and incremental costs, (Δ_e, Δ_c), for each comparator. The dashed diagonal line is obtained in correspondence with the willingness-to-pay threshold, set to £50,000.

Model validation

B29. PRIORITY. CS, Section B.3.10, Figure 30, page 125. Figure 30 presents a comparison of the modelled OS curves versus smoothed Kaplan-Meier curves from ADMIRAL. However, the smoothed curves do not reflect the Kaplan-Meier curves for the ADMIRAL ITT population given in CS Figure 4 (page 36). Instead, they appear to reflect the Kaplan-Meier curves for OS censored for HSCT analysis shown in CS Figure 5 (page 37).

(a) Please clarify whether this comparison was intentional.

(b) Please comment on the extent to which the modelled OS curves reflect the ITT Kaplan-Meier curves in ADMIRAL, explaining any deviations.

The comparison in CS Figure 30 is intentional. The figure presents the best-fitting (log-logistic) models for gilteritinib and salvage chemotherapy (weighted comparator) for patients without HSCT. These have been plotted against the observed K-M curves for OS without HSCT (in ADMIRAL, these correspond to “OS censored for HSCT” as shown in CS Figure 5). CS Figure 30 was produced to evaluate the internal validity of the model.

Consider not stratifying by HSCT status and comparing the modelled OS curve overall with the overall OS K-M curves for the ADMIRAL ITT population (in CS Figure 4). As the model structure uses efficacy data from multiple sources and separately estimates efficacy stratified by HSCT, this involves weighting the patients by HSCT rates and transitioning them through the decision-tree.

Figure 9 in this document presents the overall model fit for OS against the overall ITT K-M curves for gilteritinib and salvage chemotherapy.

Generally, the predicted curves fit the observed data well upon visual inspection. The curves provide a very good fit earlier in the trial where there are more observations and where the curves fitted to patients without HSCT provide more “weight”. For instance, the observed OS

percentages after 6 months for gilteritinib and salvage chemotherapy are roughly 65% and 48%, respectively, while the predicted OS percentages are 64.6% and 45.4%.

Notwithstanding, there are deviations from the ITT Kaplan-Meier towards the end of follow-up, with the modelled curves overestimating survival. In the case of gilteritinib, the deviation occurs approximately after 20 months (after 30 months, the predicted OS percentage is 17.5% while the observed percentage is roughly 10%). In the case of salvage chemotherapy, the deviation occurs between 10 and 20 months and after the last observed events (in month 33, the observed K-M curve drops to zero but the predicted curve gives an OS percentage of 8% at month 40). The internal validity of the models fitted to the patients without HSCT is very good. Therefore, the divergences are likely to be due to the overestimation of trial OS by the external data, e.g. due to unobserved confounders. Figure 9 suggests that this bias affects both treatment arms similarly.

Figure 9 Comparison Between Predicted vs. Observed Curves – Overall Survival

Section C: Textual clarification and additional points

C1. CS, Section B.2.6, pages 33 and 34. The text on page 33 states *"The median gilteritinib dose in the trial was 119.07mg."* However, the text on page 34 states that mean dose intensity was 119.07mg. Which of these two statements is correct?

The correct statement is that the median daily gilteritinib dose in the trial was 120mg and the mean dose intensity was 119.07mg, calculated as the cumulative dose divided by duration of exposure.

C2. CS, Section B.3.3, page 55. The text refers to "HSCTF". Is this a typographical error?

Yes, this should read HSCT.

C3. CS, Figures 9, 10, 12, 13, 14 and 16. Please confirm that these figures all relate to the group of patients in ADMIRAL who did not receive HSCT.

Confirmed; these figures all relate to the group of patients in ADMIRAL who did not receive HSCT.

C4. Please provide PDF file for references for ADMIRAL - Gorcea, C.M., Burthem, J. and Tholouli, E., 2018. ASP2215 in the treatment of relapsed/refractory acute myeloid leukemia with FLT3 mutation: background and design of the ADMIRAL trial. Future Oncology, 14(20), pp.1995-2004. Also, please the provide PDF file and full details for the following reference - Perl, A.E., Cortes, J.E., Strickland, S.A., Ritchie, E.K., Neubauer, A., Martinelli, G., Naoe, T., Pigneux, A., Rousselot, P.H., Röllig, C. and Baer, M.R., 2017. An open-label, randomized phase III study of gilteritinib versus salvage chemotherapy in relapsed or refractory FLT3 mutation-positive acute myeloid leukemia.

Gorcea and Perl PDFs have been uploaded with this response. This Perl reference in the Appendix has the Reference ID of Perl 2017A; citation is Perl AE et al. Journal of Clinical Oncology 2017 35:15_suppl, TPS7067-TPS7067.

C5. Please provide the PDF file and full details for the following reference for CHRYSALIS - Perl, A. E., Altman, J. K., Cortes, J. E., Smith, C. C., Litzow, M., Baer, M. R., ... & Jurcic, J. G. (2016). Final results of the Chrysalis trial: a first-in-human phase 1/2 dose-escalation, dose-expansion study of gilteritinib (ASP2215) in patients with relapsed/refractory acute myeloid leukemia (R/R AML).

This Perl reference in the Appendix has the Reference ID of Perl 2016; citation is Perl AE et al. Blood, 128(22), 1069. A PDF has been uploaded with this response.

C6. Please, provide full details for the following reference - Dewing DRT, Tholouli DE, Dennis DM. Guidelines for the management of Acute Myeloid Leukaemia.

Reference 24 in the Company evidence submission is Dewing DRT, Tholouli DE, Dennis DM. Manchester Cancer Haemato-Oncology Pathway Guidelines for the management of Acute Myeloid Leukaemia. 2018. A PDF has been uploaded with this response.

Section D: Textual clarification and additional points

Additional data

D1. PRIORITY. Please provide Kaplan-Meier curves for EFS and OS for each treatment group for the ITT population of ADMIRAL, including information on the number of patients at risk over time.

Kaplan-Meier curves for EFS and OS were provided as Figures 4 and 6 in CS, Section B2.6. Astellas can provide Excel spreadsheets for the survival probabilities by days for EFS and OS for each treatment group for the ITT population if required.

D2. PRIORITY. Please provide Kaplan-Meier curves for post-HSCT OS for each treatment group of patients in ADMIRAL, rescaled such that time of HSCT receipt equals zero, including information on the number of patients at risk over time.

Please see below for the Kaplan-Meier curves for post-HSCT OS for each treatment group.

Figure 10 Product Limit Survival Estimates

Note: ASP-2215 is gilteritinib.

D3. PRIORITY. Please provide the distributions of time to HSCT in the gilteritinib and salvage chemotherapy groups.

Please see the time to HSCT distribution by time intervals below.

Table 16 HSCT Distribution

Time to HSCT (Months)	Gilteritinib	Salvage Chemotherapy
	n (%) N=[redacted]	n (%) N=[redacted]
1	[redacted]	[redacted]
2	[redacted]	[redacted]
3	[redacted]	[redacted]
4	[redacted]	[redacted]
5	[redacted]	[redacted]
6	[redacted]	[redacted]
7	[redacted]	[redacted]
8	[redacted]	[redacted]
9	[redacted]	[redacted]
10	[redacted]	[redacted]

D4. PRIORITY. Please provide a 2x2 crosstab detailing for each randomised group (i) number of patients with HSCT *or* number of patients without HSCT, versus (ii) number of patients dead *or* number alive at the time of the analysis.

Please see below tables for information at point of data cut off.

Table 17 Gilteritinib Survival vs HSCT Status

Transplant	Alive N(%)	Dead N(%)	Total
Yes	[redacted]	[redacted]	[redacted]
No	[redacted]	[redacted]	[redacted]

Table 18 Salvage Chemotherapy Survival vs HSCT Status

Transplant	Alive N(%)	Dead N(%)	Total
Yes	[redacted]	[redacted]	[redacted]
No	[redacted]	[redacted]	[redacted]

Additional analyses

D5. PRIORITY. CS, Section B.3.2, page 55. If cure is assumed, please analyse the data using either cure mixture or cure non-mixture models and provide information on the estimate of the cure fraction. Also, estimate overall survival by incorporating information about survival in the general population into the cure models as appropriate. Please present these analyses:

(a) For the subset of ADMIRAL patients who underwent HSCT in each treatment group

(b) For the ITT population of ADMIRAL.

Astellas is still exploring the provision of this data and will provide an update in due course.

Additional question asked during call on 17th July

D6. Please note the Gompertz function was applied incorrectly leading to inconsistent results between DSA and PSA

An updated model with this corrected will be provided in due course.

D7. Please provide clarification as to how the Hazard Ratio from the Sarkozy study been calculated i.e. what data was used to generate the HR of 2.8

The calculation of 2.86 for OS between gilteritinib and BSC was based on naïve comparison. Specifically, we used the following steps:

- Estimate HR between LDAC and BSC by comparing the ratio of reported median in Sarkozy 2013: 1.75 [calculated as 5.6/3.2]
- Estimate HR between gilteritinib and salvage chemotherapy: 1.637
- Estimate HR between gilteritinib and BSC: 1.75X1.637 = 2.86

No survival curve for BSC treatment was reported in Sarkozy 2013[40], therefore the HR method was used to inform the effectiveness of this comparator with gilteritinib as the reference arm.

[1] Cortes JE et al. Efficacy and safety of single-agent quizartinib (Q), a potent and selective FLT3 inhibitor (FLT3i), in patients (pts) with FLT3-internal tandem duplication (FLT3- ITD)-mutated relapsed/refractory (R/R) acute myeloid leukemia (AML) enrolled in the global, phase 3, randomized controlled QuANTUM-R trial. British Journal of Haematology. 2019; 185 (Supplement 1):7

[2] Cortes JE et al. Quizartinib versus salvage chemotherapy in relapsed or refractory FLT3-ITD acute myeloid leukaemia (QuANTUM-R): a multicentre, randomised, controlled, open-label, phase 3 trial. The Lancet Oncology 2019; 20(7):984-997

[3] Wang J et al. A phase III randomized study of gilteritinib versus salvage chemotherapy in FLT3 mutation-positive subjects with relapsed or refractory acute myeloid leukemia. Annals of Oncology. 2018. 29 (Supplement 9)

[4] Kang D et al. Concentration-QT analysis of quizartinib in patients with relapsed/refractory AML. Journal of Pharmacokinetics and Pharmacodynamics 2018; 45 (Supplement 1): S23-S24

[5] Foran JM et al. FLT3-ITD mutations are prevalent and significantly impact outcome after intensive therapy in elderly adults with acute myeloid leukemia (AML): Analysis of the north american intergroup E2906 phase III trial in patients age >=60 years. Blood 2018;132 (Suppl. 1)

[6] Hills RK et al. Outcomes in relapsed/refractory patients with FLT3-ITD mutated AML are poor when treated with non-targeted therapy with a potential role for stem cell transplantation: Results from the ncri AML trials. Blood 2018; 132 (Suppl. 1)

[7] Marconi G et al. Tyrosine kinase inhibitors (TKI) in relapsed/refractory (RR) patients with FLT3-ITD positive acute myeloid leukemia (AML) confer better survival than chemotherapy, due to a better safety profile. HemaSphere 2018; 2 (Supplement 2): 450-451

[8] LAN GKK. Discrete Sequential boundaries for clinical trials. Biometrika, Volume 70, Issue 3, December 1983, Pages 659–663. https://doi.org/10.1093/biomet/70.3.659

[9] Astellas Data on File. ADMIRAL SAP

[10] Mendoza T, Wang XS, Cleeland CS, et al. The rapid assessment of fatigue severity in cancer patients: use of the Brief Fatigue Inventory. Cancer 85: 1186-1196, 1999

[11] Astellas Data on File. ADMIRAL CSR

[12] Pickard S, Neary M, Cella D. Estimation of minimally important differences in EQ-5D utility and VAS scores in cancer. *Health Qual Life Outcomes*. 2007; 5: 70.

[13] Othus M, Appelbaum FR, Petersdorf SH, et al. Fate of patients with newly diagnosed acute myeloid leukemia who fail primary induction therapy. *Biology of Blood and Marrow Transplantation.* 2015;21(3):559-564

[14] Kurosawa S, Yamaguchi T, Miyawaki S, et al. Prognostic factors and outcomes of adult patients with acute myeloid leukemia after first relapse. *Haematologica.* 2010;95(11):1857-1864

[15] Jabbour E, Daver N, Champlin R, et al. Allogeneic stem cell transplantation as initial salvage for patients with acute myeloid leukemia refractory to high-dose cytarabine-based induction chemotherapy. *American journal of hematology.* 2014;89(4):395-398

[16] Schiller GJ, Tuttle P, Desai P. Allogeneic hematopoietic stem cell transplantation in FLT3-ITD–positive acute myelogenous leukemia: the role for flt3 tyrosine kinase inhibitors post-transplantation. *Biology of Blood and Marrow Transplantation.* 2016;22(6):982-990

[17] Lin P-H, Lin C-C, Yang H-I, et al. Prognostic impact of allogeneic hematopoietic stem cell transplantation for acute myeloid leukemia patients with internal tandem duplication of FLT3. *Leukemia research.* 2013;37(3):287-292

[18] National Institute of Health and Care Excellence (NICE). Technology appraisal guidance [ID512]. Brentuximab vedotin for treating relapsed or refractory systemic anaplastic large cell lymphoma. 2017; https://www.nice.org.uk/guidance/ta478/documents/Committee-papers-2. Accessed April 18, 2019

[19] Hettle R, Corbett M, Hinde S, et al. The assessment and appraisal of regenerative medicines and cell therapy products: an exploration of methods for review, economic evaluation and appraisal. Health Technol Assess. 2017;21(7):1-204

[20] National Institute for Health and Care Excellence (NICE). Technology appraisal guidance [ID1115]. Axicabtagene ciloleucel for treating diffuse large B-cell lymphoma and primary mediastinal B-cell lymphoma after 2 or more systemic therapies. 2018; https://www.nice.org.uk/guidance/ta559/documents/Committee-papers-2. Accessed April 18, 2019

[21] National Institute for Health and Care Excellence (NICE). Technology appraisal guidance [TA523]. Midostaurin for untreated acute myeloid leukaemia. 2018; https://www.nice.org.uk/guidance/ta523/history. Accessed August 8, 2018

[22] Shimoni A, Labopin M, Savani B, et al. Long-term survival and late events after allogeneic stem cell transplantation from HLA-matched siblings for acute myeloid leukemia with myeloablative compared to reduced-intensity conditioning: a report on behalf of the acute leukemia working party of European group for blood and marrow transplantation. *Journal of hematology & oncology.* 2016;9(1):118

[23] Bejanyan N, Weisdorf DJ, Logan BR, et al. Survival of patients with acute myeloid leukemia relapsing after allogeneic hematopoietic cell transplantation: a center for international blood and marrow transplant research study. *Biology of Blood and Marrow Transplantation.* 2015;21(3):454-459

[24] Othus M, Garcia-Manero G, Godwin J, et al. Associations between Complete Remissions (CRs) with 7+ 3 Induction Chemotherapy for Acute Myeloid Leukemia and 2-3 Year Survival (" Potential Cure") over the Past Four Decades: Analysis of SWOG Trial Data. In: Am Soc Hematology; 2017

[25] Takahashi K, Kantarjian H, Pemmaraju N, et al. Salvage therapy using FLT 3 inhibitors may improve long-term outcome of relapsed or refractory AML in patients with FLT 3-ITD. *British journal of haematology.* 2013;161(5):659-666

[26] Evers G, Beelen DW, Braess J, et al. Outcome of Patients with Acute Myeloid Leukemia (AML) Undergoing Allogeneic Hematopoietic Stem Cell Transplantation (HSCT) Beyond First Complete Remission (CR1). In: Am Soc Hematology; 2018

[27] Deol A, Sengsayadeth S, Ahn KW, et al. Does FLT3 mutation impact survival after hematopoietic stem cell transplantation for acute myeloid leukemia? A Center for International Blood and Marrow Transplant Research (CIBMTR) analysis. *Cancer.* 2016;122(19):3005-3014

[28] Steckel NK, Groth C, Mikesch JH, et al. High-dose melphalan-based sequential conditioning chemotherapy followed by allogeneic haematopoietic stem cell transplantation in adult patients with relapsed or refractory acute myeloid leukaemia. *British journal of haematology.* 2018;180(6):840-853

[29] Schmid C, Schleuning M, Schwerdtfeger R, et al. Long-term survival in refractory acute myeloid leukemia after sequential treatment with chemotherapy and reduced-intensity conditioning for allogeneic stem cell transplantation. *Blood.* 2006;108(3):1092-1099

[30] Fong CY, Grigoriadis G, Hocking J, et al. Fludarabine, cytarabine, granulocyte-colony stimulating factor and amsacrine: an effective salvage therapy option for acute myeloid leukemia at first relapse. *Leukemia & lymphoma.* 2013;54(2):336-341

[31] Frazer J, Couban S, Doucette S, Shivakumar S. Characteristics predicting outcomes of allogeneic stem-cell transplantation in relapsed acute myelogenous leukemia. *Current Oncology.* 2017;24(2):e123

[32] Ustun C, Giannotti F, Zhang M-J, et al. Outcomes of UCB transplantation are comparable in FLT3+ AML: results of CIBMTR, EUROCORD and EBMT collaborative analysis. *Leukemia.* 2017;31(6):1408

[33] Guyot P Enhanced secondary analysis of survival data:reconstructing the data from published Kaplan-Meier survival curves *BMC Medical Research Methodology* 2012, 12:9 http://www.biomedcentral.com/1471-2288/12/9

[34] National Institute for Health and Care Excellence (NICE). Technology appraisal guidance [TA554]. Tisagenlecleucel for treating relapsed or refractory B-cell acute lymphoblastic leukaemia in people aged up to 25 years. 2018 https://www.nice.org.uk/guidance/ta554/ Accessed 24 July 2019

[35] National Institute for Health and Care Excellence (NICE). Technology appraisal guidance [TA567]. Tisagenlecleucel for treating relapsed or refractory diffuse large B-cell lymphoma after 2 or more systemic therapies. 2019 https://www.nice.org.uk/guidance/ta567 Accessed 24 July 2019

[36] National Institute for Health and Care Excellence (NICE). Technology appraisal guidance [TA559]. Axicabtagene ciloleucel for treating diffuse large B-cell lymphoma and primary mediastinal large B-cell lymphoma after 2 or more systemic therapies. https://www.nice.org.uk/Guidance/TA559 Accessed 24 July 2019

[37] Duval M, Klein JP, He W, et al. Hematopoietic stem-cell transplantation for acute leukemia in relapse or primary induction failure. *Journal of clinical oncology.* 2010;28(23):3730.

[38] Ara R, Brazier JE. Populating an economic model with health state utility values: moving toward better practice. Value in Health. 2010;13(5):509-518.

[39] Nazha A, Cortes J, Faderl S, et al. Activating internal tandem duplication mutations of the fms-like tyrosine kinase-3 (FLT3-ITD) at complete response and relapse in patients with acute myeloid leukemia. *Haematologica.* 2012;97(8):1242-1245

[40] Sarkozy C, Gardin C, Gachard N, et al. Outcome of older patients with acute myeloid leukemia in first relapse. *Am J Hematol*. 2013;88(9):758-764. doi:10.1002/ajh.23498

NATIONAL INSTITUTE FOR HEALTH AND CARE EXCELLENCE

Single technology appraisal

Gilteritinib for treating relapsed or refractory acute myeloid leukaemia [ID1484]

Additional Clarification Questions

July 2019

File name	Version	Contains confidential information	Date
ID1484 Gilteritinib AML Clarification Questions2	**Redacted**	**Yes**	**6th August 2019**

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.

Additional requests

1. Regarding your response to question D2 in the clarification letter, The ERG has inputted these curves into the model and there are still some issues with the OS predictions – this seems to be caused by the assumption of a fixed timepoint at which HSCT occurs. The ERG has made the following data request:

- Please provide KM curves from randomisation to death for the HSCT patients in each treatment group? i.e. Figure 10 of your clarification response, but where time zero relates to the point of randomisation. Ideally, these should be presented as a Kaplan-Meier output in list form to allow the ERG to plot the curve in Excel. Please also comment on what you expect these curves to look like if further data were to be collected with longer follow-up.

The requested KM curve is provided below in Figure 1. The curve is based on small patient numbers and the salvage chemotherapy arm in particular is subject to low patient numbers and high levels of censoring. Astellas would expect the curves to separate, showing a favourable effect of gilteritinib if the data were more mature, however we acknowledge that the curve provided cannot substantiate this.

Because gilteritinib allowed more patients to have a HSCT, Astellas believes the gilteritinib curve is more robust than the salvage chemotherapy curve. The chemotherapy curve should be considered very immature due to the very low patient numbers and high levels of censoring driven in part by the low number of patients being eligible for HSCT and also the

time point of data cut-off. We believe if the salvage chemotherapy treated patients were followed up for longer, a steeper curve would be seen which would separate from the gilteritinib curve. Astellas acknowledges that the curve provided cannot substantiate this.

Figure 1: KM Curves from Randomisation to Death for HSCT Patients in each Treatment Group

Data to construct the KM curve is given in Table 1 for gilteritinib and Table 2 for salvage chemotherapy.

Table 1: Product-Limit Survival Estimates for gilteritinib Arm

Product-Limit Survival Estimates for gilteritinib arm						
Time since randomisation (months)		Survival	Failure	Survival Standard Error	Number Failed	Number Left
[redacted]						

Product-Limit Survival Estimates for gilteritinib arm						
Time since randomisation (months)		Survival	Failure	Survival Standard Error	Number Failed	Number Left
[redacted]						

Table 2: **Product-Limit Survival Estimates for Salvage Chemotherapy Arm**

Product-Limit Survival Estimates for Salvage Chemotherapy						
Time since randomisation (months)		Survival	Failure	Survival Standard Error	Number Failed	Number Left
[redacted]						

The marked survival times are censored observations.

Clarification of prior response to D2. [*Please provide Kaplan-Meier curves for post-HSCT OS for each treatment group of patients in ADMIRAL, rescaled such that time of HSCT receipt equals zero, including information on the number of patients at risk over time*]

While preparing the above analysis, we reviewed the original clarification question D2 and identified an error in the SAS code where the wrong censoring was used. This means that Figure 10 supporting D2 in our original clarification response was wrong. Please accept our apologies and replace with Figure 2 below, showing the Kaplan-Meier curves for post-HSCT OS for each treatment group.

Figure 2: Kaplan-Meier curves for Post-HSCT OS for each Treatment Group

2. Regarding your response to question B21 in the clarification letter,

- Please provide either: (i) Time to treatment discontinuation curves for the gilteritinib group, with separate curves for the HSCT and no HSCT subgroups (from the point of randomisation, provided in list form to allow the ERG to plot the data in Excel) *or* (ii) Mean number of cycles received in: (a) the No HSCT

subgroup; (b) the HSCT subgroup who did not receive maintenance therapy and (c) the HSCT subgroup who did receive maintenance therapy.

Figure 3 below, shows the time to treatment discontinuation curves for the gilteritinib group, with separate curves for the HSCT and no HSCT subgroups (from the point of randomisation). Tables of supporting data are available in Excel upon request.

Figure 3: Time to Treatment Discontinuation Curves for gilteritinib Treated Patients with and without HSCT

3. Regarding your response to question D1 in the clarification letter:

- The ERG would like to request the following KM curves ([1] overall ITT EFS & OS, [2] No HSCT EFS and OS, and [3] With HSCT EFS and OS). Note that [3] is the same as the first request in this section and the KM outputs should be provided in list form, as above.

In response to [1], Figure 4 shows the overall ITT EFS curves, and Figure 5 shows the overall ITT OS curves.

Figure 4: Overall ITT EFS Curves from ADMIRAL

Figure 5: Overall ITT OS curves from ADMIRAL

In response to [2] Figure 6 shows the EFS curve for patients not receiving HSCT and Figure 7 shows the OS curve for patients not receiving HSCT.

Figure 6: EFS of Patients not receiving HSCT

Figure 7: OS of Patients Not Receiving HSCT

In response to [3] Figure 8 shows the EFS of patients receiving HSCT, and Figure 9 shows the OS of patients receiving HSCT.

Figure 8: EFS of Patients Receiving HSCT

Figure 9: OS of Patients Receiving HSCT

Note: the response to [3] regarding OS is the same as that presented for [1]

Additional clarification

4. Regarding your response to question B20 in the clarification letter:

- In response to question B20, you have stated that "It is expected that gilteritinib will be prescribed 3 x 28 days treatment." Does this mean they will be given 1 months supply (3 packs of 28 x 40mg gilteritinib) or 3 months supply (3 x 28 days treatment)?

Both of the above are incorrect. It is common practice to prescribe three x 28 days of treatment and the pharmacy to dispense one 28-day pack at a time. The patient receives one 28-day pack on Day 0, one 28-day pack if they return at Day 28 and one 28-day pack if they return at Day 56. So they are given 1 month (28 days) supply at a time (1 x 28 days treatment).

Gilteritinib for treating relapsed or refractory acute myeloid leukaemia [ID1484]

Additional questions from company and ERG at clarification by email

From:	
To:	

Question:

Clarification Question A18. CS, Section B.2.7, page 39. Please provide subgroup analyses of OS with the stratification variables included in models irrespective of their statistical significance and with additional baseline characteristics (particularly sex, race, baseline ECOG and region) included in full and reduced models assessed using likelihood ratio tests. Also, unless there is a justification for the risk of death changing when patients become 65 years of age, please include age as a continuous variable, ideally allowing for non-linear relationships using splines.

Astellas queries:

- How different is the purpose of A18 to that of A6, to address subgroup analyses of OS? Is A18 supposed to repeat for the same subgroups listed in Figure 7 (Forest plot for subgroup analysis of OS in ITT population)?

- For the likelihood ratio tests, does the full model include all stratification variables (i.e., response to first-line AML therapy and preselected salvage chemotherapy) + all baseline characteristics (particularly sex, race, baseline ECOG and region)? And is the reduced model the stratification variables only? Or is the reduced model removing one subgroup analysis factor each time?

Response:

There are four problems with subgroup analyses as commonly applied as in this submission: multiplicity; not borrowing information from the overall effect; the use of improper groups by categorising continuous variables; and subgroup effects being explained by an ignored covariate. Clarification questions A6 and A18 are related but not identical.

A fundamental principle of the analysis of a clinical trial is to analyse the data according to the way the treatments were randomised and to account for know prognostic factors irrespective of whether they are statistically significant in the particular analysis. Thus, clarification question A6 is asking for the co-primary outcomes to be analysed according to this principle.

Investigating whether further covariates are prognostic factors or treatment effect modifiers for OS is much more exploratory and involves model building based on measured baseline covariates. Clarification question A18 is asking for a model to be built that does not suffer from any of the four problems with subgroup analyses as described above. The reduced model should include all stratification variables and known prognostic factors irrespective of whether they are statistically significant. The full model should include any

potential prognostic factors and treatment effect modifiers that were measure at baseline, and should not categorise continuous variable or assume that any relationship between a continuous variable and outcome is linear. Consideration should also be given to interactions between variable and with treatment; the latter being an assessment of whether a variable is a treatment effect modifier.

From:	
To:	

Question:

The CS states that [redacted] of patients are still on treatment in ADMIRAL, whilst the clarification response quotes a value of [redacted] Which value is correct?

Response:

The number of people continuing to receive gilteritinib at the final analysis of ADMIRAL was [redacted] as previously stated [Astellas Data on File, ADMIRAL CSR]. Of these, [redacted] had maintained gilteritinib treatment without discontinuation since their enrollment in the study.

From:	
To:	

Question:

The ERG would like to request the KM list data for all figures presented in the additional analysis document. Please can you provide these?

In addition to the below, can NICE check with you where the company is at in relations to clarification question D5? In your clarification response, you have stated "Astellas is still exploring the provision of this data and will provide an update in due course". Could Astellas please respond to this question before 5pm Friday 9 August 2019.

We understand the error made in the post-HSCT OS data in D2 of the clarification response. However, the number of OS events for the gilteritinib group was previously listed in D4 as [redacted] (with [redacted] patients censored). Please explain why the Kaplan-Meier list given in the additional data request suggests [redacted] events. Is this the same dataset?" Please provide an answer before COP tomorrow, 8 August 2019

Response:

I have uploaded a zip file for the data behind the KMs, I did this to the Clarifications section of Documents.

At this point Astellas continues to explore this with our Global statistical team, looking to align on the best way to approach this request

The reason for the discrepancy is the following:

- For the original answer to D4 given 25th July, Astellas used the ADSL dataset, which is the ADMIRAL subject level dataset which contains all key variables. This dataset has a flag for death (DTHFL) and a date of death (DTHDT)
- For the overall survival analysis, Astellas used the ADTTE dataset, which is the ADMIRAL dataset with the time to event endpoints
- There are [redacted] patients ([redacted] who had a transplant and [redacted] who did not have a transplant) who died and they are flagged in the ADSL as such, but their date of death is after the analysis cut-off (see below). This is possible because of the small lag between the date of data-cut off and the actual database lock
- So, for the survival analysis, the [redacted] patients who had an event described in the ADSL were censored at the date of the cut-off

The correct dataset to be used is the ADTTE as this is the dataset with the events at the date of the analysis cut-off. So the [redacted] events is the correct number of events among the patients with a transplant.
Table 17 provided as the response to D4 on 25th July should be to the below. There are no changes for the chemo arm.

Clarification questions, 25th July 2019; D4
Table 17: Gilteritinib Survival vs HSCT Status

Transplant	Alive N(%)	Dead N(%)	Total
Yes	[redacted]	[redacted]	[redacted]
No	[redacted]	[redacted]	[redacted]

Request for updated model (ad hoc and in Clarification Questions). In Astellas' submission of an updated model for ID1484 yesterday, the following changes were made:

Key model updates	Model tabs updated	Notes
Implement PSA for weighted comparator	All results and input tabs	• Placeholder arm was updated to replace the weighted comparator arm throughout to enable implementation of PSA for the weighted comparator • PSA results were largely consistent with the base-case results
Fixed PSA mechanism for Gompertz distribution for HSCT OS	<<PSA CEAC>>, <<PSA Setup>>	• The large discrepancies between PSA and DSA were driven by the Gompertz distribution used to inform OS with HSCT • The variance-covariance matrix was updated to allow appropriate evaluation of the Gompertz distribution in the PSA
Additional scenarios based on NICE clarification questions	<<DSA>>, <<DSA inputs>>	• Six scenarios were added to the DSA table; all the scenario results were presented in the response document for clarification questions
Model correction	<<Summary_BSC>>	• We noticed a small linking issue in the trace for BSC when incorporating the SMR-adjusted survival. It was fixed in the updated model

From:	
To:	

Question:

Q1: The ERG has tried to run the PSA using your submitted model and it doesn’t seem to work for the weighted comparator. They have requested if you could provide a version which generates results for this comparator.

Q2: The PSA does run, but the problem is in generating the results for the weighted comparator. The text at the top of the "PSA Setup" worksheet suggests that this can be done by using the "placeholder" comparator drop down box, but that doesn't seem to work. Please can you ask the company if it is possible to run the comparison for gilteritinib versus the weighted comparator directly in the model.

Response:

A1: We have checked the model and, for us, it does run correctly. It took one hour to run. May I ask, you should see a status bar in dark green with grey writing on it which says how many iterations have completed. Can you see this, or did you get an error message at some point?

A2: I have passed this feedback to the team and I will come back to you in due course.

As an update, the team are looking to fix the issue you have flagged. I have asked them for a timeline and hope to have that today for you.

From:	
To:	

Question:

We wanted to check that “at the time of analysis” means “data cut-off” and if not, would you please clarify what is meant.

D4. PRIORITY. Please provide a 2x2 crosstab detailing for each randomised group (i) number of patients with HSCT *or* number of patients without HSCT, versus (ii) number of patients dead *or* number alive at the time of the analysis

Response:

Yes, we meant at the time of data cut-off.

NICE National Institute for Health and Care Excellence

Patient organisation submission

Gilteritinib for treating relapsed or refractory acute myeloid leukaemia [ID1484]

Thank you for agreeing to give us your organisation's views on this technology and its possible use in the NHS.

You can provide a unique perspective on conditions and their treatment that is not typically available from other sources.

To help you give your views, please use this questionnaire with our guide for patient submissions.

You do not have to answer every question – they are prompts to guide you. The text boxes will expand as you type.

Information on completing this submission

- Please do not embed documents (such as a PDF) in a submission because this may lead to the information being mislaid or make the submission unreadable
- We are committed to meeting the requirements of copyright legislation. If you intend to include **journal articles** in your submission you must have copyright clearance for these articles. We can accept journal articles in NICE Docs.
- Your response should not be longer than 10 pages.

About you	
1.Your name	[redacted]

2. Name of organisation	Leukaemia Care
3. Job title or position	Advocacy Officer
4a. Brief description of the organisation (including who funds it). How many members does it have?	Leukaemia Care is a national blood cancer charity, first registered with the Charity Commission in 1969. We work to ensure that everybody affected by blood cancer has access to the right information, advice and support. Key services fall into 4 categories; • Patient services: such as a freephone helpline, nurse advisors, conferences and information booklets • Advocacy: individual advocacy, health technology appraisals, information and patient surveys • Campaigns: our biggest campaign is Spot Leukaemia, aiming to raise awareness of the signs and symptoms of leukaemia • Services for healthcare professionals, including conferences and online learning platforms. In 2016/17 and 2017/18, over 85% of our funding came from our own fundraising activities and those of our volunteers. This includes a wide range of activities – such as legacies, community events, marathons, recycling campaigns etc. Leukaemia Care also receives funding from a wide range of pharmaceutical companies, which in total represent approximately 15% of our annual income. Any funds received from the pharmaceutical industry are in accordance with the ABPI Code of Practice and the Leukaemia Care Code of Practice, our voluntary commitment that governs how we work with, and accept funding from, the pharmaceutical industry: www.leukaemiacare.org.uk/resources/code-of-practice

4b. Do you have any direct or indirect links with, or funding from, the tobacco industry?	n/a
5. How did you gather information about the experiences of patients and carers to include in your submission?	Information is primarily gathered through Leukaemia Care patient experience survey – 'Living with Leukaemia' (www.leukaemiacare.org.uk/living-with-leukaemia). The latest survey, run in 2017, had 2884 responses (including 443 acute myeloid leukaemia patients). We have also have a patient advisory panel, where we hold focus groups to gather in depth qualitative data about patient experiences. We also use patient stories, as written or spoken by patients themselves, which we publish on our website in written and video format. Additionally, we have gathered information through our helpline, support groups, communication with our membership and one to one discussion with patients. We also work closely with other patient groups and share expertise.
Living with the condition	
6. What is it like to live with the condition? What do carers experience when caring for someone with the condition?	Acute myeloid leukaemia (AML) accounts for around a third of cases of leukaemia in adults. 2662 people were newly diagnosed with AML in England in 2016. Approximately two thirds of patients are diagnosed aged 65 and over; old age is associated with poorer prognosis. Due to the rapidly progressing nature of AML, 54% of patients in our survey said they had experienced symptoms for less than a month before visiting their GP. Common symptoms experienced prior to

	diagnosis include fatigue (69% of respondents), weakness/breathlessness (55%) and bruising or bleeding easily (35%). The NCIN 'Routes to Diagnosis' report shows that 53% of AML patients are diagnosed via emergency presentation, compared to a cancer average of 22%, and emergency diagnosis is correlated with poor prognosis. In addition to the impact of an AML diagnosisgenerally, the FLT3-ITD mutation is associated with poor prognosis. A large proportion of patients with this mutation relapse and the patients have also been shown to relapse quicker (Levis 2015). In our survey, patients who had relapsed were more likely to report a negative emotional or psychological impact. Therefore, patients with FLT3-ITD positive AML, and their families, can face a particularly distressing time. Being told you have cancer can be very upsetting. It can be especially difficult with acute leukaemia as you often get ill suddenly and must start treatment quickly (55% of AML patients surveyed started treatment within a week of diagnosis). AML patients experience a considerable emotional impact as a result of their emergency diagnosis. 47% of AML patients surveyed report being depressed or anxious more often since diagnosis. Symptoms continue after diagnosis. The most common ones reported, post-diagnosis, by patients in our survey include fatigue (72%), weakness or breathlessness (48%), nausea or vomiting (36%) and sleeping problems (33%). These symptoms contribute to the practical impact on patients also identified in our survey, with 50% of patients also experiencing pain as a direct result of their condition. Of those in work or education before their diagnosis, 84% have been impacted (52% reduced their hours, 32% no longer able

	to work or continue education). Consequently, 56% of patients reported a negative financial impact as a result of having AML (through increased costs or reduced income).The financial impact is a particular issue for acute patients, compared with other leukaemia types in our survey, so we plan to present data relating to this at the European Haematology Association Annual meeting in 2019, to raise awareness of the problem among treating clinicians and other stakeholders. The additional impacts of being diagnosed with AML may also affect the wider family. Family and friends will be understandably affected by their close one receiving a diagnosis. If patients are less able to look after themselves, the burden of care often falls upon their family, at a time which is already stressful. This may also increase the financial impact, if those caring are not able to work either. "I've not got a partner to take me in but I've got my two sons. If you were on your own, I don't quite know how you would cope to be honest with you."
Current treatment of the condition in the NHS	
7. What do patients or carers think of current treatments and care available on the NHS?	For relapsed/refractory patients with FLT3-ITD mutations, treatment options are limited to salvage chemotherapy currently (although quizartinib is currently being appraised) The most common used salvage chemotherapy in England and Wales are FLAG-Ida and low dose cytarabine (LDAC), both of which were comparators in the ADMIRAL trial. In the trial, gilteritinib was shown to increase overall survival and the percentage of patients in complete remission, compared to those assigned to standard chemotherapy. The performance of standard chemotherapy was low, with an overall survival time of 5.6 months and only 17% of patients alive at 12 months; this demonstrates the need for improved treatments in this particularly difficult to treat population with FLT3-ITD mutations.

8. Is there an unmet need for patients with this condition?	As previously mentioned, survival is low in this particular group of patients, so there is a clear need for treatments that increase patient survival.
Advantages of the technology	
9. What do patients or carers think are the advantages of the technology?	80% of AML patients surveyed identified improved survival as a priority for new treatments; gilteritinib is shows an overall survival benefits for all patients in the ADMIRAL study, regardless of other mutations they might have (which can affect the way the patient respond to other treatments). Gilteritinib is also an oral therapy. Although patients in our survey preferred an intravenous therapy, oral therapy was the second most preferred option. Patients may have this preference for IV therapy as they associate it with more potent treatments. However, oral therapy has benefits including potentially less pain or fear from injections, so it is a welcome option.
Disadvantages of the technology	
10. What do patients or carers think are the disadvantages of the technology?	Gilteritinib has some side effects as it targets kinases other than the one affected by the FLT3-ITD mutations. The most common side effects were problems with the production of other blood cells, leading to anaemia, febrile neutropenia and thrombocytopenia, but also problems with enzyme function, such as increased alanine aminotransferase and increased aspartate aminotransferase. However, it is improvement upon 1st generation FLT3-ITD inhibitors.

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.	Patients with a high allelic burden of mutations have a smaller overall survival benefit with gilteritinib than those with a low allelic burden. However, there was still an overall survival in the group vs. salvage chemotherapy, so these patients should still have access to gilteritinib.
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:

- Acute myeloid leukaemia (AML) accounts for around a third of cases of leukaemia in adults. Approximately two thirds of patients are diagnosed aged 65 and over; old age is associated with poorer prognosis.
- For relapsed/refractory patients with FLT3-ITD mutations, treatment options are limited to salvage chemotherapy currently.
- In the trial, gilteritinib was shown to increase overall survival and the percentage of patients in complete remission. 80% of AML patients surveyed identified improved survival as a priority for new treatments.
- Gilteritinib is also an oral therapy. Oral therapy has benefits, including potentially less pain or fear from injections
- There is a need for improved treatments in this particularly difficult to treat population.

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

Gilteritinib for treating relapsed or refractory acute myeloid leukaemia [ID1484]

Thank you for agreeing to give us your organisation's views on this technology and its possible use in the NHS.

You can provide a unique perspective on the technology in the context of current clinical practice that is not typically available from the published literature.

To help you give your views, please use this questionnaire. You do not have to answer every question – they are prompts to guide you. The text boxes will expand as you type.

Information on completing this submission

- Please do not embed documents (such as a PDF) in a submission because this may lead to the information being mislaid or make the submission unreadable
- We are committed to meeting the requirements of copyright legislation. If you intend to include **journal articles** in your submission you must have copyright clearance for these articles. We can accept journal articles in NICE Docs.
- Your response should not be longer than 13 pages.

About you	
1. Your name	[redacted]
2. Name of organisation	**Royal College of Pathologists**

3. Job title or position	**Consultant Haematologist**
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).	The Acute Myeloid Leukaemia NCRI Clinical Studies Group ('AML Working Group') is funded by the National Cancer Research Institute. It comprises clinical haematologists with a specialist interest in AML, representing the majority of large AML-treating centres in the UK (including all 3 devolved nations), specialist laboratory scientists and trial designers / statisticians. Over the last 30-40 years the AML Working Group has designed and overseen the MRC/NCRI national AML trials.
5b. Do you have any direct or indirect links with, or funding from, the tobacco industry?	**No**
The aim of treatment for this condition	
6. What is the main aim of treatment? (For example, to stop progression, to improve mobility, to cure the condition,	In the setting of relapsed or refractory FLT3 mutated AML, the subject of this appraisal, the main aim of Gilteritinib therapy is to re-establish disease control, which then allows suitably fit patients to be 'bridged' to undergo potentially-curative allogeneic stem cell transplantation. The overall aim is

or prevent progression or disability.)	therefore curative. In the Admiral study 63/246 patients were successfully bridged to transplant compared to 19/124 in the control arm When the drug is used in isolation in frailer / older patients where there is no prospect of using additional stem cell transplantation, clinical responses tend to be of relatively short duration principally due to the development of drug resistance The median duration of CR/CRh in the Admiral trial was 4.6 months in this group the aim will be prolongation of overall survival and improved quality of life in comparison to what could be achieved with existing treatment options such as LDAC which are generally ineffective in this population.
7. What do you consider a clinically significant treatment response? (For example, a reduction in tumour size by x cm, or a reduction in disease activity by a certain amount.)	For 'transplantable' patients a clinically significant response would be a sufficient reduction in disease activity to allow an allogeneic stem cell transplant to take place. This would generally be defined as achievement of some type of complete remission: either a standard 'complete remission' (CR) in which bone marrow blasts are reduced to below 5% with recovery of blood neutrophil and platelet counts, or 'completion remission with incomplete count recovery' (CRi or CRp) in which marrow blast numbers are reduced but neutropenia / low platelets persist. With Gilteritinib, the majority of significant clinical responses fall into the 'CRi' category, with persistently low white blood cell counts persisting while the patient remains on treatment, but thus allowing bridging to transplant. In the Admiral trial 56% of patients achieved a CRc (CR+CRi+CRp) compared to just 27% in the control arm. For 'non transplantable patients' the reductions in disease activity described above may also be clinically significant as they are associated with extension of overall survival in comparison with patients who fail to respond.
8. In your view, is there an unmet need for patients and	Relapsed / refractory *FLT3* mutated (either *FLT3 ITD* or *TKD*) AML is an extremely difficult-to-treat clinical scenario. These patients have a desperately poor prognosis with median survivals in the order of 12-13 weeks. Existing salvage chemotherapy options such as FLAG-Ida have high levels of associated toxicity are generally unsuccessful in achieving sufficiently stable levels of disease response to allow a successful allogeneic stem cell transplant to take place. Non-intensive approaches such as LDAC have are even less

healthcare professionals in this condition?	effective and are purely palliative. There is currently a big unmet need for effective novel therapeutic options in this area.
What is the expected place of the technology in current practice?	
9. How is the condition currently treated in the NHS?	Patients with relapsed / refractory *FLT3* mutated AML who are considered fit enough for intensive therapy, including allogeneic stem cell transplant generally receive 'salvage chemotherapy' with a combination regimen such as FLAG-Ida (fludarabine, cytarabine, idarubicin, G-CSF) or equivalent. If patients achieve complete remission they are then candidates for an allogeneic transplant as a curative intervention. In individual cases it may additionally be possible for AML-treating clinicians to access a FLT3 inhibitor via a pharma compassionate access scheme or to obtain individual patient funding to add a multi-kinase inhibitor (eg sorafenib). In our experience IFR applications in this setting are invariably unsuccessful Older, frailer patients who are not suitable for intensive chemotherapy will generally be treated supportively (transfusions, hydroxycarbamide, to temporarily control the white blood cell count) or receive palliative chemotherapy with low dose cytarabine (LDAC).
• Are any clinical guidelines used in the treatment of the condition, and if so, which?	There are no relevant UK-based guidelines. The ELN (European Leukaemia Net) Guidelines published in 2017 recommend no specific salvage chemotherapy regimen in relapsed / refractory AML but highlight that the priority is to stabilise the disease, generally with salvage chemotherapy, prior to consolidation with allogeneic stem cell transplant. There are no specific recommendations for relapsed *FLT3-ITD* positive AML within the ELN Guidelines which pre-date the most recent trial results and current regulatory applications regarding Quizartinib in patients with a *FLT3-ITD* and Gilteritinib in *FLT3* mutated patients
• Is the pathway of care well defined? Does it vary or are there	Patients are treated in larger centres with experience of treating AML and managed through the MDT structure; generally these are centres that participate in the NCRI AML trials. In general, treatment pathways are less well-defined in the setting of relapsed AML than in newly-presenting disease. The AML

differences of opinion between professionals across the NHS? (Please state if your experience is from outside England.)	19 trial offers a treatment option for younger patients with relapsed AML who have previously entered the trial. A number of UK sites have participated in phase 3 clinical trials with Gilteritinib and the drug has been made available more widely via a CU programme. There is clear agreement that allogeneic stem cell transplant is the overarching therapeutic goal as it offers the potential of long term cure, but salvage chemotherapy regimens and other measures to achieve disease control prior to transplant will likely vary depending on local preferences, the individual patient's prior chemo exposures and the local availability of suitable clinical trials.
• What impact would the technology have on the current pathway of care?	It is likely that, were Gilteritinib approved for this indication, then a substantial proportion of patients would receive this oral monotherapy as an outpatient instead of conventional inpatient-delivered salvage chemotherapy regimens with the aim of achieving a complete remission. The longer term goal would remain allogeneic stem cell transplantation and the Admiral result demonstrated that a larger proportion of patients would reach transplant – due to a combination of better disease responses to Gilteritinib, and less treatment-associated toxicity.
10. Will the technology be used (or is it already used) in the same way as current care in NHS clinical practice?	Many AML-treating centres are already familiar with Gilteritinib through clinical trials and the company CU programme
• How does healthcare resource use differ between the technology and current care?	As discussed in Q9. The main difference is that current care is chiefly based around admitting patients to an inpatient haematology unit to receive intensive salvage chemotherapy (eg. FLAG-Ida). In that model, patients are frequently admitted for 4-5 weeks, with associated costs and extensive need for transfusional support (red cells, platelets), treatment of neutropenic infections etc. Gilteritinib is an oral therapy that patients will generally take at home with frequent monitoring via visits (usually 2 per week) to a haematology day unit. The overall need for blood product support and rates of

	neutropenic infections are likely to be significantly lower that with FLAG-Ida. This was clearly demonstrated in the Admiral trial where Gilteritinib was associated with less overall toxicity and a lower risk of febrile neutropenia. There was also less requirement for blood products particularly blood transfusions with 35% of patients becoming transfusion independent.
• In what clinical setting should the technology be used? (For example, primary or secondary care, specialist clinics.)	Secondary / tertiary care. Most frequently in an outpatient, day unit setting within a large AML-treating centre.
• What investment is needed to introduce the technology? (For example, for facilities, equipment, or training.)	We think little is required
11. Do you expect the technology to provide clinically meaningful benefits compared with current care?	Yes. Gilteritinib is active in patients with a *FLT3-ITD* and *TKD.* We know from NCRI trials that the outcome for this group of patients is extremely poor compared to other relapsed patients with just 17% of patients achieving a CR with a median survival of just 86 days and 1 year OS of 13%. However in the 53 patients (25% overall) who received a transplant the median survival was 301 days with 42% alive at one year. So transplant is beneficial. In the Admiral trial Gilteritinib significantly improved response rates (56% vs 27%) and improved OS (9.3 vs 5.6 months, HR 0.63)) and importantly bridged more patients to BMT (26% vs 15%). Post-transplant use of Gilteritinib also appeared to prolong survival in transplanted patients.
• Do you expect the technology to increase length of life more than current care?	Yes significant increases in median overall survival in comparison with current care were seen in the Admiral study (9.3 vs 5.6m, p=0.0177). The successful subsequent delivery of allogeneic stem cell transplant remains fundamental to the extension of overall survival.

• Do you expect the technology to increase health-related quality of life more than current care?	Yes. In the short term there is considerably improved QoL through receiving oral therapy as an outpatient rather than a prolonged inpatient stay for salvage chemotherapy. Gilteritinib appeared generally well-tolerated and was associated with fewer side effects than standard salvage chemotherapy which brings with it a combination of expected toxicities including emesis, hair loss, prolonged bone marrow suppression including potential for life-threatening neutropenic infections.
12. Are there any groups of people for whom the technology would be more or less effective (or appropriate) than the general population?	Due to its targeted FLT3-inhibitory activity, benefits are restricted to patients with relapsed AML which is *FLT3* mutation positive.. Thgis includes patients with a *FLT3 ITD* and a *FLT3-TKD*
The use of the technology	
13. Will the technology be easier or more difficult to use for patients or healthcare professionals than current care? Are there any practical implications for its use (for example, any concomitant treatments needed, additional	As discussed above, much easier due to oral administration. Can be given at home, not needing to be given through indwelling central venous access lines

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 any additional testing?	Gilteritinib is administered orally in 28-day cycles. It will usually be relatively quickly evident, based on the results of blood count and bone marrow monitoring, whether a patient is deriving clinical benefit by the end of 2-3 cycles of therapy. This does not involve any additional testing to what the patient would be receiving if they were to be treated with current conventional care (salvage chemotherapy).
15. Do you consider that the use of the technology will result in any substantial health-related benefits that are unlikely to be included in the quality-adjusted life year (QALY) calculation?	No
16. Do you consider the technology to be innovative in its potential to make a	Yes. Gilteritinib is highly innovative as a FLT3-targeted small molecule inhibitor to treat 'FLT3-driven' relapsed / refractory AML. It will potentially remove the need, in many cases, to use non-targeted traditional salvage chemotherapy regimens to bridge the patient to allogeneic stem cell transplant. It competes

significant and substantial impact on health-related benefits and how might it improve the way that current need is met?	however with Quizartinib which is seeking approval in broadly the same patient population. Gilteritinib (unlike Quizartinib) is also applicable to patients with *FLT3* tyrosine kinase domain (TKD) point mutations (as well as patients *FLT3* ITD mutations). We understand that in the Admiral trial there were no significant differences in response rates and survival between patients with ITD and TKD mutations
• Is the technology a 'step-change' in the management of the condition?	Yes. A switch from relatively inefficient non-targeted intensive chemotherapy to an oral targeted treatment approach.
• Does the use of the technology address any particular unmet need of the patient population?	Yes. The patient population don't currently have a meaningful effective therapy – median survival is around 2-3 months. With Gilteritinib approximately 50% gain a meaningful response and in around 30% this is sufficient to enable a potentially-curative stem cell transplant to be delivered which will improve survival
17. How do any side effects or adverse effects of the technology affect the management of the condition and the patient's quality of life?	The drug seems well tolerated with a reduced number of toxicities compared to standard salvage chemotherapy

Sources of evidence	
18. Do the clinical trials on the technology reflect current UK clinical practice?	The control arm of the admiral trial is consistent with the UK experience.
• 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?	As discussed above. The most important clinical outcomes with Gilteritinib are response rate (CR and CRi), prolongation of overall survival and, particularly, the rates of bridging to allogeneic stem cell transplant. These were compared with standard therapy in a randomised phase 3 study. Also patients benefitted by having less toxic therapy that could be given as ambulatory care
• If surrogate outcome measures were used, do they adequately predict long-term clinical outcomes?	Not applicable. The trial met its primary endpoint and showed prolonged overall survival.
• Are there any adverse effects that were not apparent in clinical trials but have come to light subsequently?	Not to our knowledge

19. Are you aware of any relevant evidence that might not be found by a systematic review of the trial evidence?	no
20. How do data on real-world experience compare with the trial data?	The real world UK experience in NCRI trials has been detailed in section 11 and will not be repeated here.
Equality	
21a. Are there any potential equality issues that should be taken into account when considering this treatment?	no
21b. Consider whether these issues are different from issues with current care and why.	
Key messages	

24. In up to 5 bullet points, please summarise the key messages of your submission.

- Relapsed *FLT3-mutated* AML has a desperately poor prognosis with median survival of 2-3 months with conventional therapies
- Gilteritinib is an orally-delivered FLT3-targeted agent that can be administered at home without recourse to long inpatient stays
- In a phase 3 RCT, in comparison with conventional salvage chemotherapy, the drug met the primary endpoint of prolonging overall survival, there were higher rates of initial response and more patients were 'bridged' to potentially-curative allogeneic stem cell transplant
- Gilteritinib is well-tolerated in comparison to conventional salvage chemotherapy
- UK experience confirms the benefit of transplant in relapsed *FLT3-ITD* AML however the poor response to existing salvage chemotherapy means that currently few patients benefit,

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.

……

NICE National Institute for Health and Care Excellence

Clinical expert statement

Gilteritinib for treating relapsed or refractory acute myeloid leukaemia [ID1484]

Thank you for agreeing to give us your views on this technology and its possible use in the NHS.

You can provide a unique perspective on the technology in the context of current clinical practice that is not typically available from the published literature.

To help you give your views, please use this questionnaire. You do not have to answer every question – they are prompts to guide you. The text boxes will expand as you type.

Information on completing this expert statement

- Please do not embed documents (such as a PDF) in a submission because this may lead to the information being mislaid or make the submission unreadable
- We are committed to meeting the requirements of copyright legislation. If you intend to include **journal articles** in your submission you must have copyright clearance for these articles. We can accept journal articles in NICE Docs.
- Your response should not be longer than 13 pages.

About you	
1. Your name	**Mike Dennis**
2. Name of organisation	**RCPath, RCPhysicians, BSH**

3. Job title or position	**Consultant Haematologist**
4. Are you (please tick all that apply):	☐ an employee or representative of a healthcare professional organisation that represents clinicians? ☒ a specialist in the treatment of people with this condition? ☐ a specialist in the clinical evidence base for this condition or technology? ☐ other (please specify):
5. Do you wish to agree with your nominating organisation's submission? (We would encourage you to complete this form even if you agree with your nominating organisation's submission)	☒ yes, I agree with it ☐ no, I disagree with it ☐ I agree with some of it, but disagree with some of it ☐ other (they didn't submit one, I don't know if they submitted one etc.)
6. If you wrote the organisation submission and/ or do not have anything to add, tick here. (If you tick this box, the rest of this form will be deleted after submission.)	☒ yes

Thank you for your time.

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

.....

Your privacy

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

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

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

.....

NICE National Institute for Health and Care Excellence

Clinical expert statement

Gilteritinib for treating relapsed or refractory acute myeloid leukaemia [ID1484]

Thank you for agreeing to give us your views on this technology and its possible use in the NHS.

You can provide a unique perspective on the technology in the context of current clinical practice that is not typically available from the published literature.

To help you give your views, please use this questionnaire. You do not have to answer every question – they are prompts to guide you. The text boxes will expand as you type.

Information on completing this expert statement

- Please do not embed documents (such as a PDF) in a submission because this may lead to the information being mislaid or make the submission unreadable
- We are committed to meeting the requirements of copyright legislation. If you intend to include **journal articles** in your submission you must have copyright clearance for these articles. We can accept journal articles in NICE Docs.
- Your response should not be longer than 13 pages.

About you	
1. Your name	[redacted]
2. Name of organisation	The Acute Myeloid Leukaemia NCRI Clinical Studies Group ('AML Working Group') is funded by the National Cancer Research Institute. It comprises clinical haematologists with a specialist interest in AML, representing the majority of large AML-treating centres in the UK (including all 3 devolved nations),

	specialist laboratory scientists and trial designers / statisticians. Over the last 30-40 years the AML Working Group has designed and overseen the MRC/NCRI national AML trials.
3. Job title or position	Professor of Haematology, Nottingham University Hospital
4. Are you (please tick all that apply):	☐ an employee or representative of a healthcare professional organisation that represents clinicians? **x**☐ a specialist in the treatment of people with this condition? ☐ a specialist in the clinical evidence base for this condition or technology? ☐ other (please specify):
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)	**x**☐ yes, I agree with it ☐ no, I disagree with it ☐ I agree with some of it, but disagree with some of it ☐ other (they didn't submit one, I don't know if they submitted one etc.)
6. If you wrote the organisation submission and/ or do not have anything to add, tick here. (If you tick this box, the	☐ yes

rest of this form will be deleted after submission.)	
The aim of treatment for this condition	
7. What is the main aim of treatment? (For example, to stop progression, to improve mobility, to cure the condition, or prevent progression or disability.)	In the setting of relapsed or refractory FLT3 mutated AML, the subject of this appraisal, the main aim of Gilteritinib therapy is to re-establish disease control, which then allows suitably fit patients to be 'bridged' to undergo potentially-curative allogeneic stem cell transplantation. The overall aim is therefore curative. In the Admiral study 63/246 patients were successfully bridged to transplant compared to 19/124 in the control arm When the drug is used in isolation in frailer / older patients where there is no prospect of using additional stem cell transplantation, clinical responses tend to be of relatively short duration principally due to the development of drug resistance The median duration of CR/CRh in the Admiral trial was 4.6 months in this group the aim will be prolongation of overall survival and improved quality of life in comparison to what could be achieved with existing treatment options such as LDAC which are generally ineffective in this population.
8. What do you consider a clinically significant treatment response? (For example, a reduction in tumour size by x cm, or a reduction in disease activity by a certain amount.)	For 'transplantable' patients a clinically significant response would be a sufficient reduction in disease activity to allow an allogeneic stem cell transplant to take place. This would generally be defined as achievement of some type of complete remission: either a standard 'complete remission' (CR) in which bone marrow blasts are reduced to below 5% with recovery of blood neutrophil and platelet counts, or 'completion remission with incomplete count recovery' (CRi or CRp) in which marrow blast numbers are reduced but neutropenia / low platelets persist. With Gilteritinib, the majority of significant clinical responses fall into the 'CRi' category, with persistently low white blood cell counts persisting while the patient remains on treatment, but thus allowing bridging to transplant. In the Admiral trial 56% of patients achieved a CRc (CR+CRi+CRp) compared to just 27% in the control arm.

	For 'non transplantable patients' the reductions in disease activity described above may also be clinically significant as they are associated with extension of overall survival in comparison with patients who fail to respond.
9. In your view, is there an unmet need for patients and healthcare professionals in this condition?	Relapsed / refractory *FLT3* mutated (either *FLT3 ITD* or *TKD*) AML is an extremely difficult-to-treat clinical scenario. These patients have a desperately poor prognosis with median survivals in the order of 12-13 weeks. Existing salvage chemotherapy options such as FLAG-Ida have high levels of associated toxicity are generally unsuccessful in achieving sufficiently stable levels of disease response to allow a successful allogeneic stem cell transplant to take place. Non-intensive approaches such as LDAC have are even less effective and are purely palliative. There is currently a big unmet need for effective novel therapeutic options in this area.
What is the expected place of the technology in current practice?	
10. How is the condition currently treated in the NHS?	Patients with relapsed / refractory *FLT3* mutated AML who are considered fit enough for intensive therapy, including allogeneic stem cell transplant generally receive 'salvage chemotherapy' with a combination regimen such as FLAG-Ida (fludarabine, cytarabine, idarubicin, G-CSF) or equivalent. If patients achieve complete remission they are then candidates for an allogeneic transplant as a curative intervention. In individual cases it may additionally be possible for AML-treating clinicians to access a FLT3 inhibitor via a pharma compassionate access scheme or to obtain individual patient funding to add a multi-kinase inhibitor (eg sorafenib). In our experience IFR applications in this setting are invariably unsuccessful Older, frailer patients who are not suitable for intensive chemotherapy will generally be treated supportively (transfusions, hydroxycarbamide, to temporarily control the white blood cell count) or receive palliative chemotherapy with low dose cytarabine (LDAC).

• Are any clinical guidelines used in the treatment of the condition, and if so, which?	There are no relevant UK-based guidelines. The ELN (European Leukaemia Net) Guidelines published in 2017 recommend no specific salvage chemotherapy regimen in relapsed / refractory AML but highlight that the priority is to stabilise the disease, generally with salvage chemotherapy, prior to consolidation with allogeneic stem cell transplant. There are no specific recommendations for relapsed *FLT3-ITD* positive AML within the ELN Guidelines which pre-date the most recent trial results and current regulatory applications regarding Quizartinib in patients with a *FLT3-ITD* and Gilteritinib in *FLT3* mutated patients
• 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.)	Patients are treated in larger centres with experience of treating AML and managed through the MDT structure; generally these are centres that participate in the NCRI AML trials. In general, treatment pathways are less well-defined in the setting of relapsed AML than in newly-presenting disease. The AML 19 trial offers a treatment option for younger patients with relapsed AML who have previously entered the trial. A number of UK sites have participated in phase 3 clinical trials with Gilteritinib and the drug has been made available more widely via a CU programme. There are national guidelines from the BSBMT that allogeneic stem cell transplant in second remission is the overarching therapeutic goal as it offers the potential of long term cure, but salvage chemotherapy regimens and other measures to achieve disease control prior to transplant will likely vary depending on local preferences, the individual patient's prior chemo exposures and the local availability of suitable clinical trials.
• What impact would the technology have on the current pathway of care?	It is likely that, were Gilteritinib approved for this indication, then a substantial proportion of patients would receive this oral monotherapy as an outpatient instead of conventional inpatient-delivered salvage chemotherapy regimens with the aim of achieving a complete remission. The longer term goal would remain allogeneic stem cell transplantation and the Admiral result demonstrated that a larger proportion of patients would reach transplant in CR2 – due to a combination of better disease responses to Gilteritinib, and less treatment-associated toxicity.
11. Will the technology be used (or is it already used) in	Many AML-treating centres are already familiar with Gilteritinib through clinical trials and the company CU programme. Unlike current treatment strategies the drug will be used primarily in the out-patient and day case setting

the same way as current care in NHS clinical practice?	
• How does healthcare resource use differ between the technology and current care?	As discussed above. The main difference is that current care is chiefly based around admitting patients to an inpatient haematology unit to receive intensive salvage chemotherapy (eg. FLAG-Ida). In that model, patients are frequently admitted for 4-5 weeks, with associated costs and extensive need for transfusional support (red cells, platelets), treatment of neutropenic infections etc. Gilteritinib is an oral therapy that patients will generally take at home with frequent monitoring via visits (usually 2 per week) to a haematology day unit. The overall need for blood product support and rates of neutropenic infections are likely to be significantly lower that with FLAG-Ida. This was clearly demonstrated in the Admiral trial where Gilteritinib was associated with less overall toxicity and a lower risk of febrile neutropenia. There was also less requirement for blood products particularly blood transfusions with 35% of patients becoming transfusion independent
• In what clinical setting should the technology be used? (For example, primary or secondary care, specialist clinics.)	Secondary / tertiary care. Most frequently in an outpatient, day unit setting within a large AML-treating centre
• What investment is needed to introduce the technology? (For example, for facilities, equipment, or training.)	Very little is required
12. Do you expect the technology to provide clinically	Yes. Gilteritinib is active in patients with a *FLT3-ITD* and *TKD.* We know from NCRI trials that the outcome for this group of patients is extremely poor compared to other relapsed patients with just 17% of patients achieving a CR with a median survival of just 86 days and 1 year OS of 13%. However in the 53 patients

meaningful benefits compared with current care?	(25% overall) who received a transplant in CR2 the median survival was 301 days with 42% alive at one year. So transplant is beneficial. In the Admiral trial Gilteritinib significantly improved response rates (56% vs 27%) and improved OS (9.6 vs 5.6 months, HR 0.63)) and importantly bridged more patients to BMT (26% vs 15%). A landmark study also showed that post-transplant use of Gilteritinib also appeared to prolong survival in transplanted patients.
• Do you expect the technology to increase length of life more than current care?	Yes significant increases in median overall survival in comparison with current care were seen in the Admiral study (9.3 vs 5.6m, p=0.0177). The successful delivery of allogeneic stem cell transplant remains fundamental to the extension of overall survival although there is evidence that no-transplanted patients had an improved OS.
• Do you expect the technology to increase health-related quality of life more than current care?	Yes, rather than a prolonged inpatient stay for salvage chemotherapy. Gilteritinib appeared generally well-tolerated and was associated with fewer side effects than standard salvage chemotherapy which brings with it a combination of expected toxicities including emesis, hair loss, prolonged bone marrow suppression including potential for life-threatening neutropenic infections. Yes. In the short term there is considerably improved QoL through receiving oral therapy as an outpatient
13. Are there any groups of people for whom the technology would be more or less effective (or appropriate) than the general population?	Due to its targeted FLT3-inhibitory activity, benefits are restricted to patients with relapsed AML which is *FLT3* mutation positive.. This includes patients with a *FLT3 ITD* and a *FLT3-TKD*
The use of the technology	

14. Will the technology be easier or more difficult to use for patients or healthcare professionals than current care? Are there any practical implications for its use (for example, any concomitant treatments needed, additional clinical requirements, factors affecting patient acceptability or ease of use or additional tests or monitoring needed.)	As discussed above, much easier due to oral administration. Can be given at home, not needing to be given through indwelling central venous access lines
15. Will any rules (informal or formal) be used to start or stop treatment with the technology? Do these include any additional testing?	Gilteritinib is administered orally in 28-day cycles. It will usually be relatively quickly evident, based on the results of blood count and bone marrow monitoring, whether a patient is deriving clinical benefit by the end of 2-3 cycles of therapy. This does not involve any additional testing to what the patient would be receiving if they were to be treated with current conventional care (salvage chemotherapy).
16. Do you consider that the use of the technology will result in any substantial health-	No

related benefits that are unlikely to be included in the quality-adjusted life year (QALY) calculation?	
17. Do you consider the technology to be innovative in its potential to make a significant and substantial impact on health-related benefits and how might it improve the way that current need is met?	Yes. Gilteritinib is highly innovative as a FLT3-targeted small molecule inhibitor to treat 'FLT3-driven' relapsed / refractory AML. It will potentially remove the need to use non-targeted traditional salvage chemotherapy regimens to bridge the patient to allogeneic stem cell transplant. It competes however with Quizartinib which is seeking approval in broadly the same patient population. Gilteritinib (unlike Quizartinib) is also applicable to patients with *FLT3* tyrosine kinase domain (TKD) point mutations (as well as patients *FLT3* ITD mutations). We understand that in the Admiral trial there were no significant differences in response rates and survival between patients with ITD and TKD mutations and whether the patient had received therapy with Midostaurin or not
• Is the technology a 'step-change' in the management of the condition?	Yes. A switch from relatively inefficient non-targeted intensive chemotherapy to an oral targeted treatment approach.
• Does the use of the technology address any particular unmet need of the patient population?	Yes. The patient population don't currently have a meaningful effective therapy – median survival is around 2-3 months. With Gilteritinib approximately 50% gain a meaningful response and in around 30% this is sufficient to enable a potentially-curative stem cell transplant to be delivered which will improve survival

18. How do any side effects or adverse effects of the technology affect the management of the condition and the patient's quality of life?	The drug seems well tolerated with a reduced number of toxicities compared to standard salvage chemotherapy
Sources of evidence	
19. Do the clinical trials on the technology reflect current UK clinical practice?	The control arm of the Admiral trial is consistent with the UK experience
• 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?	As discussed above. The most important clinical outcomes with Gilteritinib are response rate (CR and CRi), prolongation of overall survival and, particularly the larger number of patients proceeding to allogeneic stem cell transplant in CR2. These were compared with standard therapy in a randomised phase 3 Admiral study. Also patients benefitted by having less toxic therapy that could be given as ambulatory care. The response rates achieved in the Phase 3 trial were consistent with those reported in earlier Phase 2 studies in the same patient population

• If surrogate outcome measures were used, do they adequately predict long-term clinical outcomes?	Not applicable. The trial met its primary endpoint and showed prolonged overall survival.
• Are there any adverse effects that were not apparent in clinical trials but have come to light subsequently?	Not to our knowledge
20. Are you aware of any relevant evidence that might not be found by a systematic review of the trial evidence?	no
21. How do data on real-world experience compare with the trial data?	The real world UK experience in NCRI trials has been detailed above. We know from NCRI trials that the outcome for relapsed FLT3 mutated patients is extremely poor compared to other types of relapsed AML with just 17% of patients achieving a CR with a median survival of just 86 days and 1 year OS of 13% (Hills et al ASH 2017). This compares to an expected CR2 rate of >50% in patients without a FLT3 mutation.
Equality	

22a. Are there any potential equality issues that should be taken into account when considering this treatment?	No
22b. Consider whether these issues are different from issues with current care and why.	
Topic-specific questions	
23a. In the ADMIRAL trial, gilteritinib was compared with salvage chemotherapy (which consisted of low dose cytarabine, azacitidine, MEC or FLAG-Ida). Are these treatments reflective of NHS standard care for Adults with relapsed or refractory FLT3-	Broadly they are.

mutation positive acute myeloid leukaemia? 23b. Could you estimate the frequency of use (in %) of each of the comparator treatments referred to in 23a in NHS clinical practice for this population? 23c. Could you also give an estimate of the proportion of patients who would receive best supportive care in NHS clinical practice? 24. Are there other treatments that are currently part of standard NHS clinical practice that are not mentioned above? If so, could you state them and estimate the proportions of	To obtain real world data a total of 264 adult patients in the NCRI AML 15,16&17 trials were identified as having a FLT3 ITD mutation and had been treated with intensive chemotherapy, and were either refractory to two courses of induction therapy, or relapsed within six months of transplant, or did not receive a prior transplant and had a remission duration of 6 months or less. Of these 264 patients, 65% were treated intensively, 8% non-intensively, and 20% with palliation with some receiving experimental therapies (See attached abstract, (Hills at al ASH, 2017). The NCRI study however excluded later relapses (>6 months) as it was based upon the QuantumR entry criteria. As late relapses are more likely to receive intensive salvage treatment the we may underestimate the intensively treated population 20% received palliation/ BSC A few may access compassionate use FLT3 inhibitors

their use (in %) in NHS clinical practice? 25. Is survival post HSCT (haematopoietic stem cell transplantation) influenced by prior treatment before HSCT in this population? If so, please explain. Please also describe the level of uncertainty around any estimates. 26. Are you able to provide an estimate of how many patients (in %) would receive a HSCT with current NHS standard clinical care? Is this influenced by prior treatments? Please also describe the level of uncertainty around any estimates.	In the UK/NCRI experience 56 patients who were transplanted in CR2 following salvage chemotherapy the median survival was 301 days with 42% alive at one year .So survival was improved indicating that a treatment which can deliver patients to transplant can improve patient outcomes. In my opinion there is no reason to believe that the outcome of patients in CR/CRi prior to BMT would be different if they had received Gilteritinib or salvage chemotherapy. What is different is that Gilteritinib can deliver more patients to BMT. In the NCRI study of the 215 patients who had not had a prior BMT in CR1, 53 (25%) received a transplant .These patients would have received intensive salvage chemotherapy prior to BMT. Although some patients would not have been suitable/eligible for BMT the limiting factor with current approaches limiting the use of BMT is the poor response to salvage chemotherapy.

27. Are you able to provide estimates for survival post-HSCT in this population at 5,10 and 20 years? Please also describe the level of uncertainty around any estimates.	In this study 42% were alive at 1 year post BMT. By this time the major relapse risk would have passed and a large majority would be alive at 5 years

Key messages

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

- Relapsed *FLT3-mutated* AML has a desperately poor prognosis with median survival of 2-3 months with conventional therapies
- Gilteritinib is an orally-delivered FLT3-targeted agent that can be administered at home without recourse to long inpatient stays
- In a phase 3 RCT, in comparison with conventional salvage chemotherapy, the drug met the primary endpoint of prolonging overall survival, there were higher rates of initial response and more patients were 'bridged' to potentially-curative allogeneic stem cell transplant
- Gilteritinib is well-tolerated in comparison to conventional salvage chemotherapy
- UK experience confirms the benefit of transplant in relapsed *FLT3-ITD* AML however the poor response to existing salvage chemotherapy means that currently few patients benefit,

Thank you for your time.

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

.....

Your privacy

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

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

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

.....

NICE National Institute for Health and Care Excellence

Patient expert statement

Gilteritinib for treating relapsed or refractory acute myeloid leukaemia [ID1484]

Thank you for agreeing to give us your views on this technology and its possible use in the NHS.

You can provide a unique perspective on conditions and their treatment that is not typically available from other sources.

To help you give your views, please use this questionnaire with our guide for patient submissions.

You do not have to answer every question – they are prompts to guide you. The text boxes will expand as you type.

Information on completing this expert statement

- Please do not embed documents (such as a PDF) in a submission because this may lead to the information being mislaid or make the submission unreadable
- We are committed to meeting the requirements of copyright legislation. If you intend to include **journal articles** in your submission you must have copyright clearance for these articles. We can accept journal articles in NICE Docs.
- Your response should not be longer than 10 pages.

About you	
1.Your name	**Charlotte Martin**
2. Are you (please tick all that apply):	☐ a patient with the condition? ☐ a carer of a patient with the condition?

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

6. If you wrote the organisation submission and/ or do not have anything to add, tick here. (If you tick this box, the rest of this form will be deleted after submission.)	☒ yes

Thank you for your time.

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

……

Your privacy

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

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

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

……

Gilteritinib for treating relapsed or refractory acute myeloid leukaemia: A Single Technology Appraisal

Produced by	School of Health and Related Research (ScHARR), The University of Sheffield
Authors	Paul Tappenden, Professor of Health Economic Modelling, ScHARR, University of Sheffield, Sheffield, UK Emma Simpson, Senior Research Fellow, ScHARR, University of Sheffield, Sheffield, UK John Stevens, Reader in Decision Science, ScHARR, University of Sheffield, Sheffield, UK Aline Navega Biz, Research Associate, ScHARR, University of Sheffield, Sheffield, UK Ruth Wong, Information Specialist, ScHARR, University of Sheffield, Sheffield, UK Srinivas Pillai, Consultant Haematologist, University Hospitals of North Midlands, Stoke-on-Trent, UK
Correspondence Author	Paul Tappenden, Professor of Health Economic Modelling, ScHARR, University of Sheffield, Sheffield, UK
Date completed	21st August 2019

Source of funding: This report was commissioned by the NIHR HTA Programme as project number 129611.

Declared competing interests of the authors

None of the authors have any conflicts of interest to declare.

Acknowledgements

We would like to thank Becky Pennington, ScHARR, for providing comments on the draft report and Gill Rooney Programme Manager, ScHARR, for providing administrative support and in preparing and formatting the report.

Rider on responsibility for report

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

This report should be referenced as follows:

Tappenden P, Simpson E.L., Stevens J, Navega Biz A, Wong R, Pillai S. Gilteritinib for treating relapsed or refractory acute myeloid leukaemia: A Single Technology Appraisal. School of Health and Related Research (ScHARR), The University of Sheffield, 2019.

Contributions of authors

Emma Simpson summarised and critiqued the clinical effectiveness evidence reported within the company's submission. Paul Tappenden and Aline Navega Biz critiqued the company's health economic model. John Stevens critiqued the company's statistical analyses. Ruth Wong critiqued the company's search strategy. Srinivas Pillai provided clinical advice throughout the appraisal. All authors were involved in drafting and commenting on the final report.

CONTENTS

List of tables

List of Figures

List of boxes

Abbreviations

AE	Adverse event
AIC	Akaike Information Criterion
ALFA	Acute Leukemia French Association
ALT	Alanine transaminase
AML	Acute myeloid leukaemia
ANC	Absolute neutrophil count
ASCO	American Society of Clinical Oncology
ASH	American Society of Hematology
AST	Aspartate transaminase
AZA	Azacitidine
BFI	Brief Fatigue Inventory
BIC	Bayesian Information Criterion
BSA	Body surface area
BSC	Best supportive care
CEAC	Cost-effectiveness acceptability curve
CHMP	Committee for Medicinal Products for Human Use
CI	Confidence interval
CIBMTR	Center for International Blood and Marrow Transplant Research
cPAS	Comparator Patient Access Scheme
CPI	Consumer Price Index
CR	Complete remission
CR1	First complete remission
CR2	Second complete remission
CRc	Composite complete remission
CRD	Centre for Reviews and Dissemination
CRh	Complete remission with partial haematological recovery
CrI	Credible interval
CRi	Complete remission with incomplete haematological recovery
CRp	Complete remission with incomplete platelet recovery
CRUK	Cancer Research UK
CS	Company's submission
CSR	Clinical Study Report
DFS	Disease-free survival
DNA	Deoxyribonucleic acid
DoR	Duration of remission
DSA	Deterministic sensitivity analysis
DSU	Decision Support Unit
ECOG	Eastern Cooperative Oncology Group
ED	Emergency department
EFS	Event-free survival
eGFR	Estimated glomerular filtration rate
EHA	European Haematology Association
ELN	European LeukemiaNet
EMA	European Medicines Agency
eMIT	Electronic Market Information Tool
EQ-5D-3L	Euroqol 5-Dimensions (3-level)
EQ-5D-5L	Euroqol 5-Dimensions (5-level)
ERG	Evidence Review Group
ESMO	European Society for Medical Oncology
EU	European Union
EUCTR	EU Clinical Trials Register
FACIT-Dys-SF	Functional Assessment of Chronic Illness Therapy–Dyspnea-Short Form
FACT-Leu	Functional Assessment of Cancer Therapy-Leukemia

FAS	Full Analysis Set
FDA	Food and Drug Administration
FLAG-IDA	Fludarabine, cytarabine, G-CSF with idarubicin
FLT3	FMS-like tyrosine kinase 3
FLT3-ITD	FLT3 internal tandem duplication
FLT3-TKD	FLT3 tyrosine kinase domain
G-CSF	Granulocyte-colony stimulating factor
GEE	Generalised estimating equation
GP	General practitioner
GvHD	Graft versus Host Disease
HiDAC	High-dose cytarabine
HR	Hazard ratio
HRQoL	Health-related quality of life
HSCT	Haematopoietic stem cell transplant
HSE	Health Survey for England
ICER	Incremental cost-effectiveness ratio
IDAC	Intermediate-dose cytarabine
IPD	Individual patient-level data
IRT	Interactive response technology
ISPOR	International Society for Pharmacoeconomics and Outcomes Research
ISSG	InterTASC Information Specialists' Sub-Group
ITT	Intention-to-treat
IV	Intravenous
LFS	Leukaemia-free survival
LoDAC	low-dose cytarabine
LS	Least-squares
LVEF	Left ventricular ejection fraction
LYG	Life year gained
m^2	Metres squared
MEC	Mitoxantrone, etoposide and cytarabine
Mg	Milligram
MIMS	Monthly Index of Medical Specialities
N	Number
N/a	Not applicable
NCCN	National Comprehensive Cancer Network
NCI-CTCAE	National Cancer Institute - Common Terminology Criteria for Adverse Events
NE	Not evaluable
NGS	Next-generation sequencing
NICE	National Institute for Health and Care Excellence
NMA	Network meta-analysis
NYHA	New York Heart Association
OS	Overall survival
PAS	Patient Access Scheme
PCR	Polymerase chain reaction
PH	Proportional hazards
PICO	Population, interventions, comparators, outcomes
PR	Partial remission
PRIMA	Preliminary Independent Model Advice
PRO	Patient reported outcome
PSA	Probabilistic sensitivity analysis
PSS	Personal Social Services
PSSRU	Personal Social Services Research Unit
QALY	Quality-adjusted life year
R/R	Relapsed or refractory

RBC	Red blood cell
RCT	Randomised controlled trial
RDI	Relative dose intensity
SAE	Serious adverse event
SAF	Safety Analysis Set
SC	Subcutaneous
SCT	Stem cell transplant
SD	Standard deviation
SE	Standard error
SLR	Systematic literature review
SmPC	Summary of Product Characteristics
SMR	Standardised mortality ratio
STA	Single Technology Appraisal
TA	Technology Appraisal
TEAE	Treatment-emergent adverse event
TSD	Technical Support Document
TTO	Time-trade-off
UK	United Kingdom
ULN	Upper limit of normal
URD	Unrelated donor
VAS	Visual analogue scale
WTP	Willingness-to-pay

1 SUMMARY

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

The decision problem addressed in the company's submission (CS) is generally appropriate and is in line with the final scope issued by the National Institute for Health and Care Excellence (NICE) with regards to:

- Intervention - gilteritinib (oral), 120mg (3 x 40mg tablets) once daily
- Target population - adults with relapsed or refractory (R/R) FMS-like tyrosine kinase 3 (FLT3) mutation-positive acute myeloid leukaemia (AML)
- Outcomes - overall survival (OS); event-free survival (EFS); disease-free survival (DFS); response rates, including remission; haematopoietic stem cell transplant (HSCT); adverse effects of treatment, and health-related quality of life (HRQoL).

The main comparator considered in the CS reflects a physician's choice of salvage chemotherapy, based on the comparator arm of the ADMIRAL trial. There are some differences between the regimens included in ADMIRAL and the comparators listed in the final NICE scope. However, the clinical advisor to the Evidence Review Group (ERG) stated that the regimens included as comparators in the CS and the company's model are representative of the treatments currently used in clinical practice in England.

1.2 Summary of clinical effectiveness evidence submitted by the company

Aside from a dose expansion study (CHRYSALIS), the key evidence of the clinical effectiveness of gilteritinib was derived from one randomised controlled trial (RCT) - ADMIRAL. In ADMIRAL, adults with FLT3+ R/R AML were randomised to receive gilteritinib 120mg/day (n=247) or salvage chemotherapy (n=124) which was treatment of physician's choice of four alternative regimens:

(i) LoDAC (low-dose cytarabine, 20mg twice-daily SC or IV for 10 days);
(ii) AZA (azacitidine, 75mg/m^2 daily SC or IV for 7 days);
(iii) MEC (mitoxantrone 8mg/m^2 per day, etoposide 100mg/m^2 per day, cytarabine 1,000mg/m^2 per day, all IV 5 days on days 1-5), or;
(iv) FLAG-IDA (fludarabine 30mg/m^2 per day and cytarabine 2,000mg/m^2 per day, both IV for 5 days on days 2-6; granulocyte-colony stimulating factor (G-CSF) 300µg/m^2 per day SC or IV for 5 days on days 1-5, and idarubicin 10mg/m^2 per day IV for 3 days on days 2-4).

OS was statistically significantly longer for patients randomised to gilteritinib (median 9.3 months) than patients randomised to salvage chemotherapy (median 5.6 months), hazard ratio (HR) 0.637 (95% confidence interval (CI) 0.490, 0.830; p=0.0004).

The rate of patients receiving HSCT during the study period was 25.5% for the gilteritinib group and 15.3% for the salvage chemotherapy group (treatment difference *p*=0.0333).

The rate of complete remission or complete remission with partial haematological recovery (CR/CRh) was statistically significantly higher in the gilteritinib arm compared with the salvage chemotherapy arm (34.0% versus 15.3%; *p*=0.0001).

For HRQoL measures in ADMIRAL, [redacted].

Adverse events (AEs) were graded according to the National Cancer Institute Common Terminology Criteria for Adverse Events (NCI-CTCAE). The most common Grade 3 or higher AEs experienced in the gilteritinib group were: febrile neutropenia (45.9%); anaemia (40.7%); thrombocytopenia (22.8%), and platelet count decreased (22.0%).

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

The ERG believes that all RCTs with available data on the clinical effectiveness of gilteritinib in adults with FLT3+ R/R AML were included in the CS.

The study selection criteria applied in the company's review of clinical effectiveness were consistent with the decision problem in the final NICE scope.

The quality of the ADMIRAL RCT was assessed using well-established and recognised criteria. ADMIRAL was an open-label trial, but was of otherwise good methodological quality.

A literature review of gilteritinib and comparator studies identified 38 studies, of which 21 were non-comparative studies, and the other studies did not provide results, or in one instance was a dose-finding study. Only one RCT was identified which could have been included in an indirect comparison with ADMIRAL; however, this was an RCT of quizartinib which was not included in the final NICE scope.

1.4 Summary of cost effectiveness submitted evidence by the company

The CS presents the methods and results of a *de novo* partitioned survival model developed by the company to assess the cost-effectiveness of gilteritinib versus a blended comparator of four salvage chemotherapy regimens (azacitidine, LoDAC, MEC and FLAG-IDA) for the treatment of FLT3+ R/R AML. The blended comparator is based on the overall outcomes from the comparator arm of ADMIRAL and includes weighted costs for each regimen. This comparison represents the company's base case analysis. The company's model also allows for pairwise comparisons of gilteritinib versus

each individual salvage chemotherapy regimen (using regimen-specific costs and outcomes data for the overall ADMIRAL salvage chemotherapy arm) and gilteritinib versus BSC. Incremental health gains, costs and cost-effectiveness are evaluated over a 40-year time horizon from the perspective of the NHS and Personal Social Services (PSS). The company's economic analysis includes a Patient Access Scheme (PAS) for gilteritinib which takes the form of a simple price discount. A comparator PAS is available for azacitidine; the impact of this comparator PAS on the cost-effectiveness of gilteritinib is presented in a separate confidential appendix to this ERG report.

The company's model structure subdivides the overall patient population into two discrete groups according to whether or not they receive HSCT after initiating treatment with gilteritinib or salvage chemotherapy. Based on ADMIRAL, a higher proportion of patients treated with gilteritinib are assumed to receive HSCT compared with salvage chemotherapy (25.5% versus 15.3%, respectively). For each treatment option, separate partitioned survival sub-models are used to estimate health outcomes and costs for the With HSCT and No HSCT groups. Each sub-model includes the same three health states: (i) event-free; (ii) post-event and (iii) death. Within the No HSCT group, EFS and OS were informed by data for patients who did not receive HSCT in ADMIRAL. EFS and OS for the With HSCT group were informed by external data reported by Evers *et al*. The company fitted parametric survival models to the time-to-event data. The company's model includes an assumption of a fixed cure point which applies to all patients who survive up to 3 years; after this timepoint, survival is modelled using general population mortality rates uplifted using a standardised mortality ratio (SMR) of 2.0. With the exception of EFS in the No HSCT subgroup, this cure assumption overrides the event risks predicted by the fitted parametric models after the 3-year timepoint for all surviving patients. Prior to the 3-year cure timepoint, health utility values are based on a generalised estimating equation (GEE) model fitted to EQ-5D-3L data (mapped from the EQ-5D-5L) collected in ADMIRAL. After the cure point, the ERG believes that the company intended to apply general population utility estimates; however, this aspect of the model is subject to a significant error. Resource use estimates were derived from ADMIRAL, standard costing sources, literature and assumptions. With the exception of ongoing post-progression treatment costs for patients leaving the EFS state, after the 3-year cure point, the model assumes that surviving patients incur no further disease management costs.

The probabilistic version of the company's updated model suggests that the incremental cost-effectiveness ratio (ICER) for gilteritinib versus salvage chemotherapy is £46,716 per quality-adjusted life year (QALY) gained. The company's deterministic sensitivity analyses (DSAs) indicate that the probability that patients receive HSCT is a key driver of the cost-effectiveness of gilteritinib; the ERG notes that the benefits accrued by these patients is also a key determinant of the ICER.

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

The ERG critically appraised the company's health economic analyses and double-programmed the deterministic version of the company's original model. The ERG's critical appraisal identified several issues relating to the company's model and the evidence used to inform its parameters. The most pertinent of these include: (i) the presence of model errors; (ii) concerns regarding the company's model structure; (iii) uncertainty surrounding the cure point and the use of external evidence to inform OS for patients receiving HSCT; (iv) inconsistencies between model-predicted OS and observed OS in ADMIRAL; (v) issues relating to the company's indirect comparisons, particularly the estimated treatment effect for gilteritinib maintenance therapy, and (vi) the underestimation of gilteritinib drug costs.

The ERG notes that the company's decision to use external evidence from Evers *et al* to inform post-HSCT OS, rather than using the available data on OS from randomisation in ADMIRAL, has significant implications for the proportion of patients who are assumed to receive the benefits of cure. The company's model suggests that approximately [redacted] of gilteritinib-treated patients who receive HSCT will remain alive at the assumed 3-year cure point. The available final data cut-off of ADMIRAL (17th September 2018) suggests a considerably less favourable survival prognosis for FLT3+ R/R AML patients who have received HSCT (3-year OS probability based on the ERG's preferred standard log normal model=[redacted]). As such, the ERG believes that the results of the company's model are likely to be optimistic.

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

1.6.1 Strengths

- The company's search for gilteritinib studies was comprehensive. The ERG believes that no relevant available RCTs of gilteritinib were excluded.
- With the exception of its open-label design, the ADMIRAL trial was of good methodological quality.
- ADMIRAL included an active comparator (not placebo).
- According to clinical advice, the comparators used in ADMIRAL were reflective of clinical practice in England.
- The ERG's clinical advisor believed that the demographics of the ADMIRAL trial population are sufficiently representative of the target population for the results of the trial to be applicable for patients in England.
- The company's survival modelling was well presented within the CS. The description of the company's model was clear.

1.6.2 Weaknesses and areas of uncertainty

- Apart from a dose expansion study, there was only one RCT of gilteritinib.
- ADMIRAL adopted an open-label design.
- In ADMIRAL, the majority of salvage chemotherapy patients finished treatment by cycle 2, and there was a low completion rate of HRQoL instruments in the comparator group.
- The available evidence for effectiveness of comparator interventions did not allow an indirect comparison between gilteritinib and all of the comparators listed in the final NICE scope.
- The company's model excludes data for patients who received HSCT in ADMIRAL trial.
- There is uncertainty regarding long-term outcomes for FLT3+ R/R AML patients who receive HSCT. The company's clarification response states that the latest data cut-off of ADMIRAL is final; further analyses are not anticipated. The ERG believes that additional follow-up would have helped to resolve uncertainty surrounding the OS benefits associated with gilteritinib.

1.7 Summary of exploratory and sensitivity analyses undertaken by the ERG

The ERG undertook eight exploratory analyses. These included: (i) correcting model errors; (ii) applying parametric models fitted to data on OS from randomisation for patients with HSCT in ADMIRAL; (iii) the inclusion of utility values based on Ara and Brazier; (iv) removing double-counting of health losses and costs associated with progressive AML; (v) removing the treatment effect for gilteritinib maintenance therapy; (vi) the inclusion of drug wastage for gilteritinib; (vii) amending costs associated with long-term follow-up, post-relapse/progression treatment and FLT3 mutation re-testing rates. The ERG's preferred base case combines all of these model amendments. The ERG undertook additional sensitivity analyses using the ERG-preferred model to explore the impact of: (i) using an alternative external source of evidence for post-HSCT outcomes (Poiré *et al*); (ii) applying alternative survival models for patients without HSCT; (iii) applying alternative survival models for patients with HSCT, and (iv) assuming alternative cure timepoints. The ERG also fitted mixture-cure models to the OS data for patients with HSCT in ADMIRAL in order to estimate cure fractions; this analysis was used to explore whether the results obtained from company's model and the ERG's preferred model are likely to be optimistic.

The ERG's preferred base case analysis suggests that the ICER for gilteritinib versus salvage chemotherapy is £102,234 per QALY gained. This is considerably higher than the company's base case ICER of £47,695 per QALY gained. This difference is largely a consequence of the lower expected 3-year survival rate for patients receiving HSCT in ADMIRAL compared with Evers *et al*. All of the ERG's analyses which use OS data for patients with HSCT in ADMIRAL lead to ICERs which are higher than the company's base case estimate.

The ERG believes that gilteritinib is likely to meet NICE's end of life criteria.

2 BACKGROUND

This chapter presents a brief critique of the company's description of the disease and the current treatment pathway in England.

2.1 Critique of the company's description of the underlying health problem

The company's submission[1] (CS) contains a brief but accurate overview of acute myeloid leukaemia (AML). AML is a cancer of the white blood cells, characterised by the uncontrolled proliferation and infiltration of bone marrow and blood by abnormally or poorly differentiated leukaemic blasts of the myeloid cell lineage. The CS highlights that AML is rare, and cites an age-standardised crude incidence rate of 4.8 per 100,000 people in the UK, based on estimates reported by Cancer Research UK (CRUK).[2] Data reported by CRUK indicate that in 2016 there were 2,543 new cases of AML. The incidence of AML increases sharply with age, and the disease is most common in patients aged 60 years and above. The CS highlights that FMS-like tyrosine kinase 3 (FLT3) mutations (particularly FLT3-ITD and FLT3-TKD) are common and occur in around 30% of patients with AML. Risk classifications published by the National Comprehensive Cancer Network (NCCN) and the European LeukemiaNet (ELN) both classify patients with FLT3-ITD mutations as "poor-risk".[3, 4]

The CS[1] notes that the prognosis for patients with AML is poor, with typical survival rates of around 25% at 5-years. Survival prognosis is strongly influenced by age, with younger patients having a more favourable outlook. The indication for gilteritinib within this Single Technology Appraisal (STA) relates to patients with relapsed/refractory (R/R) AML with a FLT3 mutation. The CS highlights that survival outcomes are worse for patients with relapsed disease and for patients with FLT3+ mutations.

According to the CS,[1] only patients with a FLT3+ mutation will be eligible for treatment with gilteritinib. FLT3 mutations can evolve over the patient's lifetime,[5] hence FLT3 mutation re-testing will be required for all AML patients with R/R disease, regardless of their previous FLT3 mutation status.

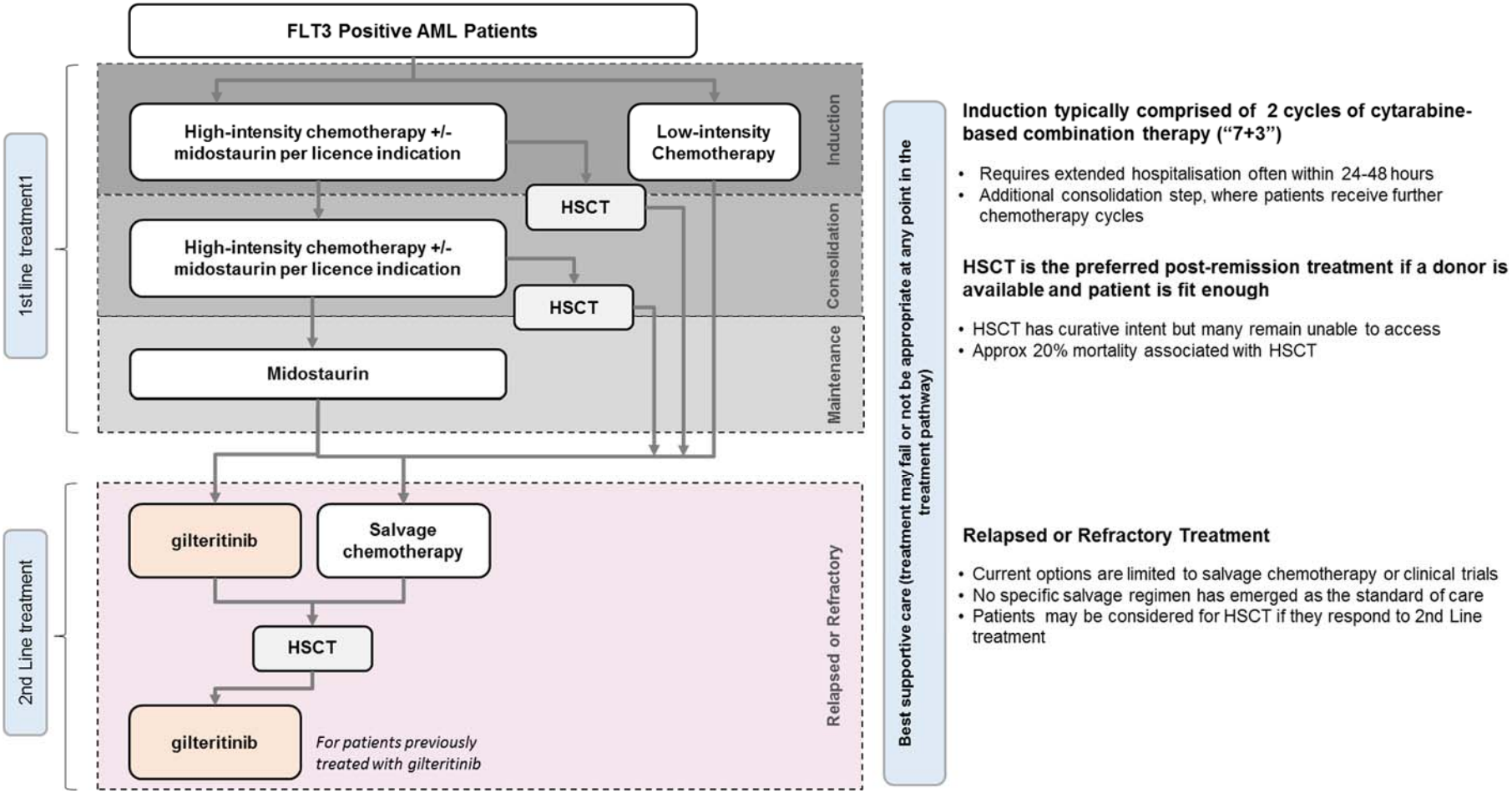
2.2 Critique of the company's overview of current service provision

The company's view of the current treatment pathway is shown in Figure 1. The clinical advisor to the Evidence Review Group (ERG) considered this to be a generally reasonable representation of the current treatment pathway for patients with AML. The relevant part of the treatment pathway for this appraisal relates to the lower portion of the figure (second-line treatments for patients with R/R disease). The ERG's clinical advisor commented that the diagram of the treatment pathway should have included low intensity chemotherapy and best supportive care (BSC) more clearly as potential options. The aim of treatment for patients with FLT3+ R/R AML is to achieve complete remission (CR) followed by haematopoietic stem cell transplant (HSCT) in suitable patients for long-term disease control. The CS[1]

highlights that whilst many patients receiving salvage chemotherapy may achieve remission, this tends to be of a limited duration and comes at the cost of significant toxicity. The only treatment which has a curative potential is HSCT, but this is not an option for the majority of patients due to the low probability of achieving CR in FLT3+ R/R AML and lack of fitness. The CS highlights that there is no standard of care chemotherapy regimen and a wide range of alternative regimens may be considered. The ERG's clinical advisor noted that higher intensity regimens (e.g. FLAG-IDA - fludarabine, cytarabine, and granulocyte-colony stimulating factor [G-CSF] with idarubicin) may be offered to fitter patients who have an increased likelihood of being eligible for transplant, whilst lower intensity options may be used for patients who are unlikely to be able to proceed to HSCT. The ERG's clinical advisor also noted that BSC is a relevant option, as a proportion of patients with FLT3+ R/R AML may choose not to receive active therapy due to toxicity associated with chemotherapy or the burden associated with hospital attendances and prolonged inpatient stays for infusional therapies.

The CS[1] highlights that gilteritinib is the first FLT3 inhibitor available for the treatment of R/R AML and that it provides benefits in terms of survival (with an increased probability of bridging to HSCT) as well as reduced toxicity compared with salvage chemotherapy. The company's intended positioning of gilteritinib is as an alternative to salvage chemotherapy, and potentially BSC. As shown in Figure 1, it is anticipated that gilteritinib may be used in all patients irrespective of whether they will subsequently undergo HSCT, and after HSCT as a maintenance therapy. The ERG notes that the figure does not include re-testing for the FLT3 mutation; this would be required for all patients following failure of first-line therapy. The CS states that most centres in England currently test for FLT3 mutations using polymerase chain reaction (PCR) based assays using deoxyribonucleic acid (DNA) isolated from patient samples. However, the ERG's clinical advisor commented that there is variation in the methodology used, as some laboratories are currently using next-generation sequencing (NGS) whilst others use the technique referred to in the CS.

Figure 1: AML treatment pathway (reproduced from CS, Figure 1)



AML – acute myeloid leukaemia; FLT3 – FMS-like tyrosine kinase 3; HSCT – haematopoietic stem cell transplant.

3 CRITIQUE OF COMPANY'S DEFINITION OF THE DECISION PROBLEM

This chapter presents a summary and critique of the decision problem addressed by the CS.[1] A summary of the decision problem as outlined in the final scope issued by the National Institute for Health and Care Excellence[6] (NICE) and addressed in the CS is presented in Table 1.

Table 1: Company's statement of the decision problem (reproduced from CS, Table 1)

	Final scope issued by NICE	Decision problem addressed in the CS	Rationale if different from the final NICE scope
Population	Adults with relapsed or refractory FLT3 mutation-positive acute myeloid leukaemia	Adults with relapsed or refractory FLT3 mutation-positive acute myeloid leukaemia	N/a
Intervention	Gilteritinib	Gilteritinib	N/a
Comparator(s)	Established clinical management without gilteritinib, for example: • Intermediate dose cytarabine (IDAC) • Fludarabine, cytarabine, granulocyte-colony stimulating factor (G-CSF) with idarubicin (FLAG-IDA) • Best supportive care • Hydroxycarbamide (for people who cannot have chemotherapy or stem cell transplant)	Established clinical management without gilteritinib including, but not limited to cytarabine or azacitidine based chemotherapy. For some patients, best supportive care may be their only option currently	The comparators used in the model are those included within the pivotal Phase III trial (ADMIRAL). These were considered commonly used agents across the geographies for the trial
Outcomes	The outcome measures to be considered include: • Overall survival • Event-free survival • Disease-free survival • Response rates, including remission • Stem cell transplant • Adverse effects of treatment • Health-related quality of life	The outcome measures to be considered include: • Overall survival • Event-free survival • Disease-free survival • Response rates, including remission • Stem cell transplant • Adverse effects of treatment • Health-related quality of life	N/a
Special considerations including issues related to equity or equality		Gilteritinib represents an end-of-life (EoL) treatment based on NICE criteria: it is indicated in a population with a life expectancy less than 24 months and offers a survival extension of greater than 3 months. AML is an orphan condition, with an incidence of approximately 4.8 per 100,000. Relapsed or refractory patients are estimated to be 57% of these and FLT3 mutation occurs in approximately 30% of patients	Scope did not include such commentary, but did include the commentary around the life expectancy

CS - company's submission; FLT3 - FMS-like tyrosine kinase 3; N/a – not applicable

3.1 Population

The patient population in the CS[1] relates to people with R/R AML with a FLT3 mutation. This is in line with the population defined in the final NICE scope[6] and the main source of clinical evidence included in the CS – the ADMIRAL trial.[7] This indication is also in line with the draft Summary of Product Characteristics (SmPC) for gilteritinib (CS Appendix C[8]).

The ADMIRAL trial was conducted in 107 sites in a total of 14 countries including North America, Europe, Asia and the rest of the world.[7] Of these, four centres were based in the UK.[7] The ERG's clinical advisor suggested that the population recruited into this trial broadly reflects the population who would be eligible for treatment with gilteritinib in England.

As gilteritinib has not yet received a European/UK marketing authorisation for any indication, it is not yet clear whether certain medical conditions or patient groups may be contraindicated for treatment. The draft SmPC[8] states that the safety and efficacy of gilteritinib has not yet been established in children aged below 18 years and that no data are available.

3.2 Intervention

The intervention considered in the CS[1] is 120mg gilteritinib (3 x 40mg tablets) once daily. Gilteritinib (ASP2215, XOSPATA™) is a FMS-like tyrosine kinase-3 (FLT3) and AXL inhibitor manufactured by Astellas Pharmaceuticals Ltd. Gilteritinib was granted an orphan designation (EU/3/17/1961) by European Commission in January 2018. According to the CS, Astellas applied for a licence with the European Medicines Agency (EMA) on 28th February 2019 and expects a recommendation to be made by the Committee for Medicinal Products for Human Use (CHMP) in [redacted], with a licence granted in [redacted]. The CS states that gilteritinib is currently being assessed under Accelerated Assessment criteria. The CS states that the expected indication of gilteritinib is for the treatment of adult patients who have R/R AML with a FLT3 mutation.

The anticipated list price per pack of 84 x 40mg gilteritinib tablets (28 days' supply) is [redacted].[1] The company has proposed a Patient Access Scheme (PAS) which takes the form of a simple price discount of [redacted]; the discounted cost per pack of gilteritinib is [redacted].

The draft SmPC[8] states that treatment should continue until the patient is no longer clinically benefiting from gilteritinib. Within the ADMIRAL trial,[7] criteria for discontinuing treatment with gilteritinib included intolerable or unacceptable toxicity and disease progression.

3.3 Comparators

The final NICE scope[6] lists four comparators: (i) intermediate-dose cytarabine (IDAC); (ii) fludarabine, cytarabine, and G-CSF with idarubicin (FLAG-IDA); (iii) BSC, and (iv) hydroxycarbamide (for people who cannot have chemotherapy or stem cell transplant).

The main comparator considered within the clinical section of the CS[1] and the company's health economic model reflects the salvage chemotherapy arm of the ADMIRAL trial.[7] The salvage chemotherapy arm of ADMIRAL included four active treatment regimens: (i) azacitidine; (ii) low-dose cytarabine (LoDAC); (iii) mitoxantrone, etoposide and cytarabine (MEC), and (iv) FLAG-IDA. Within the company's model, these regimens form a blended comparator, based on pooled outcomes data for the ADMIRAL comparator group with weighted regimen-specific drug acquisition, administration and hospitalisation costs. In addition, the company's model includes BSC as a further comparator.

Despite the differences between the company's model and the NICE scope,[6] the ERG's clinical advisor stated that the regimens included as comparators in the CS[1] and the model are representative of treatments currently used in clinical practice in England. They also noted that these include regimens which would be offered to patients for whom HSCT may be considered (FLAG-IDA and MEC) as well as regimens which would be offered to less fit patients for whom the treatment intent is to achieve maximum disease control where transplant is unlikely (azacitidine and LoDAC).

The ERG notes that a comparator Patient Access Scheme (cPAS) is available for one of the technologies included in the salvage chemotherapy arm of the company's model (azacitidine). The results of the company's model including this cPAS are presented in a separate confidential appendix to this ERG report.

3.4 Outcomes

Outcomes included in the final NICE scope[6] include:

- Overall survival (OS)
- Event-free survival (EFS)
- Disease-free survival (DFS, referred to as leukaemia-free survival [LFS])
- Response rates, including remission
- Stem cell transplant (SCT)
- Adverse effects of treatment
- Health-related quality of life (HRQoL).

The CS[1] reports clinical results for all of these endpoints. The company's model includes data relating to OS, EFS, HSCT rates, AEs and HRQoL (see Section 5).

3.5 Other relevant factors

The CS[1] states that the company is not aware of any issues of equality relevant to this appraisal. The CS states that gilteritinib should be considered as an end-of-life treatment; this is discussed further in Section 6.

4 CLINICAL EFFECTIVENESS

4.1 Review methods

The company conducted a systematic literature review (SLR) to identify all studies of gilteritinib and its comparators for the treatment of FLT3 mutation-positive R/R AML in adults. HRQoL was investigated within a separate systematic review (company's SLR report,[9] Section 6).

Searches

The company performed one clinical effectiveness search to identify all studies of gilteritinib and currently licensed drugs (azacitidine, cytarabine, decitabine, FLAG-IDA, G-CSF, idarubicin, mitoxantrone with etoposide and cytarabine and sorafenib) and drugs in Phase III development (crenolanib, ibrutinib and sorafenib) for the treatment of adults with FLT3+ R/R AML.

The company searched several electronic bibliographic databases in October 2018, including: MEDLINE [via Ovid]; MEDLINE in Process [via Ovid]; EMBASE [via Ovid]; Cochrane Database of Systematic Reviews [via Wiley]; Cochrane Central Register of Controlled Trials [via Wiley], and Database of Abstracts of Reviews of Effects [via Wiley].

The company's search was completed more than six months prior to the date of the submission. During the clarification process, the ERG requested that the company update their search to confirm that no further relevant studies had been published since the date of the original search (see clarification response,[7] question A1). In response to this request, the company updated their search up to July 2019; this resulted the identification of seven additional relevant publications.[7]

Whilst the company's clinical search strategy (company's SLR report,[9] Appendix I: 8.1.2 statements #11-18) lists all the interventions and comparators, the reasons for the omission of the free-text trade names of interventions and comparators (e.g. XOSPATA) in the database searches, and their impact on search recall, are unclear.

The company searched the Clinicaltrials.gov trials registry in October 2018. Supplementary searches conducted by the company included searching several conference abstract websites (since 2016): American Society of Clinical Oncology (ASCO); American Society of Hematology (ASH); European Haematology Association (EHA); European Society for Medical Oncology (ESMO) and the International Society for Pharmacoeconomics and Outcomes Research (ISPOR).[8]

Table 1 of CS Appendix D[8] lists the inclusion and exclusion criteria for the company's review. This includes observational studies as a study design of interest. However, based on the search strategies

given in Appendix I of the company's SLR report,[9] the randomised controlled trial (RCT) and non-RCT search filters applied to the clinical search strategy for both Medline and Embase (statements #19-22) were not adequate for the retrieval of observational studies. In their clarification response[7] (question A1), the company agreed that the terms for observational studies should have been included. The company revised the search (to include observational*.ti,ab. or exp observational study/ or exp observational studies as topic/), although the impact on the number of records retrieved was minimal and no additional relevant studies were identified. The ERG believes that the company's additional search terms were not sensitive for the purpose of retrieving observational studies, and that the use of the observational studies search filter from the InterTASC Information Specialists' Sub-Group (ISSG) resource[10] would have been more appropriate.

Inclusion criteria

The eligibility criteria applied in the company's clinical effectiveness review (CS Appendix D,[8] Table 1 and company's SLR report,[9] Section 4.2.3) were considered by the ERG to be reasonable and consistent with the decision problem outlined in the final NICE scope.[6] Study selection was conducted by one reviewer and checked by another reviewer (CS Appendix D,[8] Section D1.1.3), as is good practice in systematic reviews.

The review inclusion criteria specified the population as adult patients aged 18 years or older with FLT3+ R/R AML; this is in line with the final NICE scope.[6] The intervention included was gilteritinib monotherapy; this is also consistent with the final NICE scope.

The final NICE scope[6] lists the following comparators: *"Established clinical management without gilteritinib including but not limited to: IDAC; FLAG-IDA; best supportive care; hydroxycarbamide."* With the exception of hydroxycarbamide (for people who cannot have chemotherapy or HSCT), all of these treatments were listed as comparators in the company's review (CS Appendix D,[8] Table 1, and company's SLR report,[9] Table 4-1). Hydroxycarbamide was not listed as a comparator sought by the search; however, it was considered by the company as part of BSC. The following additional comparators were also included in the company's review: azacitidine; crenolanib; cytarabine (low-dose [LoDAC]; high-dose [HiDAC]); decitabine; G-CSF; ibrutinib; idarubicin; MEC, and sorafenib. The company included quizartinib, which was appropriate as it had been included as a comparator in the draft NICE scope at the time at which the searches were undertaken, but was subsequently removed from the final NICE scope.

The company included the following outcomes in line with the final NICE scope: overall survival (OS); event-free survival (EFS); disease-free survival (leukaemia-free survival [LFS]); response rates including remission (ORR, rates of complete remission with partial haematological recovery [CRh],

incomplete haematological recovery [CRi], and incomplete platelet recovery [CRp] and duration of remission [DoR]); adverse effects of treatment, and HSCT rate.[9] The company additionally included transfusion conversion rate (the proportion of patients who were transfusion-dependent at baseline but became transfusion-independent during gilteritinib treatment); and transfusion maintenance rate (the proportion of patients who were transfusion-independent at baseline and remained transfusion-independent during gilteritinib treatment).[9] The final NICE scope[6] also listed HRQoL as an outcome; the company's review of HRQoL evidence was undertaken separately (reported in the company's SLR report,[9] Section 6).

Data extraction

A data extraction form was designed for the review. Items for extraction included information about the population, interventions, comparators, outcomes (PICO) and study characteristics.[9] Data were extracted by one reviewer and checked by another reviewer,[8] as is good practice in systematic reviews. The ERG checked information reported within the CS[1] against the ADMIRAL Clinical Study Report[7] (CSR) and trial publications[11-15] where possible, and found the data provided to be accurate.

Quality assessment

According to CS Appendix D1.1.5,[8] study quality was assessed by one reviewer and checked by a second reviewer, as is good practice in systematic reviews. The items included in the company's quality assessment were taken from the Centre for Reviews and Dissemination (CRD) guidelines for undertaking reviews in health care.[16] These are standard and appropriate criteria for assessing the risk of bias in RCTs, and are applicable to the ADMIRAL trial.[7] These criteria would not usually be applied to a dose-finding trial (CHRYSALIS) in which the primary outcomes relate to safety, tolerability and pharmacokinetics. Applying these criteria to the CHRYSALIS trial (CS Appendix D,[8] see ERG Appendix 1) indicated that the randomised part of the trial was well-conducted in terms of randomisation, with a computer generated randomisation sequence and allocation by interactive response technology (IRT).[7] A modified intention-to-treat (ITT) analysis was presented, with the Full Analysis Set (FAS) comprising patients who received at least one dose of study drug and had at least one post-treatment data point (CS Appendix D1.2[8]).

The company's quality assessment of ADMIRAL is provided in CS Appendix D1.3.[8] This was checked by the ERG against information provided by the company, the CSR[7] and clinical trial registries[17, 18] (see Table 2).

Randomised sequence generation and allocation concealment were conducted by IRT (see clarification response,[7] question A5 and company's SLR report,[9] Table 12-1), giving a low risk of selection bias. Randomisation in ADMIRAL was stratified according to: (1) pre-selected salvage therapy: high

(FLAG-IDA, MEC); or low intensity (LoDAC, AZA), and (2) patients' response to first-line AML therapy: primary refractory disease without HSCT; relapse within or after 6 months of chemotherapy alone, no HSCT; relapse within or after 6 months of allogeneic HSCT.[7, 13]

There was also a low risk of bias with respect to balance between groups, as baseline characteristics were similar and there were no unexpected imbalances in drop-outs between groups.[1, 8] For effectiveness measures, an ITT analysis was presented of all randomised patients.

The use of a comparator involving physician's choice of treatment is common in cancer trials in which no single treatment is favoured at a particular point in the care pathway, and selecting treatment prior to randomisation, as was the case in ADMIRAL, could reduce the risk of selection bias. All comparator treatments were sufficiently representative of current clinical practice in England. The ERG notes that the use of a comparator involving treatment of physician's choice means that attempting to compare the intervention with one of the comparators would break randomisation and would lead to a smaller sample size than using one comparator drug.[19]

The ADMIRAL trial[7] was open-label. Lack of blinding can lead to a high risk of performance and detection bias. Patient-reported outcome measures are more likely to be biased than objective measures such as OS.[16] Due to differences between the intervention and comparators in terms of route of administration, blinding would require a double-dummy trial design. This would reduce bias for objective measures, but would disguise potential HRQoL benefits associated with the mode of administration.

The results of the ADMIRAL trial[7] have not yet been published; as such, it cannot be assessed if the authors measured more outcomes than they reported. However, data for outcomes of relevance to this review were all provided in the CS and the accompanying appendices, the company's SLR report and the references supplied.

Table 2: Quality assessment of ADMIRAL (adapted from CS Appendix D, Table 24)

ADMIRAL EudraCT 2015-000140-42 NCT02421939	Company's quality assessment	ERG's quality assessment
Was randomisation carried out appropriately?	Yes	Yes Randomisation sequence generated by IRT (Clarification response,[7] question A5)
Was the concealment of treatment allocation adequate?	N/a, open-label design	Yes Allocation by IRT (Clarification response,[7] question A5; company's SLR report,[9] Table 12-1; ADMIRAL CSR,[7] Section 5.3.3)
Were the groups similar at the outset of the study in terms of prognostic factors?	Yes	Yes (CS[1] Table 9)
Were the care providers, participants and outcome assessors blind to treatment allocation?	Open-label design: investigators and participants were not blinded to allocation, but sponsor statisticians were blinded	No (CS[1] Section B.2.2)
Were there any unexpected imbalances in drop-outs between groups?	No	No (CS Appendix D1.2[8])
Is there any evidence to suggest that the authors measured more outcomes than they reported?	No	N/a Results for ADMIRAL have not yet been published.
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 (CS Appendix D1.2[8] and company's SLR report,[9] Table 12-1)

CS - company's submission; CSR - clinical study report; IRT - interactive response technology; N/a - not applicable; SLR - company's systematic literature review

4.2 Included trials of gilteritinib

The systematic review of clinical effectiveness of gilteritinib and comparators identified 870 unique records.[9] At the full text sift stage, 68 articles were excluded,[9] with the majority of these being rejected for not providing data (company's clarification response,[7] question A17). Thirty-eight studies of gilteritinib or comparators were included in the review.[9] Several of these also did not provide data (company's clarification response,[7] question B2). Of these, two studies of gilteritinib were included and reported on in the CS[1] (CHRYSALIS and ADMIRAL).

Seven ongoing studies of gilteritinib were identified. Six of these relate to studies of gilteritinib at a different position in the treatment pathway than that considered within this appraisal. One ongoing study is investigating gilteritinib monotherapy in FLT3+ R/R AML (2215-CL-0303, NCT03182244). This is an open-label RCT comparing gilteritinib with salvage chemotherapy (LoDAC, MEC, FLAG), with centres in China, Malaysia, Russia, Singapore and Thailand. The estimated date for collecting final data for the primary outcome (OS) is March 2020.[18]

4.2.1 Included trials

Two gilteritinib trials were included in the company's review (see Table 3). One of these was a dose-expansion study (CHRYSALIS), whilst the other study (ADMIRAL) provided the key evidence for the clinical effectiveness of gilteritinib. Both trials were registered on the EU Clinical Trials Register[17] (EUCTR) and ClinicalTrials.gov.[18] CHRYSALIS was published as an abstract (Perl *et al*, 2016[11]) and a full paper (Perl *et al*, 2017[12]). At the time of writing, the protocol of the ADMIRAL trial had been published (Gorcea, 2018;[13] Perl, 2017;[14] Wang, 2018[15]), but the results had not (CS Appendix D[8] and clarification response,[7] question A1). Both ADMIRAL and CHRYSALIS were international multicentre trials. ADMIRAL had four centres from the UK (clarification response,[7] question A11); CHRYSALIS had none.[1]

Table 3: Included gilteritinib trials

Study name	Study design	References
CHRYSALIS **2215-CL-0101** **NCT02014558** **EudraCT 2014-002217-31**	Phase I-II RCT open-label Multicentre, international	Perl *et al* (2016). Final results of the Chrysalis trial: a first-in-human phase 1/2 dose-escalation, dose-expansion study of gilteritinib (ASP2215) in patients with relapsed/refractory acute myeloid leukemia (R/R AML). Blood, 128(22), 1069.[11] Perl *et al* (2017). Selective inhibition of FLT3 by gilteritinib in relapsed or refractory acute myeloid leukaemia: a multicentre, first-in-human, open-label, phase 1–2 study. The Lancet Oncology, 18(8), 1061-1075.[12] US National Library of Medicine[18] https://clinicaltrials.gov/ct2/show/NCT02014558 EU Clinical Trials Register[17] https://www.clinicaltrialsregister.eu/ctr-search/trial/2014-002217-31/results
ADMIRAL **2215-CL-0301** **NCT02421939** **EudraCT 2015-000140-42**	Phase III RCT open-label Multicentre, international	ADMIRAL CSR[7] US National Library of Medicine[18] https://clinicaltrials.gov/ct2/show/NCT02421939 EU Clinical Trials Register[17] https://www.clinicaltrialsregister.eu/ctr-search/trial/2015-000140-42/DE Gorcea *et al* 2018. ASP2215 in the treatment of relapsed/refractory acute myeloid leukemia with FLT3 mutation: background and design of the ADMIRAL trial. Future Oncology, 14(20), pp.1995-2004.[13] Perl *et al* 2017. An open-label, randomized phase III study of gilteritinib versus salvage chemotherapy in relapsed or refractory FLT3 mutation-positive acute myeloid leukemia. Journal of Clinical Oncology 2017 35:15_suppl, TPS7067-TPS7067[14] Wang J *et al*. A phase III randomized study of gilteritinib versus salvage chemotherapy in FLT3 mutation-positive subjects with relapsed or refractory acute myeloid leukemia. Annals of Oncology. 2018. 29 (Supplement 9)[15]

CSR - clinical study report; FLT3 - FMS-like tyrosine kinase 3; RCT - randomized controlled trial
Based on information provided in CS Appendix D Table 2 and company's clarification response (question A1).

Eligibility criteria for both studies were provided in CS Table 6 (see ERG Appendix 1). Both trials included patients aged 18 years or older, with R/R AML (see Table 4 and Table 5). Within the

ADMIRAL trial, patients were required to have a confirmed FLT3 mutation. Within CHRYSALIS, patients were not required to have a FLT3 mutation, but the study did require at least 10 patients with FLT3 in each dose expansion group.[11, 12] Eligibility criteria for both studies specified an Eastern Cooperative Oncology Group (ECOG) performance status (PS) of ≤2, alanine transaminase (ALT) or aspartate transaminase (AST) ≤2.5 X the upper limit of normal (ULN), total bilirubin ≤1.5 X ULN and serum creatinine ≤1.5 X ULN or estimated glomerular filtration rate (eGFR) >50 mL/min, and excluded patients with Long QT syndrome or New York Heart Association (NYHA) class 3 or 4 heart failure (unless a screening echocardiogram resulted in a left ventricular ejection fraction [LVEF] of ≥45%).

Table 4: CHRYSALIS study characteristics

Study acronym and ID	Population	Sample size	Intervention	Primary outcomes
CHRYSALIS NCT02014558 2215-CL-0101 EudraCT 2014-002217-31	Adults with R/R AML (Perl *et al*, 2017)[12]	Total N=252: comprising N=23 dose escalation and N=229 dose expansion N=191 had FLT3 mutations of whom N=169 received a dose of ≥80mg/day (CS,[1] Section B.2.5) (Perl *et al* 2017)[12]	Gilteritinib: 7 doses (20mg, 40mg, 80mg, 120mg, 200mg, 300mg, 450mg /day) (Perl *et al* 2017)[12]	Tolerability (first dose to end of cycle, 30 days) Safety (up to 30 days after last dose of study drug) Pharmacokinetics (during treatment) (Perl *et al*, 2016)[11] (Perl *et al*, 2017)[12]

FLT3 - FMS-like tyrosine kinase 3; GIL - gilteritinib; R/R AML - relapsed or refractory acute myeloid leukaemia.
Based on information from CS Appendix D Table 2, CS Table 5 and CS Section B.2.2

In CHRYSALIS, the Safety Analysis Set (SAF) included all 252 patients who received at least 1 dose of the study drug. The FAS included 249 patients with at least one post-treatment data point;[1] three patients were excluded *"due to concerns with the site's compliance to good clinical practice."*[8, 12]

An ITT analysis was presented for ADMIRAL. The ADMIRAL FAS included all randomised patients with a FLT3 mutation (n=366; gilteritinib n=243, salvage chemotherapy n=123).[1] The SAF included all patients who took at least one dose of study treatment (n=355; gilteritinib n=246, salvage chemotherapy n=109).[1]

The ADMIRAL trial allowed concomitant treatment with hydroxycarbamide daily for up to 2 weeks.[1] Patients in the gilteritinib arm were prohibited from receiving drugs that were strong inducers of CYP3A, strong inhibitors or inducers of P-gp and concomitant drugs that target serotonin 5HT1R or 5HT2BR or sigma nonspecific receptor, and precaution was advised for drugs that are known to prolong QT or QTc intervals and drugs that are substrates of breast cancer resistance protein.[1]

The outcomes used in ADMIRAL and CHRYSALIS is defined in Table 6.

Table 5: ADMIRAL study characteristics

Study	Population	Sample size	Intervention	Comparator	Primary outcomes
ADMIRAL 2215-CL-0301 NCT02421939 EudraCT 2015-000140-42	Adults with R/R FLT3 mutation-positive AML	371 total 247 randomised to gilteritinib, of whom n=246 received gilteritinib; 124 randomised to comparator of whom n=109 received salvage chemotherapy	Gilteritinib 120mg/day continued until patient meets a treatment discontinuation criterion	Salvage chemotherapy, investigator's choice LoDAC (low-dose cytarabine, 20mg twice-daily SC or IV for 10 days), continued until patient meets a treatment discontinuation criterion, n=17 randomised, n=16 received AZA (azacitidine, 75mg/m^2 daily SC or IV for 7 days), continued until patient meets a treatment discontinuation criterion, n=32 randomised, n=25 received MEC (mitoxantrone 8mg/m^2 per day, etoposide 100mg/m^2 per day, cytarabine 1,000mg/m^2 per day, all IV 5 days on days 1-5), maximum 2 cycles, n=33 randomised, n=28 received FLAG-IDA (fludarabine 30mg/m^2 per day and cytarabine 2,000mg/m^2 per day, both IV for 5 days on days 2-6; G-CSF 300µg/m^2 per day SC or IV for 5 days on days 1-5; idarubicin 10mg/m^2 per day IV for 3 days on days 2-4), maximum 2 cycles, n=42 randomised, n=40 received	• Overall survival (OS) • Rate of complete remission (CR) and complete remission with partial haematological recovery (CR_h)

AML - acute myeloid leukaemia; FLT3 - FMS-like tyrosine kinase 3; G-CSF - granulocyte-colony stimulating factor; IV - intravenous; R/R - relapsed or refractory; SC - subcutaneous
Based on information provided in CS Appendix D Table 2, CS Table 5 Gorcea et al (2018) and ADMIRAL CSR

Table 6: Outcome definitions for CHRYSALIS and ADMIRAL (reproduced from company's clarification response, question A9)

Definition	Description
Complete remission (CR)	For subjects to be classified as being in CR at a post-baseline visit, they must have bone marrow regenerating normal hematopoietic cells and achieve a morphologic leukaemia-free state and must have an ANC $\geq 1 \times 10^9$/L and platelet count $\geq 100 \times 10^9$/L and normal marrow differential with < 5% blasts, and they will be RBC and platelet transfusion independent (defined as 1 week without RBC transfusion and 1 week without platelet transfusion). There should be no evidence of extramedullary leukaemia
Complete remission with partial haematologic recovery (CRh)	At a post-baseline visit, subjects will be classified as CRh if they have marrow blasts < 5%, partial haematologic recovery ANC $\geq 0.5 \times 10^9$/L and platelets $\geq 50 \times 10^9$/L, no evidence of extramedullary leukaemia and cannot be classified as CR
Complete remission with incomplete platelet recovery (CRp)	For subjects to be classified as being in CRp at a post-baseline visit, they must achieve CR except for incomplete platelet recovery ($< 100 \times 10^9$/L)
Complete remission with incomplete haematologic recovery (CRi)	For subjects to be classified as being in CRi at a post-baseline visit, they must fulfill all the criteria for CR except for incomplete haematological recovery with residual neutropenia $< 1 \times 10^9$/L with or without complete platelet recovery. RBC and platelet transfusion independence is not required
Composite complete remission (CRc)	For subjects to be classified as being in CRc at a post-baseline visit, they must either achieve CR, CRp or CRi at the visit
Partial remission (PR)	For subjects to be classified as being in PR at a post-baseline visit, they must have bone marrow regenerating normal haematopoietic cells with evidence of peripheral recovery with no (or only a few regenerating) circulating blasts and with a decrease of at least 50% in the percentage of blasts in the bone marrow aspirate with the total marrow blasts between 5% and 25%. A value of less or equal than 5% blasts is also considered a PR if Auer rods are present
Best response	Best response was defined as the best measured response to treatment for all post-baseline visits (in the order of CR, CRp, CRi, PR, NR and not evaluable). Patients with best responses of CR, CRp, CRi or PR were considered responders. Patients who did not achieve at least a best response of PR were considered non-responders
Duration of remission (DoR)	Time from achieving remission to relapse. Duration of remission included duration of CRc, duration of CR/CRh, duration of CRh, duration of CR and duration of response (CRc + PR)

Definition	Description
Overall survival (OS)	OS was defined as the time from the date of randomisation until the date of death from any cause (death date – randomisation date + 1). For a patient who was not known to have died by the end of study follow-up, OS was censored at the date of last contact (date of last contact – randomised date + 1). The date of last contact was the latest date that the patient was known to be alive by the cut-off date. The last contact date was derived for patients alive at the analysis cut-off date. Patients with a last contact date beyond the analysis cut-off date were censored at the analysis cut-off date
Leukaemia-free survival (LFS)	LFS was defined as the time from the date of first CRc until the date of documented relapse (excluding relapse from PR) or death for patients who achieved CRc (relapse date or death date – first CRc disease assessment date + 1). For a patient who was not known to have relapsed or died, LFS was censored on the date of last relapse-free disease assessment date (last relapse-free disease assessment date – first CRc disease assessment date + 1)
Event-free survival (EFS)	EFS was defined as the time from the date of randomisation until the date of documented relapse (excluding relapse after PR), treatment failure or death from any cause within 30 days after the last dose of study drug, whichever occurred first (earliest of [relapse date, treatment failure date, death date] – randomisation date + 1). If a patient experienced relapse or death within 30 days after the last dose of study drug, the patient was defined as having an EFS event related to either "relapse" or "death", and the event date was the date of relapse or death
Transfusion conversion rate; transfusion maintenance rate	Transfusion conversion rate and transfusion maintenance rate were only defined for the patients in the gilteritinib arm. For the purpose of defining transfusion conversion rate and transfusion maintenance rate, transfusion status (independent vs. dependent) during the baseline period and during the post-baseline period was defined as follows for patients who took at least 1 dose of study drug: Baseline transfusion status: • The baseline period was defined as the period from 28 days prior to the first dose to 28 days after the first dose. For patients who were on treatment < 28 days, the baseline period was from 28 days prior to the first dose until the end of treatment. • Patients were classified as baseline transfusion independent if there were no RBC or platelet transfusions within the baseline period; otherwise, the patient was baseline transfusion dependent. Post-baseline transfusion status: • The post-baseline period was defined as the period from 29 days after the first dose until the last dose. • For patients who were on treatment ≥ 84 days, they were classified as post-baseline transfusion independent if there was 1 consecutive period of 56 days without any RBC or platelet transfusion within the post-baseline period.

Definition	Description
	• For patients who were on treatment > 28 days but < 84 days, if there was no RBC or platelet transfusion within the post-baseline period, then post-baseline transfusion status was not evaluable. • For patients who were on treatment ≤ 28 days, post-baseline transfusion status was not evaluable. • Otherwise, the patient was considered post-baseline transfusion dependent. Both transfusion conversion rate and maintenance rate were defined for patients who had evaluable post-baseline transfusion status. Transfusion conversion rate was defined as the number of patients who were transfusion dependent during the baseline period but became transfusion independent during the post-baseline period divided by the total number of patients who were transfusion dependent during the baseline period. Transfusion maintenance rate was defined as the number of patients who were transfusion independent during the baseline period and still maintained transfusion independence during the post-baseline period divided by the total number of patients who were transfusion independent during the baseline period
Transplantation rate	The transplantation rate was defined as the percentage of patients who underwent HSCT during the study period
Brief Fatigue Inventory (BFI) patient-reported fatigue.[20]	The BFI was developed to assess the severity of fatigue and the impact of fatigue on daily functioning in patients with fatigue due to cancer and cancer treatment. The BFI short form has 9 items and a 24-hour recall. A global fatigue score is computed by averaging the 9 items. The BFI was administered at site visits directly to the patients via an electronic PRO device. A higher BFI fatigue score indicates a more unfavorable outcome

ANC - absolute neutrophil count; HSCT - haematopoietic stem cell transplant; PRO - patient reported outcome; RBC - red blood cell.

Additional outcomes from the ADMIRAL trial, as exploratory objectives, were the Functional Assessment of Chronic Illness Therapy–Dyspnea-Short Form (FACIT-Dys-SF), the Cancer Therapy-Leukemia (FACT-Leu) and EuroQol 5-Dimensions 5-Level instrument (EQ-5D-5L), and resource utilisation including hospitalisation, blood transfusion and intravenous antibiotic use.[13] [redacted]

In the ADMIRAL trial, the majority of salvage chemotherapy patients finished the study by cycle 2 of treatment. This led to high levels of censoring for the DoR and LFS endpoints.

Baseline characteristics of patients in CHRYSALIS were provided in Table 8 of the CS.[1] In the ITT population, 70% of patients had ≥2 prior AML therapies, 29% had a prior SCT, and 25% had prior tyrosine kinase inhibitor treatment.[11] Across the study, 194 patients had a locally confirmed FLT3 mutation (ITD, n=159; D835, n=13; ITD-D835, n=16; other, n=6),[11] of whom 191 patients were included in the FAS.[1, 12]

Table 9 of the CS[1] presents the baseline characteristics of patients in ADMIRAL (see ERG Appendix 1).

The mean patient age was [redacted] in the gilteritinib group and [redacted] years in the salvage chemotherapy group.

Within the gilteritinib group, [redacted] of patients tested positive for FLT3 by central testing by LeukoStrat CDx FLT3 Mutation Assay, of which the majority (87%) had FLT3-ITD alone (CS Table 9).[1] In the salvage chemotherapy group, [redacted] of patients tested positive and 91.1% had FLT3-ITD alone [redacted] patients tested negative for FLT3 by central testing (CS Table 9).[1] [redacted]

All patients had prior chemotherapy for AML, and prior FLT3 inhibitor was received by 13% of the gilteritinib group and 11.3% of the salvage chemotherapy group. There had been relapse following HSCT for 19.4% of the gilteritinib group and 21.0% of the salvage chemotherapy group. 39.7% of the gilteritinib group and 38.7% of the salvage chemotherapy group had primary refractory disease without HSCT.[1]

At the time of the final analysis of CHRYSALIS, 88% of patients had discontinued treatment.[12] Reasons for discontinuation included progressive disease (n=75), lack of efficacy (n=44), adverse events (AEs,

n=34), and death (n=29).[11] Twelve percent of patients remained on treatment, with a median treatment duration of 25.9 weeks (quartiles 15, 50 weeks).[1, 12]

In ADMIRAL, 247 patients were randomised to gilteritinib, of whom 246 received gilteritinib. [redacted] completed gilteritinib treatment, whilst [redacted] patients discontinued gilteritinib treatment.[8] One hundred and twenty four patients were randomised to salvage chemotherapy, of whom 109 received salvage chemotherapy.[8] [redacted] patients completed study chemotherapy, defined as *"patients on high dose chemotherapy who either completed 1 cycle of treatment with a CRc and were taken off treatment, or completed 2 cycles of treatment."*[7, 8] Primary reasons for treatment discontinuation are given in Table 7. Overall treatment-emergent adverse events (TEAEs) leading to withdrawal of treatment rates were [redacted] in the gilteritinib arm, and [redacted] in the salvage chemotherapy arm.[1]

Table 7: Treatment discontinuations (adapted from CS Appendix D and ADMIRAL CSR)

Treatment discontinuation at final analysis cut-off	**Gilteritinib (N=247) N (%)**	**Salvage chemotherapy (N=124) N(%)**
	[redacted]	[redacted]
Continued gilteritinib at final analysis cut-off	[redacted]	[redacted]
Completed study chemotherapy	[redacted]	[redacted]
Primary end of treatment reason		
Progressive disease	[redacted]	[redacted]
Lack of efficacy	[redacted]	[redacted]
Death	[redacted]	[redacted]
Disease relapse	[redacted]	[redacted]
Adverse event	[redacted]	[redacted]
Withdrawal by patient	[redacted]	[redacted]
Physician decision	[redacted]	[redacted]
Other	[redacted]	[redacted]
Protocol deviation	[redacted]	[redacted]

N – number; N/a – not applicable

The results of the CHRYSALIS dose-finding trial are summarised in Section 4.2.2. The results of the ADMIRAL trial, which provides the key effectiveness evidence for gilteritinib, are reported in Section 4.2.3.

4.2.2 CHRYSALIS - effectiveness results

The CHRYSALIS study included arms with doses ranging from 20mg to 450mg, and included patients with or without the FLT3 mutation. Results reported below are from the FLT3 mutation-positive patients who received doses of ≥80mg/day (see Table 8).

Eighteen of 169 patients (11%) achieved CR.[1, 12]

Median OS was 31 weeks (95% confidence interval (CI) 24 to 59 weeks).[1] Thirty two out of 169 patients (19%) had HSCT following an overall response.[12] Median OS in these 32 patients was 47 weeks (95% CI 32 to 61 weeks).[12] For 61 patients who achieved an overall response but did not undergo HSCT, median OS was 42 weeks (95% CI 31 to 48 weeks).[12]

Table 8: CHRYSALIS outcomes (adapted from CS, Table 10)

Outcomes	**FLT3 mutation positive, ≥80mg gilteritinib/day n=169**
Response, %, (95% CI)	
CR	11% (6% to 16%)
CR_p	6% (3% to 11%)
CR_i	24% (18% to 31%)
CR_c	41% (33% to 49%)
Median DoR, weeks (95% CI)	20 (14 to 33) Range 1.1 to 55 weeks[11]
Median OS, weeks (95% CI)	31 (24 to 59) Range 1.7 to 61 weeks[11]

CI - confidence interval; CR - complete remission; CR_c - composite complete remission; CR_i - complete remission with incomplete hematologic recovery; CR_p - complete remission with incomplete platelet recovery; DoR - duration of remission; FLT3 - FMS-like tyrosine kinase 3; OS - overall survival

4.2.3 ADMIRAL – effectiveness results

Outcome definitions are summarised in Table 6. Time-to-event outcomes in ADMIRAL are presented in Table 9. Subgroup analysis of OS is presented in Table 10. A Kaplan-Meier plot of OS for the ITT population is shown in Figure 2. OS was statistically significantly longer for patients randomised to gilteritinib (median 9.3 months) than patients randomised to salvage chemotherapy (median 5.6 months), hazard ratio (HR) 0.637 (95% CI 0.490 to 0.830; p=0.0004).[1, 7] The OS rate was higher for gilteritinib than for salvage chemotherapy at 6 months (65.5% versus 48.9%, respectively), and also at 12 months (37.1% and 16.7% respectively; p-values not reported).[4]

Randomisation was stratified by response to first-line therapy and pre-selected salvage chemotherapy. The CS[1] presents several subgroup analyses according to an approach commonly performed when reporting subgroup analyses in the literature and submissions to NICE. However, there are four problems with subgroup analyses as commonly applied: (i) multiplicity; (ii) not borrowing information from the overall effect; (iii) the use of improper groups by categorising continuous variables, and (iv) subgroup effects being explained by an ignored covariate. The ERG's clarification letter[21] (questions A6 and A18) asked the company for additional analyses to address these limitations. Firstly, a fundamental principle of the analysis of a clinical trial is to analyse the data according to the way the treatments were randomised and to account for known prognostic factors in an analysis, irrespective of

whether they are statistically significant. Investigating whether further covariates are prognostic factors or treatment effect modifiers is much more exploratory and involves model building based on measured baseline covariates. The ERG asked the company to compare a reduced model that included all stratification variables and known prognostic factors irrespective of whether they are statistically significant with a full model including any potential prognostic factors and treatment effect modifiers that were measured at baseline and without categorising continuous variables or assuming that any relationship between a continuous variable and outcome is linear. The ERG asked for consideration to be given to interactions between different variables and between variables and treatment, the latter being an assessment of whether a variable is a treatment effect modifier. However, the company did not provide these analyses.

The results of the company's subgroup analysis for OS according to response to first-line therapy and pre-selected salvage chemotherapy are shown in Table 10 (adapted from CS Appendix E,[8] Table 25). For both high and low intensity salvage chemotherapy, median OS was [redacted] for gilteritinib than for salvage chemotherapy: gilteritinib versus high intensity chemotherapy HR= [redacted] (95% CI, [redacted] gilteritinib versus low intensity chemotherapy HR= [redacted] (95% CI [redacted]

For primary refractory without HSCT patients, the difference between groups was [redacted] (HR [redacted] 95% CI [redacted] For the subgroup of patients with relapse more than 6 months after HSCT, there was [redacted] however, there were only [redacted] patients in this subgroup. [redacted]

Other subgroups tested (age, sex, race, ECOG, region, FLT3 mutation type, prior FLT3 inhibitor, cytogenic risk status) which had not been stratified at randomisation, had [redacted] for gilteritinib than for salvage chemotherapy, [redacted] The HR for unfavourable cytogenetic risk status was [redacted] (95% CI [redacted] however, this subgroup included only [redacted] patients [redacted] For patients with no prior FLT3 inhibitor, gilteritinib was [redacted] (HR [redacted] 95% CI [redacted] For the [redacted] patients with prior use of a FLT3 inhibitor, the treatment difference was [redacted] (HR [redacted] 95% CI [redacted] Section B.2.7 of the CS[1] points out that for racial subgroups, the most pronounced survival advantage was seen in Asian patients (HR= [redacted], 95% CI [redacted]

Median EFS was 2.8 months in the gilteritinib group, and 0.7 months in the salvage chemotherapy arm; this difference was not statistically significant [redacted] - see Table 9).[1, 7, 8]

Section B.2.6 of the CS[1] reports a modified analysis of EFS that included the salvage chemotherapy patients through subsequent AML therapies (*"events at initiation of new anti-leukaemic treatments reported in long term follow-up"*). This modification of EFS showed a statistically significant benefit for gilteritinib (2.3 months versus 0.7 months: HR=0.499, 95% CI 0.387 to 0.643; $p<0.0001$).[1]

LFS was measured from time of first CRc [redacted]; see Table 9).[7] The ERG notes that assessment of treatment effects measured from any timepoint other than randomisation is likely to be subject to bias. Median time to CRc was [redacted] months for the gilteritinib group and [redacted] months for the salvage chemotherapy group (CS,[1] Table 11). Median LFS was [redacted] months for patients randomised to gilteritinib, and [redacted] months) for patients randomised to salvage chemotherapy. Comparative statistics were reported as unreliable [redacted]

Table 9: ADMIRAL Survival outcomes

Outcomes	**Gilteritinib (N=247)**	**Salvage chemotherapy (N=124)**
Deaths, n (%)	[redacted]	[redacted]
Overall survival, median months (95% CI)	9.3 (7.7 to 10.7)	5.6 (4.7 to 7.3)
HR (95% CI) *p*-value 1-sided	0.637 (0.490 to 0.830) *p*=0.0004	
Overall survival rate % (95% CI)		
6 months	65.5 (59.2 to 71.1)	48.9 (39.3 to 57.8)
12 months	37.1 (30.7 to 43.6)	16.7 (9.9 to 25.0)
24 months	19.0 (12.8 to 26.0)	13.8 (7.5 to 22.0)
Duration of EFS, median months (95% CI)	2.8 [redacted]	0.7 [redacted]
Duration of LFS, median months (95% CI)	[redacted]	[redacted]

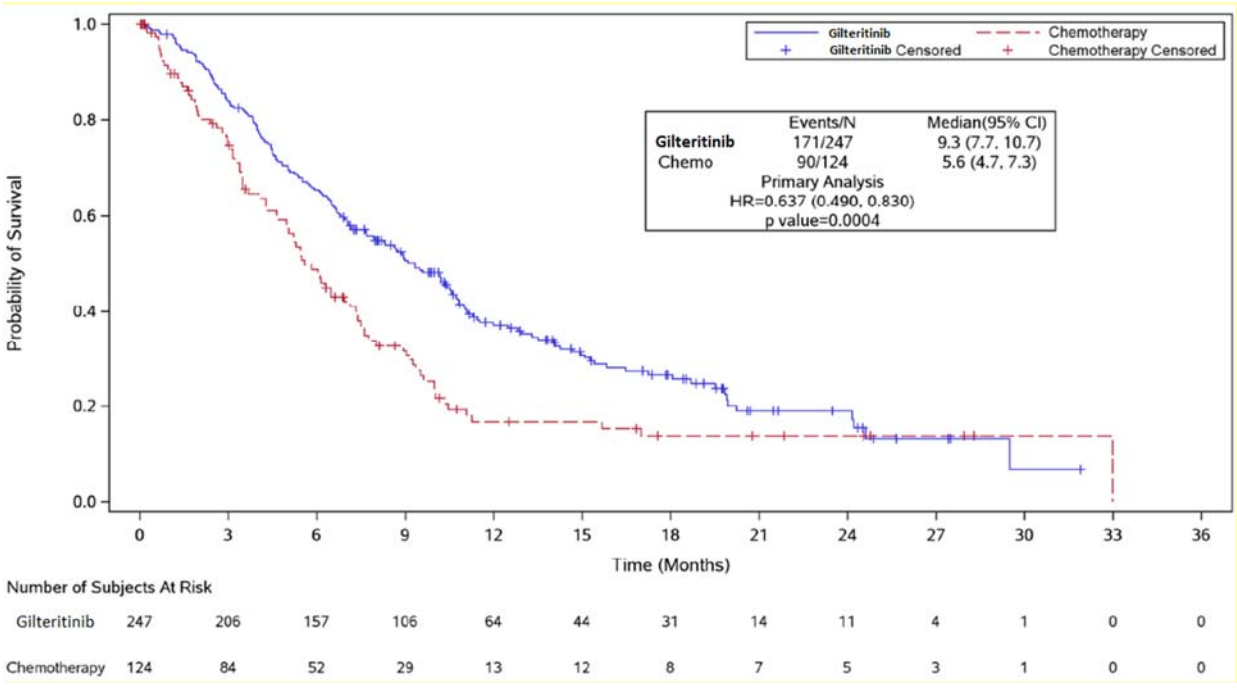
CI - confidence interval; EFS - event-free survival; LFS - leukaemia-free survival; NE - not evaluable
Based on information reported in CS Table 11, CS Appendix C Table 2 and ADMIRAL CSR

Table 10: Overall survival subgroup analysis

	Gilteritinib	Salvage chemotherapy	HR	*p*-value
Response to first-line therapy				
n/N (%) [Median months]				
Primary refractory without HSCT	[redacted]	[redacted]	[redacted]	[redacted]
Relapse within 6 months after CRc and no HSCT	[redacted]	[redacted]	[redacted]	[redacted]
Relapse after 6 months after CRc and no HSCT	[redacted]	[redacted]	[redacted]	[redacted]
Relapse within 6 months after allogeneic HSCT	[redacted]	[redacted]	[redacted]	[redacted]
Relapse after 6 months after allogeneic HSCT	[redacted]	[redacted]	[redacted]	[redacted]
Preselected salvage chemotherapy				
n/N (%) [Median months]				
High intensity	[redacted]	[redacted]	[redacted]	[redacted]
Low intensity	[redacted]	[redacted]	[redacted]	[redacted]

CR_c - composite complete remission; HSCT - haematopoietic stem cell transplant
Based on information reported in CS Appendix E Table 25 and ADMIRAL CSR Table 16

Figure 2: Kaplan-Meier plot of overall survival by treatment arm in ADMIRAL (reproduced from CS, Figure 4)



The CR rate was 21.1% for gilteritinib and 10.5% for the salvage chemotherapy group (see Table 11).[7, 8] As EFS was not statistically significant, and given the pre-planned hierarchical testing method, statistical significance of CR rate was not assessed.[7]

The CR/CR_h rate was statistically significantly higher in the gilteritinib group than in the salvage chemotherapy group (34.0% versus 15.3%; p=0.0001).[1, 8] DoR could not be reliably estimated for the salvage chemotherapy arm due to most patients finishing the study by cycle 2, and duration of CR for gilteritinib could not be reliably estimated as median duration was not reached for the gilteritinib arm.[1, 8]

For CR/CR_h, the median time to remission was [redacted] months [redacted] for the gilteritinib group, and [redacted] months [redacted] for the salvage chemotherapy group.[1] The ERG's clinical advisor commented that the difference in time to remission is due to the differences in the mechanism of action of the treatments – if FLAG-IDA and MEC produce a response, they will do so quickly (i.e. in the first cycle).

Table 11: Response outcomes (adapted from CS, Table 11)

Outcomes	Gilteritinib (N=247)	Salvage chemotherapy (N=124)
Patients achieving CR/CR_h (%)	84 (34.0%)	19 (15.3%)
Best response rate, n (%)		
CR	52 (21.1)	13 (10.5)
CR_p	19 (7.7)	0
CR_i	63 (25.5)	14 (11.3)
CR_c	134 (54.3)	27 (21.8)
CR_h	32 (13.0)	6 (4.8)
Duration of remission, median months (95% CI)		
CR	[redacted]	[redacted]
CR_c	[redacted]	[redacted]
CR_h	[redacted]	[redacted]
CR/CR_h	[redacted]	[redacted]
Time to remission, median months (95% CI)		
CR	[redacted]	[redacted]
CR_c	[redacted]	[redacted]
CR_h	[redacted]	[redacted]
CR/CR_h	[redacted]	[redacted]

CI - confidence interval; CR - complete remission; CR_c - composite complete remission; CR_i - complete remission with incomplete haematologic recovery; CR_h - complete remission with partial haematological recovery; CR_p - complete remission with incomplete platelet recovery; NE - not evaluable

The proportion of patients receiving HSCT during the study period was 25.5% for the gilteritinib group and 15.3% for the salvage chemotherapy group (treatment difference 10.2%, 95% CI 1.2% to 19.1%; p=0.0333, see Table 12).[1, 7] In the gilteritinib group, 197 patients were dependent on transfusion at baseline, and 34.5% of these patients became independent of transfusion during treatment.[1]

Table 12: Other outcomes (adapted from CS, Tables 11 and 12)

Outcomes	Gilteritinib (N=247)	Salvage chemotherapy (N=124)
Transplantation rate, n (%)	63 (25.5%)	19 (15.3%)
Transfusion conversion rate*, n/N (%)	68/197 (34.5%)	Not reported
Transfusion maintenance rate**, n/N (%)	[redacted]	Not reported
Patients with post-baseline evaluable transfusion status transfusion conversion rate*, n/N (%)	[redacted]	[redacted]
Patients with post-baseline evaluable transfusion status transfusion maintenance rate**, n/N (%)	[redacted]	[redacted]

**Transfusion conversion rate: The number of subjects who were transfusion dependent at baseline period but become transfusion independent at post-baseline period divided by the total number of subjects who are transfusion dependent at baseline period.*
***Transfusion maintenance rate: The number of subjects who were transfusion independent at baseline period and still maintain transfusion independent at post-baseline period divided by the total number of subjects who are transfusion independent at baseline period.*

HRQoL outcomes reported in ADMIRAL

The completion rate for HRQoL instruments was low in the salvage chemotherapy arm, dropping below 30% at cycle 2.[22] For the gilteritinib arm, there was a completion rate of >80% for the first 14 cycles.[22]

The Brief Fatigue Inventory[20] (BFI) is a 9-item, 11-point (0-10) rating scale, with higher scores indicating more unfavourable outcomes.[7, 8] At baseline, mean BFI score was 3.0 for the gilteritinib group (n=227) and 2.7 for the salvage chemotherapy group (n=97).[7]

Changes in BFI [redacted] (see Table 13).[7] At cycle 1 day 8, the mean change in BFI was [redacted] for the gilteritinib group and [redacted] for the salvage chemotherapy group. At cycle 2 day 1, mean change in BFI was [redacted], respectively.[1]

Table 13: Change from baseline brief fatigue inventory (BFI) Global Fatigue Score (adapted from CS Table 11 and ADMIRAL PRO report)

Outcomes	Gilteritinib	Salvage chemotherapy
Cycle 1 day 8, Mean (SD)	[redacted]	[redacted]
Cycle 2 day 1, Mean (SD)	[redacted]	[redacted]

SD - standard deviation.

The FACIT-Dys-SF assesses dyspnea severity and related functional limitations, with higher scores indicating more unfavourable outcomes (ADMIRAL CSR).[7] [redacted].[7] At cycle 2 day 1, the mean change in the dyspnea subscale was [redacted] for gilteritinib patients (n=115), and [redacted] for salvage chemotherapy patients (n=11). Mean changes in the functional limitation subscale were [redacted] and [redacted], respectively.[7]

The FACT-Leu measures leukaemia-specific signs, symptoms and the impact of AML on patients. For the FACT-Leu total score, at cycle 2 day 1, the median change from baseline was -0.1 for the gilteritinib arm (n=198), and 9.0 for the salvage chemotherapy arm (n=15).[7] Two additional questionnaires of leukaemia-specific signs were administered; dizziness; and mouth sores.[7] [redacted]

The EQ-5D visual analogue scale (VAS) was measured, with a scale of 0-100 with higher scores being more favourable. At cycle 2 day 1, change from baseline EQ-5D VAS was small and not clinically meaningful (see Table 14 and company's clarification response,[7] question A14). The mean change from baseline was [redacted] for gilteritinib patients (n=193), and [redacted] for salvage chemotherapy patients (n=15). [redacted] The company's clarification response[7] (question A14) reports that clinically meaningful changes from baseline [redacted] were observed at cycle 24 (least-squares [LS] mean: 15.022, 95% CI: 3.231 to 26.814) and cycle 27 (LS mean: 13.360, 95% CI: -1.054 to 27.773).[7]

Table 14: EQ-5D VAS score (adapted from ADMIRAL CSR, Table 37)

	Gilteritinib	**Salvage chemotherapy**
Baseline n	N= 218	N=96
Mean (SD)	[redacted] (24.6)	[redacted] (24.2)
Median (min, max)	[redacted]	[redacted]
Change from baseline at cycle 2 day 1	N=193	N=15
Mean (SD)	[redacted] (22.4)	[redacted] (32.4)
Median (min, max)	0.0 (-71, 69)	-3.0 (-41, 89)

The ADMIRAL patient-reported outcomes (PRO) report[22] summarises HRQoL measures at the end of treatment. [redacted]

4.2.4 Adverse events

The draft SmPC for gilteritinib[8] provides adverse reactions by frequency categories; these data are reproduced in Table 15. Data were based on patients from three clinical trials: ADMIRAL, CHRYSALIS and 2215-CL-0102 (NCT02181660 - a phase I dose-finding study in 24 Japanese patients with R/R AML). Whilst ADMIRAL was restricted to FLT3+ R/R AML, the other two studies were not.

The draft SmPC[8] provides the following special warnings and precautions for use: posterior reversible encephalopathy syndrome; prolonged cardiac ventricular repolarisation (QT interval) advising that hypokalaemia or hypomagnesaemia should be corrected prior to gilteritinib treatment; pancreatitis, although association with gilteritinib not confirmed; potential risk to a foetus and breast-feeding not recommended.

Table 15: Frequency of adverse events (reproduced from CS Appendix C1.1, Table 1)

Adverse drug reaction	**Gilteritinib 120mg daily (N=319)**		
	All grades (%)	**Grades 3/4 (%)**	**Frequency category**
Cardiac disorders			
Pericardial effusion	4.1	0.9	Common
Pericarditis	1.6	0	Common
Cardiac failure	1.3	1.3	Common
Gastrointestinal disorders			
Diarrhoea	35.1	4.1	Very common
Nausea	29.8	1.9	Very common
Constipation	28.2	0.6	Very common
General disorders and administration site conditions			
Fatigue	30.4	3.1	Very common
Peripheral oedema	24.1	0.3	Very common
Asthenia	13.8	2.5	Very common
Malaise	4.4	0	Common
Immune system disorders			
Anaphylactic reaction	1.3	1.3	Common
Investigations			
Blood creatine phosphokinase increased*	93.4	3.1	Very common
Alanine aminotransferase increased*	82.1	12.9	Very common
Aspartate aminotransferase increased*	80.6	10.3	Very common
Blood alkaline phosphatase increased*	68.7	1.6	Very common
Electrocardiogram QT prolonged	8.8	2.5	Common
Musculoskeletal and connective tissue disorders			
Pain in extremity	14.7	0.6	Very common
Arthralgia	12.5	1.3	Very common
Myalgia	12.5	0.3	Very common
Musculoskeletal pain	4.1	0.3	Common
Nervous system disorders			
Dizziness	20.4	0.3	Very common
Posterior reversible encephalopathy syndrome	0.6	0.6	Uncommon
Respiratory, thoracic and mediastinal disorders			
Cough	28.2	0.3	Very common
Dyspnea	24.1	4.4	Very common
Vascular disorders			
Hypotension	17.2	7.2	Very common

Very common (≥1/10); common (≥1/100 to <1/10); uncommon (≥1/1,000 to <1/100). Preferred term in MedDRA (v. 19.1).
**Frequency is based on central laboratory values. Median duration of exposure to XOSPATA™ was 111 days (range 4 to 1320 days).*

Adverse events reported in CHRYSALIS

AEs in CHRYSALIS are reported in CS Appendix F.[8] CHRYSALIS included patients without the FLT3 mutation, and doses of gilteritinib were not restricted to licensed doses (see ERG Appendix 1).

In CHRYSALIS, seven deaths were considered to be possibly related to gilteritinib treatment: pulmonary embolism; respiratory failure; haemoptysis; intracranial haemorrhage; ventricular fibrillation; septic shock; and neutropenia.[12]

Serious adverse events (SAEs) that occurred in 5% or more of patients (whether or not judged as caused by the treatment) were: febrile neutropenia (31%); progressive disease (17%); sepsis (14%); pneumonia (11%); acute renal failure (10%); pyrexia (8%); bacteraemia (6%); and respiratory failure (6%).[12]

The most common AEs of any grade were: diarrhoea (16%) and fatigue (15%).[11] Eleven of 252 patients (4.4%) had a maximum post-baseline QT interval corrected by Fredericia's formula (QTcF interval) >500 msec.[11]

Adverse events reported in ADMIRAL

The SAF of ADMIRAL included all patients who received at least one dose of study treatment.[1] An AE was classed as "serious", by the investigator or sponsor, if it was life threatening, required hospitalisation, prolonged a period of hospitalisation, resulted in death, persistent/significant disability, or birth defect, or was an "other medically important" event (ADMIRAL CSR).[7] Grades of events were classed according to the National Cancer Institute - Common Terminology Criteria for Adverse Events (NCI-CTCAE) guidelines (version 4.03).[23] AEs are summarised in Table 16; AEs occurring in 10% or more of patients are shown in ERG Appendix 1 (based on data from CS Appendix F[8]).

[redacted] in the gilteritinib group and [redacted] of patients in the salvage chemotherapy group experienced an AE during the study: these were considered to be drug-related in [redacted] of gilteritinib patients and [redacted] of salvage chemotherapy patients. NCI-CTCAE Grade 3 or higher AEs were experienced by [redacted] of the gilteritinib group and [redacted] of the salvage chemotherapy group: these were considered drug-related in [redacted] of patients, respectively.

[redacted] of gilteritinib patients and [redacted] of salvage chemotherapy patients experienced a drug-related TEAE leading to death.[1] [redacted]

Within both groups, more than 20% patients experienced Grade 3 or higher AEs of febrile neutropenia (gilteritinib 45.9%; salvage chemotherapy 36.7%), anaemia (40.7% versus 30.3%, respectively), and

platelet count decreased (22.0% versus 24.8%, respectively). For the gilteritinib group, 22.8% of patients experienced thrombocytopenia. The AEs of Grade 3 or higher ALT increased and AST increased occurred more frequently in the gilteritinib group than the salvage chemotherapy group (ALT increased 13.8% versus 4.6%, respectively; AST increases 14.6% versus 1.8%, respectively).[8]

The most common AEs of any grade occurring in the gilteritinib group were: anaemia (47.2%); febrile neutropenia (46.7%); pyrexia (42.7%); ALT increased (41.9%); AST increased (40.2%); [redacted]

AEs occurring more frequently in the gilteritinib arm compared with the salvage chemotherapy arm were ALT increased (41.9% versus 9.2%); AST (40.2% versus, 11.9%); [redacted]

Table 16: Adverse events from ADMIRAL (adapted from CS, Table 13)

AEs	Gilteritinib (N=246)	Salvage chemotherapy (N=109)
	n (%) patients	n (%) patients
TEAE	[redacted]	[redacted]
Drug-related TEAE	[redacted]	[redacted]
Serious TEAE	[redacted]	[redacted]
Drug-related serious TEAE	[redacted]	[redacted]
TEAE leading to death	[redacted]	[redacted]
Drug-related TEAE leading to death	[redacted]	[redacted]
TEAE leading to withdrawal of treatment	[redacted]	[redacted]
Drug-related TEAE leading to withdrawal of treatment	[redacted]	[redacted]
NCI-CTCAE Grade 3 or higher TEAE	[redacted]	[redacted]
Drug-related Grade 3 or higher TEAE	[redacted]	[redacted]
Death	[redacted]	[redacted]

NCI-CTCAE - National Cancer Institute common terminology criteria for adverse events; TEAE - treatment emergent adverse events

4.3 Trials identified for potential indirect comparison

Whilst CHRYSALIS was a dose-finding study, the ADMIRAL RCT did have potential to be used in an indirect comparison. One study (QuANTUM-R[24]) was identified with the potential to provide an

indirect comparison with ADMIRAL (company's SLR report,[9] Section 4.4). However, this study compared quizartinib with salvage chemotherapy, and quizartinib was excluded from the final NICE scope.[6]

In response to a request for clarification from the ERG[7] (question B2), the company provided reasons why the 36 studies of comparators could not contribute to a network meta-analysis (NMA). The table provided within the company's clarification response does not include the QuANTUM-R trial, and also excludes two studies of chemotherapy (Schnetzke 2014 and Takahashi 2011). These two studies were both retrospective observational studies. Ten studies were excluded for providing no results, 21 for being single-arm studies, and one for being a dose-finding study (see ERG Appendix 1).

4.4 Conclusions of the clinical effectiveness section

The ERG does not believe that any RCTs of gilteritinib which are relevant to the decision problem set out in the final NICE scope[6] have been missed by the CS. Two trials of gilteritinib were identified: CHRYSALIS - a dose-finding study, and ADMIRAL - an RCT which provides the key evidence for the clinical effectiveness of gilteritinib.

ADMIRAL was an open-label trial, but was of otherwise good methodological quality. Because the majority of salvage chemotherapy patients finished study treatment by cycle 2 of treatment, there was high censoring for DoR and LFS endpoints. There was a low rate of completion for HRQoL instruments in the salvage chemotherapy arm.

According to clinical advice received by the ERG, the demographics of the ADMIRAL trial population are sufficiently representative for the results to be applicable to the population of patients who would be considered eligible for gilteritinib treatment in England.

The search for clinical evidence reflected the decision problem in the final NICE scope;[6] however, the evidence for comparators presented in the CS was limited to those in the ADMIRAL trial (LoDAC, AZA, FLAG-IDA, MEC). None of the studies identified by the search for comparator evidence allowed for an indirect comparison with ADMIRAL.

In the ADMIRAL trial, median OS was 9.3 months in the gilteritinib arm and 5.6 months in the salvage chemotherapy arm. The difference in OS was statistically significantly longer for patients randomised to gilteritinib than patients randomised to salvage chemotherapy (HR 0.637, 95% CI 0.490 to 0.830; p=0.0004).

The rate of patients receiving HSCT during the study period was 25.5% for the gilteritinib group and 15.3% for the salvage chemotherapy group (treatment difference [redacted]).

The rate of CR/CRh was statistically significantly higher in the gilteritinib arm compared with the salvage chemotherapy arm (34.0% versus 15.3%; *p*=0.0001). The rate of patients receiving HSCT during the study period was 25.5% for the gilteritinib group and 15.3% for the salvage chemotherapy group (treatment difference [redacted]).

In ADMIRAL, [redacted] in the gilteritinib group and [redacted] of patients in the salvage chemotherapy group experienced an AE during the study; these were considered drug-related in [redacted] of gilteritinib patients and [redacted] of salvage chemotherapy patients. NCI-CTCAE Grade 3 or higher AEs were experienced by [redacted] of patients in the gilteritinib group and [redacted] of patients in the salvage chemotherapy group; these were considered to be drug-related in [redacted] of patients, respectively. The most common Grade 3 or higher AEs experienced in the gilteritinib group were febrile neutropenia ([redacted]), anaemia ([redacted]), thrombocytopenia ([redacted]) and platelet count decreased ([redacted]).

For HRQoL measures in ADMIRAL, [redacted].

5 COST EFFECTIVENESS

This chapter presents a summary and critique of the company's health economic analyses of gilteritinib for the treatment of adult patients with FLT3+ R/R AML. Section 5.1 presents a critique of the company's review of existing health economic analyses. Section 5.2 summarises the methods and results of the company's model. Sections 5.3 and 5.4 present a detailed critique of the model and additional exploratory analyses undertaken by the ERG, respectively. Section 5.5 presents a discussion of the available economic evidence.

5.1 Company's review of published cost-effectiveness studies

5.1.1 Summary and critique of the company's search strategy

The company performed two searches: (i) a search for economic analyses of gilteritinib and other therapies for the treatment of R/R AML and cost and healthcare resource use studies (CS Appendices G and I), and (ii) HRQoL studies in R/R AML (CS Appendix H).[8]

The company's review of economic studies and cost/resource use studies included searches in the following databases: MEDLINE [via Ovid]; MEDLINE In-Process [via Ovid]; Embase [via Ovid]; PsycINFO [via Ovid]; Cochrane Database of Systematic Reviews [via Wiley]; EconLit [via Ovid], and NHS EED [via Wiley] in October 2018. The company's HRQoL review involved searching the following databases: MEDLINE [via Ovid]; MEDLINE In-Process [via Ovid]; Embase [via Ovid]; PsycINFO [via Ovid]; CDSR [via Wiley]; CENTRAL [via Wiley]; DARE [via Wiley], and the HTA Accelerator. In addition, the company also searched for economic, cost and HRQoL studies in the following websites for the last three years: ASCO; ASH; EHA; ESMO, and ISPOR. The company also searched the Clinicaltrials.gov clinical trials registry. These searches were conducted in December 2018.

With respect to the company's economic and cost/resource use search, the ERG notes that since August 2018, the NHS EED database is no longer hosted in the Cochrane Library. Therefore, the ERG suggests that the NIHR CRD NHS EED database should have been searched. In addition, the ERG did not recognise the economic filter applied to the MEDLINE and Embase searches (CS Appendix I: 8.2.3 Economic search strategy statements #6-13). A comparison of the company's subject headings and free-text terms against existing economic evaluation search filters (e.g. the CADTH Economic Evaluations/Cost/Economic Models filter) suggests that the company's filter is less sensitive. The translation of the economic search filter from MEDLINE and Embase to the Cochane Library (CS Appendix I: 8.2.3 statements #13-18) was not consistent; this could impact on search recall. Given these limitations, the ERG cannot confirm if the economic review searches are comprehensive and sufficiently sensitive to retrieve all the eligible studies.

The ERG also did not recognise the company's HRQoL filter (CS Appendix I: 8.2.4 Utilities search strategy statements #7-8, SLR Appendix I: 8.2.5 HRQoL search strategy statement #7) applied to the search in MEDLINE and Embase. A comparison of the company's subject headings and free-text terms suggests less sensitivity compared with published health state utility search filters (e.g. Arber *et al*. 2017[25] and the ISSG[10]). The translation of the MEDLINE and Embase health utility filters to the Cochrane Library (CS Appendix I: 8.1.5 to 8.2.5 statement #7-8) was not consistent; this could impact search recall. Given the above limitations, the ERG cannot confirm if the company's HRQoL searches are comprehensive and sufficiently sensitive to retrieve all eligible studies.

5.1.2 Summary of company's review findings

The company's searches did not identify any economic analyses of gilteritinib for the treatment of FLT3+ R/R AML. CS Appendix I[8] provides further details regarding three economic analyses of treatments for AML which were identified from the company's searches: (1) a Canadian analysis of arsenic trioxide versus all-trans-retinoic acid with chemotherapy in patients with relapsed Acute Promyelocytic Leukaemia (APL - a rare sub-type of AML) (Lachaine *et al*[26]); (2) a UK-based economic analysis of histamine dihydrochloride plus low dose interleukin-2 versus standard care as maintenance therapy for patients with AML in first complete remission (CR1, Magar *et al*[27]), and (3) the health economic model developed to inform the NICE appraisal of midostaurin for the treatment of patients with FLT3+ treatment-naïve AML (Novartis[28]). None of these analyses are directly relevant to the current appraisal as they relate to patient populations and treatments which are not included in the final NICE scope.[6] The CS[1] (Table 18) and its appendices[8] include some discussion of key issues raised by the ERG and the Appraisal Committee during the midostaurin appraisal (relating to model structure, utility assumptions, cure modelling and associated mortality assumptions) and how these issues have been addressed within the company's model for this appraisal.

The company's review of HRQoL studies identified 16 studies reporting health utility values for patients with AML (company's SLR report,[9] Section 6). The results of this review are not presented in the CS;[1] however, the company notes that there is no published literature that reports utility estimates for patients with FLT3+ R/R AML. One study identified within the review (Joshi *et al*[29]) is used to inform health losses associated with HSCT; the other key AML-related utility parameters are based on the ADMIRAL trial.[7]

5.2 Description of company's health economic analysis

This section provides a detailed description of the methods and results of the company's *de novo* health economic analysis.

5.2.1 Model scope

As part of its submission to NICE,[1] the company submitted a fully executable health economic model programmed in Microsoft Excel®. The scope of the company's model is summarised in Table 17.

Table 17: Summary of company's model scope

Population	Adults with FLT3+ R/R AML
Time horizon	40 years
Intervention	Gilteritinib
Comparator	Blended comparator of salvage chemotherapy options used in the ADMIRAL trial: • Azacitidine (AZA) • Low-dose cytarabine (LoDAC) • Mitoxantrone, etoposide, cytarabine (MEC) • Fludarabine, cytarabine, G-CSF and idarubicin (FLAG-IDA)
Outcome	Incremental cost per QALY gained
Perspective	NHS and PSS
Discount rate	3.5% for health outcomes and costs
Price year	2018

AML - acute myeloid leukaemia; FLT3 - FMS-like tyrosine kinase 3; QALY - quality-adjusted life year; PSS - Personal Social Services

The company's economic analysis assesses the incremental cost-effectiveness of gilteritinib versus a blended comparator of alternative salvage chemotherapy regimens for the treatment of FLT3+ R/R AML. The economic analysis was undertaken from the perspective of the National Health Service (NHS) and Personal Social Services (PSS) over a 40-year (lifetime) horizon. Cost-effectiveness is expressed in terms of the incremental cost per quality-adjusted life year (QALY) gained. Unit costs are valued at 2018 prices. Health outcomes and costs are discounted at a rate of 3.5% per annum.

The blended comparator included in the company's economic analysis is comprised of four salvage chemotherapy regimens: (i) azacitidine; (ii) low-dose cytarabine (LoDAC); (iii) mitoxantrone, etoposide, cytarabine (MEC); and (iv) fludarabine, cytarabine, G-CSF and idarubicin (FLAG-IDA). This comparator, and the proportionate use of each regimen within the "blend", reflect the regimens used in the salvage chemotherapy arm of the ADMIRAL study.[7] This comparator is modelled using pooled salvage chemotherapy outcomes data from ADMIRAL and weighted regimen-specific costs. The CS[1] also reports the results of pairwise comparisons of gilteritinib versus each individual chemotherapy regimen (based on effectiveness and AE data for the whole ADMIRAL salvage chemotherapy group, but using regimen-specific costs), as well as a separate pairwise comparison of gilteritinib versus BSC. Whilst not entirely clear from the CS,[1] the company's clarification response[7] (question B3) confirms that the comparison of gilteritinib versus this blended comparator reflects the company's base case analysis.

Population

The modelled population relates to adult patients with R/R AML who have a FLT3 mutation. This is consistent with the ITT population of the ADMIRAL study,[7] the final NICE scope[6] and the anticipated marketing authorisation for gilteritinib.[1, 8] At model entry, patients are assumed to have a mean age of [redacted] years.[7] The economic analysis subdivides the patient population into two mutually exclusive groups: (i) patients who receive HSCT following the initiation of gilteritinib/salvage chemotherapy, and (ii) patients who do not receive HSCT. Based on data from ADMIRAL, the proportion of patients who receive HSCT is assumed to differ according to treatment group (proportion of patients receiving HSCT: gilteritinib=25.5% [63/247]; salvage chemotherapy=15.3% [19/124]; BSC=0% [assumed]). All economic analyses presented within the CS[1] relate to the overall FLT3+ R/R AML population; no subgroup analyses are presented.

Interventions

The intervention evaluated within the model is gilteritinib administered orally as a monotherapy.[1] Within the model, gilteritinib is assumed to be administered at a dose of 120mg daily (3 x 40mg tablets) during each 28-day dosing cycle. The model does not include an explicit treatment discontinuation rule and drug acquisition costs are calculated independently of patients' health state. Drug acquisition costs for gilteritinib over the patient's lifetime are based on a mean of [redacted] 28-day dosing cycles per patient, based on the final analysis of ADMIRAL (data cut-off 17th September 2018).[7] The model assumes that a proportion of patients who undergo HSCT will continue to receive gilteritinib as maintenance therapy after the transplant. According to the CS (page 49), [redacted] of patients were still receiving gilteritinib at the time of the final analysis.

Comparators

The CS[1] (page 15) notes that there is no standard of care for patients with FLT3+ R/R AML and highlights differences in treatment guidelines across England. The four salvage chemotherapy regimens included in the company's health economic analysis (azacitidine, FLAG IDA, MEC, and LoDAC) are assumed to be administered according to the dosing schedules listed in Table 18. These regimens are not fully consistent with the comparators listed in the final NICE scope,[6] which includes: (i) intermediate-dose cytarabine (IDAC); (ii) FLAG-IDA; (iii) BSC, and (iv) hydroxycarbamide (for people who cannot have chemotherapy or SCT). According to the CS (page 48), the individual regimens which make up the salvage chemotherapy blended comparator were selected for inclusion in the model on the basis of their relevance for the treatment of FLT3+ R/R AML and because these regimens were included in the ADMIRAL study,[7] thereby providing direct head-to-head evidence relative to gilteritinib. Issues relating to the relevance of the model comparators to the decision problem are discussed further in Section 5.3.

The weights applied to each salvage chemotherapy regimen in the blend are based on the proportions of patients in ADMIRAL who received each regimen (azacitidine - [redacted]; LDAC - [redacted], MEC - [redacted]; FLAG-IDA - [redacted]).[7] The model does not include a separate arm for the blended comparator; instead, lifetime costs for the blended comparator are estimated by applying these weights to the costs estimated for the individual regimens, each of which is represented using a separate model arm. Health outcomes for all chemotherapy options are based on data for the overall salvage chemotherapy arm of ADMIRAL.

Table 18: Summary of treatment options included in company's model

Regimen	Administration route	Dosing schedule (per 28-day dosing cycle)	No. of 28-day dosing cycles	Weights for blended comparator
Gilteritinib	Oral	120mg daily	[redacted]	-
Azacitidine	SC/IV*	75mg/m^2 daily, days 1-7	2.24	[redacted]
LoDAC	SC/IV*	Cytarabine: 20mg twice daily, days 1-10	1.68	[redacted]
MEC	IV	Mitoxantrone: 8mg/m^2 daily, days 1-5 Etoposide: 100mg/m^2 daily, days 1-5 Cytarabine: 1,000mg/m^2 daily, days 1-5	1.13	[redacted]
FLAG-IDA	SC/IV	G-CSF: 300µg/m^2 daily, days 1-5 Fludarabine: 30mg/m^2 daily, days 2-6 Cytarabine: 2,000mg/m^2 daily, days 2-6 Idarubicin: 10mg/m^2 daily, days 2-4	1.02	[redacted]
BSC	N/a	N/a	N/a	-

SC – subcutaneous, IV – intravenous; BSC – best supportive care; m – metre; mg – milligram; N/a - not applicable
** In clinical practice in England, azacitadine and LoDAC are given subcutaneously*

In addition, the company's model allows for pairwise economic comparisons between gilteritinib and each salvage chemotherapy regimen. All individual chemotherapy regimens are assumed to generate the same number of QALYs, but differ in terms of the drug acquisition, administration and hospitalisation costs applied in the event-free health state.

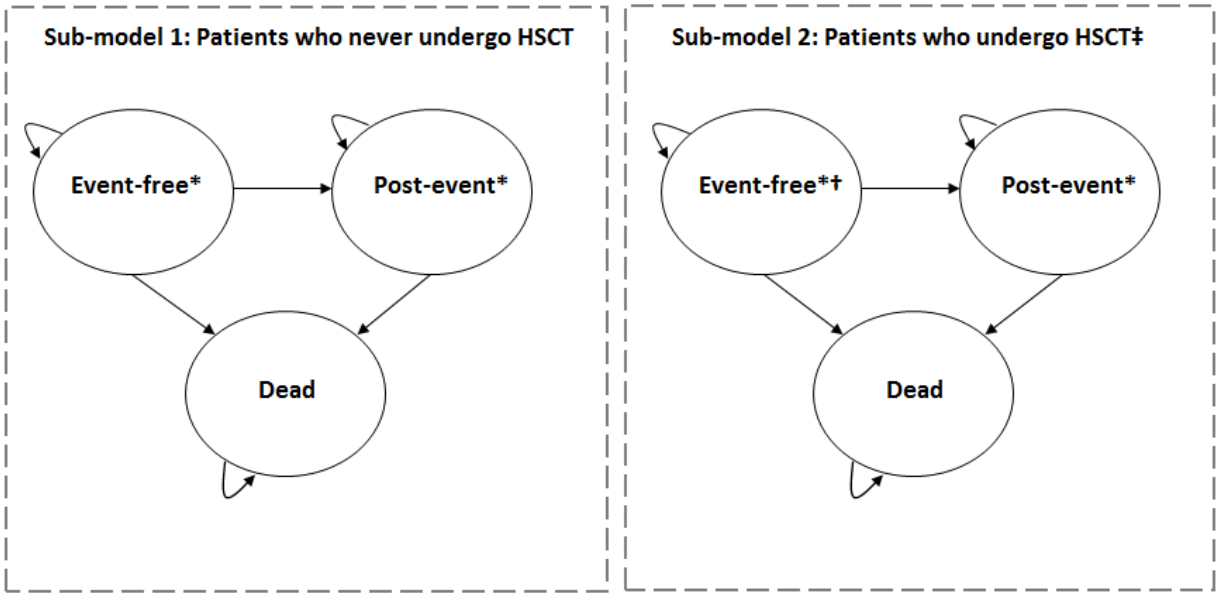
The model also includes BSC as an additional comparator. BSC is defined in the CS[1] (page 48) as *"supportive medication or procedures that do not include any active anti-leukemic treatments."* The CS states that this may include hydroxycarbamide (also known as hydroxyurea), blood transfusions, growth factors, and anti-infective treatments. As the ADMIRAL trial did not include a BSC arm, the effectiveness of BSC is estimated through an indirect comparison using data from ADMIRAL[7] and a retrospective cohort study of AML patients in first relapse (Sarkozy *et al*[30]). The CS[1] (page 65) notes that BSC is *"unlikely to be the comparator of greatest interest."* However, the ERG's clinical advisor noted that around 25-30% of the FLT3+ R/R AML population currently receive BSC, and that some patients who would otherwise elect not to undergo active treatment (for example, due to treatment burden associated with inpatient IV chemotherapy), may opt for treatment with gilteritinib due to its

oral administration route and toxicity profile. Despite the lack of direct head-to-head evidence, the ERG considers BSC to be a relevant comparator for a subset of the target population.

5.2.2 Model structure and logic

The CS[1] (page 49) describes the company's economic model structure as a combination of a decision-tree followed by a partitioned survival model. Within the company's model, patients are subdivided into two groups according to whether or not they receive HSCT after initiating treatment with gilteritinib or salvage chemotherapy. The ERG notes that membership of either group is not a baseline patient characteristic, but instead reflects what happened after randomisation in the ADMIRAL trial. Hereafter, these two modelled populations are referred to as the "With HSCT" and "No HSCT" groups. For each treatment option, separate partitioned survival sub-models are used to estimate health outcomes and costs for the With HSCT and No HSCT groups. Each sub-model includes the same three health states: (i) event-free; (ii) post-event and (iii) death (see Figure 1). Whilst each sub-model adopts the same general model structure, the proportions of patients entering into each are assumed to differ according to treatment group, thereby reflecting the increased HSCT rate in the gilteritinib group.

Figure 3: Company's model structure (drawn by the ERG)



HSCT - haematopoietic stem cell transplant
Notes: Patients receiving active treatment who do not experience early treatment failure (within the first 30 days) enter the model in the event-free states; BSC-treated patients enter in the No HSCT post-event state
** Patients who remain alive at 3-years are assumed to be "cured" (subsequent mortality risk is uplifted using an SMR of 2.0*
† HSCT is assumed to occur at a fixed timepoint (gilteritinib - [redacted]; salvage chemotherapy - [redacted]). All patients in the HSCT sub-models remain alive and event-free until this timepoint (i.e. OS = fixed time to transplant + post-transplant survival time)
‡ The proportion of patients entering the HSCT sub-model is dependent on treatment group

The model operates as follows. Upon entry into the model, a proportion of patients are allocated to the With HSCT sub-model based on HSCT rate of patients in ADMIRAL (proportion of patients receiving HSCT: gilteritinib 25.5%; salvage chemotherapy 15.3%; BSC 0% [assumed]). The remainder are allocated to the No HSCT sub-model for each treatment group.

For those patients in the No HSCT sub-model, EFS and OS are modelled using treatment group-specific parametric (log logistic) distributions fitted to time-to-event data relating to those patients in ADMIRAL who did not receive HSCT during the trial follow-up period.[7] For any time *t*, the probability of being alive and event-free is given by the cumulative probability of EFS, the probability of being alive is given by the cumulative probability of OS, and the probability of being alive in the post-event state is given by the cumulative probability of OS minus the cumulative probability of EFS.

All patients in the With HSCT sub-model are assumed to remain alive and event-free until a fixed timepoint when the transplant is assumed to occur (gilteritinib - [redacted]; salvage chemotherapy - [redacted], based on ADMIRAL[7]). Subsequently, OS is modelled using a parametric (Gompertz) distribution fitted to post-HSCT OS outcomes based on external data (Evers *et al*[31] – relating to patients with R/R AML in second complete remission [CR2]). This OS survivor function is applied to all patients in the salvage chemotherapy group who receive HSCT, and to the proportion of patients in the gilteritinib group who do not receive gilteritinib maintenance therapy after HSCT. The model assumes that [redacted] of patients who receive gilteritinib and undergo HSCT will subsequently continue to receive (and obtain additional benefit from) gilteritinib as maintenance therapy. For these patients, an HR is applied to the post-HSCT OS Gompertz survivor function, based on a naïve arm-based indirect comparison of OS data for patients receiving gilteritinib maintenance therapy in ADMIRAL[7] and post-HSCT OS data reported by Evers *et al*.[31] For all patients receiving HSCT, the hazard of EFS is assumed to be proportional to the hazard of OS, and is modelled using an HR derived from a comparison of LFS (used as a proxy for EFS) and OS reported within a separate retrospective observational study (Uston *et al*[32] – relating to patients with FLT3+ AML who underwent unrelated donor [URD] HSCT). The general approach used to estimate health state occupancy at each timepoint in the With HSCT sub-model is the same as that applied in the No HSCT sub-model.

The No HSCT and With HSCT sub-models are combined into a single model trace in the first cycle. This trace forms the basis for the QALY and cost calculations for each treatment group.

Within the BSC group, all patients are assumed to enter the model in the post-event state. Survival outcomes for these patients are modelled based on an indirect comparison between ADMIRAL[7] and a retrospective cohort study (Sarkozy *et al*[30]). An HR of 2.86 is applied to the gilteritinib No HSCT OS

survivor function (fitted to ADMIRAL data) in order to reflect a less favourable survival prognosis for patients receiving BSC.

The model assumes that all patients who remain alive at 36 months (3 years) are "cured", irrespective of whether they are event-free and irrespective of whether they have received HSCT. The subsequent survival prognosis for these patients is modelled using general population mortality rates[33] uplifted using a standardised mortality ratio (SMR) of 2.0.

HRQoL is determined by the patient's health state (event-free or post-event), whether they receive HSCT, and whether the cure point has been reached. During the first 36 months of the time horizon (before the cure point), health utilities are based on the results of a generalised estimating equation (GEE) model fitted to EQ-5D-5L data collected in ADMIRAL (mapped to the EQ-5D-3L using the algorithm reported by Van Hout *et al*[34]), adjusted for age using data reported by Janssen *et al.*[35] The ERG believes that following the cure point (from month 36 onwards), the company intended to assume that health utilities reflect those of the general population; however, the implementation of this assumption in the model is subject to a significant error (Section 5.3.3). The model also includes a once-only QALY loss associated with Grade 3/4 AEs (dependent on treatment group) which is applied in the first model cycle and a disutility associated with HSCT which is applied for 6 months after the procedure.

The model includes costs associated with: (i) drug acquisition and administration; (ii) the HSCT procedure (including stem cell harvesting); (iii) re-testing for the FLT3 mutation; (iv) disease management (health state costs); (v) treatments following disease relapse/progression; (vi) management of AEs, and (vii) end of life care costs.

Drug acquisition costs are modelled as a function of the number of dosing cycles received, the dose per cycle (adjusted by relative dose intensity [RDI]) and unit costs (including a PAS for gilteritinib). Drug administration costs are assumed to reflect the treatment schedules reported in Table 18. Lifetime drug acquisition and administration costs are applied as once-only costs in the first model cycle. Costs related to the management of AEs, FLT3 mutation re-testing and the HSCT procedure are also applied as once-only costs in the first model cycle. Disease management costs (medical visits, hospitalisations, diagnostic procedures, and blood transfusions) are assumed to be dependent on the patient's health state, whether they have had HSCT and the treatment received (hospitalisation costs are assumed to differ between gilteritinib and the individual salvage chemotherapy regimens). These disease management costs are applied in each model cycle until the cure point is reached, and are subsequently assumed to be zero. Costs associated with post-progression chemotherapy are applied as a lump sum cost to patients leaving the event-free state; these costs are applied during all model cycles (both before and after the

cure point, irrespective of whether the patient progressed/relapsed or died). End of life care costs are applied as a once-only cost at the point of death.

Costs and health outcomes are evaluated over a total of 480 monthly cycles (40 years). Half-cycle correction is applied to account for the timing of events.

Incremental cost-effectiveness is calculated in a pairwise fashion based on the difference in costs divided by the difference in QALYs for gilteritinib versus the salvage chemotherapy blended comparator. Secondary analyses are presented for comparisons against individual chemotherapy regimen and BSC.

5.2.3 Key assumptions employed in the company's model

The company's model employs the following key assumptions:

- Within the pairwise comparisons of gilteritinib versus individual chemotherapy regimens, each regimen is assumed to be equivalent in terms of safety and efficacy (including the likelihood of each regimen providing a "bridge" to HSCT), based on effectiveness outcomes and AEs for the salvage chemotherapy arm of ADMIRAL.
- OS for patients receiving BSC is modelled by applying an HR derived from a naïve indirect comparison of data from ADMIRAL and Sarkozy *et al*[30] to the gilteritinib OS function. Upon model entry, all BSC-treated patients are assumed to have immediate relapse/progression.
- Gilteritinib is assumed to increase the proportion of patients who undergo HSCT compared with salvage chemotherapy. Health outcomes, and to some degree costs, are modelled separately for patients who receive HSCT and for those who do not, based on a partitioned survival approach.
- During the first 3 years of the modelled time horizon, risks of relapse/progression and death are modelled using data from ADMIRAL[7] and external sources.[30-32] After this 3-year timepoint, surviving patients are assumed to be "cured", irrespective of their treatment group, current health state (event-free or post-event) and whether they have received HSCT. Cured patients are assigned a mortality risk which reflects that of the general population,[33] uplifted using an SMR of 2.0. With the exception of EFS in the No HSCT group, this assumption of cure overrides the event risks predicted by the fitted parametric models after the 3-year timepoint for all surviving patients.
- HSCT is assumed to occur at a fixed timepoint after initiating treatment with gilteritinib or salvage chemotherapy (gilteritinib - [redacted]; salvage chemotherapy - [redacted]). All patients in the With HSCT sub-model remain alive and event-free until this timepoint.

- EFS and OS for the No HSCT group are modelled using log logistic survivor functions fitted to time-to-event data for patients who did not receive HSCT in ADMIRAL.[7]
- OS for patients who undergo HSCT is modelled using a Gompertz model fitted to post-HSCT OS data reported by Evers *et al.*[31]
- A proportion of gilteritinib-treated patients who undergo HSCT are assumed to continue to receive gilteritinib maintenance therapy after the procedure. OS for these patients is modelled using an indirect comparison based on data from ADMIRAL and Evers *et al*.[31] The resulting HR is applied indefinitely, but only has an impact on the results prior to the 3-year cure point.
- A proportional relationship between the hazards for relapse/progression and death is assumed for patients in the With HSCT group. This is modelled using an HR estimated using data reported by Ustun *et al.*[32] After the cure point, the cumulative probability of EFS is assumed to remain constant until it reaches the cumulative probability of OS (hence, all deaths during this time interval are in progressed/relapsed patients).
- Prior to the 3-year cure timepoint, HRQoL for patients with FLT3+ R/R AML is assumed to be dependent on the patient's health state (pre-/post-event) and whether the patient has received or will receive HSCT. Patients surviving beyond 3 years are assumed to have perfect health (utility = 1.0) prior to the application of age-related utility adjustments. As a consequence of an implementation error, the company's model assumes that all patients who survive beyond 3 years experience a level of HRQoL which is considerably better than that of the general population; this is discussed further in Section 5.3.3.
- HSCT is assumed to be associated with a disutility which is applied for 6-months after the procedure. No long-term health losses are included in the model.
- AEs are assumed to be associated with a QALY loss applied in the first model cycle.
- Drug acquisition and administration costs for all treatments are applied as once-only costs in the first model cycle; within the model, these costs are structurally independent of the patient's health state.
- HSCT is not assumed to be associated with any additional AE-related costs over and above those included in the NHS tariff cost.[36]
- Following the assumed 3-year cure point, disease management costs are assumed to be zero.
- Irrespective of the cure point, once-only costs associated with treatments for relapse/progression are applied to incident patients who relapse/progress or die in each cycle. The same cost is applied irrespective of previous treatment received.

5.2.4 Evidence used to inform the company's model parameters

The sources of evidence used to inform company's model parameters are summarised in Table 19. These are discussed in detail in the subsequent sections.

Table 19: Summary of evidence used to inform the company's base case analyses

Parameter group	Source
Patient characteristics (age, BSA, weight, proportion of females)	ADMIRAL ITT population[7]
Probability of receiving HSCT - gilteritinib and salvage chemotherapy	ADMIRAL ITT population[7]
Time to HSCT - gilteritinib and salvage chemotherapy	Mean time to HSCT observed in each arm of ADMIRAL[7]
No HSCT, EFS - gilteritinib	Log logistic models fitted to time-to-event data for patients without HSCT in ADMIRAL.[7] Models fitted separately to each treatment group. EFS is modelled as a function of the probability of early treatment failure (≤30 days) and parametric distributions fitted to data for patients without early treatment failure (>30 days).
No HSCT, EFS - salvage chemotherapy	
No HSCT, OS - gilteritinib	
No HSCT, OS - salvage chemotherapy	
No HSCT, OS - HR for BSC versus salvage chemotherapy	Naïve arm-based indirect comparison using data from ADMIRAL[7] and Sarkozy *et al*[30]
With HSCT, OS – salvage chemotherapy and no gilteritinib maintenance therapy	Gompertz model fitted to data reported by Evers *et al*[31] (R/R AML patients in CR2)
With HSCT, HR for OS for gilteritinib maintenance therapy	Indirect comparison using data from ADMIRAL[7] (patients who received HSCT and gilteritinib maintenance therapy) and patients in CR2 in Evers *et al*[31]
Probability of receiving gilteritinib maintenance therapy	ADMIRAL ITT population[7]
With HSCT, HR for OS to EFS - gilteritinib and chemotherapy	HR estimated from LFS and OS curves for URD group in Ustun *et al*[32]
Survival following cure	Mortality rates from interim life tables[33] uplifted with SMR=2.0
Cure point	Assumption based on TA523,[37] published literature[38-40] and clinical advice[1]
HRQoL - event-free and post-event, both sub-models (prior to cure point)	GEE model fitted to EQ-5D-5L data collected in ADMIRAL[7] (mapped to EQ-5D-3L using Van Hout *et al*[34])
HRQoL for long-term AML survivors	Assumption (perfect health prior to age-adjustment)
HRQoL age-adjustment	Based on Janssen *et al*[35]
AE disutility	AE frequencies taken from ADMIRAL;[7] disutilities taken from Swinburn *et al*,[41] Lloyd *et al*,[42] Nafees *et al*,[43] Doyle *et al,*[44] ADMIRAL[7] and assumptions
HSCT disutility	Joshi *et al.*[29] 6-month duration assumed
Drug acquisition costs	Drug usage based on ADMIRAL.[7] Unit costs taken from MIMS and eMIT[45, 46]
Drug administration costs	Unit costs taken from NHS Reference Costs 2017/18.[36] Administration frequency and RDI based on ADMIRAL.[7]
FLT3 mutation re-testing cost	NICE TA523[37]
Costs related to HSCT procedure	NHS Reference Costs 2017/18[36]
Disease management costs – event-free, No HSCT (monthly)	Resource use based on AML chart review.[47] Hospitalisation duration based on No HSCT patients in ADMIRAL.[7] Unit costs taken from the PSSRU[48] and NHS Reference Costs 2017/18.[36]

Parameter group	Source
Disease management costs – event-free, With HSCT (monthly)	Monthly follow-up duration based on Tremblay *et al*.[49] Unit costs taken from PSSRU 2018.[48]
Disease management costs – post-event (monthly)	Resource use based on AML chart review.[47] Unit costs taken from PSSRU[48] and NHS Reference Costs 2017/18.[36]
Post-progression treatment costs	Proportion of patients receiving subsequent treatment and average number of cycles from ADMIRAL.[7] Chemotherapy cost per cycle based on from Wang *et al*.[50]
Costs incurred by long-term AML survivors	Assumption (zero cost applied after 3 years)
Costs associated with AEs	AE frequencies based on ADMIRAL.[7] Unit costs taken from NHS Reference Costs 2017/18[36] and assumptions[1]
End of life care costs	Georghiou and Bardsley[51]

AE - adverse event; AML - acute myeloid leukaemia; BSA - body surface area; BSC - best supportive care; CR2 - second complete remission; EFS - event-free survival; EQ-5D - Euroqol 5-dimensions; FLT3 - FMS-like tyrosine kinase 3; GEE - generalised estimating equation; HRQoL - health-related quality of life; HSCT - haematopoietic stem cell transplant; ITT - intention-to-treat; LFS - leukaemia-free survival; MIMS - Monthly Index of Medical Specialities; eMIT - Electronic Market Information Tool; OS - overall survival; PSSRU - Personal Social Services Research Unit; RDI - relative dose intensity; R/R - relapsed or refractory; SMR - standardised mortality ratio; TA – technology appraisal; URD – unrelated donor group

Initial patient characteristics at model entry

The model assumes that patients enter the model aged [redacted] years and approximately [redacted] of the modelled cohort is assumed to be male. Patients are assumed to have a mean body surface area (BSA) of [redacted]. These characteristics reflect those of the ADMIRAL ITT population.[7]

Time-to-event parameters

The model uses separate data sources to model time-to-event outcomes (EFS and OS) for patients in the With HSCT and No HSCT sub-models. Within the No HSCT group, EFS and OS outcomes are modelled using data from ADMIRAL.[7] EFS and OS outcomes for the With HSCT group are modelled using external data (Evers *et al*[31]). The key features of the company's survival analysis approach and its application within the health economic model are summarised in Box 1. It should be noted that the survival analyses described within this section do not account for the company's 3-year cure assumption.

Box 1: Key features of the company's survival analysis

- Time-to-event outcomes for the No HSCT and With HSCT groups are modelled separately
- EFS and OS for the No HSCT group are based on data from ADMIRAL[7]
- OS for the With HSCT subgroup is based on Evers *et al*[32]
- For the With HSCT group, the hazards for relapse/progression and death are assumed to be proportional, based on data reported by Ustun *et al*[32]
- An additional OS and EFS benefit is applied for patients who receive gilteritinib maintenance therapy following HSCT, based on a naïve arm-based indirect comparison using data from ADMIRAL (patients who underwent HSCT and received gilteritinib maintenance therapy) and Evers *et al*.[31]
- Company's selected models:
    - No HSCT, EFS, gilteritinib – log logistic
    - No HSCT, OS, gilteritinib – log logistic
    - No HSCT, EFS, salvage chemotherapy – log logistic
    - No HSCT, OS, salvage chemotherapy – log logistic
    - With HSCT, OS, salvage chemotherapy and those not receiving gilteritinib maintenance – Gompertz
    - With HSCT, OS, patients receiving gilteritinib maintenance therapy – Gompertz uplifted using HR
    - With HSCT, EFS, both treatments – HR applied to Gompertz model (including maintenance therapy HR)
- All surviving patients are assumed to be "cured" after 3 years, irrespective of whether they have relapsed/progressed and irrespective of whether they have previously received HSCT. An SMR of 2.0 is applied relative to general population mortality rates.[33] With the exception of EFS in the No HSCT subgroup, this cure assumption overrides the risks predicted by the company's parametric survival models for the remaining 37 years of the time horizon.

No HSCT – overall survival

Within ADMIRAL, OS was defined as the time from randomisation to death from any cause.[7] Within the No HSCT group, OS was modelled by fitting parametric survival models to the available data for patients who did not receive HSCT in ADMIRAL (gilteritinib n=184; salvage chemotherapy n=105 patients). The company fitted six parametric models to the individual patient-level data (IPD) from the trial; these included the exponential, Weibull, Gompertz, log logistic, log normal and generalised gamma models. Models were fitted separately to data for each treatment group. The CS[1] (page 58) states that the company selected their preferred OS survivor function through consideration of: relative goodness-of-fit statistics (the Akaike Information Criterion [AIC] and the Bayesian Information

Criterion [BIC]); visual inspection of the fitted distributions; examination of log cumulative hazard plots and testing the proportional hazards (PH) assumption, and clinical input and external validation.

Figure 4 and Figure 5 present comparisons of the model-predicted and observed cumulative survival probabilities for OS without HSCT for gilteritinib and salvage chemotherapy, respectively. The AIC and BIC statistics for each of the candidate models are presented in Table 20. Log cumulative hazard plots for three of the parametric OS models (log normal, log logistic and generalised gamma) are presented in Figure 25 and Figure 26 in ERG Appendix 2.

Figure 4: No HSCT, OS, gilteritinib - Kaplan-Meier plot and company's survival models (redrawn by the ERG)

Figure 5: No HSCT, OS, salvage chemotherapy - Kaplan-Meier plot and company's survival models (redrawn by the ERG)

Table 20: No HSCT, OS, both groups - AIC and BIC statistics (adapted from CS Tables 20 and 21)

Survivor function	Gilteritinib		Salvage chemotherapy	
	AIC	BIC	AIC	BIC
Exponential	943.39	946.61	485.93	488.59
Weibull	938.75	945.18	485.73	491.04
Gompertz	944.71	951.14	487.89	493.20
Log normal	933.22	939.65	485.31	490.62
Log logistic	**930.70**	**937.13**	**482.73**	**488.04**
Generalised gamma	933.21	942.85	483.89	491.85

AIC – Akaike Information Criterion; BIC – Bayesian Information Criterion.
Lowest AIC/BIC values shown in bold.

The company selected the log logistic models for both treatment groups within their base case analysis. According to the CS[1] (page 64), this decision was made on the basis that the log logistic model had the lowest AIC and BIC, the log cumulative hazard plot indicates non-monotonic hazards over time, and the model provides a good visual fit to the observed Kaplan-Meier survivor function. The selected log logistic survivor functions are presented in Figure 6.

Figure 6: No HSCT, OS, both groups - log logistic models (redrawn by the ERG)

No HSCT – event-free survival

EFS was defined as the time from the date of randomisation until the date of documented relapse (excluding relapse after partial remission [PR]), treatment failure or death from any cause within 30 days after the last dose of the study drug, whichever occurred first.[7] As with the OS analysis, the analysis of EFS was informed by IPD for those patients who did not receive HSCT during the follow-up period of ADMIRAL (gilteritinib n=[redacted]; salvage chemotherapy=[redacted]). Treatment failures occurring within the first 30 days of follow-up were re-coded to day 0, based on discussion and agreement with the US Food and Drug Administration ([FDA]; see clarification response,[7] question A8). As discussed in the CS,[1] the recoding of early treatment failures leads to an immediate drop in the EFS survivor functions for the ADMIRAL ITT population (gilteritinib - [redacted]; salvage chemotherapy - [redacted]). The CS comments that the shape of the recoded EFS distributions makes standard parametric modelling problematic. In order to address this problem, the company adopted a two-stage approach: (i) early treatment failures occurring in the first month were assumed to occur immediately (on day 0), and (ii) only patients without treatment failure within the first 30 days were included in the survival analysis for EFS. The company fitted six standard parametric models to the EFS data for patients without day 0 treatment failure; these included the exponential, Weibull, Gompertz, log logistic, log normal and generalised gamma models. As with the analysis of OS for the No HSCT group, the company selected their preferred base case EFS survivor functions through consideration of: relative goodness-of-fit statistics

(AIC and BIC); visual inspection of the fitted models; examination of log cumulative hazard plots and testing the PH assumption, and clinical input and external validation.

Figure 7 and Figure 8 present comparisons of the model-predicted and observed cumulative survival probabilities for EFS without HSCT for gilteritinib and salvage chemotherapy, respectively. The AIC and BIC statistics for the candidate models are presented in Table 21. Log cumulative hazard plots for three of the parametric EFS models (log normal, log logistic and generalised gamma) are presented in Figure 27 and Figure 28 in ERG Appendix 2.

Figure 7: No HSCT, EFS, gilteritinib - Kaplan-Meier plot and company's survival models (redrawn by the ERG)

Figure 8: No HSCT, EFS, salvage chemotherapy - Kaplan-Meier plot and company's survival models (redrawn by the ERG)

Table 21: No HSCT, EFS, both groups - AIC and BIC statistics (adapted from CS Tables 23 and 24)

Survivor function	Gilteritinib		Salvage chemotherapy	
	AIC	BIC	AIC	BIC
Exponential	452.60	455.19	63.37	65.48
Weibull	454.23	459.40	65.33	69.55
Gompertz	453.18	458.35	64.11	68.33
Log normal	447.62	452.79	61.14	**65.36**
Log logistic	**445.29**	**450.46**	62.46	66.68
Generalised gamma	449.38	457.13	**60.87**	67.20

AIC – Akaike Information Criterion; BIC – Bayesian Information Criterion.
Lowest AIC/BIC values shown in bold.

According to the CS,[1] the company selected the log logistic function for the gilteritinib group on the basis that this model: (i) has the lowest AIC and BIC values; (ii) demonstrated good visual fit against the observed data, and (iii) makes assumptions which are consistent with the non-monotonic pattern seen in the log cumulative hazard plot. The company selected the same distributional form for the salvage chemotherapy group based on guidance on curve selection given in NICE Decision Support Unit (DSU) Technical Support Document (TSD) 14.[52] The log logistic functions used in the company's model are shown in Figure 9.

Figure 9: No HSCT, EFS, log logistic model applied to both treatment groups – Kaplan-Meier plot and company's survival models (redrawn by the ERG)

With HSCT – overall survival

Survival outcomes for patients who undergo HSCT were not based on data from ADMIRAL.[7] According to the CS,[1] (page 73), follow-up for these patients was limited (the median post-HSCT follow-up was [redacted] months) and only [redacted] patients ([redacted]) and [redacted] patients ([redacted]) had survival data beyond years 1 and 2, respectively.[1] For this reason, the company used external data to inform OS and EFS outcomes for patients receiving HSCT. The ERG notes that data from ADMIRAL are still used to estimate the relative benefit of gilteritinib maintenance therapy versus no maintenance therapy following HSCT; this is discussed in further detail in Section 5.3.3.

The company undertook a targeted literature review to identify studies which could be used to inform OS and EFS outcomes for patients following HSCT. The company's review identified ten potentially relevant studies.[31, 32, 53-60] According to the CS[1] (page 74) the study reported by Evers *et al*[31] was selected for inclusion in the model *"because it included the largest sample size and had the longest follow-up time."* Evers *et al* report a *post hoc* analysis of outcomes data for 498 subjects enrolled into the AMLCG1999 clinical trial aged 17 to 74 years of age who underwent allogeneic HSCT after achieving first CR. The company used data on the subset of patients with CR2. These patients have R/R AML, but do not specifically have FLT3+ disease. The company's clarification response[7] (question B9) provides further justification for the selection of Evers *et al* as the source of post-HSCT OS outcomes,

noting that the population included R/R patients after first-line treatment and better reflects the ADMIRAL trial population compared with other studies included in the company's review. Evers *et al*[31] reports only on OS; data on EFS are not available from this study.

The company digitised the OS Kaplan-Meier survivor function for the CR2 subgroup in Evers *et al*[31] and replicated the underlying pseudo-IPD using the algorithm reported by Guyot *et al*.[61] The company fitted exponential, Weibull, Gompertz, log logistic, log normal and generalised gamma models to the available data. The company selected their preferred OS survivor function through consideration of: relative goodness-of-fit statistics (AIC and BIC); visual inspection of the fitted distributions, and examination of log cumulative hazard plots. The CS[1] does not mention whether the clinical plausibility of the selected survivor function was one of the criteria for OS model selection; the company's clarification response[21] (question B16) briefly mentions the use of expert opinion, but does not provide details. As noted earlier, the assumed 3-year cure point overrides the long-term OS projection for these patients.

Figure 10 presents a comparison of the model-predicted and observed cumulative survival probabilities for post-HSCT OS; this is applied to both groups in the model (but excludes the impact of the HR for gilteritinib maintenance therapy). The AIC and BIC statistics for these models are presented in Table 22. The log cumulative hazard plot for post-HSCT OS is presented in
Figure 29 in ERG Appendix 2.

Figure 10: With HSCT, OS, both treatment groups- Kaplan-Meier plot and company's survival models (redrawn by the ERG)

Table 22: OS with HSCT - AIC and BIC statistics (adapted from CS, Table 27)

Survivor function	OS post-HSCT (both treatment groups)	
	AIC	BIC
Exponential	942.45	945.30
Weibull	849.02	854.73
Gompertz	840.88	846.59
Log normal	828.56	834.26
Log logistic	835.35	841.05
Generalised gamma	**800.35**	**808.91**

AIC - Akaike Information Criterion; BIC - Bayesian Information Criterion; HSCT - haematopoietic stem cell transplant; OS – overall survival.
Lowest AIC/BIC values shown in bold.

As shown in Table 22, the AIC and BIC statistics indicate that the generalised gamma distribution provides the best statistical fit to the data. However, the company did not select this distribution because it is not a PH model, and the company had elected to estimate EFS by applying an HR (estimated from another study – Uston *et al*[32]) to the OS survivor function from Evers *et al*.[31] Instead, the company selected the Gompertz OS model for use in the health economic analysis. The CS[1] states that this decision was made on the basis that: the Gompertz model has a lower AIC and BIC compared with the other PH models; it provides a reasonable visual fit, and the log cumulative hazard plot indicates a non-monotonic pattern.

Within the company's health economic model, this Gompertz survival model is applied to all patients within the salvage chemotherapy group (from [redacted] onwards), and to the proportion of patients in the gilteritinib group who do not receive maintenance gilteritinib following HSCT (from [redacted] onwards). For those patients who receive gilteritinib maintenance therapy following HSCT, the baseline Gompertz OS model is uplifted using an HR (also applied from [redacted] onwards). This HR was estimated by performing a naïve arm-based indirect comparison using OS outcomes for patients receiving gilteritinib maintenance therapy in ADMIRAL[7] (n=[redacted]) versus OS outcomes for patients with CR2 HR in Evers *et al*[31] (estimated HR=[redacted]). The application of this HR also uplifts the EFS function for these patients (because EFS is modelled as a function of OS). The HR-adjusted post-HSCT OS survivor function and the methods through which this HR was derived are not described in the CS.[1] In response to a request for clarification from the ERG,[7] (question B11), the company stated that they used a Cox PH model to implement the analysis. Additional information provided in the company's clarification response indicates that the PH assumption is violated.

With HSCT – event-free survival

Evers *et al*[31] do not report data on EFS. Instead, the company assumed that the hazard of relapse/progression is proportional to the hazard of death for patients who have undergone HSCT. This HR was estimated using data reported in Ustun *et al*[32] - this was one of the ten studies identified within

the company's targeted review of studies reporting on HSCT outcomes. Ustun *et al*[32] is a retrospective observational analysis of outcomes for sibling and URD HSCT based on data from the Center for International Blood and Marrow Transplant Research (CIBMTR). All patients in this study have FLT3+ AML, but the majority are in CR1 (they do not have relapsed disease). The overall HR for OS to EFS was calculated by estimating the ratio between the log cumulative hazards for OS and EFS for each month up to year 5; the overall HR for OS to EFS was calculated as the average of the cumulative HRs. Within the company's model, this overall HR is applied up to year 3; beyond this timepoint, EFS is assumed to remain constant until it reaches the OS survival function, thereby assuming that all deaths during this interval are in patients with prior relapse/progression.

Cure assumption

The model assumes that all patients who remain alive after 3 years are "cured"; this is equivalent to a piecewise survival model with a turning point in the hazard of death at year 3. From this timepoint onwards, the hazard of death for all surviving patients is based on general population mortality rates[33] uplifted using an SMR of 2.0. With the exception of EFS in the No HSCT subgroup (for whom EFS risk is assumed to continue indefinitely), after 3 years, the SMR-adjusted general population survivor function overrides all of the event risks predicted by the parametric survivor functions described above.

Time-to-event outcomes for BSC

ADMIRAL[7] did not include a BSC arm; effectiveness outcomes for BSC-treated patients were instead based on external data. The company undertook a targeted literature review to identify studies which reported outcomes for patients receiving BSC. The company selected a study reported by Sarkozy *et al*.[30] This study is a pooled analysis of outcomes for patients enrolled in the Acute Leukemia French Association (ALFA) 9801 and 9803 trials. A proportion of the patients in these trials were known to have FLT3 mutation-positive disease. Sarkozy *et al* only report median OS by treatment type and do not report an OS Kaplan-Meier survivor function for patients receiving BSC alone. According to the CS,[1] an HR for BSC versus gilteritinib was estimated by comparing the ratio between the median OS for LDAC and BSC in Sarkozy *et al*[30] versus the HR for salvage therapy versus gilteritinib in ADMIRAL.[7] The CS reports that this produced an HR of 2.86. However, the CS does not provide any details of the values used to inform this analysis and the ERG was not able to replicate this value (see Section 5.3.3).

HRQoL

Health state utility values (applied prior to the 3-year cure timepoint)

ADMIRAL[7] included the collection of HRQoL data from patients using the EQ-5D-5L questionnaire. The questionnaire was administered at cycle 1 day 1 pre-dose, cycle 2 day 1 (± 2 days) and all subsequent cycles' day 1 (± 2 days) and during the pre-HSCT/end of treatment visit and at the 30-day

follow-up visit; during long-term follow-up, patients were contacted by site personnel via telephone to provide responses to the questionnaire.[1] The company mapped the EQ-5D-5L data to the EQ-5D-3L using the algorithm reported by Van Hout *et al*[34] and fitted a GEE model to the available data, including a robust variance estimator to account for correlation within patients' repeated assessments.[1] The characteristics of the utility data and the estimates applied in the company's model are summarised in Table 23. These utilities are applied only for the first 3 years of the model time horizon (until the assumed cure point) and are assumed to be independent of treatment group.

Table 23: Mapped EQ-5D-3L estimates used in company's model (adapted from CS, Table 28)

Health state	No. of patients	No. of assessments	Mean utility	SD	SE
No HSCT, event-free	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
No HSCT, post-event	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
With HSCT, event-free	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
With HSCT, post-event	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]

HSCT - haematopoietic stem cell transplantation; SD - standard deviation; SE - standard error

Age-adjustment of utilities (applied after the 3-year cure timepoint)

The model applies age-adjustment to the health state utilities based on UK general population norms reported by Janssen and Szende.[35] These population norms are based on Kind *et al*[62] and are reported crudely across three age bands: age 55-64 years (utility=0.799); 65-74 years (utility=0.779) and 75+ years (utility=0.726). Within the company's model, age-adjustment is applied as a relative decrease from the age 55-64 band utility: for example, a surviving patient aged 77 years old is assumed to have an age-adjustment multiplier of 0.91 (calculated as 0.726 divided by 0.779). Prior to the cure timepoint, the age-adjustment has no effect as surviving patients are still in the lowest age band. Subsequent to the cure timepoint, all patients are assumed to have a utility score of 1.0 (a notional state of perfect health); this utility is then age-adjusted using the multiplier approach described above. The ERG notes that the company's utility calculations are subject to a significant error; this is discussed in Section 5.3.3.

Disutility associated with HSCT

The model applies a disutility of [redacted] to patients who undergo HSCT, based on a general population time-trade-off (TTO) study of alternative health states for patients with newly diagnosed AML (Joshi *et al*[29]). The model assumes that this disutility applies for a duration of 6 months, based on expert clinical input.[1] Within the company's model, this disutility is subtracted from the health state utilities for all patients in the With HSCT sub-model during the first 6 monthly cycles after the HSCT timepoint. No further HSCT-related health losses are assumed in subsequent model cycles.

QALY losses due to AEs

The model includes a once-only QALY loss associated with Grade 3/4 AEs (see Table 24). AE frequency data were based on the ITT population of ADMIRAL,[7] whilst disutilities associated with each AE were based on estimates from ADMIRAL[7] and other literature.[41-44] The ERG notes that it is unclear whether these disutilities would already be accounted for in the utility estimates derived from the GEE model. The AE rate for BSC was assumed to be zero. The ERG notes that progressive AML is included as an AE and an associated disutility for the event is calculated as the difference between the event-free and post-event utility scores (without HSCT) from the GEE model; the ERG notes that this is double-counting the impact of progression on HRQoL. The ERG also notes that several AEs are assumed to be associated with zero disutility; it is unclear from the CS[1] whether this reflects an assumption that these AEs do not have detrimental impact on patients' health, or a lack of evidence through which to quantify a non-zero health impact. The duration of all AEs is assumed to be one month; based on the data reported in Table 24, this leads to overall QALY losses of [redacted] for gilteritinib, [redacted] for all salvage chemotherapy regimens and zero for BSC. This QALY loss is applied in the first model cycle.

Table 24: Frequency of Grade 3/4 AEs and associated disutilities (from company's model)

Adverse event	GILT	AZA	FLAG-IDA	MEC	LDAC	BSC	Disutility
Anaemia	40.7%	30.3%	30.3%	30.3%	30.3%	0.0%	-0.119
Dyspnoea	[redacted]	2.8%	2.8%	2.8%	2.8%	0.0%	-0.050
Elevated ALT	13.8%	4.6%	4.6%	4.6%	4.6%	0.0%	0.000
Elevated AST	14.6%	1.8%	1.8%	1.8%	1.8%	0.0%	0.000
Elevated blood phosphocreatine kinase	[redacted]	0.0%	0.0%	0.0%	0.0%	0.0%	0.000
Fatigue	[redacted]	1.8%	1.8%	1.8%	1.8%	0.0%	-0.115
Febrile neutropenia	45.9%	36.7%	36.7%	36.7%	36.7%	0.0%	-0.150
Hyperglycaemia	[redacted]	8.3%	8.3%	8.3%	8.3%	0.0%	0.000
Hypertension	[redacted]	3.7%	3.7%	3.7%	3.7%	0.0%	-0.153
Hypokalaemia	13.0%	11.0%	11.0%	11.0%	11.0%	0.0%	0.000
Hyponatraemia	[redacted]	2.8%	2.8%	2.8%	2.8%	0.0%	0.000
Hypophosphatemia	[redacted]	3.7%	3.7%	3.7%	3.7%	0.0%	0.000
Hypotension	[redacted]	2.8%	2.8%	2.8%	2.8%	0.0%	-0.153
Leukopenia	[redacted]	0.0%	0.0%	0.0%	0.0%	0.0%	-0.090
Neutropenia	[redacted]	13.8%	13.8%	13.8%	13.8%	0.0%	-0.090
Neutrophil count decreased	[redacted]	11.0%	11.0%	11.0%	11.0%	0.0%	0.000
Platelet count decreased	22.0%	24.8%	24.8%	24.8%	24.8%	0.0%	0.000
Pneumonia	[redacted]	4.6%	4.6%	4.6%	4.6%	0.0%	-0.153
Progressive AML	[redacted]	3.7%	3.7%	3.7%	3.7%	0.0%	[redacted]
Sepsis	[redacted]	0.0%	0.0%	0.0%	0.0%	0.0%	-0.090
Thrombocytopenia	22.8%	16.5%	16.5%	16.5%	16.5%	0.0%	-0.090
White blood cell count decreased	[redacted]	17.4%	17.4%	17.4%	17.4%	0.0%	0.000

AML - acute myeloid leukaemia; AZA - azacitidine; BSC - best supportive care; FLAG-IDA - fludarabine, cytarabine, G-CSF and idarubicin ; GILT - gilteritinib; LDAC - low-dose cytarabine; MEC - mitoxantrone, etoposide and cytarabine

Resource costs

The model includes the costs associated with: (i) drug acquisition and administration; (ii) the HSCT procedure (including stem cell harvesting); (iii) re-testing for the FLT3 mutation; (iv) disease management (monthly health state costs); (v) treatments following disease relapse/progression; (vi) management of AEs and (vii) end of life care costs. The costs associated with drugs, tests, events and heath states applied in the company's model are summarised in Table 25. With the exception of the once-only cost of treating disease progression, the model assumes that all disease management costs are zero after the assumed 3-year cure point.

Table 25: Summary of drug, test, event and heath state costs applied in the company's model

Cost parameter	Base case analysis		Secondary comparisons				
	Gilteritinib	Salvage chemotherapy	Azacitidine	FLAG-IDA	MEC	LDAC	BSC
Drug and administration costs (once-only)	[redacted]	£5,712†	£13,698†	£3,336	£1,849	£4,048	N/a
HSCT procedure (once-only)	£40,774	£40,774	£40,774	£40,774	£40,774	£40,774	N/a
Re-testing for the FLT3 mutation (once-only)	£309	N/a	N/a	N/a	N/a	N/a	N/a
Disease management – event-free, No HSCT (monthly)	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	N/a
Disease management (hospitalisation) – event-free, No HSCT (monthly)	[redacted]	£2,543	£690	£3,754	£3,754	£690	N/a
Disease management – event-free disease, With HSCT (monthly)	£170	£170	£170	£170	£170	£170	N/a
Disease management – progressed disease (monthly)	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
Post-progression treatment costs (once-only, applied when patient leaves the event-free state)	[redacted]	£8,264	£8,264	£8,264	£8,264	£8,264	N/a
End of life care (once-only)	£2,553	£2,553	£2,553	£2,553	£2,553	£2,553	£2,553
Grade 3+ AEs (once-only)	[redacted]	£1,828	£1,828	£1,828	£1,828	£1,828	N/a

AE - adverse event; BSC - best supportive care; FLAG-IDA - fludarabine, cytarabine, G-CSF and idarubicin; FLT3 - FMS-like tyrosine kinase; HSCT - haematopoietic stem cell transplant; MEC - mitoxantrone, etoposide and cytarabine; LDAC - low-dose cytarabine

** Includes PAS for gilteritinib.*

† Excludes cPAS for azacitadine.

(i) Drug acquisition and administration costs

Based on its list price, the cost per pack of 84 x 40mg gilteritinib tablets (28 days' supply) is [redacted]. The company has proposed a PAS which takes the form of a simple price discount of [redacted]; the discounted cost per pack of gilteritinib is [redacted]. Within the model, the lifetime acquisition cost of gilteritinib is estimated as a function of the unit cost per pack, the planned treatment schedule, the amount of planned treatment received (RDI=[redacted]) and the mean treatment duration observed in ADMIRAL[7] ([redacted] 28-day dosing cycles). The model does not include wastage associated with prescribed tablets that were supplied but not consumed by patients in the trial (see clarification response,[7] question B20). The model assumes that administration of gilteritinib will require one outpatient attendance every 3 months; the unit cost for gilteritinib administration was taken from NHS Reference Costs 2017/18[36] (SB11Z - "Deliver Exclusively Oral Chemotherapy - Outpatient", cost=£131.61). The total cost of gilteritinib acquisition and administration is applied as a once-only cost in the first model cycle, and is structurally independent of the patient's health state.

Drug acquisition and administration costs for each individual salvage chemotherapy regimen are calculated as function of the patient's mean BSA, the RDI for each regimen component, and the treatment duration observed in ADMIRAL.[7] Drug prices were taken from the Monthly Index of Medical Specialities[45] (MIMS) and the electronic Market Information Tool (eMIT),[46] whilst unit costs associated with drug administration were taken from the NHS Reference Costs 2017/18.[36] For the salvage chemotherapy comparator, the weights for each individual regimen were estimated from observed data in ADMIRAL.[7] Drug acquisition and administration costs for salvage chemotherapy are applied as once-only costs in the first model cycle. The ERG notes that a comparator PAS (cPAS) is available for azacitidine; the impact of this cPAS on the cost-effectiveness of gilteritinib is presented in a confidential appendix to this ERG report.

(ii) Costs related to HSCT procedure

Costs related to HSCT include stem cell harvesting and the stem cell transplant procedure; these were taken from the NHS Reference Costs 2017/18.[36] The total cost of HSCT is estimated to be £40,774; this is applied as a once-only cost to the proportion of patients undergoing HSCT in the first model cycle.

(iii) Costs for FLT3 re-testing

The model assumes that the cost of FLT3 mutation re-testing is £154.54, based on the test cost used in NICE TA523[28] (uplifted to 2018 prices using the Consumer Price Index [CPI][48]). The model assumes a testing rate of 200% i.e. two R/R AML patients would need to be tested in order to identify one R/R AML patient with FLT mutation-positive disease. This cost is applied in the first model cycle (for all patients in the gilteritinib treatment group only).

(iv) Disease management costs

Disease management costs include: outpatient visits to general practitioners (GPs); haematologists and nurses; emergency department (ED) visits; hospitalisations; diagnostic imaging procedures; laboratory tests, and blood transfusions. Monthly resource use and cost assumptions by health state are summarised in Table 26.

The disease management costs for patients in the event-free state correspond to one of the two following sets of costs:

(i) With HSCT group: A follow-up cost of £170.02 is applied to patients from the timepoint at which HSCT is assumed to be undertaken (estimated from Tremblay *et al*[49] and Curtis *et al*[48]), or;

(ii) No HSCT group: Costs include a fixed number of visits and procedures (independent of treatment option) and hospitalisation costs (dependent on treatment option, based on ADMIRAL) per cycle. These costs were estimated based on an AML chart review,[47] ADMIRAL,[7] Curtis *et al*[48] and NHS Reference Costs 2017/18.[36]

Disease management costs for patients in the post-progression state are assumed to be the same for all patients, regardless of treatment group or HSCT status. These include visits and procedures, and hospitalisation costs, and are based on the AML chart review,[47] Curtis *et al*[48] and NHS Reference Costs 2017/18.[36]

All disease management costs (for event-free [including hospitalisation costs] and post-event states) are estimated on a monthly basis, and are applied in every cycle prior to the assumed cure timepoint. From this point onwards, disease management costs are assumed to be zero.

Table 26: Summary of health state resource use and costs (monthly, prior to assumed cure timepoint)

Resource type	Resource component	Treatment	Frequency – event-free (monthly)	Frequency – post-event (monthly)	Unit cost	Total – event-free	Total – post-event
Outpatient visits	Haematologist visits	All treatments	[redacted]	[redacted]	£108.00	[redacted]	[redacted]
	Nurse visits		[redacted]	[redacted]	£37.00	[redacted]	[redacted]
	GP visits		[redacted]	[redacted]	£93.75	[redacted]	[redacted]
ED visits	ED visits		[redacted]	[redacted]	£202.15	[redacted]	[redacted]
Diagnostic procedures and tests	Imaging procedures		[redacted]	[redacted]	£42.30	[redacted]	[redacted]
	Bone marrow biopsy		[redacted]	[redacted]	£519.82	[redacted]	[redacted]
	Lumbar puncture		[redacted]	[redacted]	£519.82	[redacted]	[redacted]
Blood transfusions	Red blood cells		[redacted]	[redacted]	£128.99	[redacted]	[redacted]
	Platelets		[redacted]	[redacted]	£208.68	[redacted]	[redacted]
	Plasma		[redacted]	[redacted]	£28.46	[redacted]	[redacted]
Hospitalisations	ICU days	Gilteritinib	[redacted]	0.22	£1,049.23	[redacted]	[redacted]
		Azacitidine or LDAC	0.00			[redacted]	
		FLAG-IDA or MEC	0.54			[redacted]	
	Hospitalisation days (day case)	Gilteritinib	[redacted]	2.13	£396.30	[redacted]	[redacted]
		Azacitidine or LDAC	1.03			[redacted]	
		FLAG-IDA or MEC	6.59			[redacted]	
	Both ICU and hospitalisation days*	Gilteritinib	[redacted]	N/a	£722.76	[redacted]	N/a
		Azacitidine or LDAC	0.39			[redacted]	
		FLAG-IDA or MEC	0.80			[redacted]	
Follow-up after HSCT (minutes)	-	All treatments	101	-	£1.68	£170.02	-
Total (With HSCT, all treatments)		All treatments	-	-	-	**£170**	[redacted]
Subtotal (No HSCT, all treatments)		All treatments	-	-	-	[redacted]	[redacted]
Subtotal (EFS hospitalisations – No HSCT, by treatment)		Gilteritinib	-	-	-	[redacted]	-
		Azacitidine or LDAC	-	-	-	[redacted]	-
		FLAG-IDA or MEC	-	-	-	[redacted]	-

ED - emergency department; EFS - event-free survival; FLAG-IDA - fludarabine, cytarabine, G-CSF and idarubicin; HSCT - haematopoietic stem cell transplant; ICU - intensive care unit; LDAC - low-dose cytarabine; MEC - mitoxantrone, etoposide and cytarabine; GP - general practitioner

(v) Costs of treatments following disease relapse/progression

The costs associated with treatments for relapse/progression are applied as a lump sum cost to all patients who leave the event-free state, irrespective of whether they have relapsed/progressed or died. These costs vary according to treatment group, and are based on the proportions of patients who received post-relapse/progression treatment in ADMIRAL (gilteritinib – [redacted]; salvage chemotherapy – 61%).[7] The costs of post-relapse/progression treatment per cycle are based on Wang *et al*[50] and include drug acquisition, administration, hospitalisation and the management of AEs.

(vi) AE management costs

Costs related to the management of AEs are applied as once-only costs during the first model cycle, based on the frequency of individual Grade 3/4 AEs observed in ADMIRAL (see Table 24),[7] NHS Reference Costs 2017/18[36] and assumptions. The ERG notes that the costs of progressive AML are included in the model both as an AE and through the use of different costs for the event-free/post-event states; hence, this represents double-counting.

(vii) End of life care costs

The cost of end of life care was estimated to be £2,553 based on Georghiou and Bardsley[51] (uplifted to 2018 prices using the CPI[48]). This is applied as a once-only cost at the point of death.

5.2.5 Model evaluation methods

The CS[1] presents base case incremental cost-effectiveness ratios (ICERs) for gilteritinib versus the salvage chemotherapy blended comparator. Pairwise ICERs are also presented for gilteritinib versus each of the individual salvage chemotherapy regimens and versus BSC. Results are presented using the deterministic and probabilistic versions and of the model; the probabilistic ICERs are based on 1,000 Monte Carlo simulations. Table 27 summarises the distributions used to characterise uncertainty around the model parameters within the company's probabilistic sensitivity analysis (PSA). The results of the PSA are presented as a cost-effectiveness plane and as cost-effectiveness acceptability curves (CEACs) for gilteritinib versus the salvage chemotherapy blended comparator only; cost-effectiveness planes and CEACs are not presented for the comparison of gilteritinib versus BSC.

Deterministic sensitivity analyses (DSAs) are presented for gilteritinib versus the blended comparator and versus each individual salvage chemotherapy regimen using tornado plots. Some of these analyses involve varying parameters according to their 95% CIs where available, or using +/- 25% of the expected value where 95% CIs were not available. The plots include scenario analyses as well as one-way sensitivity analyses; the reasons for this are unclear.

The CS[1] also reports the results of five scenario analyses undertaken to explore impact of assumptions around post-event utilities following HSCT, comparator regimen weightings, the cure point, the SMR following cure, and the proportion of patients assumed to undergo HSCT in each treatment group. These analyses are reported only for gilteritinib versus the salvage chemotherapy blended comparator.

Table 27: Distributions used in the company's PSA

Parameter group	Parameter / parameter group	Distribution	ERG comment
Patient characteristics	Initial age	Normal	-
	BSA	Normal	-
	Proportion female	Fixed	-
EFS	No HSCT, gilteritinib	MVN	-
	No HSCT, chemotherapy	MVN	-
	With HSCT, HR for OS to EFS	Log normal	SEs assumed to be 25% of the log mean
OS	General population mortality	Fixed	No uncertainty included around SMR
	No HSCT, gilteritinib	MVN	-
	No HSCT, chemotherapy	MVN	-
	With HSCT, gilteritinib and chemotherapy	MVN	Sampling for the Gompertz model produces volatile results. This was resolved in the company's updated model after the clarification round.
	With HSCT, HR for gilteritinib maintenance versus no maintenance	Fixed	No uncertainty included
	HR for BSC versus gilteritinib	Log normal	Source of SE unclear
HRQoL	Health state utility values (event-free, post-event and long-term AML survivors)	Beta	Mean values as specified in base case and SEs obtained from ADMIRAL
	HSCT and AE-related disutilities	Beta	Mean values as specified in base case and SEs obtained from ADMIRAL
	HRQoL age adjustment	Fixed	-
	Duration HSCT disutility	Fixed	No uncertainty included
Resource use and costs	Weights for blended comparator	Fixed	No uncertainty included
	Treatment duration	Gamma	Mean values as specified in base case and SEs obtained from ADMIRAL
	HSCT procedure	Gamma	SE arbitrarily assumed to be equal to 25% of mean
	Disease management	Gamma	
	FLT3 test	Gamma	
	Post-event treatment costs	Gamma	
	Costs associated with AEs	Gamma	
	End of life care costs	Gamma	
Other	HSCT rate	Beta	Mean values as specified in base case and SEs obtained from ADMIRAL
	Time to HSCT	Fixed	No uncertainty included.
	Gilteritinib maintenance therapy probability	Fixed	

AE - adverse event; AML - acute myeloid leukaemia; BSA - body surface area; BSC - best supportive care; HRQoL - health-related quality of life; HSCT - haematopoietic stem cell transplant; OS - overall survival; EFS - event-free survival; QALY - quality-adjusted life year; SE - standard error; SMR - standardised mortality ratio; MVN - multivariate normal

5.2.6 Company's model validation and verification

The CS[1] (pages 124-125) describes the company's model validation activities. This involved comparing the model-predicted OS against the observed OS in ADMIRAL (CS, Figure 30); however, the ERG notes that this exercise only compares OS for the modelled No HSCT group against the Kaplan-Meier OS functions relating to the ITT population of ADMIRAL censored for HSCT (shown in Figure 5 of the CS). The ERG notes that this analysis does not relate to the overall FLT3+ R/R AML population, as outcomes for patients receiving HSCT are not included in the comparison.

The CS also states that the long-term OS projections for the salvage chemotherapy options are consistent with the data reported in the literature, citing Sarkozy *et al,*[30] the study used to inform the indirect comparison against BSC (see Section 5.2.4), and an abstract (Rowe *et al*[63]) which reports median OS estimates post-relapse for 2,441 patients enrolled into eight consecutive ECOG studies. The CS states that further validation of long-term OS for gilteritinib-treated patients is difficult due to limited external evidence. The ERG notes that arm-specific long-term projections will depend on the characteristics of patients included in the particular study; this is likely to be of particular concern for comparisons of OS across studies of patients with different ages.

The CS[1] describes various uses of clinical input to inform the structure and assumptions used within the model. The CS also states that the model structure was based on a scientific review undertaken through the NICE Preliminary Independent Model Advice (PRIMA) initiative.

The CS[1] does not present a comparison of the OS predictions for the modelled gilteritinib and salvage chemotherapy groups against the survivor function for the ITT population of ADMIRAL,[7] and does not provide a commentary on whether the overall extrapolations were deemed to be plausible by clinical experts. In addition, whilst PRIMA is mentioned, it is unclear what efforts were made to ensure that the model was internally valid. Further issues relating to model errors and the extent to which the model predictions are consistent with the ADMIRAL ITT population are discussed in further detail in Section 5.3.3.

5.2.7 Company's model results

Following the clarification process, the company provided an updated version of their model which included some minor modifications to the PSA. In particular, the covariance matrix for the Gompertz post-HSCT OS distribution was updated in order to address the problem described in Table 27. The probabilistic results presented in this section are based on the updated version of the company's model; the deterministic ICER is the same for both the original and updated models.

Table 28 presents the central estimates of cost-effectiveness generated using the company's model for the comparison of gilteritinib versus the salvage chemotherapy blended comparator. The probabilistic version of the updated model suggests that gilteritinib is expected to generate an additional [redacted] QALYs at an additional cost of [redacted] per patient compared with salvage chemotherapy; the corresponding ICER is £46,716 per QALY gained. The deterministic version of the model produces a higher ICER of £47,695 per QALY gained.

Table 28: Company's base case results - gilteritinib versus salvage chemotherapy blended comparator (based on the company's updated model)

Option	LYGs	QALYs	Costs	Inc. LYGs	Inc. QALYs	Inc. Costs	ICER
Probabilistic model*							
Gilteritinib	3.07	[redacted]	[redacted]	1.30	[redacted]	[redacted]	**£46,716**
Salvage chemotherapy	1.77	[redacted]	[redacted]		-	-	-
Deterministic model							
Gilteritinib	3.03	[redacted]	[redacted]	1.28	[redacted]	[redacted]	**£47,695**
Salvage chemotherapy	1.75	[redacted]	[redacted]	-	-	-	-

ICER - incremental cost-effectiveness ratio; Inc. - incremental; LYG - life year gained; QALY - quality-adjusted life year
** Based on random sampling of salvage chemotherapy comparator costs and outcomes undertaken by the ERG. The company's updated model indicates a slightly different ICER.*

Table 29 presents the results for gilteritinib versus each individual comparator, including BSC. Whilst the CS[1] presents only pairwise comparisons between gilteritinib and individual comparators, the ERG believes that a fully incremental analysis between all options is more informative, although this is still inherently flawed due to the inappropriate assumption that all salvage regimens are equally effective (based on data for the overall salvage chemotherapy arm in ADMIRAL[7]). This analysis indicates that BSC is the least effective option. Azacitidine, FLAG-IDA and MEC are ruled out of the analysis due to simple dominance (by LDAC – the least expensive salvage chemotherapy regimen). The ICER for LDAC versus BSC is estimated to be £20,049 per QALY gained. The ICER for gilteritinib versus LDAC is estimated to be £52,954 per QALY gained. Given the limitations of this analysis, the ERG report does not include further results for gilteritinib versus individual salvage chemotherapy regimens.

Table 29: Company's base case results - gilteritinib versus individual chemotherapy regimens, fully incremental analysis, deterministic

Option	LYGs	QALYs	Costs	Inc. LYGs	Inc. QALYs	Inc. costs	ICER
Gilteritinib	3.03	[redacted]	[redacted]	1.28	[redacted]	[redacted]	**£52,954**
Azacitidine	1.75	[redacted]	[redacted]	-	-	-	**Dominated**
FLAG-IDA	1.75	[redacted]	[redacted]	-	-	-	**Dominated**
MEC	1.75	[redacted]	[redacted]	-	-	-	**Dominated**
LDAC	1.75	[redacted]	[redacted]	1.42	[redacted]	[redacted]	**£20,049**
BSC	0.33	[redacted]	[redacted]	-	-	-	-

BSC - best supportive care; FLAG-IDA - fludarabine, cytarabine, G-CSF and idarubicin; ICER - incremental cost-effectiveness ratio; Inc. - incremental; LDAC - low-dose cytarabine; LYG - life year gained; MEC - mitoxantrone, etoposide, cytarabine; QALY - quality-adjusted life year

The company's tornado plot is shown in Figure 11. The plot indicates that the probability that patients receive HSCT is a key driver of the model: setting the probability of HSCT equal to zero for all treatments increases the ICER to [redacted] per QALY gained. The plot also indicates that the ICER for gilteritinib versus salvage chemotherapy is greater than £50,000 per QALY gained for many of the scenarios.

Figure 11: Company's deterministic sensitivity analysis results - tornado plot for gilteritinib versus salvage chemotherapy blended comparator (reproduced from CS, Figure 21)

AML - acute myeloid leukaemia; CI - confidence interval; CR_c - composite complete remission; DSA - deterministic sensitivity analysis; HSCT - haematopoietic stem cell transplant; ICER - incremental cost-effectiveness ratio; OS - overall survival; PAS - Patient Access Scheme; QALY - quality-adjusted life year; SMR – standardised mortality ratio

Figure 12 presents the CEACs for gilteritinib versus salvage chemotherapy generated by the ERG. Assuming willingness-to-pay (WTP) thresholds of £30,000 and £50,000 per QALY gained, the updated model suggests that the probability that gilteritinib generates more net benefit than salvage chemotherapy is [redacted] and [redacted], respectively.

Figure 12: Company's probabilistic sensitivity analysis results - CEACs for gilteritinib and salvage chemotherapy blended comparator (generated by the ERG using the company's updated model*)

** Based on random sampling of salvage chemotherapy comparator costs and outcomes undertaken by the ERG.*

Table 30 summarises the results of the company's scenario analyses for gilteritinib versus the salvage chemotherapy blended comparator. With the exception of Scenario 1 (post-event utility set equal to [redacted]), these analyses produce ICERs which are lower than the company's base case ICER. Uncertainty surrounding the timing of the assumed cure point was not explored within the CS.[1]

Table 30: Company's scenario analysis results – gilteritinib versus salvage chemotherapy blended comparator, deterministic (generated by the ERG using the company's model)

Scenario	Inc. QALYs	Inc. costs	ICER
Company's base case	[redacted]	[redacted]	**£47,695**
Scenario 1: All post-event utility reduced from [redacted] to [redacted]	[redacted]	[redacted]	**£47,707**
Scenario 2: BSC included in weighted comparator (20% of weight). Relative weightings between salvage chemotherapy options same as base case*	[redacted]	[redacted]	**£43,601**
Scenario 3: Cure point = 2 years	[redacted]	[redacted]	**£38,769**
Scenario 4: Year 4 SMR=3.0, year 5 SMR=2.0, subsequent years SMR=1.0*	[redacted]	[redacted]	**£42,229**
Scenario 5: HSCT rate reduced to those with observed CRc in ADMIRAL (gilteritinib = [redacted] salvage chemotherapy = [redacted])	[redacted]	[redacted]	**£44,549**

BSC - best supportive care; CR_c - composite complete remission; ICER - incremental cost-effectiveness ratio; Inc. incremental; QALY - quality-adjusted life year; SMR - standardised mortality ratio
** The ERG's analysis indicates a different result to that reported in the CS*

5.3 Critical appraisal of the company's health economic analysis

The ERG adopted a number of approaches to explore, interrogate and critically appraise the company's submitted economic analyses and the underlying health economic model upon which this was based. These included:

- Consideration of key items contained within published economic evaluation and health economic modelling checklists.[64, 65]
- Scrutiny of the company's model by health economic modellers and discussion of issues identified amongst the members of the ERG.
- Double-programming of the deterministic version of the company's model to fully assess the logic of the company's model structure, to draw out any unwritten assumptions and to identify any apparent errors in the implementation of the model.
- Examination of the correspondence between the description of the model reported in the CS[1] and the company's executable model.
- Replication of the base case results, PSA, DSAs and scenario analyses presented within the CS.[1]
- Where possible, checking of key parameter values used in the company's model against their original data sources.
- The use of expert clinical input to judge the credibility of the company's economic evaluation and the assumptions underpinning the model.

5.3.1 Model verification

The ERG rebuilt the deterministic version of the company's original base case model in order to verify its implementation. As shown in Table 31, the ERG's results are similar to those generated using the company's model. During the process of rebuilding the model, the ERG identified several programming errors as well as conceptual issues relating to the model structure and its use of evidence; these are discussed in Section 5.3.3. The ERG is confident that all significant implementation errors within the company's model have been identified.

Table 31: Comparison of company's base case model and ERG's rebuilt model results, deterministic (results exclude the correction of model errors identified by the ERG)

	Company's model			**ERG's rebuilt model***		
Gilteritinib versus salvage chemotherapy blended comparator						
Model outcome	**Gilteritinib**	**Comparator**	**Inc.**	**Gilteritinib**	**Comparator**	**Inc.**
LYGs	3.03	1.75	1.28	3.01	1.72	1.29
QALYs	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
Costs	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
ICER	-	-	**£47,695**	-	-	**£48,342**
Gilteritinib versus BSC						
Model outcome	**Gilteritinib**	**Comparator**	**Inc.**	**Gilteritinib**	**Comparator**	**Inc.**
LYGs	3.03	0.33	2.70	3.01	0.32	2.69
QALYs	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
Costs	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
ICER	-	-	**£35,778**	-	-	**£36,356**

ERG - Evidence Review Group; ICER - incremental cost-effectiveness ratio; Inc. – incremental; LYG - life year gained; QALY - quality-adjusted life year

** Excludes minor correction to BSC group applied in company's updated model*

5.3.2 Adherence to the NICE Reference Case

The company's economic analysis is generally in line with the NICE Reference Case[66] (see Table 32). The most notable deviation from the scope relates to the comparators included in the company's economic analysis; this is discussed in Section 5.3.3.

Table 32: Adherence of the company's economic analyses to the NICE Reference Case

Element	Reference case	ERG comments
Defining the decision problem	The scope developed by NICE	The company's health economic analysis is generally in line with the final NICE scope.[6] The economic analyses relate to the ADMIRAL ITT population.[7] The ERG's clinical advisor believes that the ADMIRAL trial broadly represents the patient population seen in clinical practice in England.
Comparator(s)	As listed in the scope developed by NICE	The comparators considered within the CS[1] are not fully consistent with the final NICE scope:[6] • The model includes low-dose cytarabine, whereas the NICE scope refers to intermediate-dose cytarabine • The model includes hydroxycarbamide as part of BSC, whereas the NICE scope lists hydroxycarbamide and BSC as distinct comparators • The model includes azacitidine and MEC, but neither of these regimens are listed in the NICE scope. The company's base case analysis uses a blended comparator of salvage chemotherapies as the main comparator. Pairwise comparisons of gilteritinib versus individual chemotherapy regimens and gilteritinib versus BSC are included as secondary analyses.
Perspective on outcomes	All direct health effects, whether for patients or, when relevant, carers	Health gains accrued by patients are valued in terms of QALYs gained. Impacts on caregivers are not included.
Perspective on costs	NHS and PSS	The analysis adopts an NHS and PSS perspective.
Type of economic evaluation	Cost-utility analysis with fully incremental analysis	The results of the company's base case analysis are presented in terms of the incremental cost per QALY gained for gilteritinib versus the salvage chemotherapy blended comparator.
Time horizon	Long enough to reflect all important differences in costs or outcomes between the technologies being compared	The model adopts a 40-year time horizon. At this point, approximately 100% of patients in the model have died.

Element	Reference case	ERG comments
Synthesis of evidence on health effects	Based on systematic review	Prior to the assumed 3-year cure point, time-to-event outcomes (EFS and OS) for patients without HSCT, HRQoL estimates and AE frequencies for patients receiving gilteritinib and salvage chemotherapy are based on data from ADMIRAL;[7] this was the key study included in the company's systematic review of clinical evidence. Outcomes for patients who receive HSCT after initiating gilteritinib or chemotherapy are based on external literature.[31, 32] The assumed 3-year cure point is not based on analyses of ADMIRAL. Instead, this is based on TA523,[37] published literature[38-40] and clinical advice.[1]
Measuring and valuing health effects	Health effects should be expressed in QALYs. The EQ-5D is the preferred measure of HRQoL in adults.	Prior to the cure point, health state utility values are based on EQ-5D-5L data collected in ADMIRAL[7] (mapped to the EQ-5D-3L). HRQoL for cured patients is also based on EQ-5D-3L data.[35] HRQoL losses due to HSCT and AEs are based on literature,[41-44] and ADMIRAL.[7] Not all of the published AE disutility sources report EQ-5D valuations.
Source of data for measurement of HRQoL	Reported directly by patients and/or carers	
Source of preference data for valuation of changes in HRQoL	Representative sample of the UK population	
Equity considerations	An additional QALY has the same weight regardless of the other characteristics of the individuals receiving the health benefit	No additional equity weighting is applied to estimated QALY gains.
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	Resource costs include those relevant to the NHS and PSS. Unit costs were valued at 2017/18 prices.
Discount rate	The same annual rate for both costs and health effects (currently 3.5%)	Costs and health effects are discounted at a rate of 3.5% per annum.

AE - adverse event; BSC - best supportive care; CS - company's submission; EFS - event-free survival; ERG - Evidence Review Group; EQ-5D - Euroqol 5-dimensions; HRQoL - health-related quality of life; HSCT - haematopoietic stem cell transplant; ITT - intention-to-treat; MEC - mitoxantrone, etoposide and cytarabine; OS - overall survival; PSS - Personal Social Services; QALY - quality-adjusted life year; TA - technology appraisal

5.3.3 Main issues identified within the critical appraisal

Box 2 summarises the main issues identified within the ERG's critical appraisal of the company's economic analyses. These issues are discussed in further detail in the subsequent sections.

Box 2: Main issues identified within the critical appraisal undertaken by the ERG

(1) Identification of model errors
(2) Concerns regarding comparators
(3) Concerns regarding the company's model structure
(4) Uncertainty surrounding the cure assumption and the use of external data
(5) Issues relating to the company's survival modelling
(6) Issues relating to the company's indirect comparisons
(7) Concerns regarding model-predicted OS
(8) Inappropriate approach used to estimate drug costs
(9) Issues relating to adverse events
(10) Issues relating to health state costs
(11) Approach to handling uncertainty

(1) Identification of model errors

(i) Incorrect estimation of general population mortality risk

The company's model estimates mortality risk for patients who are cured at 3-years by applying an SMR to general population mortality rates obtained from interim life tables[33] (model worksheet "Life Table", cells I19:S121). The model estimates mortality rates for women and men separately and applies a constant weighting for each age based on the proportions of women and men at baseline in ADMIRAL.[7] The overall death rate for each age is then calculated as the weighted mean of these sex-specific rates; this overall mortality rate is then uplifted by an SMR of 2.0 and converted back to a monthly probability, assuming a constant event rate.

The ERG considers this approach to be incorrect because it assumes that the ratio of women to men will remain constant with increasing age, yet the life tables show that women and men have different mortality risks at each age. The ERG believes that a more appropriate approach would involve generating separate SMR-adjusted survivor functions for women and men, applying an initial women-to-men weighting at time t_0 (age 59 years), and then calculating the overall mortality risk in each period based on the sum of survival probabilities from each weighted model. Under this approach, the proportions of surviving women and men do not remain constant as they are influenced by sex-specific mortality risks. The ERG's correction slightly increases the expected survival for patients reaching the

cure point in the model; this reduces the deterministic ICER for gilteritinib versus the salvage chemotherapy blended comparator from £47,695 to £46,928 per QALY gained.

The ERG notes that applying this same change to the company's updated model produces a different result (a lower ICER of £46,138 per QALY gained). This is unexpected given that the original and updated models produce the same base case ICER (reported in Table 28), and the company's additional clarification response[67] does not indicate that any other changes to the deterministic model have been made. In the absence of an explanation for this difference, all subsequent analyses presented by the ERG relate to the original version of the model.

(ii) Incorrect application of HR for OS in patients receiving gilteritinib maintenance therapy

The company's model applies an HR (based on an indirect comparison of ADMIRAL[7] and Evers *et al*[31]) to the OS functions for patients in the With HSCT sub-model; this is used to reflect an additional survival benefit associated with gilteritinib maintenance therapy versus no maintenance therapy post-HSCT (model worksheet "Effectiveness_calculation", column AJ). The company's model calculates the cumulative survival probability in each cycle for patients in the gilteritinib With HSCT group as a function of the cumulative survival probability in the previous cycle multiplied by one minus the probability of death in the current cycle, which in turn, is calculated as:

$$1\text{-p(maintenance)} \cdot \text{p(deathHSCT)} + \text{p(maintenance)} \cdot \text{p(deathHSCT)}^{HR\_OS} \quad [i]$$

Where: p(maintenance) is the probability of receiving gilteritinib maintenance therapy [probability=[redacted]]; p(deathHSCT) is the probability of dying after HSCT within current cycle t_j, and HR_OS is the estimated gilteritinib maintenance therapy treatment effect [HR=[redacted]]

Notwithstanding the ERG's concerns regarding the reliability of this estimated treatment effect for gilteritinib maintenance therapy (discussed in critical appraisal point [6]), the ERG believes that the calculations in which it is applied in the model are incorrect. This is because the company's approach assumes a constant weighting over time – the proportion of surviving patients with HSCT who will receive maintenance therapy is assumed to be [redacted] at every time *t*. However, given that gilteritinib maintenance therapy is assumed to be associated with an OS benefit, it then follows that over time, the surviving cohort will include more patients who are receiving maintenance therapy (because their survival prognosis is assumed to be better) and fewer patients who are not receiving maintenance therapy (because their survival prognosis is assumed to be worse). Assuming it is reasonable to include this treatment effect, the ERG's preferred approach would involve modelling the two distributions

separately with an initial weighting of [redacted] at the time of HSCT, thereby reflecting the proportion of patients who initiate gilteritinib maintenance therapy. This correction has only a minimal impact on the model-predicted survival for patients undergoing HSCT in the gilteritinib group. Incorporating this correction into the company's model reduces the deterministic ICER for gilteritinib versus salvage chemotherapy from £47,695 to £47,172 per QALY gained.

As a separate point, the ERG notes that the company's Gompertz model indicates a mortality rate of close to zero after around 12 years. This lacks face validity, although as noted previously, the survival probabilities from this model are not used after the 3-year cure point. This is discussed in further detail in critical appraisal point [5].

(iii) Sum of health state occupancies is greater than 1.0

Within the model traces for the No HSCT sub-models, the sum of the probabilities of being in any health state is greater than 1.0 from year 35 onwards. This error indicates a logical inconsistency, whereby after year 35 the probability of being alive and event-free is greater than the probability of being alive (i.e. $S(t_{35})_{PFS}>S(t_{35})_{OS}$). This also implies that even after the cure point, patients can still experience relapse/progression; this is inconsistent with the company's notion of a fixed timepoint at which patients are cured. This error affects the overall model traces for all active treatment options; the BSC comparator is unaffected as these patients enter the model in the post-event state. This error could have been avoided by constraining the cumulative survival probability for EFS at time *t* by the maximum of the cumulative survival probabilities for EFS and OS at each time *t*. Incorporating this correction into the company's model has only a negligible impact on the model results (ICER increased by £0.03).

This logical inconsistency does not apply to the With HSCT group: after the 3-year cure point in this sub-model, the model calculations force the cumulative probability of EFS to remain constant until it reaches the cumulative probability of OS. Whilst not discussed in the CS,[1] this approach implies that only those patients who have previously relapsed/progressed are at risk of death; patients who are event-free cannot die during this interval. Whilst the validity this assumption is questionable, this approach protects against the logical inconsistency error described above.

(iv) Incorrect application of the disutility for HSCT

The company's model applies a disutility of [redacted] for HSCT for six months following HSCT, based on a TTO valuation study reported by Joshi *et al*;[29] this health loss is applied to all surviving patients in the With HSCT sub-model from months [redacted] for the gilteritinib group and from months [redacted] in the salvage

chemotherapy group. Elsewhere in the company's utility calculations, these same patients gain QALYs during this 6-month post-HSCT interval by virtue of being alive and residing in the event-free/post-event health states. The ERG notes that the value reported by Joshi *et al*[29] is describing a health state utility value associated with SCT – valued as being worse than death – rather than a disutility relative to some non-zero baseline utility. The company's model is therefore subject to an error, as the company should have applied the value of [redacted] as an absolute utility value. This error could be corrected by setting the utility values for the With HSCT health states equal to zero during the 6-month post-HSCT interval. Applying this correction increases the ICER for gilteritinib versus salvage chemotherapy from £47,695 to £49,231 per QALY gained. However, the ERG notes that based on the company's utility estimates obtained from their GEE model and the estimates reported in studies included in their HRQoL review, the utility value reported by Joshi *et al*[29] appears to be an outlier. Other literature reporting on post-HSCT states in patients with AML (e.g. Grulke *et al*[68] and Kurosawa *et al*[69]) report positive values which are similar to those obtained from the company's GEE model. As such, the ERG has not applied this correction to the company's model in its exploratory analyses. The ERG also notes that this is not a key driver of the cost-effectiveness of gilteritinib.

(v) Utility for long-term AML survivors inconsistent between CS and model

Within the first 3 years of the time horizon, health state utility values are based on the estimates derived from the company's GEE model fitted to the ADMIRAL data.[7] As shown in Table 23, the utilities for the event-free and post-event health states differ between the With HSCT and No HSCT sub-models. After this 3-year timepoint, health utilities for all states are based on age-adjusted values estimated using Janssen *et al*.[35]

Figure 13 shows the company's modelled utility profile together with the estimates of general population utility by age band reported by Janssen *et al*[35] (the grey stepped function) and the model-based estimates of general population utility reported by Ara and Brazier[70] (the smooth decreasing function, not used in the company's model). As shown in the figure, utility values are state-dependent during the first 3 years of the time horizon. At all subsequent timepoints, modelled utility for all health states increases to 1.0, before dropping to [redacted] and [redacted] at age 65 and age 75, respectively; between these age intervals, the model assumes that health utility is constant.

Figure 13: Company's modelled utility profile versus general population EQ-5D-3L estimates

HSCT - haematopoietic stem cell transplant

With respect to the company's utility assumptions, the ERG notes the following:

- Within the company's base case model, patients who have FLT3+ R/R AML and survive for 3 years subsequently experience a level of HRQoL which is considerably higher than that of the general population for the remainder of their lives; this is shown as the difference between the red line and either of the dashed lines in Figure 13. This error is the consequence of two factors: (i) before adjusting for age, all cured patients are assigned a utility value of 1.0, and; (ii) the company's model uses the utility value for the youngest age band in Janssen *et al*[35] (55-64 years) as the denominator for the calculation of age-adjustment utility multipliers applied to each age band. This error could have been avoided by using the general population utility values directly as reported (without applying multipliers to some age-band specific baseline).
- The CS[1] (page 84) states that the utility value for long-term AML survivors is *"assumed equal to EFS with HSCT."* This suggests that after 3-years, the company should have applied a utility value of [redacted] rather than 1.0 to all surviving patients. However, if the company had applied this lower value in the model, the long-term utility projection would have remained higher than that of the general population.

- The UK values from Janssen *et al* used in the model have been derived from the York Valuation Study of the EQ-5D-3L (Kind *et al*[62]). These data were collected more than 25 years ago. In addition, as the data are reported only by age band, they fail to fully account for changes in population health with increasing age (e.g. the utility value for patients aged 75+ is applied for 25 years of the 40-year model time horizon). In contrast, the analysis reported by Ara and Brazier[70] uses more recently collected data from the Health Survey for England (HSE, years 2003 and 2006), includes a large number of respondents (n=26,679) and provides utility estimates for each individual age; the ERG believes that this represents a more appropriate source.

The ERG believes that given the company's assumption of a fixed cure point for all patients surviving to 3-years, HRQoL after this point should be valued using the model reported by Ara and Brazier[70] directly. Applying these values to the model increases the company's base case ICER from £47,695 to £55,949 per QALY gained. The ERG notes that this approach leads to a situation whereby the utility value for patients with prior progression "jumps" at 3-years – this is unlikely to be realistic but is a consequence of the company's assumption of cure despite prior relapse/progression.

(2) Concerns regarding comparators

The salvage chemotherapy regimens included in the company's model are not fully consistent with those listed in the final NICE scope[6] (see Table 33). The main differences are:

- The model includes low-dose cytarabine, whereas the NICE scope includes intermediate-dose cytarabine;
- The model includes hydroxycarbamide as part of BSC, whereas the NICE scope lists hydroxycarbamide and BSC as distinct comparators;
- The model includes azacitidine and MEC, yet neither of these regimens are listed in the NICE scope.

Table 33: Comparators included in final NICE scope, the ADMIRAL trial and the company's model

Comparator	NICE scope[6]	ADMIRAL[7]	Company's model[1]
Azacitidine	No	Yes	Yes*
IDAC	Yes	No	No
LoDAC	No	Yes	Yes*
FLAG-IDA	Yes	Yes	Yes*
MEC	No	Yes	Yes*
BSC	Yes	No	Yes (included as separate comparator)
Hydroxycarbamide	Yes	No	Yes (included as part of BSC)

BSC - best supportive care.
** Regimen included in company's salvage chemotherapy blended comparator*

Despite the differences between the model and the NICE scope,[6] the ERG's clinical advisor stated that the regimens included as comparators in the model are broadly representative of current clinical practice in England. The ERG's clinical advisor also noted that ADMIRAL included chemotherapy regimens which would be offered to patients who may be fit enough to undergo HSCT (FLAG-IDA and MEC) as well as regimens which would be offered to patients for whom the clinical intent is to achieve disease control without proceeding to HSCT (azacitidine and LoDAC).

The ERG also notes that the interpretation of the pairwise comparisons of gilteritinib versus individual salvage chemotherapy comparators presented in the CS is problematic as these comparisons assume equivalent effectiveness between regimens (including the probability that the patient will subsequently undergo HSCT) and are not based on data specific to each regimen. This assumption of equivalence is unlikely to be appropriate, as the choice of salvage chemotherapy is typically guided by the patient's fitness, the toxicity of the regimen and whether the patient is considered eligible for HSCT. In response to a request for clarification from the ERG[7] (question B14), the company provided additional information regarding the proportion of patients receiving each regimen who subsequently received HSCT. These data show that of the [redacted] patients in the salvage chemotherapy group who subsequently received HSCT, [redacted] of these ([redacted]) had received either FLAG-IDA or MEC; only [redacted] of [redacted] patients ([redacted]) who received LoDAC or azacitidine subsequently received HSCT. The ERG also notes that the EFS and OS survivor functions and AE profiles are likely to be different for each salvage chemotherapy regimen. As such, the company's pairwise comparisons of gilteritinib versus individual salvage chemotherapy regimens are unlikely to be reliable. Whilst the ERG believes that some caution should be exercised in interpreting the results of economic analyses based on blended comparisons, the ERG agrees that this represents the most appropriate approach in this case.

(3) Concerns regarding the company's model structure

The ERG has some concerns regarding the company's modelling approach. Conceptually, the company's approach bears some similarities to a landmark or response-based model, with the exception that the landmark timepoint (time of HSCT) is artificially assumed to occur at a fixed timepoint and is assumed to differ between the intervention and comparator groups. The ERG takes the general view that the most appropriate model structure should make the best use of the available evidence and should not impose inappropriate structural constraints which lead to bias in the model results. With respect to the company's model approach, the ERG believes that three structural issues warrant some discussion:

1. Are the sub-models (With HSCT and No HSCT) and the health states (event-free and post-event) appropriate?
2. How should time-to-HSCT be modelled?
3. Which approach should be used to estimate health state occupancy over time?

The ERG has no concerns regarding the definition of the model health states. Clinical advice received by the ERG suggests that patients who are alive and event-free will experience a higher level of HRQoL compared with those who have relapsed/progressed disease. This view is supported by the estimates derived from the company's GEE utility model (see Table 23). The ERG's clinical advisor also commented that treatment with gilteritinib would be discontinued upon relapse/progression. Furthermore, the ERG also believes that it is reasonable for the model to distinguish between patients who have HSCT and those who do not, because: (a) patients undergoing HSCT are expected to have different outcomes compared with those who do not receive HSCT, and (b) this makes it easier to estimate relevant outcomes and costs according to receipt of HSCT.

The company's model makes the structural assumption that patients who undergo HSCT do so at a fixed timepoint in the With HSCT sub-model (gilteritinib – [redacted]; salvage chemotherapy – [redacted]). These timepoints are based on the mean time to HSCT observed in ADMIRAL.[7] During the clarification process, the ERG requested additional information from the company regarding the distribution of time to HSCT within ADMIRAL (see clarification response,[7] question D3). The data provided by the company are presented in Figure 14; these data show that time to HSCT in ADMIRAL ranged from [redacted] in the gilteritinib arm and from [redacted] in the salvage chemotherapy arm.

Figure 14: Time to HSCT in ADMIRAL

[redacted]

HSCT - haematopoietic stem cell transplant.
Based on data provided in company's clarification response,[21] question D3, Table 16

Given the observed variability around the timing of HSCT, the ERG believes that applying a fixed HSCT timepoint represents an inappropriate structural assumption which is inconsistent with the available evidence, and which is unnecessary because ADMIRAL provides OS data from the point of randomisation. This aspect of the model may contribute to the apparent inconsistencies between the model-predicted OS and the observed OS in ADMIRAL; this inconsistency is further discussed in critical appraisal point [7].

With respect to the overall modelling approach adopted by the company, the CS[1] (page 49) notes that partitioned survival models avoid the need for assumptions regarding transitions between health states and allow the direct use of time-to-event outcomes from ADMIRAL to estimate health state occupancy. The CS also argues that state transition approaches (e.g. Markov or semi-Markov) *"require further assumptions to estimate transition probabilities, cannot incorporate time-varying transition probabilities for all considered health states and have stricter individual-level data requirements to use treatment arms not directly evaluated in ADMIRAL"* (CS,[1] page 50). In their clarification response[7] (question B4), the company provided further justification for their modelling approach. The company's response covers similar arguments to those presented in the CS, and argues that the only major assumption in the partitioned survival model structure is around the timing of HSCT.

The ERG notes the following regarding the company's choice of model structure:

- Contrary to the company's arguments regarding the benefits of partitioned survival models, the company's model does not directly use data from ADMIRAL[7] to inform outcomes for patients receiving HSCT.
- It is unclear which assumptions would be required to estimate transition probabilities, or why these might be considered inappropriate. If the company had adopted a state transition rather than a partitioned survival approach, this would have required a similar parametric model fitting exercise to that presented in the CS, albeit for individual transitions between health states, rather than for outcomes from the point of randomisation. The ERG notes that had a state transition approach been pursued, the survival analysis would have needed to account for competing risks between events (e.g. using a multistate model[71, 72]). This introduces an additional layer of complexity and there may be limited numbers of events with which to estimate certain transitions.
- It is possible to reflect time-varying transition probabilities in any health state within a state transition-based model (e.g. through the use of a semi-Markov structure).
- The only treatment arm included in the model which was not evaluated in ADMIRAL is BSC. For this treatment group, the company's partitioned survival model assumes that all patients

have relapsed/progressed disease at the point of model entry, hence the only possible transition is from the post-event state to death. It is unclear what additional individual-level data would be required to incorporate this option or why this would pose a problem.

- The application of a fixed cure point within a partitioned survival model structure means that all patients who survive up to that timepoint are assumed to be cured, irrespective of whether they have previously relapsed/progressed; this can be seen in the proportions of modelled patients who are estimated to achieve cure (see Table 34). The ERG considers this to be clinically unrealistic, but notes that this type of assumption is unavoidable in a partitioned survival model structure with a fixed cure point, as mortality risk is not specified separately for each health state. This assumption could have been avoided through the use of a state transition model through the inclusion of a separate "cured" state and by only permitting transitions into that state from the event-free health state(s).
- The company's assumption of a fixed HSCT timepoint is not a consequence of the company's decision to adopt a partitioned survival approach. Rather, it appears to be a consequence of the company's decision to use external data relating to time from HSCT to death (from Evers *et al*[31]). The observed distribution around time to HSCT could have been reflected in a partitioned survival structure by modelling OS outcomes for the With HSCT group in ADMIRAL from the point of randomisation, or by modelling a separate transition between "event-free" and "post-HSCT" states within a state transition-based model.

Table 34: Probability of being cured at 3-years in the company's model

Model health state	Gilteritinib	Salvage chemotherapy
No HSCT, event-free	[redacted]	[redacted]
No HSCT, post-event	[redacted]	[redacted]
With HSCT, event-free	[redacted]	[redacted]
With HSCT, post-event	[redacted]	[redacted]

HSCT - haematopoietic stem cell transplant

Overall, the ERG considers that the company's model 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. The former assumption could have been avoided through the use of a state transition model; however, this has not been done and the impact of using such an approach on the ICER is unclear. The latter assumption is inconsistent with the data from ADMIRAL (see Figure 14) and could have been avoided by using data on OS from randomisation in ADMIRAL. The impact of this is assessed within the ERG's exploratory analyses (see Section 5.4).

(4) Uncertainty surrounding the cure assumption and the use of external data

The ERG believes that there is uncertainty surrounding the potential for gilteritinib to provide a cure and the way in which this assumption is applied in the company's model.

ADMIRAL is not used to inform assumptions of cure

The company's model includes a structural assumption of cure for all patients who remain alive at the 3-year timepoint. The CS[1] does not present any evidence from ADMIRAL to suggest that patients in the target population may achieve cure. Rather, the source of the cure assumption is cited as the Appraisal Committee's comments in the previous NICE midostaurin appraisal (TA523[37]), the company's interpretation of published literature[38-40] and clinical advice.[1] The company have not modelled cure (e.g. using a mixture-cure model) using data from ADMIRAL or any other evidence source.

Evidence for the ADMIRAL With HSCT group is not used in the company's model

Whilst the CS[1] (page 14) states that HSCT is *"the only treatment with curative intent"*, data for these patients are not presented in the CS and they are specifically excluded from the model as they were considered to be "*immature"*, with limited follow-up (median follow-up post HSCT [redacted]) and a limited sample size. In response to a request for clarification from the ERG[7] (question D2), the company provided additional data on the proportions of patients in ADMIRAL who received HSCT and who were alive or dead at the final data cut-off (see Table 35). Following a further data request from the ERG,[67] the company also provided Kaplan-Meier survival functions for patients who received HSCT in ADMIRAL; these are presented here in terms of survival from the point of randomisation (Figure 15) and survival from the point of HSCT receipt (Figure 16).

Table 35: Survival status amongst those receiving HSCT in ADMIRAL at final data cut-off (adapted from company's clarification response, question D4, with corrections by the ERG)

Survival status	Gilteritinib	Percent	Salvage chemotherapy	Percent
Alive	[redacted]	[redacted]	[redacted]	[redacted]
Dead	[redacted]	[redacted]	[redacted]	[redacted]
Total	[redacted]	[redacted]	[redacted]	[redacted]

Figure 15: Time from randomisation to death in patients who received HSCT in ADMIRAL (reproduced from company's additional data request, Figure 1)

Figure 16: Time from HSCT to death in patients who received HSCT in ADMIRAL (reproduced from company's additional data request, Figure 2*)

HSCT - haematopoietic stem cell transplant.
** The ERG notes that the data provided by the company on post-HSCT OS in ADMIRAL (Figure 10) was incorrectly analysed and should be disregarded*

The ERG notes that the data underlying the Kaplan-Meier survivor functions for the salvage chemotherapy group are sparse and include only [redacted] events. However, the Kaplan-Meier survivor functions for the gilteritinib group include more events; as shown in the Kaplan-Meier functions, [redacted] gilteritinib-treated patients who received HSCT in ADMIRAL[7] had already died at the data cut-off. The ERG notes that the Kaplan-Meier survivor function for the gilteritinib group does not indicate the classic pattern of survival that is indicative of the presence of a cure (e.g. a plateau at the end of the survival function) and much of the censoring applies across the whole time period. Whilst it is possible that further long-term follow-up may change the shape of these survivor functions, one would not necessarily expect the shape of the distributions to shift systematically. The ERG also notes that despite the very limited number of events in the salvage chemotherapy group, there does not appear to be conclusive evidence to support the hypothesis that gilteritinib maintenance therapy is associated with an additional OS benefit. As part of the company's response to the ERG's additional data request[67] (question 1), the company stated: *"We believe if the salvage chemotherapy treated patients were followed up for longer, a steeper curve would be seen which would separate from the gilteritinib curve. Astellas acknowledges that the curve provided cannot substantiate this."* The company did not comment on whether they expected the shape of the distribution for the gilteritinib group to change with additional follow-up. Disappointingly, the company's clarification response[7] (question A10) states that the latest data-cut of ADMIRAL is final; further analysis of longer-term follow-up data is not anticipated.

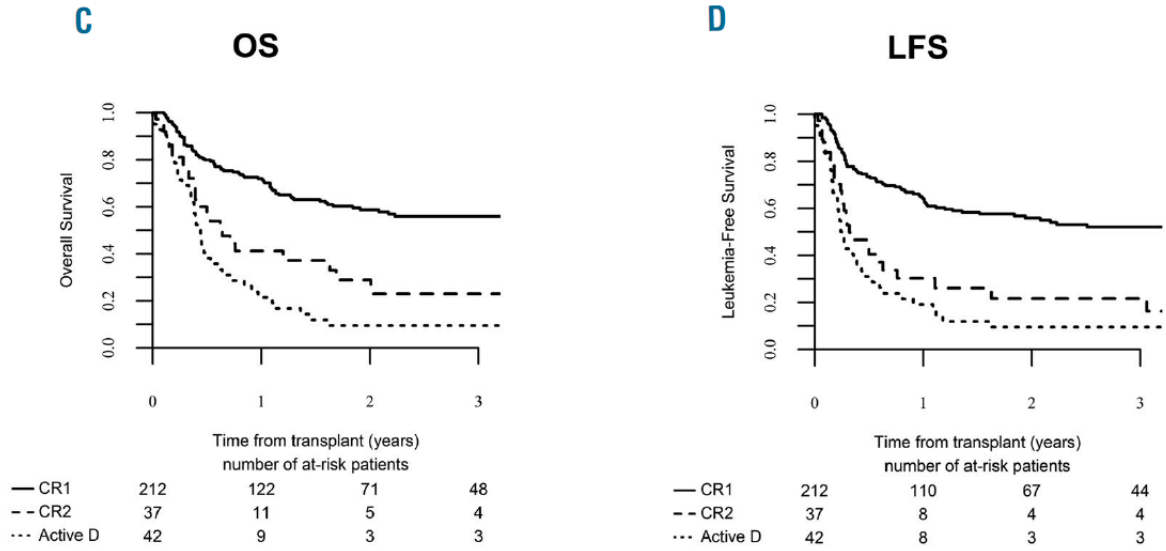
Concerns regarding the relevance of external evidence used to inform post-HSCT outcomes

Within the company's model, post-HSCT OS is based on a survival model fitted to data from Evers *et al*.[31] These data are used instead of ADMIRAL for patients undergoing HSCT. The data from Evers *et al* relate to patients with R/R AML in CR2; however, these patients did not specifically have FLT3 mutation-positive disease. The data used to inform the HR for post-HSCT EFS are based on Ustun *et al;*[32] these patients had FLT+ disease, but 80% of the URD group were in CR1 rather than CR2. The CS argues that FLT3 mutation status does not impact survival outcomes post-HSCT, although the ERG notes that this view is based on a study in which the majority (79%) of patients were in CR1 (Deol *et al*[60]). None of the studies included in the company's targeted review of post-HSCT outcomes directly relate to the target population of interest (FLT3+ R/R AML).

The ERG's clinical advisor identified a further published study which reports on post-HSCT outcomes (LFS and OS) for patients with FLT3+ R/R AML in CR2 (Poiré *et al*[73]). The ERG notes however that the sample size for the CR2 subgroup in Poiré *et al* is small (n=37) and follow-up is limited to around

3 years. This subgroup better reflects the post-HSCT population of ADMIRAL, but was not included in the company's targeted review.

Figure 17: LFS and OS outcomes reported in Poiré *et al* (FLT3+ R/R AML patients)



CR1 - first complete remission; CR2 - second complete remission; LFS - leukaemia-free survival; OS - overall survival

Implications for the company's economic analysis

The ERG notes the following implications for the company's economic analysis:

- The ERG's clinical advisor stated that most relapses occur in the first 2 years after transplant. For a patient who remains event-free at 3-years, the probability of subsequent relapse is unlikely. This provides some justification for the company's assumption of cure; however, this has not been explored using statistical (e.g. mixture-cure) models. The ERG also notes that cure fractions, if present, are likely to be specific to the population under consideration.
- Post-HSCT OS is a key driver of the ICER for gilteritinib.
- Overall, the ERG does not consider it appropriate to ignore evidence from the ADMIRAL trial,[7] irrespective of the number of events or the duration of post-HSCT follow-up. If external information is considered relevant for inclusion in the model, then it should supplement the evidence from ADMIRAL, not replace it. If multiple sources of evidence are considered relevant, the company could have synthesised these and, in the event of heterogeneity, made use of the predictive distribution.
- The company's model suggests that approximately [redacted] of gilteritinib-treated patients who undergo HSCT will be cured. This is a consequence of the use of Evers *et al* together with an assumption of a fixed 3-year cure point.

- The available data from gilteritinib-treated patients who received HSCT in ADMIRAL suggest a cumulative OS probability of [redacted] at around [redacted] (the last observed event).
- At the final data cut-off of ADMIRAL, [redacted] of all gilteritinib-treated patients who received HSCT had already died; in order to achieve the 3-year cure rate estimated by the company's model, this would require the vast majority of surviving (censored) patients in the ADMIRAL gilteritinib-treated HSCT group to be cured. The ERG considers this to be unlikely. The inclusion of the observed data for the HSCT group in ADMIRAL within the model would substantially increase the ICER for gilteritinib relative to the company's base case estimate.
- The study reported by Poiré *et al*[73] appears to better reflect the target population than all 10 studies included in the company's review, as it relates specifically to patients with FLT3+ R/R AML. However, the CR2 population in this study is small and there may be other differences in patient characteristics compared with ADMIRAL. This study indicates a lower 3-year survival probability of around 22% (see Figure 17, Panel C). The inclusion of these data in the model would also increase the ICER for gilteritinib relative to the company's base case estimate.

Overall, the ERG believes that ADMIRAL reflects the most relevant source of evidence on the expected outcomes for gilteritinib-treated patients who undergo HSCT. The impact of including these data in the company's model is assessed within the ERG's exploratory analyses (see Section 5.4).

(5) Issues relating to the company's survival modelling

Survival modelling for No HSCT EFS and OS

The ERG considers that the description of the analysis of EFS and OS for the No HSCT group is well written and is generally in line with TSD 14.[52] A Gompertz distribution and standard parametric distributions that are members of the Generalised F distribution were fitted. Separate distributions were fitted to data for each treatment arm. Although this is a common approach to survival modelling in HTA, it ignores correlation between parameters across treatments.

The OS data are reasonably mature and require very little extrapolation of survival functions. The ERG is satisfied with the choice of the log logistic distribution for the base case based on BIC for OS for both gilteritinib and salvage chemotherapy. As noted in Section 5.2.4, extrapolation of the OS survival distributions plays a minor role within the model due to the assumption of a fixed cure point.

The company's base case model for gilteritinib was the log logistic distribution; this model had the lowest BIC, provided a reasonable representation of the sample survival function based on visual

inspection and the log cumulative hazard plot indicated a non-monotonic hazard pattern over time. The company also chose to model the data from the salvage chemotherapy arm using a log logistic distribution. The ERG notes that whilst this is the same model form as the gilteritinib arm, there is relatively little sample evidence with which to model long-term EFS and there is structural uncertainty about which is the most appropriate model.

Gompertz for OS post-HSCT

The best fitting model based on BIC was a generalised gamma distribution; this also appeared to provide the best representation of the sample data based on visual inspection. However, EFS data were not available from Evers *et al*.[31] In order to generate an EFS survival function, the company made the assumption that the hazard for OS was proportional to the hazard for EFS, and estimated the HR using evidence from Ustun *et al*.[32] Given that a generalised gamma distribution is not a PH model, the company chose to model OS from Evers *et al* using a Gompertz distribution. The ERG does not consider that a Gompertz distribution provides a reasonable representation of the data even within the first three years. Furthermore, the proportional relationship between the hazards for OS and EFS is a modelling assumption that has been justified based on clinical inputs and evidence that EFS is highly correlated with OS. The ERG notes that no information was provided from Ustun *et al*[32] in the CS to support the assumption of PH between EFS and OS, and that the HR was calculated using data from this study up to five years rather than three years. Furthermore, the application of a constant HR to estimate EFS (using Uston *et al*[32]) does not appear to be appropriate based on Figure 18.

Figure 18: Log-cumulative hazard plot for HSCT efficacy sources (reproduced from company's clarification response, Figure 1)

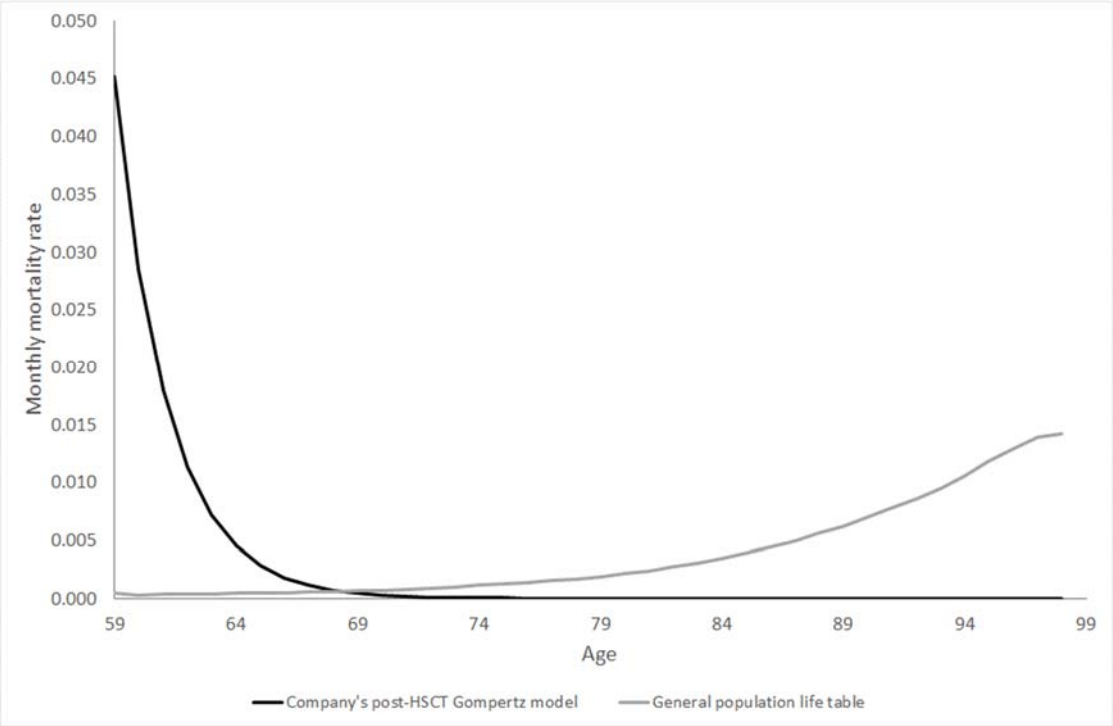
EFS - event-free survival; OS - overall survival

The ERG notes that the company has not attempted to model cure using any source (the ADMIRAL ITT population,[7] the subset of patients in ADMIRAL who underwent HSCT, or Evers *et al*[31]). The ERG suggests that there is visually little evidence of a cure in the Evers *et al* data until at least beyond month 50 rather than month 36 as assumed by the company. This may reflect important differences in patient characteristics between the published literature to which the company refers and the evidence from Evers *et al*. Furthermore, the prognosis may be worse in FLT3+ patients; this may affect the long-term survival of patients.

In response to the ERG's clarification letter[7] (question B16), the company suggested that it is the extrapolated survival function that is more important when assessing relative cost-effectiveness, and that external data and clinical expert opinion suggested that a Gompertz distribution provides a plausible extrapolation. The ERG agrees that the plausibility of the extrapolated survivor function is important, but contests the company's view that the Gompertz provides a plausible extrapolation. As previously shown in Figure 10, the modelled post-HSCT Gompertz function suggests that the hazard of death decreases with increasing age, and the survivor function becomes almost flat after around 9 years.

Figure 19 presents a comparison of the monthly mortality rates predicted by the company's post-HSCT Gompertz model and the approximate equivalent rates from general population life tables. As shown in the figure, the mortality rate from the Gompertz becomes lower than that for the general population after 9 years (around age 68 in the model). The extrapolated post-HSCT Gompertz function (excluding the 3-year cure assumption) therefore suggests a survival prognosis for R/R AML patients which is better than the survival prognosis in the general population; the ERG does not consider this to be plausible.

Figure 19: Comparison of monthly mortality rates from post-HSCT Gompertz model and general population life tables



HSCT - haematopoietic stem cell transplant

Given that the Gompertz distribution does not provide a good representation of the data reported by Evers *et al*[31] prior to the assumed cure point, or a plausible extrapolation after it, the ERG believes that the use of this model is inappropriate.

(6) Issues relating to the company's indirect comparisons (estimation of outcomes for BSC and estimation of the relative benefit of gilteritinib maintenance therapy)

HR for gilteritinib versus BSC

The ERG has concerns regarding the company's approach to comparing gilteritinib with BSC. The company conducted a literature review to identify studies that reported outcomes for BSC in a comparable patient population to the target population. The company selected Sarkozy *et al*[30] *"as the most relevant publication because it evaluated efficacy of BSC in a comparable population (i.e. AML patients in first relapse including patients with and without FLT3 mutation positive) and included a large sample size of patients who received BSC (N=124)"* (CS,[1] page 64). It is not clear whether the company found any other evidence that might also have been relevant.

Sarkozy *et al*[30] does not report a survival function for BSC and so the company derived an HR in an indirect comparison with ADMIRAL.[7] In response to a request for clarification from the ERG,[7]

(question D7), the company stated that *"the calculation of 2.86 for OS between gilteritinib and BSC was based on naïve comparison. Specifically, we used the following steps:*

- *Estimate HR between LDAC and BSC by comparing the ratio of reported median in Sarkozy 2013: 1.75 [calculated as 5.6/3.2]*
- *Estimate HR between gilteritinib and salvage chemotherapy:* [redacted]
- *Estimate HR between gilteritinib and BSC: 1.75X1.637 = 2.86".*

The ERG notes the following:

- Estimating an HR based on a ratio of medians implicitly assumes that the times-to-event follow an exponential distribution
- The indirect comparison assumes that additivity equations apply and that LoDAC is equivalent to salvage chemotherapy
- The source of the HR comparing gilteritinib and salvage chemotherapy is unclear
- The use of an HR applied to a log logistic distribution is inappropriate because this is not a PH model
- Assuming PH is a modelling assumption that may not be appropriate.

Gilteritinib maintenance post-HSCT versus no maintenance post-HSCT

The company's model assumes that gilteritinib maintenance therapy is associated with additional survival benefits following HSCT. However, the CS[1] notes that the post-HSCT OS data from ADMIRAL[7] are immature. Consequently, the company estimated the OS benefit associated with post-HSCT gilteritinib maintenance therapy by the application of an HR. This was estimated using OS data from patients receiving gilteritinib maintenance after HSCT in ADMIRAL and OS data of CR2 patients from Evers *et al*.[31] The HR was applied to the predicted post-HSCT OS Gompertz model (fitted to data for the CR2 group from Evers *et al*) to reflect an additional benefit of gilteritinib maintenance therapy. The ERG has several concerns with this approach:

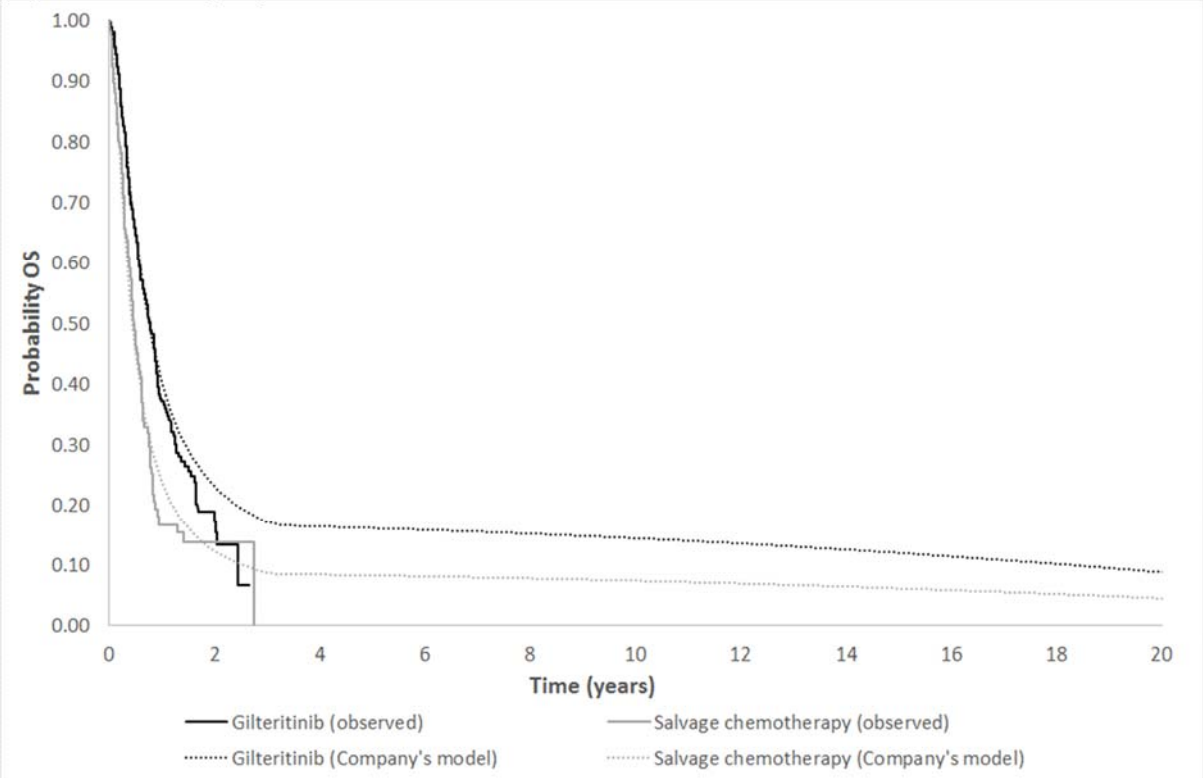
- In response to the ERG's clarification letter[7] (question B11), the company acknowledges that there is little evidence to support the PH assumption and that what evidence there is suggests that the assumption is violated
- The comparison is an unanchored naïve comparison between groups with no adjustment for differences in patient characteristics between studies
- Given that the CS argues that the OS data for patients receiving HSCT in ADMIRAL are immature and limited by sample size, hence their decision to use external data, it seems somewhat inconsistent to use a smaller subset of these patients to estimate an additional treatment effect associated with gilteritinib maintenance therapy

- The available data from ADMIRAL (see Figure 16) do not suggest conclusive evidence of a difference in post-HSCT OS between the gilteritinib and salvage chemotherapy groups
- Whilst this estimated treatment effect is subject to considerable uncertainty, this is not included in the PSA.

(7) Concerns regarding model-predicted OS

Figure 20 presents a comparison of the model-predicted and observed OS for the gilteritinib and salvage chemotherapy groups in the overall ITT population of ADMIRAL.[7] As shown in the figure, the company's OS projections appear to reflect the observed Kaplan-Meier survivor functions for the initial 8 months; at all subsequent timepoints, the company's model appears to overestimate survival in both treatment groups. In addition, the observed OS survivor functions intersect at around month 25; this is not reflected in the company's model predictions which instead suggest an indefinite separation between the treatment groups. This gap between the modelled OS survivor functions is maintained after year 3 as a consequence of the company's application of an assumed cure at this timepoint.

Figure 20: Overall survival, model-predicted versus observed, overall ADMIRAL population - company's model



The ERG believes that the apparent discrepancy between the observed Kaplan-Meier survivor functions and the company's modelled OS projections may be a consequence of the three factors:

(a) The inappropriate use of a fixed timepoint for HSCT instead of the use of a time-to-event distribution (as discussed in critical appraisal point [3]).
(b) The use of external data to model post-HSCT OS (Evers *et al*[31]) which are not consistent with the outcomes observed for patients undergoing HSCT in ADMIRAL[67] (as discussed in critical appraisal point [4]).
(c) The use of the Gompertz survivor function which provides a poor representation of the data prior to the assumed cure point (as discussed in critical appraisal point [5]).

(8) Inappropriate approach used to estimate drug costs

Drug acquisition and administration costs for gilteritinib and the chemotherapy comparators are applied as once-only costs during the first model cycle. The ERG notes that this approach is unconventional and is subject to several limitations:

(i) Discounting cannot be applied appropriately. As the costs are not modelled according to the distribution of the drugs consumed over time, it is not possible to apply discounting appropriately. As a consequence, the total drug costs will be overestimated. This will bias against gilteritinib as this treatment is expected to be given over a longer time period and because it is more expensive (per cycle) than salvage chemotherapy.
(ii) Gilteritinib treatment duration is underestimated. The number of cycles of gilteritinib included in the cost calculations is based on the final data-cut of ADMIRAL (data cut-off September 17, 2018). The CS states that [redacted] of patients receiving gilteritinib were still receiving treatment at the data cut-off. In reality, these patients will consume further doses of gilteritinib for an unknown duration, yet the model assumes that this will not lead to any additional costs over and above the mean estimate at the time of the final data cut-off. This leads to a bias in favour of gilteritinib. The ERG also notes that if it is reasonable to assume that cure occurs at 3 years, there should be no clinical rationale to continue gilteritinib treatment beyond this timepoint.
(iii) Treatment duration is not structurally linked to patients' modelled health status. The ERG's clinical advisor stated that gilteritinib would be discontinued upon disease relapse/progression. However, this is not explicitly reflected in the company's approach, and it is unclear which patients are receiving treatment or for how long. In addition, the ERG believes that patients receiving gilteritinib maintenance therapy after HSCT are likely to receive treatment over a longer duration than those patients who do not undergo HSCT; this is evident from the available data from ADMIRAL (see Figure 21), but is not reflected in the company's model.
(iv) The model excludes wastage. This is likely to underestimate costs, particularly within the gilteritinib group as treatment is expected to be prescribed in packs providing 28 days' supply.

Figure 21: Time to discontinuation of treatment discontinuation in gilteritinib-treated patients with and without HSCT (reproduced from company's additional clarification response, question 2, Figure 3)

The ERG believes it is likely that the joint impact of these issues are that the costs of gilteritinib are underestimated. The magnitude of this bias and its impact on the ICER is unclear.

(9) Issues relating to adverse events

Double-counting of progressive AML as an AE and a health state

As noted in Section 5.2.4, progressive AML is counted as an AE, resulting in a disutility and an additional cost. In addition, the model structure includes different costs and utilities for the event-free and post-event health states. This represents double-counting. In response to a request for clarification from the ERG,[7] (question B22), the company commented that progressive AML was included as an AE *"to be comprehensive and consistent with the CSR"*[7] and that setting the frequency of this event to zero decreases the ICER by only £68. The ERG considers that it is not appropriate to include disease progression as an AE but agrees that its impact is minor.

Potential underestimation of impact of HSCT on health outcomes and costs

The ERG has some concerns that the company's model may not fully account for the potential long-term health impacts associated with complications of HSCT. The model does not include any further

costs or health losses relating to AEs or complications resulting from the procedure, for example, graft versus host disease (GvHD). As noted in the ERG report for NICE TA523, GvHD is a life-threatening adverse event which affects approximately 40% of SCT recipients.[74] In response to a request for clarification from the ERG[7] (question B17), the company explained that the absence of these complications from the model was a consequence of these events not being commonly reported in the literature, and commented that *"some of the GvHD costs could be included in the HSCT procedure costs, and that the inclusion of a separate cost for GvHD would lead to double-counting"* (company's clarification response,[7] question B17). The company also commented that the median time to onset of GvHD is typically short (19 days). The ERG's clinical advisor agreed that the costs associated with treating GvHD occurring within the first three months after the procedure are likely to be included in the HSCT tariff cost.[36] However, a small proportion (10% to 15%) of patients may experience GvHD later and would require additional immunosuppression and other procedures; these costs are unlikely to be captured in the NHS tariff cost.

The ERG also notes that additional health losses associated with longer-term GvHD may not be included in the company's QALY estimates. The company could have used utility values for this AE directly from Joshi *et al*.[29] It is however unclear how much of this health loss would already be captured through the company's existing assumptions.

The ERG believes that this is likely to be a minor issue which will not markedly impact on the ICER for gilteritinib.

(10) Issues relating to health state costs

The ERG has concerns regarding some of the cost assumptions used in the model, although these are not likely to be key drivers of the ICER of gilteritinib:

- As discussed in Section 5.2.4, the company's model assumes that after the 3-year cure point, follow-up costs are zero. The ERG's clinical advisor commented that patients would continue to require long-term follow-up. For patients undergoing HSCT, long-term follow-up (6-months post-HSCT) would typically involve outpatient visits every 2 to 3 months in first year, every 3 to 6 months in the second year and subsequently annual review would be required indefinitely. The ERG's clinical advisor noted that whilst it is unlikely that patients without HSCT would achieve cure, long-term survivors (after 2 years) would typically require ongoing visits every 6-months.
- Post-progression treatment costs are applied after the assumed cure point. Given that the company's model assumes that all surviving patients are cured at 3-years, the ERG considers it

inconsistent to assume that these patients continue to accrue costs associated with relapse/progression after this timepoint.

- The model assumes that two R/R AML patients would need to be tested in order to identify one patient with FLT3+ disease (hence, the unit cost per FLT3+ mutation test is multiplied by 200%). However, the CS[1] (page 9) suggests that around 30% of patients have a FLT3 mutation. As all R/R AML patients will require re-testing, this suggests that more than three patients would need to be tested in order to identify one patient with FLT3+ disease, and the unit cost should be multiplied by 333%. As part of their clarification response[7] (question B24), the company provided additional analyses which indicate that FLT3+ re-testing costs have a negligible impact on the ICER for gilteritinib.
- The ERG's clinical advisor stated that patients receiving high-intensity salvage chemotherapy may incur substantial hospital costs whilst on treatment, including hospital days for drug administration, transfusion support, treatment of infections and monitoring. However, the model includes administration costs per treatment cycle and other hospitalisation costs over the longer period in which patients are event-free. The ERG would have preferred a model which separately captures hospitalisation costs incurred whilst the patient is on treatment and those incurred after they have discontinued.

(11) Approach to handling uncertainty

The ERG notes several issues with the PSA included in the company's original submitted model:

(i) The ICERs generated from the probabilistic model were consistently lower than those generated from the deterministic model. Scrutiny of the PSA by the ERG indicated that this was driven by volatile sampling of the rate parameter for the With HSCT Gompertz OS survival function. The ERG believes that this was likely to be a result of the Gompertz model parameters being poorly specified. The company's updated model included an amended covariance matrix which resolved this sampling issue, although the ERG is unclear about how the new matrix has been updated.

(ii) The probabilistic results presented in the CS are based on 1,000 Monte Carlo samples. It would have been prudent to explore whether this represents a sufficient number of samples to minimise Monte Carlo sampling error.

(iii) As illustrated in Table 27, a number of uncertain model parameters are assumed to be fixed within the PSA. These include:

- The probability of receiving each salvage chemotherapy regimen (uncertainty around these probabilities was included in the company's updated model)
- The proportion of patients who are female

- The general population survival function and the SMR
- The probability of patients receiving gilteritinib maintenance therapy treatment after HSCT
- The HR used to reflect the treatment effect of gilteritinib maintenance therapy
- The multipliers used for the age-related utility adjustments
- The duration of the QALY loss resulting from HSCT
- The mean time to HSCT.

As a consequence of these issues, the ERG believes that uncertainty has been underestimated within the company's model.

5.4 ERG's exploratory analyses

This section presents the methods and results of the ERG's exploratory analyses undertaken using the company's model.

5.4.1 Overview of the ERG's exploratory analyses

The ERG undertook eight exploratory analyses to address the key points identified within the critical appraisal (Section 5.3.3). These included correcting the errors identified in the company's model, using alternative evidence sources and amending assumptions. The most significant of these model amendments relates to the use of parametric survival models fitted by the ERG to data for OS from randomisation for patients receiving HSCT in ADMIRAL.[7] The exploratory analyses were combined to form the ERG's preferred base case analysis.

The ERG also undertook additional sensitivity analyses using the ERG's preferred base case model to explore the impact of: using an alternative source of post-HSCT OS data (Poiré *et al*[73]); applying alternative survival models to the With HSCT and No HSCT groups, and assuming alternative cure timepoints. The ERG also fitted mixture-cure models to the With HSCT OS data from ADMIRAL,[7] within a Bayesian framework, in order to estimate cure fractions and survivor functions for FLT3+ R/R AML patients who receive HSCT. The ERG was not able to fully apply these cure models within the company's existing model structure; instead, these cure models were used to assess whether the With HSCT OS included in the company's model and ERG's preferred model are likely to be reasonable.

All analyses were undertaken using the deterministic version of the company's model. Implementation of the ERG's exploratory analyses was repeated by a second modeller to ensure that the results are free from errors. Technical details regarding the implementation of these analyses in the company's model are presented in ERG Appendix 3.

5.4.2 ERG exploratory analysis – methods

ERG preferred base case analysis

The ERG's preferred base case analysis is comprised of seven sets of amendments to the company's model; these are detailed below.

ERG exploratory analysis 1: Correction of errors

The ERG made the following corrections to the company's model:

(i) General population mortality risk was modelled using separate life tables for men and women assuming that the proportion of women aged [redacted] years is [redacted] (based on ADMIRAL[7]). In addition, the company's 2010-2012 life tables for the UK were replaced with 2015-2017 life tables for England (see critical appraisal point [1] item 1).

(ii) Post-HSCT OS for patients who receive gilteritinib maintenance was modelled separately to patients who do not receive gilteritinib maintenance therapy (see critical appraisal point [1] item 2).

(iii) A logical consistency constraint was applied to ensure that the cumulative probability of EFS does not exceed the cumulative probability of OS at any time *t* (see critical appraisal point [1] item 3).

(iv) From month 36 onwards, health utilities for surviving patients were assumed to reflect the values reported within Janssen *et al*[35] (without age-related utility multipliers, see critical appraisal point [1] item 5).

All other exploratory analyses undertaken by the ERG are applied within the corrected version of the model.

ERG exploratory analysis 2: Use of ADMIRAL With HSCT data on time from randomisation to death

In response to an additional data request from the ERG,[67] the company provided IPD on OS for the With HSCT group in ADMIRAL (gilteritinib [redacted]; salvage chemotherapy [redacted]). The ERG fitted standard parametric models to the IPD relating to the pooled dataset (gilteritinib and salvage chemotherapy arms combined; n=[redacted]). These included the exponential, Weibull, Gompertz, log logistic, log normal and generalised gamma models. Relative goodness-of-fit statistics (AIC and BIC) were calculated for each of the candidate survival models. The ERG sought advice from the ERG's clinical advisor regarding the plausibility of the competing OS survival models. Figure 22 presents the fitted and observed survivor functions for OS for the With HSCT group. The AIC and BIC statistics for each of the candidate survival models are presented in Table 36.

Figure 22: ERG's fitted survivor functions using data for patients with HSCT in ADMIRAL (gilteritinib and salvage chemotherapy groups pooled)

Table 36: With HSCT, OS, gilteritinib group (ADMIRAL), AIC and BIC statistics

Distribution	AIC	BIC
Exponential	152.65	155.05
Weibull	141.70	146.51
Gompertz	147.68	152.50
Log normal	**136.83**	**141.64**
Log logistic	139.35	144.17
Generalised gamma	137.46	144.68

AIC - Akaike Information Criterion; BIC - Bayesian Information Criterion
** Lowest AIC/BIC values shown in bold*

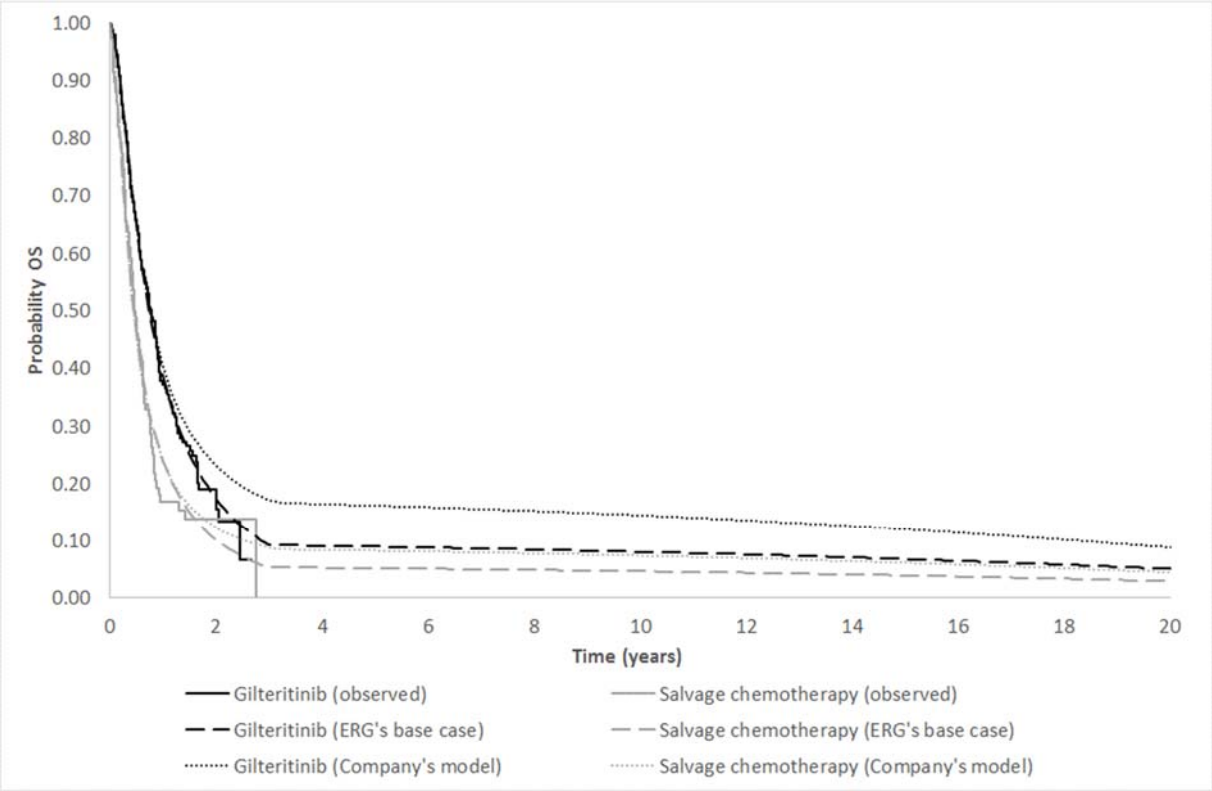
As shown in Table 36, the log normal distribution has the lowest AIC and BIC. Figure 22 suggests that the exponential, Gompertz and generalised gamma distributions do not provide a good visual fit to the data; the ERG's clinical advisor did not consider these models to be appropriate. Of the remaining survival models, the log normal and log logistic distributions produce very similar OS projections, whilst the Weibull function is comparatively less favourable. The ERG's clinical advisor noted that hazard of death for patients undergoing HSCT would be expected to remain high until the patient receives the transplant, but would subsequently decrease. On the basis of its goodness-of-fit statistics, the underlying nature of the hazard function and the plausibility of the extrapolated survivor function, the ERG selected the log normal model for inclusion in the ERG-preferred base case analysis. The

ERG's clinical advisor also considered the Weibull and log logistic survival models to be potentially appropriate; these were considered within the ERG's additional sensitivity analyses.

Within this exploratory analysis, the ERG's log normal OS model for patients receiving HSCT was applied in the company's model by overwriting the existing With HSCT sub-model trace up to the 3-year cure point. As the model no longer requires the assumption of a fixed time of transplant, the QALY loss associated with HSCT was applied in the first cycle to all patients in the With HSCT group.

Figure 23 presents a comparison of the ERG's model-predicted OS for the overall FLT3+ R/R AML population (shown as long dashed lines; including the log normal OS model fitted to data from ADMIRAL[7] for the With HSCT group) versus with the Kaplan-Meier survivor functions for the ADMIRAL ITT population (shown as solid lines). The figure also includes the company's model-predicted survival estimates (shown as dotted lines). The ERG's model produces an OS projection which appears to better reflect the observed data compared with the company's model; this is because within the ERG's model, OS for the With HSCT and No HSCT groups is based on models fitted to data from ADMIRAL[7] exclusively. Whilst the ERG's model retains the company's assumption of a fixed 3-year cure point, its impact is lessened as fewer patients are estimated to remain alive at this timepoint.

Figure 23: Overall survival, model-predicted versus observed, overall ADMIRAL population – ERG's base case model versus company's model



ERG exploratory analysis 3: Use of utilities from Ara and Brazier

Within this exploratory analysis, general population utilities after month 36 from Janssen *et al*[35] were replaced with model-based estimates derived from Ara and Brazier.[70]

ERG exploratory analysis 4: Remove double-counting of AML progression as AE

In order to avoid double-counting, the frequency of AML progression as an AE was set equal to zero.

ERG exploratory analysis 5: Remove HR for gilteritinib maintenance therapy

Within this analysis, the HR for OS for gilteritinib maintenance therapy was set equal to 1.0. This amendment was made on the basis of insufficient evidence to support an additional effect of maintenance therapy on OS. It should be noted that this amendment is only applied in analyses which use post-HSCT OS data reported by Evers *et al*.[31]

ERG exploratory analysis 6: Inclusion of wastage for gilteritinib (0.50 packs)

The company's model excludes wastage. It is expected that gilteritinib will be dispensed once every 28 days.[67] Within this analysis, the company's model was modified to assume that half a pack (14 days' supply) of wastage would be incurred by all patients who die before the 3-year cure point.

ERG exploratory analysis 7: Cost amendments (follow-up included, post-cure excluded, tests increased)

Within this analysis, four amendments were made to the company's model:

(i) After 3 years, patients who undergo HSCT were assumed to require an outpatient visit once every year; this cost is applied indefinitely.
(ii) After 3 years, patients who do not undergo HSCT were assumed to require an outpatient visit once every 6 months; this cost is applied indefinitely.
(iii) After 3 years, post-relapse/progression treatment costs were set equal to zero.
(iv) The unit cost per FLT3 test was multiplied by 3.3 instead of 2.0.

ERG exploratory analysis 8: ERG's preferred base case

The ERG's preferred base case includes ERG exploratory analysis 1-7.

Additional sensitivity analyses using the ERG preferred model

The following additional sensitivity analyses were undertaken using the ERG's preferred model.

ERG additional sensitivity analysis 1: Use of Poiré *et al* to inform post-HSCT OS

Within this analysis, cumulative probabilities of OS and EFS over the first 3 years of the model time horizon were based on cumulative OS probabilities reported by Poiré *et al*.[73] It should be noted that for simplicity, these probabilities are applied from the first cycle, rather than at fixed HSCT timepoint in each treatment group; this may introduce a small bias.

ERG additional sensitivity analysis 2: Use of alternative parametric survival models for OS within the No HSCT group

Within this analysis, all candidate OS models fitted to data for patients without HSCT in ADMIRAL[7] were applied in the No HSCT sub-model.

ERG additional sensitivity analysis 3: Use of alternative parametric survival models for OS within the With HSCT group

Within this analysis, all candidate OS models fitted to data for patients with HSCT in ADMIRAL[7] were applied in the With HSCT sub-model.

ERG additional sensitivity analysis 4: Exploration of alternative cure points

Within this analysis, alternative cure timepoints of 2, 3, 4 and 5 years were applied to all patients in the model.

ERG additional sensitivity analysis 5: Exploration of mixture-cure models

The ERG fitted mixture-cure models to the With HSCT OS data from ADMIRAL,[7] using a Bayesian framework, in order to estimate cure fractions and survival functions for patients susceptible to death from the disease for FLT3+ R/R AML patients who receive HSCT. Exponential, Weibull, log logistic, log normal, gamma, generalised gamma and Gompertz distributions were fitted to the pooled gilteritinib and salvage chemotherapy OS data. The ERG also fitted standard parametric survival models in order to assess whether the mixture-cure models provide a comparatively better fit to the data. The models were fitted using OpenBUGS software. Several models could not be fitted because of numerical problems. Some of the numerical problems might be overcome by re-parameterising the model, although the ERG was unable to address this in the time available. The relative goodness-of-fit of the models was assessed using the BIC. The mixture-cure models provided estimates of cure fractions and associated credible intervals (CrIs) around these.

The BIC statistics for the standard survival models and mixture-cure models are summarised in Table 37. The cure fractions estimated from the mixture-cure models are presented in Table 38. The ERG's analysis indicates that in general the mixture-cure models provide a better fit to the data than the standard parametric survival models. Of the models fitted, the cure model with a log normal distribution for the (uncured) patients susceptible to the disease provides the best fit; the cure fraction was estimated to be 0.17 (95% CrI 0.01, 0.42).

Table 37: With HSCT group in ADMIRAL, ERG's cure models – BIC statistics

Distribution	Standard model	Mixture-cure model
Exponential	410.8	322.5
Weibull	403.3	313.4
Log logistic	383.5	311.1
Log normal	398.3	309.1
Gamma	401.2	NE
Generalised gamma	NE	NE
Gompertz	315.6	320.1

NE - not evaluable

Table 38: With HSCT group in ADMIRAL - cure fractions estimated using cure models

Distribution	Median	2.5th percentile	97.5th percentile
Exponential	[redacted]	[redacted]	[redacted]
Weibull	[redacted]	[redacted]	[redacted]
Log logistic	[redacted]	[redacted]	[redacted]
Log normal	[redacted]	[redacted]	[redacted]
Gamma	NE	NE	NE
Generalised gamma	NE	NE	NE
Gompertz	[redacted]	[redacted]	[redacted]

NE - not evaluable

The ERG was not able to fully implement the mixture-cure models within the company's existing model structure as this would have required further statistical analyses (e.g. fitting separate mixture-cure models to data on EFS), structural amendments to model HRQoL and costs according to cured status rather than receipt of HSCT, and would likely have required several additional assumptions.

Instead, the ERG estimated the lifetime survival benefit accrued by a patient receiving HSCT based on the mixture-cure models and compared this against the estimates used in the company's model and the ERG's preferred model (using a fixed 3-year cure point). This was done to provide an indication as to whether the ICERs estimated from company's model and the ERG's preferred model are likely to be reasonable.

5.4.3 ERG exploratory analysis – results

ERG preferred base case analysis results

Table 39 presents the results of the ERG's preferred analysis. As shown in the table, correcting the errors in the company's model increases the ICER for gilteritinib versus salvage chemotherapy from £47,695 to £54,844 per QALY gained. This difference is largely attributable to the correction of the errors in the company's age-adjusted utility calculations for long-term AML survivors. The results for ERG exploratory analysis 2 show that the inclusion of the OS data from randomisation for patients undergoing HSCT in ADMIRAL[7] has a pronounced impact on the cost-effectiveness of gilteritinib; this increases the ICER to £95,642 per QALY gained. The ERG's other model amendments have a comparatively minor impact. The ERG's preferred base case ICER for gilteritinib versus salvage chemotherapy is estimated to be £102,085 per QALY gained.

Table 39: ERG exploratory analysis results, deterministic, gilteritinib versus salvage chemotherapy

Option	LYGs	QALYs	Costs	Inc. LYGs	Inc. QALYs	Inc. costs	ICER
Company's base case[1]							
Gilteritinib	3.03	[redacted]	[redacted]	1.28	[redacted]	[redacted]	**£47,695**
Salvage chemo	1.75	[redacted]	[redacted]	-	-	-	-
ERG exploratory analysis 1: Correction of errors							
Gilteritinib	3.09	[redacted]	[redacted]	1.32	[redacted]	[redacted]	**£54,844**
Salvage chemo	1.77	[redacted]	[redacted]	-	-	-	-
ERG exploratory analysis 2: Use of OS data from ADMIRAL[7] for patients receiving HSCT*							
Gilteritinib	2.10	[redacted]	[redacted]	0.74	[redacted]	[redacted]	**£95,642**
Salvage chemo	1.36	[redacted]	[redacted]	-	-	-	-
ERG exploratory analysis 3: Use of utilities from Ara and Brazier[70]							
Gilteritinib	3.09	[redacted]	[redacted]	1.32	[redacted]	[redacted]	**£54,532**
Salvage chemo	1.77	[redacted]	[redacted]	-	-	-	-
ERG exploratory analysis 4: Remove double-counting of AML progression as AE							
Gilteritinib	3.09	[redacted]	[redacted]	1.32	[redacted]	[redacted]	**£54,760**
Salvage chemo	1.77	[redacted]	[redacted]	-	-	-	-
ERG exploratory analysis 5: Remove HR for gilteritinib maintenance therapy							
Gilteritinib	2.80	[redacted]	[redacted]	1.03	[redacted]	[redacted]	**£70,515**
Salvage chemo	1.77	[redacted]	[redacted]	-	-	-	-
ERG exploratory analysis 6: Inclusion of wastage for gilteritinib (0.50 packs)							
Gilteritinib	3.09	[redacted]	[redacted]	1.32	[redacted]	[redacted]	**£58,355**
Salvage chemo	1.77	[redacted]	[redacted]	-	-	-	-
ERG exploratory analysis 7: Cost amendments (follow-up included, post-cure excluded, tests increased)							
Gilteritinib	3.09	[redacted]	[redacted]	1.32	[redacted]	[redacted]	**£54,999**
Salvage chemo	1.77	[redacted]	[redacted]	-	-	-	-
ERG exploratory analysis 8: ERG preferred base case (ERG analyses 1-7 combined)*							
Gilteritinib	2.10	[redacted]	[redacted]	0.74	[redacted]	[redacted]	**£102,085**
Salvage chemo	1.36	[redacted]	[redacted]	-	-	-	-

AE - adverse event; AML - acute myeloid leukaemia; ICER - incremental cost-effectiveness ratio; HSCT - haematopoietic stem cell transplant; Inc. - incremental; LYG - life year gained; QALY - quality-adjusted life year

For completeness, Table 40 summarises the results of the ERG's preferred analysis within a fully incremental analysis which includes both salvage chemotherapy and BSC as comparators.

Table 40: ERG's preferred base case analysis, deterministic, fully incremental analysis including BSC

Option	LYGs	QALYs	Costs	Inc. LYGs	Inc. QALYs	Inc. costs	ICER
Gilteritinib	2.10	[redacted]	[redacted]	0.74	[redacted]	[redacted]	**£102,085**
Salvage chemo	1.36	[redacted]	[redacted]	1.03	[redacted]	[redacted]	**£37,383**
BSC	0.33	[redacted]	[redacted]	-	-	-	-

BSC - best supportive care; ICER - incremental cost-effectiveness ratio; Inc. - incremental; LYG - life year gained; QALY - quality-adjusted life year

Results of additional sensitivity analyses undertaken using the ERG's preferred model

Table 41 presents the results of ERG additional sensitivity analysis 1. Within this analysis, LFS and OS outcomes from Poiré *et al*[73] are applied to all patients in the With HSCT sub-model. This analysis leads to an ICER for gilteritinib versus salvage chemotherapy of £105,071 per QALY gained. This is markedly higher than the company's base case ICER, and is driven largely by the lower estimated proportion of patients who remain alive at 3-years in the model.

Table 41: ERG additional sensitivity analysis 1 – use of cumulative EFS and OS probabilities reported in Poiré *et al*

Option	LYGs	QALYs	Costs	Inc. LYGs	Inc. QALYs	Inc. costs	ICER
Gilteritinib	2.08	[redacted]	[redacted]	0.73	[redacted]	[redacted]	**£105,071**
Salvage chemo	1.35	[redacted]	[redacted]	-	-	-	-

ICER - incremental cost-effectiveness ratio; Inc. - incremental; LYG - life year gained; QALY - quality-adjusted life year

Table 42 presents the results of ERG additional sensitivity analysis 2. Within this analysis, the alternative parametric survival functions for OS are applied to the No HSCT group. This analysis suggests that the model is sensitive to the choice of OS distribution for patients without HSCT. The ICER is markedly higher when the Weibull or Gompertz models are applied; these two distributions give the lowest predicted proportions of additional gilteritinib-treated patients remain alive at the 3-year cure point in the No HSCT group.

Table 42: ERG additional sensitivity analysis 2 – alternative parametric survival models applied in the No HSCT group

OS model for No HSCT group	Inc. LYGs	Inc. QALYs	Inc. costs	ICER
Exponential	0.75	[redacted]	[redacted]	**£102,377**
Weibull	0.61	[redacted]	[redacted]	**£124,319**
Log logistic (company's base case choice)	0.74	[redacted]	[redacted]	**£102,085**
Log normal	0.80	[redacted]	[redacted]	**£96,301**
Gompertz	0.61	[redacted]	[redacted]	**£124,455**
generalised gamma	0.78	[redacted]	[redacted]	**£98,595**

ICER - incremental cost-effectiveness ratio; HSCT - haematopoietic stem cell transplant; Inc. - incremental; LYG - life year gained; OS - overall survival; QALY - quality-adjusted life year

Table 43 presents the results of ERG additional sensitivity analysis 3. Within this analysis, the alternative parametric survival functions for OS are applied to the With HSCT group. This analysis suggests that the model is sensitive to the choice of OS distribution for patients receiving HSCT; the use of the exponential and generalised gamma distributions produces ICERs which are lower than the ERG's preferred base case (ICERs based on the exponential and generalised gamma With HSCT OS models - £85,182 and £91,322 per QALY gained, respectively). The ERG's clinical advisor preferred the log normal, log logistic and Weibull functions; the use of these survival models leads to ICERs ranging from £102,085 to £115,949 per QALY gained.

Table 43: ERG additional sensitivity analysis 3 – alternative parametric survival models applied in the With HSCT group

OS model for With HSCT group	Inc. LYGs	Inc. QALYs	Inc. costs	ICER
Exponential	0.89	[redacted]	[redacted]	**£85,182**
Weibull	0.65	[redacted]	[redacted]	**£115,949**
Log logistic	0.73	[redacted]	[redacted]	**£103,763**
Log normal (ERG's base case choice)	0.74	[redacted]	[redacted]	**£102,085**
Gompertz	0.61	[redacted]	[redacted]	**£123,357**
Generalised gamma	0.83	[redacted]	[redacted]	**£91,322**

ICER - incremental cost-effectiveness ratio; HSCT - haematopoietic stem cell transplant; Inc. - incremental; LYG - life year gained; OS - overall survival; QALY - quality-adjusted life year

Table 44 presents the results of ERG additional sensitivity analysis 4. Within this analysis, the assumed cure point was set equal to 2, 3, 4 or 5 years. The results of this analysis suggest that reducing the assumed cure timepoint to 2-years reduces the ICER to £66,123 per QALY gained. Conversely, extending the cure timepoint to 4 years increases the ICER to £133,111 per QALY gained. Extending the cure point to 5-years increases the ICER further to £156,225 per QALY gained.

Table 44: ERG additional sensitivity analysis 4 – use of alternative cure points

Option	LYGs	QALYs	Costs	Inc. LYGs	Inc. QALYs	Inc. costs	ICER
2 years							
Gilteritinib	3.05	[redacted]	[redacted]	1.12	[redacted]	[redacted]	**£66,123**
Salvage chemo	1.93	[redacted]	[redacted]	-	-	-	-
3 years (company's base case assumption)							
Gilteritinib	2.10	[redacted]	[redacted]	0.74	[redacted]	[redacted]	**£102,085**
Salvage chemo	1.36	[redacted]	[redacted]	-	-	-	-
4 years							
Gilteritinib	1.70	[redacted]	[redacted]	0.57	[redacted]	[redacted]	**£133,111**
Salvage chemo	1.12	[redacted]	[redacted]	-	-	-	-
5 years							
Gilteritinib	1.50	[redacted]	[redacted]	0.49	[redacted]	[redacted]	**£156,225**
Salvage chemo	1.01	[redacted]	[redacted]	-	-	-	-

ICER - incremental cost-effectiveness ratio; Inc. - incremental; LYG - life year gained; QALY - quality-adjusted life year

Figure 24 and Table 45 present the results of ERG additional sensitivity analysis 5. The figure presents a comparison of the company's Gompertz model (fitted to data reported by Evers et al[31]), the ERG's standard log normal model fitted to the OS data for patients receiving HSCT in ADMIRAL, and the ERG's mixture-cure models fitted to the same data from ADMIRAL. Table 45 presents estimates of mean OS derived from each of these models (calculated over a 40-year time horizon). Irrespective of whether the survival benefits of HSCT on OS are modelled assuming cure point or a cure fraction, this part of the model is a key driver of the ICER. With respect to this comparison, the ERG notes the following:

- The exponential mixture-cure model provides a poor fit to the data and is not considered further
- The log normal, log logistic and Weibull mixture-cure models produce broadly similar OS projections compared with the ERG's standard log normal model (including the 3-year cure point assumption)
- Whilst the ERG's preferred standard log normal model appears to underestimate the cumulative probability of OS at 3 years compared with the mixture-cure models, the lifetime mean OS for patients receiving HSCT predicted by the ERG's preferred model (5.18 LYGs) is similar to that predicted by the mixture-cure models (4.72 to 5.29 years).
- The company's Gompertz survivor function produces considerably more optimistic estimates of OS compared with all of the ERG's standard models and mixture-cure models. The mean OS derived from the company's Gompertz model is approximately twice as high as the ERG's estimates.
- The comparison indicates that irrespective of how cure is modelled (using cure fractions or fixed cure timepoints), a model based on the ADMIRAL With HSCT OS data will produce ICERs which are considerably higher than the company's base case.

Figure 24: Comparison of OS predictions between the company's Gompertz model, the ERG's standard log normal model and the ERG's mixture-cure models (With HSCT group)

Table 45: Comparison of mean OS between the company's Gompertz model, the ERG's log normal model and the ERG's mixture-cure models (With HSCT group)

Model applied to With HSCT group	Estimated mean OS
ERG's mixture-cure* - (log normal, ADMIRAL)	4.72
ERG's mixture-cure* - (log logistic, ADMIRAL)	5.28
ERG's mixture-cure* - (exponential, ADMIRAL)	3.81
ERG's mixture-cure* - (Gompertz, ADMIRAL)	4.97
ERG's mixture-cure* - (Weibull, ADMIRAL)	5.29
ERG's preferred model - (log normal, ADMIRAL)	5.18
Company's model - (Gompertz, Evers *et al*)	[redacted]

ERG - Evidence Review Group; OS - overall survival

** Cumulative OS probabilities in the non-cured groups were <1% at 10-years for the log normal, log logistic, Weibull and Gompertz models and 1.7% for the exponential model. The mixture-cure models assume that these patients accrue no further survival gains beyond this timepoint*

5.5 Discussion

The company's systematic review did not identify any existing economic evaluations of treatments for patients with FLT3+ R/R AML.

The CS[1] presents the methods and results of a *de novo* partitioned survival model developed by the company to assess the cost-effectiveness of gilteritinib versus a blended comparator of four salvage chemotherapy regimens (azacitidine, LoDAC, MEC and FLAG-IDA) for the treatment of FLT3+ R/R AML. The blended comparator reflects the comparator arm of the ADMIRAL trial,[7] and is based on overall outcomes and weighted costs for each regimen. This comparison represents the company's base case analysis. The company's model also allows for pairwise comparisons of gilteritinib versus each individual salvage chemotherapy regimen (using regimen-specific costs and outcomes data for the overall ADMIRAL salvage chemotherapy arm) and gilteritinib versus BSC. Incremental health gains, costs and cost-effectiveness are evaluated over a 40-year time horizon from the perspective of the NHS and PSS.

The company's model structure subdivides the overall population into two discrete groups according to whether or not they receive HSCT after initiating treatment with gilteritinib or salvage chemotherapy; based on ADMIRAL,[7] a higher proportion of patients treated with gilteritinib are assumed to receive HSCT compared with salvage chemotherapy (25.5% versus 15.3%, respectively). For each treatment option, separate partitioned survival sub-models are used to estimate health outcomes and costs for the With HSCT and No HSCT groups. Each sub-model includes the same three health states: (i) event-free; (ii) post-event and (iii) death. Within the No HSCT group, EFS and OS were informed by data for patients who did not receive HSCT in ADMIRAL. EFS and OS for the With HSCT group were informed by external data reported by Evers *et al.*[31] The company fitted parametric survival models to the time-to-event data. The company's model includes an assumption of a fixed cure point which applies to all patients who survive up to 3 years. With the exception of EFS in the No HSCT subgroup, this cure assumption overrides the event risks predicted by the fitted parametric models after the 3-year timepoint for all surviving patients. Prior to the 3-year cure timepoint, state-dependent utility values are based on a GEE model fitted to EQ-5D-3L data (mapped from the EQ-5D-5L) collected in ADMIRAL.[1] After the cure point, the ERG believes that the company intended to apply general population utility estimates; however, this aspect of the model is subject to a significant error. Resource use estimates were derived from ADMIRAL, standard costing sources, literature and assumptions. With the exception of ongoing post-progression treatment costs for patients leaving the EFS state, after the 3-year cure point, the model assumes that surviving patients incur no further disease management costs.

The probabilistic version of the company's updated model suggests that gilteritinib is expected to generate an additional [redacted] QALYs at an additional cost of [redacted] per patient compared with salvage chemotherapy; the corresponding ICER is £46,716 per QALY gained. The deterministic version of the model produces a higher ICER of £47,695 per QALY gained. The ERG notes that the company's use of external evidence to inform post-HSCT OS, and its implications for the proportion of patients who achieve a cure, is a key driver of the cost-effectiveness of gilteritinib. The company's base case model suggests that approximately [redacted] of gilteritinib-treated patients who undergo HSCT will remain alive at the assumed 3-year cure point. At the final data cut-off of ADMIRAL,[7] [redacted] of all gilteritinib-treated patients who received HSCT had already died; in order to achieve the 3-year cure proportion estimated by the company's model, this would require the vast majority of surviving (censored) patients in the gilteritinib group of ADMIRAL who received HSCT to be cured. The ERG considers this to be unlikely.

The ERG critically appraised the company's health economic analyses and double-programmed the deterministic version of the company's original model. The ERG's critical appraisal identified several issues relating to the company's model and the evidence used to inform its parameters. The most pertinent of these include: (i) the presence of model errors; (ii) concerns regarding the company's model structure; (iii) uncertainty surrounding the cure point and the use of external evidence to inform OS for patients receiving HSCT; (iv) inconsistencies between model-predicted OS and observed OS in ADMIRAL; (v) issues relating to the company's indirect comparisons, particularly the estimated treatment effect for gilteritinib maintenance therapy, and (vi) the underestimation of gilteritinib drug costs.

The ERG undertook eight exploratory analyses. These included: (i) correcting model errors; (ii) applying parametric models fitted to data on OS from randomisation for patients with HSCT in ADMIRAL;[7] (iii) the inclusion of utility values based on Ara and Brazier;[70] (iv) removing double-counting of health losses and costs associated with progressive AML; (v) removing the treatment effect for gilteritinib maintenance therapy; (vi) the inclusion of drug wastage for gilteritinib and (vii) amending costs associated with long-term follow-up, post-relapse/progression treatment and FLT3 mutation re-testing rates. The ERG's preferred base case combines all of these model amendments. The ERG undertook additional sensitivity analyses using the ERG-preferred model to explore the impact of: (i) using an alternative external source of evidence for post-HSCT outcomes; (ii) applying alternative survival models for patients without HSCT; (iii) applying alternative survival models for patients with HSCT; and (iv) assuming alternative cure timepoints. The ERG also fitted mixture-cure models to the With HSCT OS data from ADMIRAL in order to estimate cure fractions and survival functions; this

analysis was used to explore whether the results obtained from company's and the ERG's base case models are likely to be optimistic.

The ERG's preferred base case analysis suggests that the ICER for gilteritinib versus salvage chemotherapy is £102,234 per QALY gained. This analysis suggests that around 9% of gilteritinib-treated patients will remain alive at 3-years. This is considerably lower than the estimate generated by the company's model (company's estimate=17%). In addition, the ERG's preferred model produces an overall OS projection which better represents the observed data from ADMIRAL compared with the company's model.

The ERG's additional sensitivity analyses suggest that the parametric functions selected for the With HSCT and No HSCT groups are important: these analyses produced ICERs for gilteritinib versus salvage chemotherapy ranging from £85,182 to £124,455 per QALY gained. Reducing the assumed timepoint of cure to 2-years reduced the ICER to £66,123 per QALY gained; extending the cure point to 5 years increased the ICER to £156,225 per QALY gained. The ERG's exploration of mixture-cure models indicates that irrespective of how cure is represented (using a cure fraction or a fixed cure timepoint), a model based on the ADMIRAL With HSCT OS data will produce cost-effectiveness estimates which are considerably higher than the company's base case ICER.

6 END OF LIFE

NICE end of life supplementary advice should be applied in the following circumstances and when both the criteria referred to below are satisfied:

- The treatment is indicated for patients with a short life expectancy, normally less than 24 months and;
- There is sufficient evidence to indicate that the treatment offers an extension to life, normally of at least an additional 3 months, compared to current NHS treatment.

Table 46 summarises the evidence presented in the CS to support the company's argument that gilteritinib meets NICE's end of life criteria.

Table 46: Evidence presented in support of NICE's end of life criteria for gilteritinib in the CS (reproduced from CS, Table 15)

Criterion	Data available
The treatment is indicated for patients with a short life expectancy, normally less than 24 months	Median survival is reported to be two months or less in patients receiving supportive care alone[75, 76] and the pivotal ADMIRAL Phase III trial showed the median OS in the comparator salvage chemotherapy arm was 5.6 months.
There is sufficient evidence to indicate that the treatment offers an extension to life, normally of at least an additional 3 months, compared with current NHS treatment	The pivotal ADMIRAL Phase III trial[7] showed median overall survival in the gilteritinib arm was 9.3 months vs 5.6 months in the salvage chemotherapy arm, a gain of 3.7 months.

The ERG does not consider the median to represent a suitable measure of central tendency to determine whether the NICE's end of life criteria are met, as this does not account for the overall shape of the survival distribution. The mean represents a more appropriate measure of expected survival in the overall target population. Given the potential for long-term survival resulting from the use of HSCT, the company's base case model and the ERG's preferred analyses produce survival distributions which feature long tails whereby a small number of patients have prolonged survival.

Short life expectancy criterion

The ERG's preferred base case analysis produces a mean survival estimate of 0.33 years for BSC, and 1.69 years for salvage chemotherapy. The ERG believes that the short life expectancy criterion is expected to be met for gilteritinib.

Life extension criterion

The ERG's preferred base case analysis suggests that gilteritinib extends survival by 2.34 years compared with BSC and by 0.98 years compared with salvage chemotherapy. The ERG believes that the life extension criterion is likely to be met for gilteritinib.

7 OVERALL CONCLUSIONS

The clinical evidence for gilteritinib was based on one placebo-controlled RCT, ADMIRAL, which was open-label but of otherwise good methodological quality. The population of the trial was considered to be sufficiently representative of the target population for the results of the trial to be applicable for patients in England. There was a statistically significant advantage for gilteritinib over salvage chemotherapy for overall survival. The most common NCI-CTCAE Grade 3 or higher AEs experienced in the gilteritinib group were [redacted].

The key driver of the ICER for gilteritinib relates to the source of data used to inform OS in patients undergoing HSCT. The company's model uses external data to inform this aspect of the model (Evers *et al*[31]) rather than ADMIRAL;[7] the company's deterministic base case ICER for gilteritinib versus salvage chemotherapy is £47,695 per QALY gained. In contrast, the ERG's preferred model uses OS data from ADMIRAL for patients who receive HSCT. The ERG's preferred base case ICER for gilteritinib versus salvage chemotherapy is £102,234 per QALY gained. The ERG's exploration of mixture-cure models indicates that irrespective of how cure is represented (using a cure fraction or a fixed cure timepoint), a model based on the OS data for patients with HSCT in ADMIRAL will produce cost-effectiveness estimates which are considerably higher than the company's base case ICER.

7.1 Implications for research

There is uncertainty regarding long-term outcomes for FLT3+ R/R AML patients who receive HSCT. The company's clarification response states that the latest data cut-off of ADMIRAL (17th September, 2018) is final; no further analyses are expected. The ERG believes that additional follow-up would have helped to resolve existing uncertainty surrounding the OS benefits associated with gilteritinib.

8 REFERENCES

1. Astellas Pharma Ltd. Gilteritinib (XOSPATA™) for treating relapsed or refractory acute myeloid leukaemia ID1484. Company's evidence submission to the National Institute for Health and Care Excellence. Document B. Surrey, UK; 2019.
2. Cancer Research UK. Acute myeloid leukaemia (AML). 2019. https://www.cancerresearchuk.org/about-cancer/acute-myeloid-leukaemia-aml (Accessed 01/08/2019).
3. O'Donnell MR, Tallman MS, Abboud CN, Altman JK, Appelbaum FR, Arber DA, *et al.* Acute Myeloid Leukemia, Version 3.2017, NCCN Clinical Practice Guidelines in Oncology. *Journal of the National Comprehensive Cancer Network* 2017;15:926-57.
4. Döhner H, Estey E, Grimwade D, Amadori S, Appelbaum FR, Büchner T, *et al.* Diagnosis and management of AML in adults: 2017 ELN recommendations from an international expert panel. *Blood* 2017;129:424-47.
5. Daver N, Schlenk RF, Russell NH, Levis MJ. Targeting FLT3 mutations in AML: review of current knowledge and evidence. *Leukemia* 2019;33:299-312.
6. National Institute for Health and Care Excellence. Gilteritinib for treating relapsed or refractory acute myeloid leukaemia: Final scope. London; 2019.
7. Astellas Pharma Ltd. A Phase 3 open-label, multicenter, randomized study of gilteritinib versus salvage chemotherapy in patients with relapsed or refractory acute myeloid leukemia (AML) with FLT3 Mutation.. ADMIRAL Clinical Study Report. Surrey, UK; 2019.
8. Astellas Pharma Ltd. Gilteritinib (XOSPATA™) for treating relapsed or refractory acute myeloid leukaemia ID1484. Company's evidence submission to the National Institute for Health and Care Excellence. Appendices. Surrey, UK; 2019.
9. Astellas Pharma Ltd. Report for the Systematic Literature Review in the FLT3+ Relapsed / Refractory Acute Myeloid Leukaemia (R/R AML) setting: Clinical Management, Clinical Effectiveness, Economic and Humanistic Burdens. Company's evidence submission to the National Institute for Health and Care Excellence. Surrey, UK; 2019.
10. InterTASC Information Specialists' Sub-Group. ISSG Search Filter Resource. 2019. https://sites.google.com/a/york.ac.uk/issg-search-filters-resource/filters-to-identify-observational-studies (Accessed 31/07/2019).
11. Perl AE, Altman JK, Cortes JE, Smith CC, Litzow M, Baer MR, *et al.* Final Results of the Chrysalis Trial: A First-in-Human Phase 1/2 Dose-Escalation, Dose-Expansion Study of Gilteritinib (ASP2215) in Patients with Relapsed/Refractory Acute Myeloid Leukemia (R/R AML). *Blood* 2016;128:1069.
12. Perl AE, Altman JK, Cortes J, Smith C, Litzow M, Baer MR, *et al.* Selective inhibition of FLT3 by gilteritinib in relapsed or refractory acute myeloid leukaemia: a multicentre, first-in-human, open-label, phase 1-2 study. *Lancet Oncol* 2017;18:1061-75.
13. Gorcea CM, Burthem J, Tholouli E. ASP2215 in the treatment of relapsed/refractory acute myeloid leukemia with FLT3 mutation: background and design of the ADMIRAL trial. *Future Oncol* 2018;14:1995-2004.
14. Perl AE, Cortes JE, Strickland SA, Ritchie EK, Neubauer A, Martinelli G, *et al.* An open-label, randomized phase III study of gilteritinib versus salvage chemotherapy in relapsed or refractory FLT3 mutation-positive acute myeloid leukemia. *Journal of Clinical Oncology* 2017;35:TPS7067-TPS.
15. Wang J, Li J, Tantiworawit A, Izuka S, Yamada S, Hill J, *et al.* A phase III randomized study of gilteritinib versus salvage chemotherapy in FLT3 mutation-positive subjects with relapsed or refractory acute myeloid leukemia. *Annals of Oncology* 2018;29.

16. Centre for Reviews and Dissemination. CRD's Guidance for Undertaking Reviews in Health Care; 2009.

17. European Medicines Agency. EU Clinical Trials Register. 2019. https://www.clinicaltrialsregister.eu/ (Accessed 31/07/2019).

18. US National Library of Medicine. ClinicalTrails.gov. 2019. https://clinicaltrials.gov/ (Accessed 31/07/2019).

19. Mbanya Z, Madrigal AM, Wright A. Physicians’ Choice As A Comparator In Clinical Trials: Challenges for Pharmacoeconomic Modelling of Innovative Treatments to Support Health Technology Assessments. *Value in Health* 2015;18:A704.

20. Mendoza TR, Wang XS, Cleeland CS, Morrissey M, Johnson BA, Wendt JK, *et al.* The rapid assessment of fatigue severity in cancer patients: use of the Brief Fatigue Inventory. *Cancer* 1999;85:1186-96.

21. Astellas Pharma Ltd. Gilteritinib (XOSPATA™) for treating relapsed or refractory acute myeloid leukaemia ID1484. Company's response to clarification questions from the ERG. Surrey, UK; 2019.

22. Astellas Pharma Ltd. Analysis of the Patient Reported Outcomes (PRO) data collected in ADMIRAL – draft results. Company's evidence submission to the National Institute for Health and Care Excellence. Surrey, UK; 2019.

23. National Cancer Institute (NCI). Common Terminology Criteria for Adverse Events (CTCAE) guidelines version 4.03: US Department of Health and Human Services, National Institutes of Health, National Cancer Institute; 2010.

24. Cortes J. QuANTUM-R EHA 2018. Oral presentation at European Hematology Association (EHA) conference 2018: Quizartinib significantly prolongs overall survival in patients with FLT3-ITD-mutated relapsed-refractory AML in the phase 3, randomized, controlled QuANTUM-R trial. 2018.

25. Arber M, Garcia S, Veale T, Edwards M, Shaw A, Glanville JM. Performance of OVID MEDLINE search filters to identify health state utility studies. *International Journal of Technology Assessment in Health Care* 2017;33:472-80.

26. Lachaine J, Mathurin K, Barakat S, Couban S. Economic evaluation of arsenic trioxide compared to all-trans retinoic acid+ conventional chemotherapy for treatment of relapsed acute promyelocytic leukemia in Canada. *European Journal of Haematology* 2015;95:218-29.

27. Magar RS, Purdy C, Hayward A, Einarson TR, Burnett AK, . Cost-effectiveness analysis of histamine dihydrochloride + low dose interleukin-2 vs standard of care for acute myeloid leukemia patients in their first complete remission: A UK perspective. *Value in Health* 2009;12:A272.

28. Novartis Pharmaceuticals Ltd. Midostaurin for newly diagnosed acute myeloid leukaemia [ID894]. Company's evidence submission to the National Institute for Health and Care Excellence. Surrey, UK; 2017.

29. 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;37:85-92.

30. Sarkozy C, Gardin C, Gachard N, Merabet F, Turlure P, Malfuson JV. Outcome of older patients with acute myeloid leukemia in first relapse. *American Journal of Hematology* 2013;88:758-64.

31. Evers G, Beelen DW, Braess J, Sauerland C, Kolb HJ, Reichle A, *et al.* Outcome of patients with acute myeloid leukemia (AML) undergoing allogeneic hematopoietic stem cell transplantation (HSCT) beyond first complete remission (CR1). *Blood* 2018;132:4649.

32. Ustun C, Giannotti F, Zhang MJ, Wang HL, Brunstein C, Labopin M*, et al.* Outcomes of UCB transplantation are comparable in FLT3+ AML: results of CIBMTR, EUROCORD and EBMT collaborative analysis. *Leukemia* 2017;31:1408-14.

33. Office for National Statistics. Interim life tables - United Kingdom. London; 2015.

34. van Hout B, Janssen MF, Feng YS, Kohlmann T, Busschbach J, Golicki D*, et al.* Interim scoring for the EQ-5D-5L: mapping the EQ-5D-5L to EQ-5D-3L value sets. *Value in Health* 2012;15:708-15.

35. Janssen B, Szende A. Chapter 3: Population norms for the EQ-5D. In: Szende A JB, Cabases J (eds), editor. *Self-Reported Population Health: An International Perspective Based on EQ-5D*Dordrecht, Netherlands: Springer; 2014:1-12.

36. Department of Health. NHS Reference Costs 2017/18. London; 2018.

37. National Institute for Health and Care Excellence. Midostaurin for untreated acute myeloid leukaemia. Technology Appraisal Guidance No. 523. London; 2018.

38. Shimoni A, Labopin M, Savani B, Volin L, Ehninger G, Kuball J*, et al.* Long-term survival and late events after allogeneic stem cell transplantation from HLA-matched siblings for acute myeloid leukemia with myeloablative compared to reduced-intensity conditioning: a report on behalf of the acute leukemia working party of European group for blood and marrow transplantation. *Journal Hematology & Oncology* 2016;9:118.

39. Bejanyan N, Weisdorf DJ, Logan BR, Wang HL, Devine SM, de Lima M*, et al.* Survival of patients with acute myeloid leukemia relapsing after allogeneic hematopoietic cell transplantation: a center for international blood and marrow transplant research study. *Biology of Blood and Marrow Transplantation* 2015;21:454-9.

40. Othus M, Garcia-Manero G, Godwin J, Weick JK, Anderson JE, Stirewalt D*, et al.* Associations between complete remissions (CRs) with 7+3 induction chemotherapy for acute myeloid leukemia and 2-3 year survival ("potential cure") over the past four decades: analysis of SWOG trial data. *Blood* 2017;130:1301.

41. Swinburn P, Lloyd A, Nathan P, Choueiri TK, Cella D, Neary MP. Elicitation of health state utilities in metastatic renal cell carcinoma. *Current Medical Research and Opinion* 2010;26:1091-6.

42. Lloyd A, Nafees B, Narewska J, Dewilde S, Watkins J. Health state utilities for metastatic breast cancer. *British Journal of Cancer* 2006;95:683-90.

43. Nafees B, Stafford M, Gavriel S, Bhalla S, Watkins J. Health state utilities for non small cell lung cancer *Health and Quality of Life Outcomes* 2008;6.

44. Doyle S, Lloyd A, Walker M. Health state utility scores in advanced non-small cell lung cancer. *Lung Cancer* 2008;2008:3.

45. Haymarket Media Group Ltd. Monthly Index of Medical Specialities (MIMS). Drug listings by brand and generic drugs. 2019, https://www.mims.co.uk/drugs.

46. Commercial Medicines Unit. Drugs and parhamceutical electronic market information tool (eMIT). London; 2019.

47. Astellas Pharma Ltd. Adelphi AML treatment pattern study [data held on file]. 2019.

48. Curtis L, Burns A. Unit Costs of Health and Social Care 2018. Kent, UK; 2018.

49. 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 Effectiveness and Resource Allocation* 2018;16.

50. Wang H. Cancer treatment pathways: AML A detailed costing study example. PhD thesis; 2010.

51. Georghiou T, Bardsley M. Exploring the Cost of Care at the End of Life. London; 2014.

52. Latimer N. NICE DSU Technical Support Document 14: Survival analysis for economic evaluations alongside clinical trials - extrapolation with patient-level data. Sheffield; 2011.

53. Steckel NK, Groth C, Mikesch JH, Trenschel R, Ottinger H, Kordelas L*, et al.* High-dose melphalan-based sequential conditioning chemotherapy followed by allogeneic haematopoietic stem cell transplantation in adult patients with relapsed or refractory acute myeloid leukaemia. *British Journal of Haematology* 2018;180:840-53.

54. Fong CY, Grigoriadis G, Hocking J, Coutsouvelis J, Muirhead J, Campbell P*, et al.* Fludarabine, cytarabine, granulocyte-colony stimulating factor and amsacrine: an effective salvage therapy option for acute myeloid leukemia at first relapse. *Leukemia & Lymphoma* 2013;54:336-41.

55. Frazer J, Couban S, Doucette S, Shivakumar S. Characteristics predicting outcomes of allogeneic stem-cell transplantation in relapsed acute myelogenous leukemia. *Current Oncology* 2017;24:e123-e30.

56. Schmid C, Schleuning M, Schwerdtfeger R, Hertenstein B, Mischak-Weissinger E, Bunjes D*, et al.* Long-term survival in refractory acute myeloid leukemia after sequential treatment with chemotherapy and reduced-intensity conditioning for allogeneic stem cell transplantation. *Blood* 2006;108:1092-9.

57. Oran B, Cortes J, Beitinjaneh A, Chen HC, de Lima M, Patel K*, et al.* Allogeneic transplantation in first remission improves outcomes irrespective of FLT3-ITD allelic ratio in FLT3-ITD-positive acute myelogenous leukemia. *Biology of Blood and Marrow Transplantation* 2016;22:1218-26.

58. Schlenk RF, Kayser S, Bullinger L, Kobbe G, Casper J, Ringhoffer M*, et al.* Differential impact of allelic ratio and insertion site in FLT3-ITD-positive AML with respect to allogeneic transplantation. *Blood* 2014;124:3441-9.

59. Song Y, Magenau J, Li Y, Braun T, Chang L, Bixby D*, et al.* FLT3 mutational status is an independent risk factor for adverse outcomes after allogeneic transplantation in AML. *Bone Marrow Transplantation* 2016;51:511-20.

60. Deol A, Sengsayadeth S, Ahn KW, Wang HL, Aljurf M, Antin JH*, et al.* Does FLT3 mutation impact survival after hematopoietic stem cell transplantation for acute myeloid leukemia? A Center for International Blood and Marrow Transplant Research (CIBMTR) analysis. *Cancer* 2016;122:3005-14.

61. Guyot P, Ades AE, Ouwens MJ, Welton NJ. Enhanced secondary analysis of survival data: reconstructing the data from published Kaplan-Meier survival curves. *BMC Medical Research Methodology* 2012;1.

62. Kind P, Hardman G, Macran S. UK population norms for EQ-5D. CHE Discussion Paper 172. York, UK; 1999.

63. Rowe JM, Li X, Cassileth PA, Appelbaum FR, Schiffer CS, Wiernik PH*, 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;106:546.

64. Drummond MF, Sculpher MJ, Torrance GW, O'Brien BJ, Stoddart GL, . Methods for the economic evaluation of health care programmes. New York: Oxford University Press; 2005.

65. Weinstein MC, O'Brien B, Hornberger J, Jackson J, Johannesson M, McCabe C*, et al. Value in Health* 2003;Principles of good practice for decision analytic modeling in health-care evaluation: report of the ISPOR Task Force on Good Research Practices--Modeling Studies:1.

66. National Institute for Health and Care Excellence. Guide to the methods of technology appraisal. London; 2013.

67. Astellas Pharma Ltd. Gilteritinib (XOSPATA™) for treating relapsed or refractory acute myeloid leukaemia ID1484. Company's response to additional data request from the ERG. Surrey, UK; 2019.

68. Grulke N, Albani C, Bailer H. Quality of life in patients before and after haematopoietic stem cell transplantation measured with the European Organization for Research and Treatment of Cancer (EORTC) Quality of Life Core Questionnaire QLQ-C30. *Bone Marrow Transplantation* 2012;47:473-82.

69. Kurosawa S, Yamaguchi T, Mori T, Kanamori H, Onishi Y, Emi N*, et al.* Patient-reported quality of life after allogeneic hematopoietic cell transplantation or chemotherapy for acute leukemia. *Bone Marrow Transplantation* 2015;50:1241-9.

70. Ara R, Brazier JE. Populating an economic model with health state utility values: moving toward better practice. *Value in Health* 2010;13:509-18.

71. Williams C, Lewsey JD, Briggs AH, Mackay DF. Cost-effectiveness analysis in R using a multi-state modeling survival analysis framework: A tutorial. *Medical Decision Making* 2017;37:340-52.

72. Putter H, Fiocco M, Geskus RB. Tutorial in biostatistics: Competing risks and multi-state models. *Statistics in Medicine* 2007;26:2389-430.

73. Poiré X, Labopin M, Polge E, Passweg J, Craddock C, Blaise D*, et al.* Allogeneic stem cell transplantation benefits for patients ≥ 60 years with acute myeloid leukemia and FLT3 internal tandem duplication: a study from the Acute Leukemia Working Party of the European Society for Blood and Marrow Transplantation. *Haematologica* 2018;103:256-65.

74. Hodgson R, Biswas M, Mebrahtu T, Morgan P, Harden M, Woolacott N. Midostaurin for untreated acute myeloid leukaemia. Evidence Review Group report submitted to the National Institute for Health and Care Excellence. York, UK; 2017.

75. Meyers J, Yu Y, Kaye JA, Davis KL. Medicare fee-for-service enrollees with primary acute myeloid leukemia: an analysis of treatment patterns, survival, and healthcare resource utilization and costs. *Applied Health Economics and Health Policy* 2013;11:275-86.

76. Oran B, Weisdorf DJ. Survival for older patients with acute myeloid leukemia: a population-based study. *Haematologica* 2012;97:1916-24.

9 APPENDICES

Appendix 1 Quality assessment of CHRYSALIS

Table 47: Quality assessment of CHRYSALIS (adapted from CS Appendix D and Astellas SLR report)

CHRYSALIS (EudraCT 2014-002217-31) NCT02014558	Quality assessment by CS (CS Appendix D Table 23)	Quality assessment by CS (CS Appendix D Table 20)	Quality assessment by ERG
Was randomisation carried out appropriately?	Yes	NA	Part of the study was randomised For the randomised part of the study, Randomisation sequence generated by Interactive Response Technology (IRT) (clarification response,[7] question A15)
Was the concealment of treatment allocation adequate?	N/A, unblinded phase I/II dose escalation / expansion study	NA	For the randomised part of the study, allocation was by Interactive Response Technology (IRT) (clarification response,[7] question A15)
Were the groups similar at the outset of the study in terms of prognostic factors?	N/A, non-comparative study	Yes	Y (CS Table 8) https://clinicaltrials.gov/ct2/show/results/NCT02014558[18]
Were the care providers, participants and outcome assessors blind to treatment allocation?	Unblinded phase I/II dose escalation/expansion study	NA	No (CS Section B.2.3) https://clinicaltrials.gov/ct2/show/results/NCT02014558[18]
Were there any unexpected imbalances in drop-outs between groups?	No	No	No (CS Appendix D1.2) [8] https://clinicaltrials.gov/ct2/show/results/NCT02014558[18]
Is there any evidence to suggest that the authors measured more outcomes than they reported?	No	Yes	No (study results reported on clinical trials gov) https://clinicaltrials.gov/ct2/show/results/NCT02014558[18]
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	No	Modified intent to treat – Full analysis set - not all allocated but those who received at least one dose of study drug and had at least one post-treatment datapoint (CS Appendix D1.2[8])

Eligibility criteria for CHRYSALIS and ADMIRAL

Table 48: Eligibility criteria for chrysalis and admiral (adapted from CS, Table 6)

CHRYSALIS	ADMIRAL
Inclusion criteria: • Age ≥18 years • Primary or secondary AML • Refractory to at least 1 cycle of induction chemotherapy or relapsed after achieving remission with a previous drug Not required to have FLT3 mutation but did require at least 10 patients with FLT3 in each expansion dose group.[11, 12] • ECOG performance status ≤2 • ALT or AST ≤2.5 X ULN • Total bilirubin ≤1.5 X ULN • Serum creatinine ≤1.5 X ULN or eGFR >50 mL/min **Exclusion criteria:** • APL • BCR-ABL-positive leukaemia • Malignant tumours other than AML or MDS • NYHA class 3 or 4 heart failure and those who had previously had NYHA class 3 or 4 heart failure, unless a screening echocardiogram done within 3 months before study entry resulted in a LVEF of ≥45% • Long QTc syndrome • Active uncontrolled infections including hepatitis B or C and HIV • GvHD requiring treatment • Pregnancy • Presence of grade ≥2 non-haematologic toxicities from prior AML treatment • Prior HSCT within 2 months of study treatment (Cycle 1, Day 1) • Persistent grade ≥2 non-haematologic toxicities related to HSCT	**Inclusion criteria:** • Age ≥18 years • Refractory or relapsed AML (after first-line therapy with or without HSCT) • Confirmed FLT3 mutation • Eligible for pre-selected salvage chemotherapy • Can receive oral therapy • ECOG performance status ≤2 • ALT or AST ≤2.5 X ULN • Total bilirubin ≤1.5 X ULN • Serum creatinine ≤1.5 X ULN or eGFR >50 mL/min • Female patients must be either of non-child bearing potential or not pregnant at study initiation and not planning to become pregnant **Exclusion criteria:** • Acute promyelocytic leukaemia (APL) • Breakpoint cluster region-Abelson murine leukaemia (BCR-ABL-positive leukaemia) • AML secondary to prior chemotherapy for other neoplasms (except for MDS) • History of another malignancy (unless disease free for ≥5 years) • Clinically active central nervous system leukaemia • Prior treatment with gilteritinib or other FLT3 inhibitors, with exception of sorafenib or midostaurin • Clinically significant abnormality of coagulation profile • Major surgery or radiation within 4 weeks of first study dose • NYHA class 3 or 4 heart failure and those who had previously had NYHA class 3 or 4 heart failure, unless a screening echocardiogram done within 1 month before study entry resulted in a LVEF of ≥45% [[CS clar respone A2]] • Long QT syndrome • Mean of triplicate QTcF >450 milliseconds • Hypokalaemia or hypomagnesaemia (lower than lower limit of normal patients excluded [CS clar response A2]) • Active uncontrolled infections including hepatitis B or C and HIV, or other uncontrolled hepatic disorder • Active clinically significant GvHD or on treatment with systemic corticosteroids for GvHD

Baseline Characteristics of Patients in ADMIRAL

Table 49: Baseline characteristics of patients in ADMIRAL (reproduced from CS, Table 9)

Characteristic	Gilteritinib (N=247)	Salvage Chemotherapy (N=124)	Overall (N=371)
Age, years			
Mean (SD)	[redacted]	[redacted]	[redacted]
Median (range)	62.0 (20-84)	61.5 (19-85)	62.0 (19-85)
Age group, n (%)			
<65 years	[redacted]	[redacted]	[redacted]
≥65 years	[redacted]	[redacted]	[redacted]
Males, n (%)	116 (47.0)	54 (43.5)	170 (45.8)
Race, n (%)			
White	[redacted]	[redacted]	[redacted]
Asian	[redacted]	[redacted]	[redacted]
Black or African American	[redacted]	[redacted]	[redacted]
Native Hawaiian or other Pacific Islander	[redacted]	[redacted]	[redacted]
Other	[redacted]	[redacted]	[redacted]
Unknown	[redacted]	[redacted]	[redacted]
Missing	[redacted]	[redacted]	[redacted]
Baseline ECOG performance status			
0-1	[redacted]	[redacted]	[redacted]
2	[redacted]	[redacted]	[redacted]
Median disease duration, months (range)	[redacted]	[redacted]	[redacted]
FLT3 mutation status (central testing), n (%)			
FLT3-ITD alone	215 (87.0)	113 (91.1)	328 (88.4)
FLT3-TKD alone	21 (8.5)	10 (8.1)	31 (8.4)
FLT3-ITD and FLT3-TKD	7 (2.8)	0	7 (1.9)
Other (negative)	[redacted]	[redacted]	[redacted]
Prior treatments received, n (%)			
Chemotherapy for AML	[redacted]	[redacted]	[redacted]
FLT3 inhibitor	32 (13.0)	14 (11.3)	46 (12.4)
HSCT	48 (19.4)	26 (21.0)	74 (19.9)
Response to first-line therapy			
Primary refractory without HSCT	98 (39.7)	48 (38.7)	146 (39.4)
Relapse within 6 months after CR_c and no HSCT	[redacted]	[redacted]	[redacted]
Relapse after 6 months after CR_c and no HSCT	[redacted]	[redacted]	[redacted]
Relapse within 6 months after allogeneic HSCT	[redacted]	[redacted]	[redacted]
Relapse after 6 months after allogeneic HSCT	[redacted]	[redacted]	[redacted]
Preselected Salvage Chemotherapy			
High intensity chemotherapy	[redacted]	[redacted]	[redacted]
Low intensity chemotherapy	[redacted]	[redacted]	[redacted]
Cytogenetic risk group, n (%)			
Favourable	4 (1.6)	1 (0.8)	5 (1.3)
Intermediate	182 (73.7)	89 (71.8)	271 (73.0)
Unfavourable	26 (10.5)	11 (8.9)	37 (10.0)
Other	35 (14.2)	23 (18.5)	58 (15.6)

CR_c - composite complete remission; ECOG - Eastern Cooperative Oncology Group; FLT3 - FMS-like tyrosine kinase-3; FLT3-ITD - FMS-like tyrosine kinase-3-internal tandem duplication; FLT3-TKD - FMS-like tyrosine kinase-3-tyrosine kinase domain; HSCT - haematopoietic stem cell transplant; SD - standard deviation

CHRYSALIS: Most Frequent (≥ 10% from doses 80-200mg/day) AEs

Table 50: CHRYSALIS: Most frequent (≥ 10% from doses 80-200mg/day) AEs (adapted from CS Appendix F, Table 26)

MedDRA V20.0 System Organ Class Preferred Term	Gilteritinib		
	80mg (N=24) n (%)	120mg (N=69) n (%)	200mg (N=103) n (%)
Overall	23 (95.8)	67 (97.1)	103 (100.0)
Blood and Lymphatic System Disorders			
Febrile neutropenia	6 (25.0)	24 (34.8)	47 (45.6)
Anaemia	9 (37.5)	26 (37.7)	37 (35.9)
Thrombocytopenia	3 (12.5)	13 (18.8)	20 (19.4)
Neutropenia	0	6 (8.7)	14 (13.6)
Cardiac Disorders			
Tachycardia	2 (8.3)	2 (2.9)	12 (11.7)
Eye Disorders			
Dry eye	1 (4.2)	9 (13.0)	6 (5.8)
Conjunctival haemorrhage	4 (16.7)	0	3 (2.9)
Gastrointestinal Disorders			
Diarrhoea	7 (29.2)	31 (44.9)	47 (45.6)
Constipation	6 (25.0)	14 (20.3)	33 (32.0)
Nausea	6 (25.0)	15 (21.7)	30 (29.1)
Vomiting	7 (29.2)	13 (18.8)	23 (22.3)
Stomatitis	0	8 (11.6)	15 (14.6)
Abdominal pain	1 (4.2)	5 (7.2)	16 (15.5)
General Disorders and Administration Site Conditions			
Fatigue	9 (37.5)	27 (39.1)	35 (34.0)
Oedema peripheral	5 (20.8)	18 (26.1)	31 (30.1)
Pyrexia	3 (12.5)	24 (34.8)	32 (31.1)
Asthenia	4 (16.7)	6 (8.7)	19 (18.4)
Chills	3 (12.5)	5 (7.2)	12 (11.7)
Mucosal inflammation	1 (4.2)	7 (10.1)	12 (11.7)
Oedema	4 (16.7)	3 (4.3)	6 (5.8)
Infections and Infestations			
Sepsis	8 (33.3)	10 (14.5)	19 (18.4)
Pneumonia	2 (8.3)	16 (23.2)	19 (18.4)
Urinary tract infection	4 (16.7)	11 (15.9)	7 (6.8)
Bacteraemia	8 (33.3)	5 (7.2)	7 (6.8)
Upper respiratory tract infection	3 (12.5)	8 (11.6)	5 (4.9)
Lung infection	0	4 (5.8)	11 (10.7)
Cellulitis	2 (8.3)	8 (11.6)	4 (3.9)
Skin infection	3 (12.5)	4 (5.8)	4 (3.9)
Septic shock	3 (12.5)	0	4 (3.9)
Injury, Poisoning and Procedural Complications			
Fall	2 (8.3)	13 (18.8)	20 (19.4)
Contusion	3 (12.5)	3 (4.3)	10 (9.7)
Investigations			
Aspartate aminotransferase increased	4 (16.7)	20 (29.0)	36 (35.0)
Alanine aminotransferase increased	4 (16.7)	16 (23.2)	26 (25.2)
Platelet count decreased	2 (8.3)	12 (17.4)	19 (18.4)

MedDRA V20.0 System Organ Class Preferred Term	Gilteritinib 80mg (N=24) n (%)	 120mg (N=69) n (%)	 200mg (N=103) n (%)
Blood creatinine increased	5 (20.8)	13 (18.8)	21 (20.4)
Blood alkaline phosphatase increased	4 (16.7)	10 (14.5)	15 (14.6)
Neutrophil count decreased	1 (4.2)	8 (11.6)	14 (13.6)
Blood creatine phosphokinase increased	0	7 (10.1)	18 (17.5)
Blood bilirubin increased	4 (16.7)	3 (4.3)	14 (13.6)
Electrocardiogram QT prolonged	1 (4.2)	11 (15.9)	9 (8.7)
White blood cell count decreased	0	4 (5.8)	13 (12.6)
Weight increased	0	2 (2.9)	12 (11.7)
Metabolism and Nutrition Disorders			
Hypokalaemia	3 (12.5)	10 (14.5)	25 (24.3)
Hypocalcaemia	3 (12.5)	11 (15.9)	24 (23.3)
Hyponatraemia	4 (16.7)	9 (13.0)	20 (19.4)
Decreased appetite	5 (20.8)	11 (15.9)	19 (18.4)
Hypoalbuminaemia	4 (16.7)	7 (10.1)	18 (17.5)
Hypomagnesaemia	3 (12.5)	13 (18.8)	17 (16.5)
Hyperglycaemia	0	6 (8.7)	14 (13.6)
Hypophosphataemia	1 (4.2)	6 (8.7)	11 (10.7)
Hyperuricaemia	0	4 (5.8)	11 (10.7)
Dehydration	0	1 (1.4)	11 (10.7)
Musculoskeletal and Connective Tissue Disorders			
Arthralgia	5 (20.8)	11 (15.9)	17 (16.5)
Pain in extremity	2 (8.3)	10 (14.5)	12 (11.7)
Back pain	2 (8.3)	9 (13.0)	10 (9.7)
Myalgia	0	5 (7.2)	14 (13.6)
Muscular weakness	0	8 (11.6)	4 (3.9)
Neoplasms Benign, Malignant and Unspecified (Incl Cysts and Polyps)			
Acute myeloid leukaemia	5 (20.8)	10 (14.5)	20 (19.4)
Nervous System Disorders			
Dizziness	6 (25.0)	17 (24.6)	26 (25.2)
Headache	3 (12.5)	11 (15.9)	14 (13.6)
Dysgeusia	2 (8.3)	9 (13.0)	11 (10.7)
Neuropathy peripheral	1 (4.2)	7 (10.1)	5 (4.9)
Syncope	0	3 (4.3)	11 (10.7)
Lethargy	4 (16.7)	1 (1.4)	3 (2.9)
Psychiatric Disorders			
Insomnia	3 (12.5)	8 (11.6)	12 (11.7)
Confusional state	3 (12.5)	4 (5.8)	10 (9.7)
Mental status changes	3 (12.5)	0	3 (2.9)
Renal and Urinary Disorders			
Acute kidney injury	3 (12.5)	5 (7.2)	14 (13.6)
Respiratory, Thoracic and Mediastinal Disorders			
Dyspnoea	5 (20.8)	19 (27.5)	30 (29.1)
Cough	7 (29.2)	18 (26.1)	27 (26.2)
Epistaxis	4 (16.7)	15 (21.7)	20 (19.4)
Hypoxia	3 (12.5)	7 (10.1)	14 (13.6)
Pleural effusion	1 (4.2)	4 (5.8)	12 (11.7)
Nasal congestion	3 (12.5)	3 (4.3)	9 (8.7) 0
Respiratory failure	1 (4.2)	1 (1.4)	11 (10.7)

MedDRA V20.0 System Organ Class Preferred Term	Gilteritinib		
	80mg (N=24) n (%)	120mg (N=69) n (%)	200mg (N=103) n (%)
Skin and Subcutaneous Tissue Disorders			
Rash	0	11 (15.9)	11 (10.7)
Petechiae	1 (4.2)	0	15 (14.6)
Hyperhidrosis	3 (12.5)	1 (1.4)	4 (3.9)
Vascular Disorders			
Hypotension	4 (16.7)	12 (17.4)	29 (28.2)
Hypertension	1 (4.2)	7 (10.1)	16 (15.5)

ADMIRAL adverse events

Table 51: ADMIRAL AEs - Most frequent treatment-emergent AEs (≥ 10% patients in either treatment arm) by preferred term and severity – Safety analysis set (adapted from CS Appendix F, Table 27)

MedDRA v19.1 Preferred Term	Gilteritinib (n=246)	Chemotherapy (n=109)	Gilteritinib (n=246)	Chemotherapy (n=109)
	All	All	≥ Grade 3	≥ Grade 3
	n=246 n (%)	n=109 n (%)	n=246 n (%)	n=109 n (%)
Overall	[redacted]	[redacted]	[redacted]	[redacted]
Anaemia	[redacted]	[redacted]	[redacted]	[redacted]
Febrile neutropenia	[redacted]	[redacted]	[redacted]	[redacted]
Pyrexia	[redacted]	[redacted]	[redacted]	[redacted]
Alanine aminotransferase increased	[redacted]	[redacted]	[redacted]	[redacted]
Aspartate aminotransferase increased	[redacted]	[redacted]	[redacted]	[redacted]
Diarrhoea	[redacted]	[redacted]	[redacted]	[redacted]
Nausea	[redacted]	[redacted]	[redacted]	[redacted]
Constipation	[redacted]	[redacted]	[redacted]	[redacted]
Cough	[redacted]	[redacted]	[redacted]	[redacted]
Hypokalaemia	[redacted]	[redacted]	[redacted]	[redacted]
Fatigue	[redacted]	[redacted]	[redacted]	[redacted]
Headache	[redacted]	[redacted]	[redacted]	[redacted]
Thrombocytopenia	[redacted]	[redacted]	[redacted]	[redacted]
Oedema peripheral	[redacted]	[redacted]	[redacted]	[redacted]
Dyspnoea	[redacted]	[redacted]	[redacted]	[redacted]
Blood alkaline phosphatase increased	[redacted]	[redacted]	[redacted]	[redacted]
Platelet count decreased	[redacted]	[redacted]	[redacted]	[redacted]
Vomiting	[redacted]	[redacted]	[redacted]	[redacted]
Dizziness	[redacted]	[redacted]	[redacted]	[redacted]
Hypocalcaemia	[redacted]	[redacted]	[redacted]	[redacted]
Decreased appetite	[redacted]	[redacted]	[redacted]	[redacted]
Hypotension	[redacted]	[redacted]	[redacted]	[redacted]
Pneumonia	[redacted]	[redacted]	[redacted]	[redacted]
Epistaxis	[redacted]	[redacted]	[redacted]	[redacted]
Neutrophil count decreased	[redacted]	[redacted]	[redacted]	[redacted]
Hypophosphataemia	[redacted]	[redacted]	[redacted]	[redacted]
Insomnia	[redacted]	[redacted]	[redacted]	[redacted]
Hypomagnesaemia	[redacted]	[redacted]	[redacted]	[redacted]
Asthenia	[redacted]	[redacted]	[redacted]	[redacted]
Abdominal pain	[redacted]	[redacted]	[redacted]	[redacted]
Hyperglycaemia	[redacted]	[redacted]	[redacted]	[redacted]
Pain in extremity	[redacted]	[redacted]	[redacted]	[redacted]
Rash	[redacted]	[redacted]	[redacted]	[redacted]
Myalgia	[redacted]	[redacted]	[redacted]	[redacted]
Hypertension	[redacted]	[redacted]	[redacted]	[redacted]
Stomatitis	[redacted]	[redacted]	[redacted]	[redacted]
White blood cell count decreased	[redacted]	[redacted]	[redacted]	[redacted]

MedDRA v19.1 Preferred Term	Gilteritinib (n=246)	Chemotherapy (n=109)	Gilteritinib (n=246)	Chemotherapy (n=109)
	All	All	≥ Grade 3	≥ Grade 3
	n=246 n (%)	n=109 n (%)	n=246 n (%)	n=109 n (%)
Acute myeloid leukaemia	[redacted]	[redacted]	[redacted]	[redacted]
Blood creatine phosphokinase increased	[redacted]	[redacted]	[redacted]	[redacted]
Hyponatremia	[redacted]	[redacted]	[redacted]	[redacted]
Neutropenia	[redacted]	[redacted]	[redacted]	[redacted]
Hypoalbuminaemia	[redacted]	[redacted]	[redacted]	[redacted]
Back pain	[redacted]	[redacted]	[redacted]	[redacted]
Blood Creatinine increased	[redacted]	[redacted]	[redacted]	[redacted]
Arthralgia	[redacted]	[redacted]	[redacted]	[redacted]
Dysgeusia	[redacted]	[redacted]	[redacted]	[redacted]

Reasons for exclusion from NMA

Table 52: Table of reasons for exclusion from an NMA (reproduced from company's clarification response, question B2 Table 7)

Study name	Reference	Reference	Treatment arms	Reason for exclusion
CHRYSALIS NCT02014558	Perl 2017B	Perl, A. E., Altman, J. K., Cortes, J., Smith, C., Litzow, M., Baer, M. R. & Jurcic, J. G. (2017). Selective inhibition of FLT3 by gilteritinib in relapsed or refractory acute myeloid leukaemia: a multicentre, first-in-human, open-label, phase 1–2 study. The Lancet Oncology, 18(8), 1061-1075.	Seven doses of gilteritinib	Dose finding study. It could not enrich the network evidence
2215-CL-0303 NCT03182244	CT.Gov	https://clinicaltrials.gov/	Gilteritinib vs salvage chemo (LoDAC, MEC, G-CSF, FLAG)	No results
2215-CL-1101 NCT02421939	CT.Gov	https://clinicaltrials.gov/	Gilteritinib + atezolizumab	No results
2215-CL-9100 NCT03070093	CT.Gov	https://clinicaltrials.gov/	Gilteritinib	No results
2215-CL-9200 NCT03409081	CT.Gov	https://clinicaltrials.gov/	Gilteritinib	No results
M16-802 NCT03625505	CT.Gov	https://clinicaltrials.gov/	Venetoclax + Gilteritinib	Single arm
2018-0608 NCT03735875	CT.Gov	https://clinicaltrials.gov/	Venetoclax + quizatinib 30 mg/d	Single arm
2689-CL-2004 NCT01565668	Cortes 2018A;	Cortes, J.E., Tallman, M.S., Schiller, G.J., Trone, D., Gammon, G., Goldberg, S.L., Perl, A.E., Marie, J.P., Martinelli, G., Kantarjian, H.M. and Levis, M.J., 2018. Phase 2b study of two dosing regimens of quizartinib monotherapy in FLT3-ITD mutated, relapsed or refractory AML. Blood, pp.blood-2018.	Quizartinib 30 mg/d vs 0 mg/d	Dose finding study. It could not enrich the network evidence
ACE NCT00989261	Martinelli 2014	Martinelli, G., Levis, M.J., Perl, A.E., Dombret, H., Steffen, B., Rousselot, P., Estey, E.H., Shah, N.P., Gammon, G., Trone, D. and Cortes, J.E., 2014, June. Treatment with quizartinib (AC220) enables a high rate of patients with relapsed or refractory FLT3-ITD	Quizartinib 200mg	Single arm

Study name	Reference	Reference	Treatment arms	Reason for exclusion
		(+) acute myeloid leukemia to be bridged to HSCT. In Haematologica (Vol. 99, pp. 35-35).		
AC220-A-J201 NCT02984995	CT.Gov	https://clinicaltrials.gov/	Quizartinib	Single arm; no results
AC220-A-U203 NCT03746912	CT.Gov	https://clinicaltrials.gov/	Quizartinib	Single arm; no results
ARO-004 ARO-005 NCT01522469 NCT01657682	Cortes 2016A	Cortes, J.E., Kantarjian, H.M., Kadia, T.M., Borthakur, G., Konopleva, M., Garcia-Manero, G., Daver, N.G., Pemmaraju, N., Jabbour, E., Estrov, Z. and Ramachandran, A., 2016. Crenolanib besylate, a type I pan-FLT3 inhibitor, to demonstrate clinical activity in multiply relapsed FLT3-ITD and D835 AML. Abstract #7008. Presented at the 2016 American Society of Clinical Oncology Annual Meeting, Chicago, IL, June 4, 2016.	Crenolanib	Single arm
ARO-007 NCT02298166	CT.Gov	https://clinicaltrials.gov/	Crenolanib + chemotherapy vs placebo + chemotherapy	No results
ARO-013 NCT03250338	CT.Gov	https://clinicaltrials.gov/	Crenolanib + chemotherapy (HAM+FLAG-IDA) vs chemotherapy	No results
N/A	Iyer 2016	Iyer, S.P., Jethava, Y., Karanes, C., Eckardt, J.R. and Collins, R., 2016. Safety study of salvage chemotherapy high-dose Ara-C/mitoxantrone (HAM) and type I FLT3-TKI crenolanib in first relapsed/primary refractory AML. Blood . 2016;128:3983.	HAM followed by crenolanib	Single arm
N/A	Randhawa 2014	Randhawa, J.K., Kantarjian, H.M., Borthakur, G., Thompson, P.A., Konopleva, M., Daver, N., Pemmaraju, N., Jabbour, E., Kadia, T.M., Estrov, Z. and Ramachandran, A., 2014. Results of a phase II study of crenolanib in relapsed/refractory acute myeloid leukemia patients (Pts) with activating FLT3 mutations. Blood 2014 124:389	Crenolanib 200 mg/m/day TID	Single arm
2010-0511 NCT01254890	Ravandi 2013	Ravandi, F., Alattar, M.L., Grunwald, M.R., Rudek, M.A., Rajkhowa, T., Richie, M.A., Pierce, S., Daver, N., Garcia-Manero, G., Faderl, S. and Nazha, A., 2013. Phase II study of azacitidine	Sorafenib (400 mg orally BID) + azacitidine	Single arm

Study name	Reference	Reference	Treatment arms	Reason for exclusion
		plus sorafenib in patients with acute myeloid leukemia and FLT-3 internal tandem duplication mutation. Blood, 2013.		
AML004 NCT03622541	CT.Gov	https://clinicaltrials.gov/	Sorafenib	Single arm
KCP-330-001 NCT01607892	Daver 2017	Daver, N., Assi, R., Garcia-Manero, G., Ravandi, F., Borthakur, G., Jabbour, E.J., DiNardo, C.D., Kadia, T., Ning, J., González, G.N. and Pierce, S., 2017. a Phase I/II Study of Selinexor (SEL) with Sorafenib in Patients (pts) with Relapsed and/or Refractory (R/R) FLT3 Mutated Acute Myeloid Leukemia (AML).	Sorafenib (400mg BID) + selinexor	Single arm
SIRA NCT02867891	CT.Gov	https://clinicaltrials.gov/	Sorafenib	No results
N/A	Fleischmann 2016	Fleischmann M.; Schrenk K.G.; Schnetzke U.; Hilgendorf I.; Hochhaus A.; Scholl S., 2016. Evaluation of tyrosine kinase inhibitor treatment in patients with FLT3-ITD positive acute myeloid leukemia. ncology Research and Treatment. Conference: Jahrestagung der Deutschen, Osterreichischen und Schweizerischen Gesellschaften fur Hamatologie und Medizinische Onkologie 2016. Germany. 39 (Supplement 3) (pp 50-51), 2016.	Sorafenib	Single arm
N/A	Fleischmann 2017	Fleischmann, M., Schnetzke, U., Schrenk, K.G., Schmidt, V., Sayer, H.G., Hilgendorf, I., Hochhaus, A. and Scholl, S., 2017. Outcome of FLT3-ITD-positive acute myeloid leukemia: impact of allogeneic stem cell transplantation and tyrosine kinase inhibitor treatment. Journal of cancer research and clinical oncology, 143(2), pp.337-345.	Sorafenib	Single arm
N/A	Freitas 2016	De Freitas, T., Marktel, S., Piemontese, S., Carrabba, M.G., Tresoldi, C., Messina, C., Lupo Stanghellini, M.T., Assanelli, A., Corti, C., Bernardi, M. and Peccatori, J., 2016. High rate of hematological responses to sorafenib in FLT 3-ITD acute myeloid leukemia relapsed after allogeneic hematopoietic stem cell transplantation. European journal of haematology, 96(6), pp.629-636.	Sorafenib 400mg BID	Single arm

Study name	Reference	Reference	Treatment arms	Reason for exclusion
N/A	Metzelder 2009	Metzelder, S., Wang, Y., Wollmer, E., Wanzel, M., Teichler, S., Chaturvedi, A., Eilers, M., Enghofer, E., Neubauer, A. and Burchert, A., 2009. Compassionate use of sorafenib in FLT3-ITD–positive acute myeloid leukemia: sustained regression before and after allogeneic stem cell transplantation. Blood, 113(26), pp.6567-6571.	Sorafenib 400 mg BID	Single arm
N/A	Metzelder 2010	Metzelder, S.K., Wollmer, E., Neubauer, A. and Burchert, A., 2010. Sorafenib in relapsed and refractory FLT3-ITD positive acute myeloid leukemia: a novel treatment option. Deutsche medizinische Wochenschrift (1946), 135(38), pp.1852-1856.	Sorafenib 400 mg BID	Single arm
N/A	Metzelder 2012	Metzelder, S.K., Schroeder, T., Finck, A., Scholl, S., Fey, M., Götze, K., Linn, Y.C., Kröger, M., Reiter, A., Salih, H.R. and Heinicke, T., 2012. High activity of sorafenib in FLT3-ITD-positive acute myeloid leukemia synergizes with allo-immune effects to induce sustained responses. Leukemia, 26(11), p.2353.	Sorafenib 400 mg BID	Single arm
N/A	Metzelder 2017	Metzelder, S.K., Schroeder, T., Lübbert, M., Ditschkowski, M., Götze, K., Scholl, S., Meyer, R.G., Dreger, P., Basara, N., Fey, M.F. and Salih, H.R., 2017. Long-term survival of sorafenib-treated FLT3-ITD–positive acute myeloid leukaemia patients relapsing after allogeneic stem cell transplantation. European journal of cancer, 86, pp.233-239.	Sorafenib	Single arm
N/A	Rautenberg 2017	Rautenberg, C., Nachtkamp, K., Dienst, A., Schmidt, P.V., Heyn, C., Kondakci, M., Germing, U., Haas, R., Kobbe, G. and Schroeder, T., 2017. Sorafenib and azacitidine as salvage therapy for relapse of FLT 3-ITD mutated AML after allo-SCT. European journal of haematology, 98(4), pp.348-354.	Sorafenib (400 mg BID) + azacitidine (75 mg/m^2 for 7 d every 28 d)	Single arm
N/A	Schroeder 2009	Schroeder, T., Saure, C., Bruns, I., Zohren, F., Czibere, A.G., Safaian, N.N., Fenk, R., Haas, R. and Kobbe, G., 2009. Clinical Efficacy of Sorafenib in Patients with Acute Myeloid Leukemia (AML) and Activating FLT3-Mutations.	Sorafenib 800mg QD	Single arm

Study name	Reference	Reference	Treatment arms	Reason for exclusion
N/A	Sharma 2011	Sharma, M., Ravandi, F., Bayraktar, U.D., Chiattone, A., Bashir, Q., Giralt, S., Chen, J., Qazilbash, M., Kebriaei, P., Konopleva, M. and Andreeff, M., 2011. Treatment of FLT3-ITD-positive acute myeloid leukemia relapsing after allogeneic stem cell transplantation with sorafenib. Biology of Blood and Marrow Transplantation, 17(12), pp.1874-1877.	Sorafenib 400mg BID or 600mg BID +/- chemotherapy	Single arm
N/A	Sid 2017	Sid, S., Rey, J., Charbonnier, A., D'incan, E., Mohty, B., Blaise, D. and Vey, N., 2017. Treatment of Post-transplant Relapse of FLT3-ITD Mutated AML Using 5-Azacitidine and Sorafenib Bitherapy. Clinical Lymphoma, Myeloma and Leukemia, 17(4), pp.241-242.	Sorafenib (400 mg BID) + azacitidine (75 mg/m^2 for 7 d every 28 d)	Single arm
N/A	Xuan 2017	Xuan, L., Wang, Y., Huang, F., Wu, B., Fan, Z., Xu, N., Ye, J., Sun, J., Huang, X. and Liu, Q., 2017. The Effect of Sorafenib Therapy on the Outcome of Acute Myeloid Leukemia with FLT3-ITD Undergoing Allogeneic Hematopoietic Stem Cell Transplantation.	Sorafenib (400mg BID) + chemo + donor lymphocyte infusions	Single arm
NCT03642236	CT.Gov	https://clinicaltrials.gov/	Ibrutinib 420mg + sorafenib 0.4mg BID	No results
N/A	Chevallier 2010	Chevallier, P., Prebet, T., Pigneux, A., Hunault, M., Delaunay, J., Perry, F., Lode, L., Richebourg, S., Blanchet, O., Vey, N. and Ifrah, N., 2010. Influence of NPM1 and FLT3-ITD status on outcome in relapsed/refractory AML patients receiving salvage therapy including gemtuzumab ozogamicin. Leukemia, 24(2), p.467.	Gemtuzumab ozogamicin (9 mg/m2 at day 4) + cytarabine (1 g/m2 BID for days 1–5) + mitoxantrone (12 mg/m2/d for days 1–3)	Single arm

Appendix 2: Log-cumulative hazard plots - No HSCT group in ADMIRAL and post-HSCT group in Evers *et al*

Figure 25: Log-cumulative hazard plot for No HSCT, OS, ADMIRAL, gilteritinib group, (reproduced from company's clarification response, question B12, Figure 4)

Figure 26: Log-cumulative hazard plot for No HSCT, OS, ADMIRAL, salvage chemotherapy group (reproduced from company's clarification response, question B12, Figure 5)

Figure 27: Log-cumulative hazard plot for No HSCT, EFS, ADMIRAL, gilteritinib group (reproduced from company's clarification response, question B12, Figure 6)

Figure 28: Log-cumulative hazard plot for No HSCT, EFS, ADMIRAL, salvage chemotherapy group (reproduced from company's clarification response, question B12, Figure 7)

Figure 29: Log cumulative hazard plot of post-HSCT OS, Evers et al (reproduced from CS, Figure 19)

Appendix 3: Technical appendix – instructions for implementing the ERG's exploratory analyses within the company's model

ERG exploratory analysis 1: Correction of errors

In the company's model:

(i) Replace cells J26:J506 in worksheets 'Summary_Gilt', 'Summary_Aza', 'Summary_FLAG_IDA', 'Summary_MEC, 'Summary_LDAC', and 'Summary_BSC' with the values in Worksheet 'ERG EA 1', cells P4:P484 of the file "ERGTechnicalappendix.xlsx".

(ii) Replace cells AJ14:AJ494 in Worksheet 'Effectiveness_calculation' of the company's model with the values in cells W4:W484 in Worksheet 'ERG EA 1' of the file "ERGTechnicalappendix.xlsx".

(iii) In worksheets 'Summary_Gilt', 'Summary_Aza', 'Summary_FLAG_IDA', 'Summary_MEC, and 'Summary_LDAC' of the company's model, copy the formulas in cells M26 and M27 to cells L26 and L27, checking that they refer to the same cell references. Drag the formula in cell L27 down to row 506. Replace the formula in cell M26 with the formula "=IF(L26>(1-O26),(1-O26),L26)". Drag the formula down to row 506.

(iv) In worksheet 'Utility', replace cells E26:E28 with, respectively, the values "55", "65", and "75". Copy the values in cells F23:F25 and paste the values in cells F26:F28. In worksheets 'Summary_Gilt', 'Summary_Aza', 'Summary_FLAG_IDA', 'Summary_MEC, 'Summary_LDAC', and 'Summary_BSC', replace the content in cells AX27:AX506 with the value 1.0. Replace the variable 'longterm_utility' in the formulae in cells AZ27, BB27, BD27 and BF27 with the formula 'VLOOKUP($G27,Utility!$E$26:$F$28,2,1)'. Drag each formula down to row 506.

All other exploratory analyses undertaken by the ERG are applied within this corrected version of the model.

ERG exploratory analysis 2: Use of ADMIRAL With HSCT data on time from randomisation to death

Go to "ERGTechnicalappendix.xlsx", Worksheet 'ERG EA2 and ASA 3'. Copy cells Q5:S41. Go to the company's model, and select worksheets 'Summary_Gilt', 'Summary_Aza', 'Summary_FLAG_IDA', 'Summary_MEC, and 'Summary_LDAC' simultaneously. Click on cell R26 and paste values.

In these same worksheets, change the content of cell P18 (HSCT start month) to '0.0', and change BM27 to '=$BM$12/2*$N$18', which corresponds to half the value of the disutility due to HSCT, applied to the treatment-specific HSCT rate. Finally, change the content of cells BM28:BM506 to '0.0'.

ERG exploratory analysis 3: Use of utilities from Ara and Brazier

In the company's model, simultaneously select worksheets 'Summary_Gilt', 'Summary_Aza', 'Summary_FLAG_IDA', 'Summary_MEC, 'Summary_LDAC', and 'Summary_BSC'. In each of cells AZ27, BB27, BD27 and BF27, replace the part of the formula 'VLOOKUP($G27,Utility!$E$26:$F$28,2,1)' with the formula '0.9508566+(0.0212126*'Life Table'!$F$11)-(0.0002587*$G27)-(0.0000332*$G27^2)'. Drag the amended formulae down to row 506.

ERG exploratory analysis 4: Remove double-counting of AML progression as AE

In the company's model, go to worksheet 'Safety', cells F35:M35. Replace values with "0.0".

ERG exploratory analysis 5: Remove HR for gilteritinib maintenance therapy

Go to "ERGTechnicalappendix.xlsx" Worksheet 'ERG EA 5'cells F5:F484. Copy selection. Go to company's model, Worksheet 'Effectiveness_calculation', cell AJ14. Paste values.

ERG exploratory analysis 6: Inclusion of wastage for gilteritinib (0.50 packs)

Go to company's model, worksheet 'Summary_Gilt', cell CC27. Add the following term to the end of the formula in this cell only '+(0.5*AI62*(cost_drug_Gilt))'

ERG exploratory analysis 7: Cost amendments (follow-up included, post-cure excluded, tests increased)

In the company's model:

(i) Create a variable in worksheet "Resource Use" called "LTcosts_WithHSCT" (for example, in cell J94). Replace the content of the cell with "=(1/12)*$F$33". In worksheets 'Summary_Gilt', 'Summary_Aza', 'Summary_FLAG_IDA', 'Summary_MEC, 'Summary_LDAC' and 'Summary_BSC', replace the variable called "longterm_cost" in the formulas in cells CO27 and CQ27 with the variable "LTcosts_WithHSCT". Drag the formulas down to row 506.

(ii) Create a variable in worksheet "Resource Use" and name it "LTcosts_NoHSCT" (for example, in cell J93). Replace the content of the cell with "=(1/6)*$F$33". In worksheets 'Summary_Gilt', 'Summary_Aza', 'Summary_FLAG_IDA', 'Summary_MEC,

'Summary_LDAC' and 'Summary_BSC', replace the variable called "longterm_cost" in the formulas in cells CK27 and CM27 with the variable "LTcosts_NoHSCT". Drag the formulas down to row 506.

(iii) In worksheets 'Summary_Gilt', 'Summary_Aza', 'Summary_FLAG_IDA', 'Summary_MEC, and 'Summary_LDAC', replace the formula in cells CI63:CI506 with the value "0.0".

(iv) In worksheet 'Resource Use', we replaced the value in cell CC27 for '=1/0.3'.

ERG preferred base case

The ERG's preferred base case includes ERG exploratory analysis 1-7; therefore, apply all the changes listed above. Please note that the action required for Exploratory Analysis 2 overrides actions in exploratory analysis 5.

ERG sensitivity analysis 1: Use of Poiré *et al* to inform post-HSCT OS

In the company's model, replace the content in cells R26:T62 in worksheets 'Summary_Gilt', 'Summary_Aza', 'Summary_FLAG_IDA', 'Summary_MEC, and 'Summary_LDAC' with the values from Worksheet 'ERG ASA 1' of the file "ERGTechnicalappendix.xlsx", cells M4:S40.

ERG sensitivity analysis 2: Use of alternative parametric survival models for OS within the No HSCT group

In the company's model, go to Worksheet 'Effectiveness' and change all the curve selections in the dropdown menu in cells T13:T17. Choose from the options "Exponential", "Weibull", "lognormal", "Gompertz" and "Generalised gamma". Choose the same option for all cells.

ERG sensitivity analysis 3: Use of alternative parametric survival models for OS within the With HSCT group

For log logistic OS function, replace the content in cells R26:T62 in worksheets 'Summary_Gilt', 'Summary_Aza', 'Summary_FLAG_IDA', 'Summary_MEC, and 'Summary_LDAC' of the company's model with the values in cells Q5:S41 from Worksheet 'ERG ASA 3' of the file "ERGTechnicalappendix.xlsx". For Weibull OS functions, replace the same cells in the company's model with the values in cells W5:Y41 from Worksheet 'ERG ASA 3'

ERG sensitivity analysis 4: Exploration of alternative cure points

Apply all of changes in the ERG's base case, except for pasting With HSCT trace in ERG exploratory analysis 2. In the company's model, set the relevant cure point for survival, utilities and costs in

worksheet “Specifications”. In each model worksheet, adjust the number of cycles over which post-progression treatments are applied manually to reflect the cure timepoint. In the file “ERGTechnicalappendix.xlsx”, extend/reduce the trace to reflect the selected cure timepoint. Copy the trace. Paste this into each model worksheet cell R26.

ERG sensitivity analysis 5: Exploration of proportion of patients achieving cure

Not applicable

National Institute for Health and Care Excellence
Centre for Health Technology Evaluation

ERG report – factual accuracy check – ERG response

Gilteritinib for treating relapsed or refractory acute myeloid leukaemia [ID1484]

You are asked to check the ERG report to ensure there are no factual inaccuracies contained within it.

If you do identify any factual inaccuracies you must inform NICE by **5pm on Tuesday 3 September 2019** 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.

The factual accuracy check form should act as a method of detailing any inaccuracies found and how and why they should be corrected.

Issue number and heading	Description of problem	Description of proposed amendment	Justification for amendment	ERG response
1. Description of inclusion of BSC in Treatment Pathway	Page 9: The report comments *'The ERG's clinical advisor commented that the pathway should have included low intensity chemotherapy and best supportive care (BSC) as potential options.'*	Amend to *'The ERG's clinical advisor commented that the pathway should have included low intensity chemotherapy more clearly as a potential option.'*	Best supportive care was included in the Treatment Pathway presented in the Company Submission and low intensity chemotherapy was part of "salvage chemotherapy"	The text has been amended as follows: "*The ERG's clinical advisor commented that the diagram of the treatment pathway should have included low intensity chemotherapy and BSC more clearly as potential options"*

2. Denotion of Trademark	Page 14: Text reads: 'ASP2215, XOSPATA®'	Amend the denotion of the Trademark to TM not ® to: 'ASP2215, XOSPATA TM'	In line with Company Submission and Astellas policy.	We have amended the text as requested, but note that the Astellas website refers to XOSPATA® not XOSPATA™
3. Confidentiality highlighting	Page 34: Academic in Confidence highlighting is not required in the following sentence [redacted]	Highlighting not required.	Not AIC data. This is in line with the Company Submission.	The confidentiality marking has been lifted as requested.
4. Confidentiality highlighting	Page 34: Academic in Confidence highlighting is not required in the following sentence: *'This modification of EFS showed a statistically significant benefit for gilteritinib* [redacted] .'	Highlighting not required.	Not AIC data. This is in line with the Company Submission.	As requested, we have lifted the confidentiality marking from this text. However, please note that the results are marked as confidential in the latest version of the company's submission available on NICEDocs (made available to the ERG 15th August 2019). **Astellas - please check that these data do not need to remain confidential.**
5. Confidentiality highlighting	Page 43: Academic in Confidence highlighting is not required in the following	Highlighting not required.	Not AIC data. This is in line with the Company Submission.	We have lifted the confidentiality marking in this sentence. Please note

	sentence: 'The company's review of HRQoL studies identified [redacted] studies reporting health utility values for patients with AML.'			that the ERG was previously informed that the company's SLR document should be treated as being commercial-in-confidence. **Astellas - please check that these data do not need to remain confidential.**
6. Confidentiality highlighting	Page 54 and 57: Academic in Confidence highlighting is required in the following sentences: '…for patients who did not receive HSCT in ADMIRAL (gilteritinib n=[redacted]; salvage chemotherapy n=[redacted] patients)' on Page 54. '…for those patients who did not receive HSCT during the follow-up period of ADMIRAL (gilteritinib n=[redacted]; salvage chemotherapy=[redacted])' on Page 57	Highlight the number of patients in the following sentences: '…for patients who did not receive HSCT in ADMIRAL (gilteritinib n=[redacted]; salvage chemotherapy n=[redacted] patients)' on Page 54. '…for those patients who did not receive HSCT during the follow-up period of ADMIRAL (gilteritinib n=[redacted]; salvage chemotherapy=[redacted])' on Page 57.	Data presented have not yet been published and should thus be marked as Academic in Confidence. This is in line with the Company Submission.	The numbers of patients have been marked as requested.
7. Error in BSC AE rates	Page 67: The costs of Grade 3+ AEs (once-only) for BSC is reported in Table 25 of the ERG report as '£1,828'.	This value in Table 25 on Page 67 should be revised to: 'N/A'	The costs associated with salvage chemotherapy appear to have been copied across to BSC. No AE costs were assumed for BSC.	We agree with the company – this was a minor reporting error. The table has been amended to state "N/a" for the costs of treating AEs for BSC. No results are affected by

				this error.
8. Inaccurate reporting of the data	Page 96 (second bullet point): The ERG report states 'At the final data cut-off of ADMIRAL, [redacted] of all patients had already died.'	The sentence should be modified to read 'At the final data cut-off of ADMIRAL, [redacted] of gilteritinib-treated patients who received HSCT had already died;…'	Whereas the prior and subsequent sentences discuss the gilteritinib-treated HSCT patients, this sentence is misleading. The amendment should be made in order to clarify that 46% refers to HSCT patients, not the whole trial population.	We agree with the company – the text has been amended as requested.
9. Incomplete description of the HSCT data in ADMIRAL trial	Page 96: The ERG report states 'The available data from gilteritinib-treated patients who received HSCT in ADMIRAL suggest a cumulative OS probability of [redacted] at around [redacted] months (the last observed event).'	Revise the statement by providing additional details to reflect a full picture: 'The available data from gilteritinib-treated patients who received HSCT in ADMIRAL suggest a cumulative OS probability of [redacted] at around [redacted] months (the last observed event). However this should be interpreted with caution due to patients being censored on the date of the final data cut.'	The original statement does not provide full details, and may be misinterpreted. Given the short follow up period for these patients, it is not appropriate to assume the survival probability based on the very low numbers of patients at risk. In the ADMIRAL trial, the median follow-up post-HSCT survival was [redacted] months with only [redacted] patients having follow-up data beyond year [redacted], and just [redacted] patients having follow-up data beyond year 2.	This is not a factual error. The cumulative survival probability reported in the text relates to the time of the last observed event (not the time point for the last censored observation). The text in the ERG report already refers to the level of censoring in the available With SHCT data from ADMIRAL, and the Kaplan-Meier survivor functions are presented in Figure 15 with numbers at risk. The text has not been amended.
10. Irrelevant comparison of mortality rates extrapolated by post-HSCT Gompertz	Page 98: The ERG report states 'Figure 19 presents a comparison of the monthly mortality rates predicted by the company's post-HSCT Gompertz model and the	Remove Figure 19 and relevant discussions.	In the model, the mortality rates extrapolated by the post-HSCT Gompertz model are applied to the first 3-year only. Afterwards, all patients follow the SMR-adjusted	This is not a factual inaccuracy. The company's clarification response (question B16) states that

model to those of general population	approximate equivalent rates from general population life tables. As shown in the figure, the mortality rate from the Gompertz becomes lower than that for the general population after 9 years (around age 68 in the model). The extrapolated post-HSCT Gompertz function (excluding the 3-year cure assumption) therefore suggests a survival prognosis for R/R AML patients which is better than the survival prognosis in the general population; the ERG does not consider this to be plausible.'		mortality risk, which is twice that of the general population. Therefore, it is not meaningful to compare the post-HSCT Gompertz model with the natural mortality beyond year 3.	drawing on external data and clinical opinion, the selected Gompertz post-HSCT OS model *"performs a plausible extrapolation."* Figure 19 of the ERG report highlights that the long-term extrapolated OS from this model is implausible as it indicates a lower risk of death compared with that for the general population (i.e. it implies a survival rate which is better than cure). The ERG notes that the company's decision to apply an assumption of a 3-year cure point which overrides the hazards predicted by the Gompertz OS model is a separate issue which is unrelated to the justification for firstly selecting the Gompertz model. The report has not been amended.
11. Overestimating impact of wastage	Page 100: In ERG exploratory analysis 6, the model was modified to assume that half a	The wastage assumption should be assigned to patients who died before the time-on-treatment	The proposed ERG modification overestimates the potential for drug	The company's criticism is not consistent with the observed data from

	pack's wastage would be incurred by all patients who died before the 3-year cure point. This statement is counterintuitive as patients are modelled to receive [redacted] cycles of gilteritinib, hence those patients who died after cycle [redacted], but before the cure point (cycle 36), would not be receiving treatment and accordingly would not waste treatment by dying.	with gilteritinib concludes i.e. the proportion who have died up to cycle [redacted]. This should also be amended on page 147 (ERG exploratory analysis 6) where the new formula should read '+(0.5*AI33*(cost_drug_Gil))'. Exploratory analysis 6 results should be updated accordingly.	wastage as it assigns half a pack's cost to all patients who have died before the cure point i.e. [redacted]%. The proportion who have died by the end of time-on-treatment (cycle [redacted]) – and whose death would be associated with drug wastage - is lower i.e. [redacted]%.	ADMIRAL. It is not the case that all gilteritinib-treated patients survive and receive gilteritinib for [redacted] cycles and then discontinue treatment. The probability of remaining on treatment follows a time-to-event distribution, as shown in Figure 21 of the ERG report. In addition, the company's submission clarifies that at the final data cut-off, [redacted] of gilteritinib-treated patients were still receiving gilteritinib. This time point is considerably later than [redacted] cycles. The ERG notes that this time to treatment discontinuation distribution, and the costs of any gilteritinib wastage, are not included in the company's model. This presents a problem for estimating the cost-effectiveness of gilteritinib. Given the limitations of

				the company's approach to modelling costs, the ERG believes that the most reasonable approach to estimating wastage would be to apply the cost of half a pack for all patients who die before the assumed cure point. The ERG believes that the approach suggested by the company in this fact check response would considerably underestimate the costs of wastage for gilteritinib.
12. Lack of acknowledgement of limitations in ERG analyses	Pages 107-109: ERG exploratory analysis 2, "Use of ADMIRAL With HSCT data on time from randomisation to death" describes how the ERG consider the selection of a log normal model fitted to the limited ADMIRAL post-HSCT data, to be the most plausible source of post-transplant survival and hence included in the ERG-preferred base-case analysis.	Astellas does not believe that it is accurate to present this data as the most plausible post-HSCT survival curve, given the short duration of follow up and high censoring. The ERG should present this analysis as exploratory and acknowledge clearly the substantial limitations of the data informing this curve.	Alternative published data sources are available, and provide more certainty in post-HSCT survival estimates due to their longer follow up time. As previously stated, the HSCT survival data from the ADMIRAL trial has limited follow-up to reflect the expected long-term cure effect. In the ADMIRAL trial, the median follow-up post-HSCT survival was [redacted] months with only [redacted] patients having follow-up data	This is not a factual inaccuracy. All of the ERG's additional analyses appear in Section 5.4 under the heading "ERG's exploratory analyses." The ERG believes that ADMIRAL reflects the most relevant source of evidence on the expected outcomes for gilteritinib-treated patients who undergo HSCT. If external information is considered relevant for

			beyond year 1, and just [redacted] patients having follow-up data beyond year 2. Therefore, the ADMRIAL data likely does not have sufficient information to reflect the cure pattern of HSCT patients. This was similarly the case in a recent appraisal in ovarian cancer where, in the absence of mature overall survival data from the pivotal trial, an external data source was selected as the most plausible input to provide an indication of expected survival outcomes in practice.	inclusion in the model, then it should supplement the evidence from ADMIRAL, not replace it. The ERG report also highlights that whilst there is uncertainty in the ADMIRAL With HSCT data, it is unlikely that any cure fraction achieved would be as high as that implied by the company's model. The ERG's commentary on the limitations of the company's model and the uncertainty in the ADMIRAL data are already clearly discussed in the ERG report.
13. Lack of acknowledgement of limitations in ERG analyses	Page 108: A scenario analyses was carried out by the ERG, where parametric functions were fit to the pooled ADMIRAL HSCT data, from randomisation to death, for predicting OS for patients with HSCT. All the fitted curves, however, do not reflect the initial flat period (for about 3 months) of the OS curves well (Figure 22	The ERG should acknowledge the limitations of this scenario by adding the bold text in the sentence below. On Page 108: 'On the basis of its goodness-of-fit statistics, the underlying nature of the hazard function and the plausibility of the extrapolated survivor function, the ERG selected the	Given the initial flat period of the OS curves are not well reflected in the parametric functions. ERG should acknowledge the limitation. Alternatively, the ERG might consider fitting the curves from the time of the HSCT, and model the time to HSCT separately in the model.	This is not a factual inaccuracy. The ERG notes that the standard log normal model used in the ERG-preferred base case predicts an overall mortality probability during this initial "flat" 3-month period of 1%. It is clear from Figure 22 and Table 36 of the ERG

	on Page 108). During this plateau period, no death occurred in the observed data; however, all the fitted curves predict death occurred during this period.	log normal model for inclusion in the ERG-preferred base case analysis. **All the survival functions, however, do not reflect the initial flat period in the OS curve from the time of randomization. An alternative approach would be fitting parametric functions to OS from the start of HSCT, and model time to HSCT separately.'**		report that the log normal model provides a good visual and a good relative statistical fit to the observed data. The text has not been amended.
14. Error in applying costs in exploratory analysis	Page 113: A scenario analyses carried out by the ERG sets the HR for OS for gilteritinib maintenance therapy to 1 to eliminate the impact of post-HSCT gilteritinib maintenance therapy. This was referred to as "ERG exploratory analysis 5: Remove HR for gilteritinib maintenance therapy". However, although the ERG removed the benefit of post-HSCT gilteritinib maintenance therapy, it appears that they did not remember to remove the cost of post-HSCT. The total costs of gilteritinib in this analysis are very close to the costs of gilteritinib in the base case (Table 39 on Page 113): Cost of gilteritinib in the base case: '[redacted]'	Both the benefits and the costs of post-HSCT gilteritinib maintenance should be removed in the ERG exploratory analysis 5. In addition, ERG exploratory analysis 8 (containing exploratory analysis 5) should also be updated by removing the costs due to post-HSCT gilteritinib maintenance treatment.	The purpose of the analysis is to test the impact of removing post-HSCT gilteritinib maintenance treatment. Both benefits and costs should be adjusted consistently. It is noted that Astellas has not had access to the XLS file created by the ERG to check this or other errors in calculations.	This is not a factual inaccuracy. This analysis, which forms part of the ERG's preferred base case, uses pooled data from ADMIRAL for any patient receiving HSCT, and excludes any potential additional OS treatment effect resulting from gilteritinib maintenance therapy. This analysis is generally consistent with the trial – the ERG-preferred model predicts overall outcomes which are consistent with those observed in the ADMIRAL trial, and

	Cost of gilteritinib in the scenario: '[redacted]'			includes costs of gilteritinib treatment consumed by patients in the trial. As noted in the ERG report (page 110), this HR was excluded from the model on the basis of insufficient evidence to support an additional effect of maintenance therapy on OS. The report and the analyses have not been amended. Please note that the ERG's base case model and the supporting Excel file to implement this were provided to NICE by the ERG on 22nd August 2019.
15. Inappropriate data was used to fit the mixture cure model	Page 116-117: The ERG report presents results of "ERG additional sensitivity analysis 5". In this analysis, mixture-cure models were fitted to the ADMIRAL data for patients receiving HSCT. In addition, the OS predictions using mixture cure models and log-normal model based on ADMIRAL data	Remove Table 45, Figure 24 and relevant discussions related to mixture cure models using the ADMIRAL trial HSCT data. We suggest ERG considers alternative data sources to establish mixture models.	Mixture cure models might be suitable in this decision problem. However the cure point for these patients is likely to be 2-3 years post HSCT. In ADMIRAL only [redacted] gilteritinb patients were at risk 2 years post HSCT and [redacted] gilteritinb patients were at risk at 3 years.	This is not a factual inaccuracy. The company claims that ADMIRAL provided insufficient sample data with which to estimate parameters in cure models. On the contrary, the ERG was able to fit

	was presented in comparison with the OS predictions using the Gompertz model based on Evers in Figure 24.		The ADMIRAL OS data has limited follow-up to reflect the expected long-term cure effect i.e. it is not sufficiently mature to inform mixture cure models. If ERG plans to use mixture cure models to reflect the trajectory of HSCT patients, then Evers et al. or other literature with long follow-up time should be used.	five cure models and estimate cure fractions in each case. The resulting survival functions incorporating background mortality led to equivalent conclusions and was robust to model choice. Neither standard models nor cure models might represent the underlying data generation process. Nevertheless, the company accepts that a cure mixture model might be appropriate in this decision problem but maintains that cure is unlikely before 2-3 years. The ERG accepts that a mixture model with an uncertain change-point might be more plausible. However, the company acknowledges that it has not generated sufficient sample data from ADMIRAL with which to estimate the change-point. The ERG is receptive to estimating parameters using external information in addition to sample data but not in preference to

				sample data. Relevant external evidence should be shown to be consistent with the ADMIRAL data. The limitations of Evers *et al* with respect to the decision problem are discussed in the ERG report (Section 5.3.3). For brevity, these have not been repeated here.
16. Confidentiality highlighting	Page 177, Table 45: The estimated mean OS based on 'Company's model - (Gompertz Evers et al)' should be highlighted as commercial in confidence.	The following number in Table 45 should be highlighted Company's model - (Gompertz, Evers *et al*) [redacted]	The number is Commercial in Confidence. This in line with the Company Submission.	We have marked this as requested. However, we note that the data from Evers *et al* are publicly available and others could replicate these data and fit parametric models to them.

NICE National Institute for Health and Care Excellence

Technical engagement response form

Gilteritinib for treating relapsed or refractory acute myeloid leukaemia [ID1484]

As a stakeholder you have been invited to comment on the technical report for this appraisal. The technical 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.

We need your comments and feedback on the questions below. You do not have to answer every question. The text boxes will expand as you type. Please read the notes about completing this form. We cannot accept forms that are not filled in correctly. Your comments will be summarised and used by the technical team to amend or update the scientific judgement and rationale in the technical report.

Deadline for comments: **5 November 2019.**

Thank you for your time.

Please log in to your NICE Docs account to upload your completed form, as a Word document (not a PDF).

Notes on completing this form

- Please see the technical report which summarises the background and submitted evidence. This will provide context and describe the questions below in greater detail.
- 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.
- Do not use abbreviations.
- Do not include attachments such as journal articles, letters or leaflets. 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.
- If you provide journal articles to support your comments, you must have copyright clearance for these articles.
- 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. 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.

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

Your name	**Sarah Crouch**
Organisation name – stakeholder or respondent (if you are responding as an individual rather than a registered stakeholder please leave blank)	**Astellas Pharma Ltd**
Disclosure Please disclose any past or current, direct or indirect links to, or funding from, the tobacco industry.	**None**

Questions for engagement

Issue 1: Comparators	
Is best supportive care (BSC) a relevant comparator for patients with relapsed or refractory FLT3+ AML in NHS clinical practice?	Astellas agrees with the NICE Technical Team that best supportive care (BSC) is currently an approach for patients with FLT3 mutation positive relapsed or refractory (R/R) AML. Clinical expert advice sought by Astellas is in line with that reported by the ERG's clinical advisor, indicating that some patients who would otherwise elect to receive BSC (for example, due to the treatment burden associated with inpatient IV chemotherapy) may opt for treatment with gilteritinib due to its effectiveness, toxicity profile and oral administration (which can be taken in an outpatient or home care setting).

What proportion of patients would receive BSC in this patient population?	Astellas is not aware of the exact proportion of patients receiving BSC however, Astellas believes it is reasonable to adopt the numbers suggested by the ERG's clinical advisor, i.e. 25-30% of the FLT3 mutation positive R/R AML population currently receive BSC. This is supported by clinical expert advice sought by Astellas, in which 20% was considered to be an under-estimate of the real-world practice in the UK (based on a discussion regarding data from Hills et al. 2018[1], and eligibility for the AML 15, 16 and 17 trials). A range of figures has been applied within the model, see Table 1 below. The effect of this is a reduction on the ICER, with the base case ICER moving to £52,979 when a 25% BSC rate is applied. When inputs from Issue 1 (BSC 25%), Issue 3 (2-year cure point), Issue 6 (Utility changes), Issue 7 (Costs) and Issue 8 (Outpatient vs inpatient) are considered, the revised base case ICER is £43,346; see Table 1 below. Different rates of BSC and cure points are also explored in Table 2 below. It can be seen that in 6 of the 9 combinations the ICER drops below the £50k threshold.
Issue 2: Prior midostaurin use	
Would gilteritinib be used after prior midostaurin?	Astellas agrees that gilteritinib would be used after midostaurin and that additional clinical benefits are seen in patients previously treated with midostaurin. This is in line with the clinical advisors' opinion described in the Technical Engagement Papers and with clinical expert advice sought by Astellas.

	Astellas believes that additional clinical benefits are seen in patients previously treated with midostaurin and as such the effectiveness is not expected to be different. This is in line with the clinical advisors' opinion described in the Technical Engagement Papers and with clinical expert advice sought by Astellas. Gilteritinib is a more potent FLT3 inhibitor and it is not expected that prior midostaurin exposure would affect the response. This is borne out clinically in a sub-group analysis of ADMIRAL. Astellas was asked to comment on the applicability of the end of life criteria with or without prior midostaurin and believes gilteritinib meets the criteria in either case.
What proportion of rrAML patients receive midostaurin in NHS clinical practice in England?	Astellas does not have access to data to support an understanding of the exact proportion of FLT3 mutation positive AML patients receiving midostaurin. In ADMIRAL 13% and 11.3% of patients treated with gilteritinib and salvage chemotherapy respectively had received a FLT3 inhibitor, and it seems likely that the actual proportion may have increased given the reimbursement of midostaurin in England and Wales from mid-2018, however it would be expected that use may vary across the country and other options may still be chosen. Astellas notes the clinical advisors in the Technical Engagement Papers commented that around 50% of the FLT3 mutation positive R/R AML population is likely to have previous exposure to midostaurin, and that discussion on the Technical Engagement call suggested only [redacted] patients a month are receiving midostaurin. This equates to [redacted] midostaurin patients per year, whereas it

	is estimated there are over 400 FLT3 mutation positive R/R AML patients a year i.e. this suggests [redacted] would have received midostaurin.
If prior midostaurin use differs between the ADMIRAL trial and the population in England, how will this affect the effectiveness results from the trial?	Astellas believes that additional clinical benefits are seen in patients previously treated with midostaurin and as such the effectiveness is not expected to be different. This is in line with the clinical advisors' opinion described in the Technical Engagement Papers and with clinical expert advice sought by Astellas. Gilteritinib is a more potent FLT3 inhibitor and it is not expected that prior midostaurin exposure would affect the response. This is borne out clinically in a sub-group analysis of ADMIRAL. Interestingly, the gilteritinib EPAR comments that *"The proportion of patients with prior use of FLT3 inhibitors was small (12%). However, also in this subpopulation results were in favour of gilteritinib in terms of CR rate (18% vs 0%) and the HR for OS 0.705 (95%CI: 0.346, 1.438). Thus, exclusion of patients with prior FLT3 inhibitors from the indication was not considered necessary."* **Mechanism of Action** The two agents are very different drugs and have different mechanisms of action. Midostaurin is a 1st generation broad/multi-kinase inhibitor[2], gilteritinib is a 2nd generation FLT3-specific kinase inhibitor i.e. midostaurin is hitting many tyrosine kinase inhibitor (TKI) targets/is not targeted whilst gilteritinib is more specific to FLT3 mutations and more potent. It is also relevant to consider that gilteritinib demonstrates efficacy as a monotherapy in FLT3 mutation positive R/R AML whilst midostaurin failed to demonstrate efficacy in this population[3], suggesting the difference in mechanism leads also to a difference in clinical efficacy. It is also relevant to consider that earlier

	use of midostaurin is in combination with other drugs and also in patients without the FLT3 mutation – supporting that efficacy of midostaurin is driven by other mechanisms. The reason the targeting and specificity is important is that FLT3-ITD is a common driver mutation that presents with a high leukaemic burden and confers a poor prognosis in patients with AML[4]. Therefore, it is important to block the FLT3 pathway, more than any other genetic mutation, for such FLT3 mutation positive patients. The two FLT3 inhibitors, midostaurin and gilteritinib, have slightly different activity on the FLT3 target. Midostaurin has a relatively low affinity to FLT3 receptor (IC50 6.3 nM)[5] while gilteritinib is a potent inhibitor for FLT3 (IC50 0.29 nM)[6]. In addition, gilteritinib has ability to inhibit Axl, another tyrosine kinase that has an important role in cell survival, apoptosis and chemo-resistance[7]. Axl is overexpressed in AML and has been shown to have a potential role in resistance to chemotherapy and to the FLT3 inhibitor midostaurin[7]. Preclinical studies have shown that inhibition of Axl blocks proliferation of FLT3 mutant and FLT3 wild type AML cells and also suppresses the leukaemic burden of FLT3-ITD mutation positive AML[6]. **Clinical Results** Astellas has conducted a review of the efficacy in patients who have been treated with midostaurin, adding in those treated with sorafenib on the basis that this is also a multi-kinase inhibitor used earlier in the treatment pathway (n=45). This analysis concludes that the clinical efficacy seen in this population is similar to the efficacy seen in the overall study, again supporting that prior TKI treatment is not an issue.

	 The Kaplin-Meier (K-M) curve above presents the survival data from the patient subgroup from ADMIRAL previously treated with midostaurin or sorafenib (n=45) and the survival data from the trial population as a whole. For both treatment arms, the survival between the populations is similar.
Issue 3: Cure assumptions	
Is it plausible to assume that all patients who remain alive at 3 years are 'cured' regardless of whether they have progressed or have had HSCT?	Astellas agrees that this is an important point of discussion, and that substantial uncertainty exists in the literature around cure assumptions. Two issues need to be addressed to answer this question –

	firstly the time after which cure should be assumed, and secondly the impact which this has on progressed or non-transplanted patients in the model. The Astellas base case used a cure assumption of three years in part to align with the prior midostaurin appraisal (TA523)[8], and in part to present a conservative assumption. During this appraisal clinical experts have consistently commented that a shorter period may be considered in clinical practice, this is in line with the clinical advisors' opinion described in the Technical Engagement Papers: *"The disease is aggressive so usually relapse would occur in 6-9 months."* Clinical experts sought by Astellas have indicated that a cure point of 2-3 years was appropriate in patients with FLT3 mutation positive R/R AML. In line with this opinion we have provided scenario analyses which explore cure points at 1 and 2 years. Regarding the impact on progressed or non-transplanted patients in the model, survival rate estimates for patients not receiving HSCT indicate that the majority of non-HSCT patients do not survive to the cure point. Therefore, Astellas took the view that applying a cure to all surviving patients (irrespective of HSCT status) would not be impactful on the model results. Progressed patients comprise a relatively small proportion of the transplanted patients alive so similarly, applying a cure to all patients (irrespective of progression status) should have a negligible impact on the model results. Uncertainty around the most appropriate cure point has been investigated in scenario analyses, with ICERs ranging from £31,807 (1-year cure point) to £46,156 (2-year cure point); see Table 1 below. When inputs from Issue 1 (BSC 25%), Issue 3 (2-year cure point), Issue 6 (Utility changes), Issue 7

<table>
<tr><td></td><td>(Costs) and Issue 8 (Outpatient vs inpatient) are considered, the revised base case ICER is £43,346; see Table 1 below.

Different rates of BSC and cure points are also explored in Table 2 below. It can be seen that in 6 of the 9 combinations the ICER drops below the £50k threshold.</td></tr>
<tr><td colspan="2">Issue 4: Gilteritinib effectiveness after HSCT</td></tr>
<tr><td>Is it more appropriate to use external data or ADMIRAL trial data to estimate the relative effectiveness of gilteritinib after HSCT?</td><td>Astellas believes it is more appropriate to use external data to inform post-HSCT survival rather than data from the ADMIRAL trial given the low numbers and low duration of follow-up in these patients.

Clinical experts have commented that if patients are to relapse then this is most likely to occur within 12 months following HSCT (with some suggesting this occurs within 6-9 months). This being the case, it is plausible that the remainder of the ADMIRAL post-HSCT patients (54%) who were alive at the data cut-off would still be alive at the three-year point. Interestingly it can be seen that the two datasets (ADMIRAL and Evers et al 2018) have comparable post-HSCT overall survival rates across this period, as presented in the table below.
<table>
<tr><th>Months</th><th>ADMIRAL</th><th>Evers et al 2018[9]</th></tr>
<tr><td>1</td><td>academic/commercial in confidence information removed</td><td>94.5%</td></tr>
<tr><td>2</td><td>academic/commercial in confidence information removed</td><td>87.5%</td></tr>
<tr><td>3</td><td>academic/commercial in confidence information removed</td><td>80.5%</td></tr>
<tr><td>4</td><td>academic/commercial in confidence information removed</td><td>68.8%</td></tr>
</table>
</td></tr>
</table>

5	academic/commercial in confidence information removed	68.8%
6	academic/commercial in confidence information removed	66.4%
7	academic/commercial in confidence information removed	61.7%
8	academic/commercial in confidence information removed	61.7%
9	academic/commercial in confidence information removed	59.4%
10	academic/commercial in confidence information removed	57.0%
11	academic/commercial in confidence information removed	57.0%
12	academic/commercial in confidence information removed	57.0%

Given the low numbers and short follow-up in ADMIRAL, and the broad comparability of the Evers dataset, Astellas considers it is more appropriate to apply the Evers data to inform the long-term post-HSCT survival.

The same approach was adopted and accepted by NICE in a previous appraisal (TA598) based on the same issues. This appraisal related to ovarian cancer, also a rare and aggressive cancer, where evidence from the Edinburgh Ovarian Cancer Database was used to indicate expected survival outcomes in current practice in the absence of sufficiently mature OS data from the trial.

To give more data behind this, in the ADMIRAL trial both the sample size of post-HSCT patients, and the duration of follow up, was very limited. In ADMIRAL, the median follow-up post-HSCT was academic/commercial in confidence information removed months with academic/commercial in confidence information removed patients having follow-up data beyond year 1, and academic/commercial in confidence information removed patients having follow-up data beyond year 2. Given

	the 40-year time horizon of our extrapolation, this data adds substantial uncertainty. Astellas believes that given the availability of more robust 'external' data, ADMIRAL does not represent the best available evidence to inform the cost-effectiveness base case. Based on the unsuitability of the ADMIRAL post-HSCT data, Astellas presented a thorough review of published alternatives from which data reported by Evers et al. 2018[9] was considered the most robust and therefore appropriate for the economic base case. The following criteria were used to assess candidate studies: • Comparable patient population to the model target population • Relevant OS and EFS measure reported in the form of K-M curves • Sufficient sample size and mature follow-up to reduce uncertainty with survival extrapolation In this evaluation, the recent Evers et al. 2018[9] publication was identified as having the largest sample size and longest follow-up duration, and presents a population similar to the ADMIRAL population. In the absence of an absolute matched population, Astellas did also consider Ustun et al[10], which is a FLT3 mutation positive AML population. If Ustun et al[10] is used to inform the long-term post-HSCT survival, the revised company base case ICER becomes £45,082; see Table 1 below. The ERG proposed an alternative source of external data (Poiré et al. 2018[11]) however the sample size of HSCT recipients in CR2 was substantially smaller than that studied by Evers et al. 2018[9]

	(n=37 vs n=128, respectively). The median follow-up in Poiré et al. 2018[11] was also noticeably shorter (Poiré 23 months vs Evers 6.5 years). Thus Poiré does not seem to be a suitable preferable choice. For these reasons, Astellas maintains Evers et al. 2018[9] represents a more robust estimate of post-HSCT survival rather than the ADMIRAL trial, and is the best of the available published 'external' evidence.
Which of the extrapolated survival models (see figure 1) appears to be more clinically plausible?	The extrapolated data from Evers et al. 2018[9] is the more appropriate survival model due to the uncertainty associated with the ADMIRAL data discussed above.
Issue 5: Gilteritinib maintenance therapy	
Would gilteritinib be used as maintenance therapy after HSCT in clinical practice?	Astellas agrees with the clinical advisors' opinion described in the Technical Engagement Papers, that *"It is important that [gilteritinib] is used as maintenance therapy and it is reasonable to expect an additional OS benefit."* This is in line with clinical expert advice sought by Astellas in which experts said that they would administer gilteritinib to selected FLT3 mutation positive R/R patients post-HSCT. Also relevant to consider here are the criteria used for re-starting post-HSCT gilteritinib in ADMIRAL, and subsequently described in the Summary of Product Characteristics (SPC)[12]: *"Treatment can be resumed 30 days after HSCT if engraftment was successful, the patient did not have grade ≥2 acute graft versus host disease and was in CRc."* Since all three criteria should be met in order to restart

	treatment Astellas expects use in this setting to be minimal compared to the broader FLT3 mutation positive R/R patient group.
Is it plausible that there is an *additional* effect of maintenance therapy on OS?	Astellas believes that is it plausible to expect an additional effect of maintenance therapy on OS, in line with the clinical advisors' opinion described in the Technical Engagement Papers, that *"It is important that it [gilteritinib] is used as maintenance therapy and it is reasonable to expect an additional OS benefit."* This view was supported by clinical expert opinion sought by Astellas. The prognosis of relapsed AML patients is very poor (~12 weeks), which has limited the opportunity to offer maintenance therapy to date. Midostaurin was evaluated as maintenance therapy in newly diagnosed FLT3 mutation positive AML patients in the RATIFY study that resulted in marketing authorisation and NICE recommendation (TA523)[8]. Sorafenib has been evaluated as maintenance therapy after stem cell transplant in the SORMAIN study; maintenance therapy post-transplant significantly reduced the risk of relapse and death in FLT3-ITD mutation positive patients[13]. These two data sets prove the principle of maintenance therapy in patients with FLT3 mutation positive AML. It is expected that a more specific and potent FLT3 inhibitor such as gilteritinib, should further deepen the remission and improve outcomes in a maintenance setting compared to 1st generation FLT3 inhibitors. Astellas has submitted post-HSCT survival curves for patients restarting, or not restarting, gilteritinib therapy post-HSCT. While we are aware of the limitations of our dataset in this phase of the trial, we believe that ADMIRAL is the only available source describing post-HSCT gilteritinib use. While not statistically significant, the results were intuitive to clinical experts.

Is the current method of deriving and applying an additional benefit of maintenance therapy appropriate?	Astellas believes that the best way to model a maintenance benefit is to include both the available costs and benefits of gilteritinib in this setting. We acknowledge that there is some uncertainty in using post-HSCT data from ADMIRAL to inform this but given that this is the only available evidence for continued gilteritinib treatment, it should be included. In their preferred base case, the ERG removed the additional benefit of maintenance therapy through the application of an alternative hazard ratio of 0. Astellas believes that this scenario is only logical if the costs of gilteritinib are also removed. A scenario analysis with the removal of costs and benefits of maintenance gilteritinib therapy produced an ICER of £57,935; see Table 1 below.
Issue 6: Utilities	
Are utility values from Janssen *et al.* or Ara and Brazier *et al.* more clinically plausible after the 3-year cure point? See figure 3.	Astellas agrees that utility values from Ara and Brazier et al[14]. appear to be clinically plausible for use in cured patients. The impact of these alternative utility values was included in the revised company base case and decreased the corrected base case ICER to £55,004; see Table 1 below. When inputs from Issue 1 (BSC 25%), Issue 3 (2-year cure point), Issue 6 (Utility changes), Issue 7 (Costs) and Issue 8 (Outpatient vs inpatient) are considered, the revised base case ICER is £43,346; see Table 1 below.
Issue 7: Costs	

In NHS clinical practice in England, would gilteritinib tablets be wasted if patients stopped taking it unexpectedly, for example because of death?	Astellas agrees that there is a potential for the wastage of gilteritinib tablets due to unexpected events, such as death, however this is hard to quantify. Astellas did review the NICE appraisal of sorafenib for advanced hepatocellular carcinoma (TA474, previously TA189)[15]. During the CDF rapid reconsideration the company approached two of the largest HCC-treating trusts to understand their practice and the implications of treatment wastage. Both trusts had a policy to issue one month of sorafenib at a time, and prescriptions were aligned with a patient's monthly follow-up appointment where a clinical decision was made regarding the patient's suitability for treatment in the following month. Patient's supply of sorafenib was actively managed by splitting packs where appropriate. The clinician, pharmacist and patient worked closely to reconcile what medicines were used in the month. Where the patient still had unused tablets, only the remainder of another month's supply would be issued to reduce wastage. The company presented cost-effectiveness results for analyses including the wastage of up to 7 days of treatment. The Committee concluded that it was appropriate for the company to use updated unit cost data and account for 7 days of drug wastage because this reflected the price relevant to the NHS. It would seem appropriate to take the same approach here and as such, the scenario of one week's wastage (0.25 packs) has been explored in the revised base case; see Table 1 below.
Should drug costs be applied as a one-off cost in the first cycle of the model?	Astellas believes that the application of costs in this way has a negligible effect on the model results. Drug costs are applied as a one-off cost in the first cycle of the model for simplicity, as the average patient will receive academic/commercial in confidence information removed cycles of therapy (including those who received

	maintenance therapy). Astellas acknowledge this prevents the correct application of discounting hence the estimated gilteritinib drug costs constitute a small overestimate of the true value.
Is it more plausible to assume that for patients alive after 3 years (after the assumed 'cure') a. patients who have had HSCT have 1.5 outpatient visits every month indefinitely, **or** have no follow-up costs, **or** have no follow-up costs? b. patients who have not had HSCT would require 1 outpatient visit every 6 months **or** have no follow-up costs?	Astellas agrees with the clinical advisors' opinion described in the Technical Engagement Papers, that one visit every year for patients who have had HSCT is more plausible (see ERG papers rather than the numbers seen in the question provided here which contain an error). This has been updated in the revised company base case by applying one outpatient visit every 12 months. Astellas believes that one outpatient visit per 6 months for patients who have not had a HSCT appears reasonable.
Is it reasonable to remove progression costs from the model after 3 years (after the assumed 'cure')?	Astellas agrees that it is reasonable to remove costs of progression and relapse from cured patients.
Is it reasonable to assume 3.3 or 2.0 FLT3 tests will be required to identify 1 patient (in other words, does FLT3 occur in around 30% of patients which would result in 3.3 tests per patient)?	Astellas agrees that, based on the reported prevalence of the FLT3 mutation, it is reasonable to assume 3.3 tests are required to identify each patient. The impact of this testing rate was explored in the revised company base case; see Table 1 below. When inputs from Issue 1 (BSC 25%), Issue 3 (2-year cure point), Issue 6 (Utility changes), Issue 7 (Costs) and Issue 8 (Outpatient vs inpatient) are considered, the revised base case ICER is £43,346; see Table 1 below.
Issue 8: Impact on cost, quality of life of outpatient vs. inpatient care	

It would be expected that patients who are managed as inpatients to receive their treatment have an inherent poorer quality of life than those who can be managed as outpatients or take their medication at home – please consider this within your submission	During the Technical Engagement call, Astellas was asked to consider the impact of inpatient care vs. outpatient/home care on quality of life. This was a comment of note from the clinical and patient expert on the call which NICE asked Astellas to address in its submission. It was also noted that the clinical advisors' opinion described in the Technical Engagement Papers, that *"In the short term there is considerably improved QoL through receiving oral therapy as an outpatient rather than a prolonged inpatient stay for salvage chemotherapy."* In line with this, Astellas has looked at the impact on the ICER of a worsened utility and an extended hospital stay for high intensity chemotherapy regimens (i.e. MEC and FLAG-IDA) in the revised base case; see Table 1 below. A limitation of the ADMIRAL trial was difficulty in collecting PRO responses in the salvage chemotherapy arm, and therefore application of a disutility from the literature seems appropriate. In the analysis a disutility of -0.044 was applied to high intensity chemotherapy (using data reported by Wehler et al[16] for *"Disutility associated with other HIC when minimum AEs"*). This value was chosen to isolate the impact of the chemotherapy and avoid double counting of disutility associated with AEs. For hospitalisation costs in patients receiving high intensity chemotherapy, it was assumed in cycle 1 that patients were hospitalised for 28 days, and from cycle 2 onwards the hospitalisation estimation from the ADMIRAL trial was applied. When inputs from Issue 1 (BSC 25%), Issue 3 (2-year cure point), Issue 6 (Utility changes) and Issue 7 (Costs) are considered the ICER is £45,442, applying both cost and utility updates to high-intensity chemotherapies (FLAG-IDA, MEC) i.e. Issue 8 (Outpatient vs inpatient), gave a revised base case ICER of £43,346; see Table 1 below.

Table 1: Technical Revisions and Impact on the Cost-Effectiveness Ratio

Issue	Description	ICER (vs. weighted comparator excl. BSC)	Change from company base case
Company base case (v1.0)	Originally submitted in June 2019	£47,695	
Company base case (v1.1)	Revised version submitted to correct running of PSA	£47,695	No change
Company ICER (ERG corrections made to re-submitted model)		£54,844	+£7,149
Company ICER (ERG corrections and update to administration cost made to re-submitted model)	Application of monthly dispensing fee for gilteritinib (based on NICE feedback on budget impact analysis)	£55,404	+£7,709
			Change from corrected base case
1. BSC as a relevant comparator	BSC is not included in the weighted comparator in the submitted base case. In the analyses below, BSC is included in the weighted comparator in the following proportions:		
	BSC = 20%	£53,408	-£1,996
	BSC = 25%	£52,979	-£2,425
	BSC = 30%	£52,573	-£2,831
2. Prior midostaurin use	Assumption of proportional hazards not satisfied (K-M curves cross at about 20 months) so not appropriate to apply HR of TKI-treated population to curve to generate new cost-effectiveness result	N/A	N/A
3. Cure assumptions	In the submitted base case, a cure point of 3 years is modelled. Based on the opinion of clinical experts, earlier cure points are possible. This is tested in the analyses below.		
	1-year cure point	£31,807	-£23,597

Issue	Description	ICER (vs. weighted comparator excl. BSC)	Change from company base case
	2-year cure point	£46,156	-£9,248
5. Gilteritinib maintenance therapy	Gilteritinib is administered to post-HSCT patients in the maintenance setting. The ERG propose that additional OS benefit should not be associated with maintenance therapy. In this analysis, the benefit is removed in line with their method but also the maintenance therapy costs.	£57,935	+£2,531
6. Utilities	In the submitted base case, age-adjusted utility from Janssen at al. is used. The ERG considers the data from Ara & Brazier to be more plausible. Furthermore, the ERG felt that progression should not be associated with a disutility as this reduction in utility is accounted for in the health state utility value. This analysis implements these changes in line with the ERG method.	£55,004	-£400
7. Costs	Some updates to the model costing approach were made in line with ERG preferences: • 0.25 of a pack's wastage cost is associated with gilteritinib • Resource use inputs were updated to include the suggestions made by the clinical experts in the technical report (1 outpatient visit per 12 months for HSCT patients, 1 outpatient visit per 6 months for non-HSCT patients) • Removal of relapse and progression costs for cured patients • Rate of FLT3 testing increased from 200% to 333%	£57,314	+£1,910

Issue	Description	ICER (vs. weighted comparator excl. BSC)	Change from company base case
8. Outpatient vs inpatient	Consideration of the impact of inpatient care vs. outpatient/home care on costs, quality of life for patients treated with high-intensity chemotherapies (FLAG-IDA, MEC):		
	Disutility applied	£54,763	-£641
	Cost updates applied	£51,816	-£3,588
	Application of both cost and disutility updates	£51,218	-£4,186
			Change from corrected base case
Revised company base case (excluding additional request from Technical Engagement call):	**Issue 1 – 25% BSC** **Issue 3 – 2-year cure point** **Issue 6 – Utilities** **Issue 7 – Costs**	**£45,442**	**-£9,962**
Revised company base case (including additional request from Technical Engagement call):	**Issue 1 – 25% BSC** **Issue 3 – 2-year cure point** **Issue 6 – Utilities** **Issue 7 – Costs** **Issue 8 – Outpatient vs inpatient**	**£43,346**	**-£12,058**
Use of Ustun 2017 as source of external data for OS post-HSCT in revised base case:		**£45,082**	**-£10,322**

Table 2: Impact on Cumulative Base Case of varying BSC Proportion in Weighted Comparator and Cure Point

	1-year cure point	2-year cure point	3-year cure point
BSC = 20% of weighted comparator	£30,547	£43,455	£51,796

BSC = 25% of weighted comparator	£30,630	£43,346	£51,589
BSC = 30% of weighted comparator	£30,708	£43,242	£51,390

*These parameters are varied in the proposed cumulative base case i.e. issue 6, 7, 8 updates are applied across all of these analyses

References:

1. Hills, R. K. *et al.* Outcomes in Relapsed/Refractory Patients with FLT3-ITD Mutated AML Are Poor When Treated with Non-Targeted Therapy with a Potential Role for Stem Cell Transplantation: Results from the NCRI AML Trials. *Blood* **132**, 1392–1392 (2018).
2. European Medicines Agency. Rydapt. *European Medicines Agency* https://www.ema.europa.eu/en/medicines/human/EPAR/rydapt (2018).
3. Fischer T. et al. Phase IIB Trial of Oral Midostaurin (PKC412), the FMS-Like Tyrosine Kinase 3 Receptor (FLT3) and Multi-Targeted Kinase Inhibitor, in Patients With Acute Myeloid Leukemia and High-Risk Myelodysplastic Syndrome With Either Wild-Type or Mutated FLT3. **28**, 4339–4345 (2010).
4. Daver N., Schlenk RF, Russell NH, Levis MJ. Targeting FLT3 mutations in AML: review of current knowledge and evidence. *Leukemia* **33**, 299–312 (2019).
5. Weisberg, E.L., E. L. Characterization of midostaurin as a dual inhibitor of FLT3 and SYK and potentiation of FLT3 inhibition against FLT3-ITD-driven leukemia harboring activated SYK kinase. *Oncotarget* **8**, 52026–52044 (2017).
6. Mori M., Kaneko N., Ueno Y. et al. Gilteritinib, a FLT3/AXL inhibitor, shows antileukemic activity in mouse models of FLT3 mutated acute myeloid leukemia. *Invest New Drugs* **35**, 556–565 (2017).
7. McMahon, C. & Perl, A. E. Gilteritinib for the treatment of relapsed and/or refractory FLT3-mutated acute myeloid leukemia. *Expert Review of Clinical Pharmacology* **12**, (2019).
8. National Institute for Health and Care Excellence (NICE). Midostaurin for untreated acute myeloid leukaemia. Technology appraisal guidance [TA523]. *NICE* https://www.nice.org.uk/guidance/ta523/history (2018).

9. Evers, G. *et al.* Outcome of Patients with Acute Myeloid Leukemia (AML) Undergoing Allogeneic Hematopoietic Stem Cell Transplantation (HSCT) Beyond First Complete Remission (CR1). *Blood* **132**, 4649–4649 (2018).
10. Ustun, C. *et al.* Outcomes of UCB transplantation are comparable in FLT3+ AML: results of CIBMTR, EUROCORD and EBMT collaborative analysis. *Leukemia* **31**, 1408–1414 (2017).
11. Poiré, Xavier *et al.* Allogeneic stem cell transplantation benefits for patients≥ 60 years with acute myeloid leukemia and FLT3 internal tandem duplication: a study from the Acute Leukemia Working Party of the European Society for Blood and Marrow Transplantation. *haematologica* **103**, 256–265 (2018).
12. Astellas Pharmaceuticals. Data on File: Summary of Product Characteristics, Xospata 40 mg film-coated tablets. (2019).
13. Burchert, Andreas *et al.* Sorafenib as maintenance therapy post allogeneic stem cell transplantation for FLT3-ITD positive AML: results from the randomized, double-blind, placebo-controlled multicentre sormain trial. *Am Soc Hematology* 661–661 (2018).
14. Ara, R. & Brazier, J. Populating an economic model with health state utility values: moving toward better practice. *Value in Health: The Journal of the International Society for Pharmacoeconomics and Outcomes Research* **13**, 509–518 (2010).
15. National Institute for Health and Care Excellence (NICE). Sorafenib for treating advanced hepatocellular carcinoma. Technology appraisal guidance [TA474]. *NICE* https://www.nice.org.uk/guidance/ta474/history (2017).
16. Wehler, E., Storm, M., Kowal, S., Campbell, C. & Boscoe, A. A Health State Utility Model Estimating the Impact of Ivosidenib on Quality of Life in Patients with Relapsed/Refractory Acute Myeloid Leukemia. *HemaSphere* **2**, 661 (2018).

NICE National Institute for Health and Care Excellence

Technical engagement response form

Gilteritinib for treating relapsed or refractory acute myeloid leukaemia [ID1484]

As a stakeholder you have been invited to comment on the technical report for this appraisal. The technical 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.

We need your comments and feedback on the questions below. You do not have to answer every question. The text boxes will expand as you type. Please read the notes about completing this form. We cannot accept forms that are not filled in correctly. Your comments will be summarised and used by the technical team to amend or update the scientific judgement and rationale in the technical report.

Deadline for comments: **5pm, 5 November 2019.**

Thank you for your time.

Please log in to your NICE Docs account to upload your completed form, as a Word document (not a PDF).

Notes on completing this form

- Please see the technical report which summarises the background and submitted evidence. This will provide context and describe the questions below in greater detail.
- 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.
- Do not use abbreviations.
- Do not include attachments such as journal articles, letters or leaflets. 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.
- If you provide journal articles to support your comments, you must have copyright clearance for these articles.
- 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. 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.

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

Your name	**Mike Dennis**
Organisation name – stakeholder or respondent (if you are responding as an individual rather than a registered stakeholder please leave blank)	**RCPath**
Disclosure Please disclose any past or current, direct or indirect links to, or funding from, the tobacco industry.	**None**

Questions for engagement

Issue 1: Comparators	
Is best supportive care (BSC) a relevant comparator for patients with relapsed or refractory FLT3+ AML in NHS clinical practice?	**Yes** Please not the treatment pathway in the document is for newly diagnosed FLT3 mutated AML (not relapse/refractory)
What proportion of patients would receive BSC in this patient population?	**10-20%**
Issue 2: Prior midostaurin use	
Would gilteritinib be used after prior midostaurin?	**Yes**
What proportion of rrAML patients receive midostaurin in NHS clinical practice in England?	**None- it is only approved for first line therapy** **Use in first line therapy is highly variable dependent upon which treatment centre the patient is attending. Although NICE approved, the FLT3 ITD status is not known for a significant proportion of patients at the time of therapy initiation. Perhaps 50% in the under 70's- of those 50% where it is known it is likely the majority receive Midostaurin. In the over 70's only a minority will have the FLT3 status identified prior to therapy, of whom only a minority would receive Midostaurin.**
If prior midostaurin use differs between the ADMIRAL trial and the population in England, how will this affect the effectiveness results from the trial?	**Unlikely to be significant as the available data suggests equivalent responses in patients with or without prior Midostaurin therapy**

Issue 3: Cure assumptions	
Is it plausible to assume that all patients who remain alive at 3 years are 'cured' regardless of whether they have progressed or have had HSCT?	**If alive at 3 years they are likely to be cured, clearly if they've progressed they shouldn't be considered cured (very rare I'd have thought). Virtually all such patients will have undergone HSCT- as the therapy is generally utilised as a bridge to transplant.**
Issue 4: Gilteritinib effectiveness after HSCT	
Is it more appropriate to use external data or ADMIRAL trial data to estimate the relative effectiveness of gilteritinib after HSCT?	**Both should be considered- as outlined both have limitations**
Which of the extrapolated survival models (see figure 1) appears to be more clinically plausible?	**Gilteritinib- company model**
Issue 5: Gilteritinib maintenance therapy	
Would gilteritinib be used as maintenance therapy after HSCT in clinical practice?	Yes
Is it plausible that there is an *additional* effect of maintenance therapy on OS?	Yes- there is data published in abstract demonstrating improved LFS and OS with Sorafenib maintenance (similar TKI therapy) for FLT3 mutated AML patients transplanted in first CR (Sormain study)- median follow up of 55 months.
Is the current method of deriving and applying an additional benefit of maintenance therapy appropriate?	Unclear
Issue 6: Utilities	
Are utility values from Janssen *et al.* or Ara and Brazier *et al.* more clinically plausible after the 3-year cure point? See figure 3.	No much in it?- Ara and Brazier et al.

Issue 7: Costs	
In NHS clinical practice in England, would gilteritinib tablets be wasted if patients stopped taking it unexpectedly, for example because of death?	During any 28 day cycle any dispensed tablets would be wasted if the patient died or decided to stop therapy- this is generally very little wastage in practice.
Should drug costs be applied as a one-off cost in the first cycle of the model?	Cost per cycle rather than per dose seems appropriate
Is it more plausible to assume that for patients alive after 3 years (after the assumed 'cure') a. patients who have had HSCT have 1.5 outpatient visits every month indefinitely, **or** have no follow-up costs? b. patients who have not had HSCT would require 1 outpatient visit every 6 months **or** have no follow-up costs?	a- Although 1.5 is more plausible, an estimated 1 visit every 2- 3 months is most plausible b- 1
Is it reasonable to remove progression costs from the model after 3 years (after the assumed 'cure')?	Yes
Is it reasonable to assume 3.3 or 2.0 FLT3 tests will be required to identify 1 patient (in other words, does FLT3 occur in around 30% of patients which would result in 3.3 tests per patient)?	This testing is already considered standard of care- not sure it should be applied to the costing model.

NICE National Institute for Health and Care Excellence

Technical engagement response form

Gilteritinib for treating relapsed or refractory acute myeloid leukaemia [ID1484]

As a stakeholder you have been invited to comment on the technical report for this appraisal. The technical 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.

We need your comments and feedback on the questions below. You do not have to answer every question. The text boxes will expand as you type. Please read the notes about completing this form. We cannot accept forms that are not filled in correctly. Your comments will be summarised and used by the technical team to amend or update the scientific judgement and rationale in the technical report.

Deadline for comments: **5pm, 5 November 2019.**

Thank you for your time.

Please log in to your NICE Docs account to upload your completed form, as a Word document (not a PDF).

Notes on completing this form

- Please see the technical report which summarises the background and submitted evidence. This will provide context and describe the questions below in greater detail.
- 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.
- Do not use abbreviations.
- Do not include attachments such as journal articles, letters or leaflets. 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.
- If you provide journal articles to support your comments, you must have copyright clearance for these articles.
- 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. 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.

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

Your name	**Nigel Russell**
Organisation name – stakeholder or respondent (if you are responding as an individual rather than a registered stakeholder please leave blank)	**RCP/ NCRI AML WG**
Disclosure Please disclose any past or current, direct or indirect links to, or funding from, the tobacco industry.	**None**

Questions for engagement

Issue 1: Comparators	
Is best supportive care (BSC) a relevant comparator for patients with relapsed or refractory FLT3+ AML in NHS clinical practice?	**Yes**
What proportion of patients would receive BSC in this patient population?	**No more than 20%**
Issue 2: Prior midostaurin use	
Would gilteritinib be used after prior midostaurin?	**Yes**
What proportion of rrAML patients receive midostaurin in NHS clinical practice in England?	**Rrpatients do not receive Midostaurin . It is nor approved inn rrAML. Patients receive Mido as first line therapy!**
If prior midostaurin use differs between the ADMIRAL trial and the population in England, how will this affect the effectiveness results from the trial?	**These is no convicing evidence that prior Midostaurin affects response to Gilteritinib. The HR CIs in the Admiral trial reflects small number of patients. In discussion with Dr Marl Levis from the US he states at least a 50% ORR in patients who have had prior Mido therapy**
Issue 3: Cure assumptions	
Is it plausible to assume that all patients who remain alive at 3 years are 'cured' regardless of whether they have progressed or have had HSCT?	**Yes it is**

Issue 4: Gilteritinib effectiveness after HSCT	
Is it more appropriate to use external data or ADMIRAL trial data to estimate the relative effectiveness of gilteritinib after HSCT?	**The trial data is the best that is available**
Which of the extrapolated survival models (see figure 1) appears to be more clinically plausible?	
Issue 5: Gilteritinib maintenance therapy	
Would gilteritinib be used as maintenance therapy after HSCT in clinical practice?	Yes
Is it plausible that there is an *additional* effect of maintenance therapy on OS?	Yes
Is the current method of deriving and applying an additional benefit of maintenance therapy appropriate?	Yes
Issue 6: Utilities	
Are utility values from Janssen *et al.* or Ara and Brazier *et al.* more clinically plausible after the 3-year cure point? See figure 3.	No comment
Issue 7: Costs	
In NHS clinical practice in England, would gilteritinib tablets be wasted if patients stopped taking it unexpectedly, for example because of death?	Yes
Should drug costs be applied as a one-off cost in the first cycle of the model?	I do not understand this question

Is it more plausible to assume that for patients alive after 3 years (after the assumed 'cure') a. patients who have had HSCT have 1.5 outpatient visits every month indefinitely, **or** have no follow-up costs? b. patients who have not had HSCT would require 1 outpatient visit every 6 months **or** have no follow-up costs?	They would have appointments every 3-4 months Every 6 months
Is it reasonable to remove progression costs from the model after 3 years (after the assumed 'cure')?	Yes
Is it reasonable to assume 3.3 or 2.0 FLT3 tests will be required to identify 1 patient (in other words, does FLT3 occur in around 30% of patients which would result in 3.3 tests per patient)?	yes

NICE National Institute for Health and Care Excellence

Technical engagement response form

Gilteritinib for treating relapsed or refractory acute myeloid leukaemia [ID1484]

As a stakeholder you have been invited to comment on the technical report for this appraisal. The technical 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.

We need your comments and feedback on the questions below. You do not have to answer every question. The text boxes will expand as you type. Please read the notes about completing this form. We cannot accept forms that are not filled in correctly. Your comments will be summarised and used by the technical team to amend or update the scientific judgement and rationale in the technical report.

Deadline for comments: **5pm, 5 November 2019.**

Thank you for your time.

Please log in to your NICE Docs account to upload your completed form, as a Word document (not a PDF).

Notes on completing this form

- Please see the technical report which summarises the background and submitted evidence. This will provide context and describe the questions below in greater detail.
- 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.
- Do not use abbreviations.
- Do not include attachments such as journal articles, letters or leaflets. 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.
- If you provide journal articles to support your comments, you must have copyright clearance for these articles.
- 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. 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.

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

Your name	**Charlotte Martin**
Organisation name – stakeholder or respondent (if you are responding as an individual rather than a registered stakeholder please leave blank)	**Leukaemia Care**
Disclosure Please disclose any past or current, direct or indirect links to, or funding from, the tobacco industry.	**n/a**

Questions for engagement

Issue 1: Comparators	
Is best supportive care (BSC) a relevant comparator for patients with relapsed or refractory FLT3+ AML in NHS clinical practice?	**Yes. Patients do choose not to continue with chemotherapy due to significant side effects, whether potential/perceived side effects or based on side effects already experienced. Therefore, other options should be considered as relevant.**
What proportion of patients would receive BSC in this patient population?	**We don't have an exact number, but this is highly likely to increase with age. In our survey, Living with Leukaemia,**
Issue 2: Prior midostaurin use	
Would gilteritinib be used after prior midostaurin?	**Yes. Clinical experts agree that this is a more potent inhibitor, and so would have activity in patients despite prior targeted treatment.**
What proportion of rrAML patients receive midostaurin in NHS clinical practice in England?	**No comment**
If prior midostaurin use differs between the ADMIRAL trial and the population in England, how will this affect the effectiveness results from the trial?	**No comment.**
Issue 3: Cure assumptions	
Is it plausible to assume that all patients who remain alive at 3 years are 'cured' regardless of whether they have progressed or have had HSCT?	**Yes.**

Issue 4: Gilteritinib effectiveness after HSCT	
Is it more appropriate to use external data or ADMIRAL trial data to estimate the relative effectiveness of gilteritinib after HSCT?	**No comment.**
Which of the extrapolated survival models (see figure 1) appears to be more clinically plausible?	**No comment.**
Issue 5: Gilteritinib maintenance therapy	
Would gilteritinib be used as maintenance therapy after HSCT in clinical practice?	No comment.
Is it plausible that there is an *additional* effect of maintenance therapy on OS?	No comment
Is the current method of deriving and applying an additional benefit of maintenance therapy appropriate?	No comment.
Issue 6: Utilities	
Are utility values from Janssen *et al.* or Ara and Brazier *et al.* more clinically plausible after the 3-year cure point? See figure 3.	No comment.
Issue 7: Costs	
In NHS clinical practice in England, would gilteritinib tablets be wasted if patients stopped taking it unexpectedly, for example because of death?	No.
Should drug costs be applied as a one-off cost in the first cycle of the model?	No comment.

Is it more plausible to assume that for patients alive after 3 years (after the assumed 'cure') a. patients who have had HSCT have 1.5 outpatient visits every month indefinitely, **or** have no follow-up costs? b. patients who have not had HSCT would require 1 outpatient visit every 6 months **or** have no follow-up costs?	Patients who had be cured, by this definition, would have these follow up appointments as described regardless of the intervention given to lead to that cure.
Is it reasonable to remove progression costs from the model after 3 years (after the assumed 'cure')?	No comment.
Is it reasonable to assume 3.3 or 2.0 FLT3 tests will be required to identify 1 patient (in other words, does FLT3 occur in around 30% of patients which would result in 3.3 tests per patient)?	All patients should receive the test to ensure they get the most appropriate and effective treatment for them, so this is a reasonable assumption.

NICE National Institute for Health and Care Excellence

Technical engagement response form

Gilteritinib for treating relapsed or refractory acute myeloid leukaemia [ID1484]

As a stakeholder you have been invited to comment on the technical report for this appraisal. The technical 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.

We need your comments and feedback on the questions below. You do not have to answer every question. The text boxes will expand as you type. Please read the notes about completing this form. We cannot accept forms that are not filled in correctly. Your comments will be summarised and used by the technical team to amend or update the scientific judgement and rationale in the technical report.

Deadline for comments: **4 November 2019.**

Thank you for your time.

Please log in to your NICE Docs account to upload your completed form, as a Word document (not a PDF).

Notes on completing this form

- Please see the technical report which summarises the background and submitted evidence. This will provide context and describe the questions below in greater detail.
- 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.
- Do not use abbreviations.
- Do not include attachments such as journal articles, letters or leaflets. 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.
- If you provide journal articles to support your comments, you must have copyright clearance for these articles.
- 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. 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.

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

Your name	[redacted]
Organisation name – stakeholder or respondent (if you are responding as an individual rather than a registered stakeholder please leave blank)	**NCRI-ACP-RCP-RCR**
Disclosure Please disclose any past or current, direct or indirect links to, or funding from, the tobacco industry.	**n/a**

Questions for engagement

Issue 1: Comparators	
Is best supportive care (BSC) a relevant comparator for patients with relapsed or refractory FLT3+ AML in NHS clinical practice?	For the majority of R/R FLT3-AML patients, salvage chemotherapy is the most relevant comparator therapy here. For patients who, following their initial diagnosis with FLT3-AML, are considered unsuitable for intensive therapy and then go on to receive non-intensive regimens, most will either prove refractory or have short remissions, after which BSC would be the likely remaining treatment option. In this group, targeted, relatively non-toxic oral therapy with gilteritinib is likely to be an attractive treatment option and BSC would be the most relevant comparator here. There may also be patients who receive intensive therapy following initial diagnosis, are no longer considered suitable for an intensive approach at the time of relapse/refractory disease due to treatment related toxicities, new medical issues etc – BSC may also be a relevant comparator to gilteritiinb in this group. In neither of these groups is 'bridge to transplant' an aim – gilteritinib would be used as a relative non-toxic option to extend survival while maintaining QoL.
What proportion of patients would receive BSC in this patient population?	Difficult to be precise. Probably in the order of 20%.
Issue 2: Prior midostaurin use	
Would gilteritinib be used after prior midostaurin?	Yes.
What proportion of rrAML patients receive midostaurin in NHS clinical practice in England?	Midostaurin is not approved for use at the time that patients become 'relapsed / refractory'. Following previous NICE approval, a substantial proportion (but not all) of newly-diagnosed FLT3+ patients now receive midostaurin with frontline induction and consolidation chemotherapy in

	standard clinical practice. Not all FLT3+ patients currently receive midostaurin, but it will probably be a majority (60%?). Reasons for NOT treating with midostaurin will include: inadequate access to timely FLT3 mutation testing or use of Mylotarg with induction chemotherapy (NICE approved Mylotarg for patients not known to have adverse cytogentics at diagnosis – a group with considerable overlap with FLT3+ patients, but there is no current safety data to allow the drugs to be used in combination).
If prior midostaurin use differs between the ADMIRAL trial and the population in England, how will this affect the effectiveness results from the trial?	Prior FLT3 inhibitors were only used in 10-15% of patients in the ADMIRAL trial and this included both midostaurin and sorafenib. There is nothing in the ADMIRAL data to suggest any lesser rates of clinical response in those patients who had received prior midostaurin (though a relatively small sub-group) but I'm not aware of any clinical data to suggest that prior exposure to a 'first generation' FLT3 inhibitor will affect the effectiveness of gilteritinib in this setting. A small minority of previously FLT3-inhibitor-treated patients will develop tyrosine kinase domain point mutations, but gilteritinib is active against the majority of FLT3 TKD mutations.
Issue 3: Cure assumptions	
Is it plausible to assume that all patients who remain alive at 3 years are 'cured' regardless of whether they have progressed or have had HSCT?	The vast majority of relapses in FLT3+ AML occur within the first 6-12 months and this will especially be true of patients who already have R/R FLT3-AML and are receiving second line therapy. If stable at 3 years the majority can be assumed to have been 'cured. (Patients surviving to 3 years having not had HSCT will be vanishingly rare),
Issue 4: Gilteritinib effectiveness after HSCT	
Is it more appropriate to use external data or ADMIRAL trial data to estimate the relative effectiveness of gilteritinib after HSCT?	Our experts were unsure which 'external data' you would be planning to access in this context. The ADMIRAL study data certainly appear robust. Additional data will be accruing from the ongoing MORPHO study of gilteritinib maintenance (vs placebo) given as maintenance post-transplant but this study is still open to recruitment so is not expected to provide meaningful information for several years.

Which of the extrapolated survival models (see figure 1) appears to be more clinically plausible?	
Issue 5: Gilteritinib maintenance therapy	
Would gilteritinib be used as maintenance therapy after HSCT in clinical practice?	If approved, then gilteritinib would certainly be used as maintenance post HSCT although there are still no clear randomised data to support using FLT3 inhibitors as maintenance post SCT – the previously-mentioned MORPHO study will be invaluable in this context. Given the fragile nature of remissions in previously R/R FLT3-AML, most clinicians believe that ongoing FLT3 inhibition is likely to suppress low level FLT3+ AML subclones that are likely to drive relapse. Certainly the available data from ADMIRAL suggest better survival in patients who are able to resume gilteritinib post-SCT (compared to those that don't) although this may be a self-selecting better prognosis group due to absence of GVHD, robust blood counts etc. There is randomised evidence from the SORMAIN study (Burchert ASH meeting 2018) that another FLT3-inhibitor, sorafenib, prolongs survival and reduces risk of relapse (in comparison to placebo) when given as maintenance for 2yrs post allo-SCT although this study has not yet been published in a peer-reviewed journal.
Is it plausible that there is an *additional* effect of maintenance therapy on OS?	Yes – this is plausible (see above) and seems to be supported by sub-group data from ADMIRAL study, but subject to the caveats above.
Is the current method of deriving and applying an additional benefit of maintenance therapy appropriate?	Our experts were not entirely clear what your current method of deriving and applying benefit from maintenance therapy is. A dedicated randomised trial (MORPHO) is likely to be the definitive test here. Comparing patient outcomes within ADMIRAL for those who continued gilteritinib post-transplant vs those who didn't has significant limitations because the groups are probably not

	equal – those not restarting drug may have active GvHD, may have insufficient blood count levels or other new post-SCT comorbidities, all of which would reduce their expected survival compared with patients who are more stable and fit post SCT.
Issue 6: Utilities	
Are utility values from Janssen *et al.* or Ara and Brazier *et al.* more clinically plausible after the 3-year cure point? See figure 3.	
Issue 7: Costs	
In NHS clinical practice in England, would gilteritinib tablets be wasted if patients stopped taking it unexpectedly, for example because of death?	Yes. Patients with R/R FLT3-AML have significantly poor prognosis and may rapidly develop complications eg. septic episodes which cause them to stop treatment unexpectedly. It may be more cost-effective to issue prescriptions initially as half courses (eg 14 days) to avoid costs associated with tablet wastage.
Should drug costs be applied as a one-off cost in the first cycle of the model?	
Is it more plausible to assume that for patients alive after 3 years (after the assumed 'cure') a. patients who have had HSCT have 1.5 outpatient visits every month indefinitely, **or** have no follow-up costs? b. patients who have not had HSCT would require 1 outpatient visit every 6 months **or** have no follow-up costs?	'b' is more plausible. By 3 years, patients will remain under active haematology follow-up but will generally be getting seen in clinic roughly every 3-4 months. It may be more frequent than this in the setting of ongoing late allograft complications (eg. chronic GvHD) but that would be unlikely to increase the visits to more than monthly.

Is it reasonable to remove progression costs from the model after 3 years (after the assumed 'cure')?	Yes – it seems reasonable to remove progression costs from the model after 3 years.
Is it reasonable to assume 3.3 or 2.0 FLT3 tests will be required to identify 1 patient (in other words, does FLT3 occur in around 30% of patients which would result in 3.3 tests per patient)?	The vast majority of patients under consideration for second line gilteritnib therapy will already be known to have a FLT3 mutation from screening performed at the time of first diagnosing their AML. It is important that the FLT3 mutation status is rechecked at the time of relapse/refractory disease but the vast majority (80-90%+) will still have the FLT3 mutation at that point – so that would probably suggest 1.1-1.2 tests to identify each patient.

Gilteritinib for treating relapsed or refractory acute myeloid leukaemia: A Single Technology Appraisal

ERG commentary on the company's technical engagement response

Produced by	School of Health and Related Research (ScHARR), The University of Sheffield
Authors	Paul Tappenden, Professor of Health Economic Modelling, ScHARR, University of Sheffield, Sheffield, UK Emma Simpson, Senior Research Fellow, ScHARR, University of Sheffield, Sheffield, UK John Stevens, Reader in Decision Science, ScHARR, University of Sheffield, Sheffield, UK Aline Navega Biz, Research Associate, ScHARR, University of Sheffield, Sheffield, UK Ruth Wong, Information Specialist, ScHARR, University of Sheffield, Sheffield, UK Srinivas Pillai, Consultant Haematologist, University Hospitals of North Midlands, Stoke-on-Trent, UK
Correspondence Author	Paul Tappenden, Professor of Health Economic Modelling, ScHARR, University of Sheffield, Sheffield, UK
Date completed	13th November 2019

1. Introduction

This document sets out the ERG's commentary on the company's technical engagement response.[1] The company's response includes discussion relating to the eight issues raised in the NICE draft technical engagement report. The main points raised in the company's response and the ERG's views regarding these are summarised in Table 1.

In addition, the company submitted a revised version of their model together with a document which details how additional amendments have been implemented within the model. The starting point for the company's revised economic analyses is the ERG-corrected version of the company's model (see ERG report,[2] Exploratory Analysis 1), including an additional monthly dispensing fee for gilteritinib (monthly rather than 3-monthly costs). The company's revised model incorporates the following amendments based on the issues raised during technical engagement:

- The proportion of patients receiving BSC in the salvage chemotherapy comparator group is set equal to 25% (Issue 1)
- The assumed cure point has been changed from 3 years to 2 years (Issue 3)
- The costs associated with wastage, outpatient visits, post-progression costs after the assumed cure point and FLT3+ mutation testing costs have been amended (Issue 7)
- Additional impacts of inpatient hospitalisation for high-intensity chemotherapy on HRQoL and costs have been included (Issue 8).

The ERG's comments on the company's response are detailed in the subsequent sections.

Table 1: Summary of company's technical engagement response and ERG's comments

No.	Issue	Summary of company's response	Summary of ERG's comments on company's response
1	Comparators - BSC	The company agrees that BSC is an option. The model has been amended to include a 25% BSC rate in the salvage chemotehrapy group, which reduces the ICER.	The ERG does not believe company's new analyses to be appropriate as they do not account for differences in characteristics between patients who were enrolled in ADMIRAL[3] and patients who would otherwise receive BSC.
2	Prior midostaurin use	Available data indicates benefit in patients previously treated with FLT3+ inhibitors. No model amendment.	The company's response appears to be reasonable. The effectiveness of gilteritinib after midostaurin is uncertain.
3	Cure assumptions	The company's base case model now includes a 2-year cure point (previously 3 years).	The ERG does not believe the company's updated analyses to be reasonable.
4	Gilteritinib effectiveness after HSCT	The company retains their original position that post-HSCT OS should be estimated using Evers *et al*.[4]	The ERG retains their view that data from ADMIRAL[3] should not be discarded. If external data are considered relevant, these should supplement the data from ADMIRAL, not replace them. The source of OS data for patients with HSCT has a substantial impact on the ICER.
5	Gilteritinib maintenance therapy	The company expects maintenance therapy to be associated with an additional OS benefit, but expects the use of maintenance therapy to be minimal. The company believes that if no additional benefit is assumed in the model, no additional cost should be included.	The company's model estimates the treatment effect for maintenance therapy using an unanchored naïve comparison using data from ADMIRAL[3] and Evers *et al*.[4] The ERG's exploratory analyses include the costs and outcomes for patients receiving HSCT (with or without maintennce therapy) based on the observed data from ADMIRAL.
6	Utilities	The company have incorporated general population utilities from Ara and Brazier[5] for patients who reach the assumed cure point.	This is in line with the ERG's preferred base case.
7	Costs	• The company has included 0.25 packs wastage • The company argues that gilteritinib acquisition costs are overestimated as they are not discounted, biasing against gilteritinib • Costs associated with outpatient vists, post-progression costs following the cure point and FLT3+ testing costs have been updated	• The amount of wastage depends on how gilteritinib will be prescribed in usual clinical practice • Gilteritinib acquisition costs are underestimated due to patients remaining on treatment at the data cut-off, biasing in favour of gilteritinib • The other updated costs are generally in line with the ERG's preferred base case.
8	Impact on cost, HRQoL of outpatient vs. inpatient care	The company has included a disutility for high-intensity chemotherapy (applied indefinitely) and higher hospitalisation costs during the first model cycle.	It may not be clinically realistic to expect AE disutilities to apply indefinitely. The company's assumptions regarding hospitalisation costs appear to be in line with clinical views expressed during the technical engagement call. Neither amendment has a marked impact on the ICER for gilteritinib.

2. ERG commentary on company's technical engagement response

Issue 1: Comparators

Within their technical engagement response,[1] the company agrees that BSC is an option for patients with FLT3+ R/R AML and has amended their model to assume that 25% of patients receive BSC. This reduces the company's ICER.

The ERG has three main concerns regarding the way in which the company has included BSC within their updated model:

1. This amendment is implemented in the company's updated model by assuming that 25% of patients in the salvage chemotherapy group instead receive BSC, whilst for the remaining 75% of patients, the original probabilities of receiving FLAG-IDA, MEC, LDAC and AZA are maintained, based on the proportionate usage of these regimens within ADMIRAL.[3] The HSCT rate in both groups is assumed to remain unchanged. The ERG believes that some patients may opt to receive BSC because of poor fitness and/or because they are unlikely to be eligible to proceed to HSCT. If the target population is comprised of some patients who elect to receive BSC alone for these reasons, the HSCT rate in the BSC group would be approximately zero, and the HSCT rate in the gilteritinib group would also be reduced. The CS[6] does not provide any evidence that gilteritinib-treated patients who would otherwise receive BSC may proceed to transplant. The ERG's clinical advisor suggested that the HSCT rate for gilteritinib in this less fit population might be roughly 10%, although given the absence of evidence, this estimate should be interpreted with caution.
2. In ADMIRAL,[3] BSC alone was not included either as a comparator arm, or as part of a comparator arm. As noted in the ERG report[2] (Section 5.3.3, critical appraisal point [6]), the company's model estimates OS for patients receiving BSC using an unadjusted indirect comparison using data from ADMIRAL[3] and Sarkozy *et al*.[7] The ERG report notes concerns regarding: (a) how the Sarkozy *et al* study was identified and whether alternative sources exist and (b) the appropriateness of the statistical comparison being made.
3. The ERG attempted to apply the company's amendment to the proportion of patients receiving BSC using the model change log provided by the company and the ERG-corrected version of the model. The resulting ICER was not the same as that reported in the company's technical engagement response; the ERG's estimated ICER was approximately £3,000 lower than the company's estimate. The reasons for this are unclear.

Whilst the ERG believes that BSC is a relevant comparator for gilteritinib, it is not appropriate to include BSC as part of the blended comparator, as these patients are likely to be less fit than the population recruited into ADMIRAL and as such they represent a different patient population. Instead, the ERG believes that it would be more appropriate to compare gilteritinib versus BSC as a separate

analysis. The ERG notes that the evidence available to inform such an analysis within this population is largely absent and any resulting cost-effectiveness estimates should be considered highly uncertain. Based on the ERG's preferred model (see ERG report,[2] exploratory analysis 8 [ERG preferred base case]), which uses data on OS for patients receiving HSCT from ADMIRAL, amending the HSCT rate for the gilteritinib group to 10% increases the ICER for gilteritinib versus BSC to £86,381 per QALY gained.

Issue 2: Prior midostaurin use

The company agrees that gilteritinib would be used after midostaurin. The company's subgroup analysis of OS in ADMIRAL (see CS,[6] Section B.2.7, Figure 7) indicates a consistent treatment effect for patients with prior FLT3 inhibitor use, although the number of patients is small (n=46) and the confidence interval is wide and crosses unity.

The ERG considers the company's response on this point to be reasonable, but notes that this remains an area of clinical uncertainty. The ERG's clinical advisor commented that they would use gilteritinib irrespective of whether the patient had previously received midostaurin.

Issue 3: Cure assumptions

The company's technical engagement response[1] states that the company's original assumption regarding the 3-year cure point was based on the previous NICE appraisal of midostaurin for untreated FLT3+ AML[8] (TA523) and was intended to be conservative. The company also comments that information provided by clinical experts in the technical engagement papers suggests that relapse would usually occur in 6-9 months. As part of their response, the company have amended their original assumption of a 3-year cure point to now reflect a 2-year cure point. The company's response also comments that the assumptions that patients who do not receive HSCT and relapsed/progressed patients with HSCT who survive to the cure point are assumed to be cured are likely to have a negligible impact on the model results.

The ERG's views regarding the company's approach to modelling cure are presented in detail in the ERG report[2] (Section 5.3.3, critical appraisal points [3] and [4]). The ERG's views on these issues remain unchanged. The ERG highlights the following key points:

- The company's decision to change the assumed cure point from 3 years (base case) to 2 years (updated base case) implies that a greater proportion of patients will achieve cure. This improves the ICER for gilteritinib. As shown in the ERG's exploratory analyses (see ERG report,[2] ERG additional sensitivity analysis 4), assuming a 2-year cure point together with the With HSCT OS data from ADMIRAL[3] reduced the ICER for gilteritinib versus salvage chemotherapy from £102,085 to £66,123 per QALY gained.

- The company's basis for changing their assumption of a cure point from 3 years to 2 years lacks justification and is not based on empirical evidence relating to the FLT3+ R/R AML population.
- Cure fractions, if present, and their implications for the proportions of patients surviving up to certain timepoints, are likely to be specific to the population under consideration.
- The company's model assumes that cure occurs at some timepoint since initiating gilteritinib/chemotherapy, rather than the time since receipt of HSCT. In ADMIRAL,[3] patients received HSCT up to [redacted] months post-randomisation (see ERG report,[2] Section 5.3.3., Figure 14). If such a cure point exists, the company's updated model implies that this is less than 2 years post-HSCT. This may not be consistent with clinical opinion.
- The company has not explored their assumptions of cure using statisitcal (e.g. mixture-cure) models using data from ADMIRAL[3] or any other source. As part of the exploratory analyses, the ERG fitted mixture-cure models to the With HSCT data from ADMIRAL. These analyses suggested that the assumption of a crude 3-year cure point produces a broadly similar OS projection to the ERG's mixture-cure models (see ERG report,[2] Section 5.4.3 Figure 24). However, all of the ERG's analyses which used OS data for patients with HSCT from ADMIRAL indicated or implied considerably less favourable OS estimates than that predicted by the company's original base case model and their revised base case model (see Issue 4). The ERG has updated this figure to allow a comparison of the company's revised model including the 2-year cure point (see ***[redacted]***1***, note - the company's revised model With HSCT OS is represented by the blue dashed line).
- The company's assumption of cure in patients who have already progressed may be unrealistic but is unavoidable given the use of a fixed cure point within a partitioned survival model structure. This assumption could have been avoided through the use of a state transition approach. However, this has not been done and its impact on the ICER remains unclear.
- It should be noted that the company's updated assumption of cure at 2-years has a further impact on the proportion of patients who do not receive HSCT but who are assumed to be cured. The company's updated model now suggests that 11% of gilteritinib-treated patients without HSCT and 5% of chemotherapy-treated patients without HSCT will remain alive at 2-years and will therefore achieve cure. In addition, around 4% of patients who receive HSCT but have progressed by 2-years are also assumed to be cured in each treatment group. The ERG does not consider these assumptions to be plausible and believes that they will bias in favour of the gilteritinib group.

[redacted]

[redacted]

Issue 4: Gilteritinib effectiveness after HSCT

The company's technical engagement response[1] maintains that post-HSCT OS should be based on Evers *et al*,[4] as the data from ADMIRAL[3] relate to a small number of patients and the median follow-up for OS post-HSCT is limited. The company's response states that given the clinical experts' comments on the technical engagement report that relapses is most likley to occur within 6-9 months, it is plausible that the [redacted] of patients who were censored at the data cut-off in ADMIRAL would still be alive at 3-years. The information provided in the company's response to this point has already been provided in the CS[6] and the company's clarification responses;[9] for the sake of brevity, this information is not repeated here.

The ERG's concerns regarding the company's approach to estimating OS for patients with HSCT is detailed in the ERG report[2] (Section 5.3.3, critical appraisal point [4]). The ERG's views remain unchanged. The ERG highlights the following key points:

- Within the company's model, OS following HSCT is a key driver of the ICER for gilteritinib.
- Within the company's original and revised models, OS for patients receiving HSCT is based on Evers *et al*;[4] data from ADMIRAL[3] are not used. The ERG does not consider it appropriate to disregard evidence from the ADMIRAL trial. If external information is considered relevant for inclusion in the model, it should supplement evidence from ADMIRAL, not replace it.
- The study reported by Evers *et al*[4] reflects a population of patients with R/R AML in CR2; these patients did not have a FLT3+ mutation. The relevance of the Evers *et al* population to the target population is unclear.

- The company has amended their model to reflect a 2-year cure point. The revised model suggests that around 60% of gilteritinib-treated patients who receive HSCT will survive up to the cure point. At the final data cut-off in ADMIRAL,[3] [redacted] of [redacted] ([redacted]) gilteritinib-treated patients who received HSCT had already died, and a further [redacted] patients who had received HSCT died during the short lag between the data cut-off and the database lock[9] (total known deaths = [redacted] of [redacted] patients i.e. [redacted] of all gilteritinib-treated patients receiving HSCT). It is therefore not possible for the ADMIRAL data to support the company's modelled prediction that 60% of gilteritinib-treated patients receiving HSCT will achieve cure at three years, as [redacted] of these patients had already died within the shorter follow-up available from ADMIRAL.
- As discussed in the ERG report,[2] the available OS data for patients with HSCT in ADMIRAL[3] do not indicate the classic pattern of survival that is indicative of the presence of a cure (e.g. a plateau at the end of the survival function). ***[redacted] 2 summarises the OS data from ADMIRAL for gilteritinib-treated patients who received HSCT (from randomisation, divided into 5-month intervals). The black bars show censored observations, whilst the grey bars show death events. As shown in the figure, many censored observations occurred early, most likely due to short follow-up. If one expects a high proportion of the group to be cured, one would also expect the majority of the censored observations to be seen towards the right hand side of the figure. Under the company's original base case model assumptions, this would have required all of the patients who are censored to be cured at 3 years. The ERG does not consider this to be likely. As noted above, the proportion of patients surviving to 2-years is known to be less than the estimate of 60% predicted by the company's model, despite censoring.

2

- The ERG's mixture-cure models fitted to the available data for patients with HSCT from ADMIRAL indicated cure fractions of between [redacted] (exponential mixture-cure model) and [redacted] (Weibull mixture-cure model). These models suggest considerably lower estimates of mean OS for patients receiving HSCT compared with the company's model (see ERG report,[2] Table 45, ERG-estimated OS using ADMIRAL: 3.81 to 5.29 years; company-estimated OS using Evers *et al*: 10.23 years).
- The company's review of studies reporting on post-HSCT OS missed the study reported by Poire *et al*.[10] This study relates to patients with FLT3+ R/R AML and includes a subgroup of patients who were in CR2; the ERG believes that this CR2 subgroup better reflects the With HSCT group in ADMIRAL.[3] Whilst the CR2 subgroup in this study is small (n=37), this indicates a less favourable OS compared with Evers *et al* (Poire *et al*:[10] cumulative survival probability at 3 years = 22%; Evers *et al*:[4] cumulative survival probability at 3 years = 46%).
- The company's response argues that the use of Evers *et al*[4] is preferable to Poire *et al*,[10] as the former has longer follow-up and included a larger sample size. The ERG considers that relevance to the target population, rather than sample size and duration of follow-up, represents a more appropriate criterion for selcting evidence.
- The ERG believes that the best way of reducing uncertainty is by collecting additional data. Within ADMIRAL,[3] the final data cut-off was the 17th September 2018. As noted in the ERG report[2] (Section 5.3.3, page 94), the company has indicated that further analysis of longer-term follow-up data is not expected. The ERG believes that the potential 14 months of additional OS follow-up for the remaining [redacted] gilteritinib-treated patients with HSCT in ADMIRAL who were censored at the data cut-off/database lock would likely resolve the key uncertainty surrounding

the cost-effectiveness of gilteritinib for FLT3+ R/R AML. The ERG questions why no further data collection is anticipated.

Issue 5: Gilteritinib maintenance therapy

The company's technical engagement response[1] states that the company agrees that gilteritinib should be used as a maintenance therapy and that it is reasonable to expect an additional OS benefit. Their response also states that they expect the use of maintenace therapy to be minimal. Further, the company also refers to the use of midostaurin and sorafenib as maintenance therapy and refers to separate Kaplan-Meier survivor functions for patients with HSCT who did or did not receive gilteritinib maintenance therapy in ADMIRAL.[3] With respect to the ERG's exploratory analyses, the company argues that those analyses which remove the additional benefit of maintenance therapy should also remove the costs of maintenance therapy; the company notes that this reduces the ICER.

The ERG's views regarding the company's approach to estimating the benefits of gilteritinib maintenance therapy are summarised in the ERG report[2] (Section 5.3.3., critical appraisal point [6]). The ERG's views have not changed. The ERG notes the following:

- Within ADMIRAL, [redacted] of patients randomised to gilteritinib who received HSCT subsequently received maintenance therapy. The ERG considers that if it is reasonable to expect an additional OS benefit from gilteritinib maintenance therapy, then it is unclear why a lower rate of maintenance therapy should be expected in clinical practice.
- If the use of maintenance therapy is lower in clinical practice than in ADMIRAL,[3] and if maintenance therapy is associated with additional OS gain, this will impact upon the ICER for gilteritinib. The company's model calculates total acquisition costs for gilteritinib using the mean number of cycles received across the whole gilteritinib group (see Issue 7); the impact of a lower rate of gilteritinib maintenance therapy on the ICER is unclear.
- The company's estimated treatment effect for gilteritinib maintenance therapy is based on an indirect comparison using OS data from patients receiving gilteritinib maintenance therapy after HSCT in ADMIRAL[3] and OS data for patients with R/R AML in CR2 from Evers *et al*.[4] This estimated treatment effect is derived through the use of an unanchored naïve comparison between the groups with no adjustment for differences in patient characteristics.
- The ERG considers it inconsistent to make arguments about the "unsuitability" of the ADMIRAL[3] data for estimating OS outcomes for patients receiving HSCT, but then to use a subset of those data to justify an assumption of an additional OS treatment effect associated with gilteritinib maintenance therapy.

- The available data for patients With HSCT in ADMIRAL[3] (see ERG report,[2] Figure 16) do not suggest conclusive evidence of a difference in post-HSCT OS between the gilteritinib and salvage chemotherapy groups.
- The ERG notes that in their exploratory analyses in which OS for all patients is based on ADMIRAL,[3] assumptions regarding additional benefits of gilteritinib maintenance therapy are generally not relevant. The ERG's base case uses a standard parametric (log normal) model fitted to the available OS data for patients who received HSCT in ADMIRAL and includes acquisitions costs for gilteritinib based on the amount of gilteritinib received within the trial. In principle, this means that both health outcomes and costs are modelled in line with the trial (although this principle is not quite met due to limitations in the company's costing approach, see Issue 7). The ERG notes that it was necessary to pool the With HSCT OS data in ADMIRAL[3] between gilteritinib and chemotherapy groups; if maintenance therapy does provide an incremental OS gain, outcomes for the chemotherapy group will be overestimated, although the ERG believes that any bias caused by this pooling of data will be small.

Issue 6: Utilities

The company agrees with the use of utilities based on Ara and Brazier.[5]

This was part of the ERG's preferred base case case; hence, no further comment from the ERG is requried.

Issue 7: Costs

Within their technical engagement response,[1] the company agrees that there is potential for wastage and cites TA474[11] (sorafenib for hepatocellular carcinoma), in which the Appraisal Committee accepted an assumption of 7 days' wastage. The company argues that the same assumption should be applied in the appraisal of gilteritinib. The company has not amended their approach to modelling the acquisition costs of gilteiritnib (all costs are applied as a once-only cost in the first model cycle) and they argue that owing to problems in applying discounting, the estimated gilteritinib drug costs *"consititute a small overestimate of the true value."* The company also accepts the ERG's amendments regarding: (a) the frequency of outpatient visits; (b) post-progression costs after 3-years and (c) the greater number of tests required to identify a patient with a FLT3+ mutation.

The ERG believes that the most appropriate assumptions regarding wastage will depend on how gilteritinib would be prescribed in usual clinical practice. The assumption of 7 days' wastage may be reasonable, although the company has not provided any empirical evidence to either support or refute this. Using the ERG's preferred base case model, applying an assumption of 0.25 months' wastage reduces the ICER for gilteritinib versus chemotherapy from £102,085 to £98,713 per QALY gained.

As noted in the ERG report[2] (Section 5.3.3, critical appraisal point [8]), the company estimated drug costs based on the number of cycles of gilteritinib consumed prior to the data cut-off in ADMIRAL[3] ([redacted] cycles). However, at the data cut-off, [redacted] of patients were still receiving gilteritinib. The additional costs of treating these patients were not included in the company's original analysis and have not been included in the company's updated analyses. This leads to a bias in favour of gilteritinib which will outweigh the impact of the company failing to discount drug costs. The ERG notes that if it is reasonable to assume that all patients are cured at 3 years (as per the company's original base case model) or even at 2 years (as per company's updated model), there should be no clinical rationale to continue treatment beyond whichever timepoint is considered to be plausible (if either is considered plausible). The ERG believes that the company should have extrapolated the available data on time to treatment discontinuation from ADMIRAL[3] in order to estimate the amount of gilteritinib consumed over time, taking explicit account of the assumed cure point. In the absence of such an analysis, the extent of this bias is unclear.

The company has amended the costs associated with outpatient visits, post-progresion treatments after 3-years, and FLT3+ mutation testing costs. The ERG notes that the post-progression treatment costs are applied up to 3-years, whilst the assumed cure point is now 2-years; this appears to reflect an error. The ERG also notes some minor differences in the long-term costs of outpatient visits between the company's revised model and the ERG's error-corrected analysis; however, these will have a negligible impact on the ICER. The amended FLT3+ mutation testing cost appears to have been applied correctly within the company's revised model.

These assumptions formed part of the ERG's exploratory analyses; hence no further comment from the ERG is required.

Issue 8: Impact on cost, quality of life of outpatient vs. inpatient care

The company has amended their model to include an additional disutility and a higher hospitalisation cost for patients receiving high-intensity chemotherapy in the salvage chemotherapy group.

The ERG notes that the disutility for high-intensity chemotherapy is applied in every model cycle over the entire time horizon; given the short duration of salvage chemotherapy (typically less than 2 months), it is unclear whether this assumption is clinically appropriate. However, the ERG notes that this has a negligible impact on the ICER. The ERG believes that the company's updated assumptions regarding additional hospital costs incurred by patients receiving high-intensity chemotherapy may be reasonable. These also have a minor impact on the ICER.

Discussion

The ERG believes that the key issue relates to the evidence source used to inform OS estimates for patients who receive HSCT. The ERG believes that the ADMIRAL data reflect the most relevant source of OS for patients with FLT3+ R/R AML who undergo HSCT. These data have not been used in the company's revised model. The ERG's exploratory analyses, which are based on ADMIRAL, indicate that the ICER for gilteritinib versus salvage chemotherapy is considerably higher than the company's original and revised estimates.

References

1. Astellas Pharma Ltd. Gilteritinib for treating relapsed or refractory acute myeloid leukaemia [ID1484]. Company's technical engagement response. Surrey, UK; 2019.
2. Tappenden P, Simpson E, Stevens J, Navega Biz A, Wong R, Pillai S. Gilteritinib for treating relapsed or refractory acute myeloid leukaemia: A Single Technology Appraisal. Sheffield; 2019.
3. Astellas Pharma Ltd. A Phase 3 open-label, multicenter, randomized study of gilteritinib versus salvage chemotherapy in patients with relapsed or refractory acute myeloid leukemia (AML) with FLT3 Mutation.. ADMIRAL Clinical Study Report. Surrey, UK; 2019.
4. Evers G, Beelen DW, Braess J, Sauerland C, Kolb HJ, Reichle A, *et al.* Outcome of patients with acute myeloid leukemia (AML) undergoing allogeneic hematopoietic stem cell transplantation (HSCT) beyond first complete remission (CR1). *Blood* 2018;132:4649.
5. Ara R, Brazier JE. Populating an economic model with health state utility values: moving toward better practice. *Value in Health* 2010;13:509-18.
6. Astellas Pharma Ltd. Gilteritinib (XOSPATA) for treating relapsed or refractory acute myeloid leukaemia ID1484. Company's evidence submission to the National Institute for Health and Care Excellence. Document B. Surrey, UK; 2019.
7. Sarkozy C, Gardin C, Gachard N, Merabet F, Turlure P, Malfuson JV. Outcome of older patients with acute myeloid leukemia in first relapse. *American Journal of Hematology* 2013;88:758-64.
8. National Institute for Health and Care Excellence. Midostaurin for untreated acute myeloid leukaemia. Technology Appraisal Guidance No. 523. London; 2018.
9. Astellas Pharma Ltd. Gilteritinib (XOSPATA) for treating relapsed or refractory acute myeloid leukaemia ID1484. Company's response to clarification questions from the ERG. Surrey, UK; 2019.
10. Poiré X, Labopin M, Polge E, Passweg J, Craddock C, Blaise D, *et al.* Allogeneic stem cell transplantation benefits for patients ≥ 60 years with acute myeloid leukemia and FLT3 internal tandem duplication: a study from the Acute Leukemia Working Party of the European Society for Blood and Marrow Transplantation. *Haematologica* 2018;103:256-65.
11. National institute for Health and Care Excellence. Sorafenib for treating advanced hepatocellular carcinoma. Technology Appraisal Guidance 474. London; 2017.

NATIONAL INSTITUTE FOR HEALTH AND CARE EXCELLENCE

Technical report

Gilteritinib for treating relapsed or refractory acute myeloid leukaemia

This document is the technical report for this appraisal. It has been prepared by the technical team with input from the lead team and chair of the appraisal committee.

The technical report and stakeholder's responses to it 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 appraisal committee meeting.

The technical report includes:

- topic background based on the company's submission
- a commentary on the evidence received and written statements
- technical judgements on the evidence by the technical team
- reflections on NICE's structured decision-making framework.

This report is based on:

- the evidence and views submitted by the company, consultees and their nominated clinical experts and patient experts and
- the evidence review group (ERG) report.

The technical report should be read with the full supporting documents for this appraisal.

1. Summary of the draft technical report

After technical engagement the technical team has collated the comments received and, if relevant, updated the judgement made by the technical team and rationale. Judgements that have been updated after engagement are highlighted in bold below.

1.1 In summary, the technical team considered the following:

Issue		Technical team's preliminary judgement
1	**BSC as a comparator**	Best supportive care is a relevant comparator.
2	**Prior midostaurin use**	It is unknown whether gilteritinib would have a different effectiveness in the clinical population in England compared with in ADMIRAL.
3	**Cure assumptions**	It is plausible that patients alive at 3 years could be considered 'cured'.
4	**Estimating effectiveness of gilteritinib after HSCT**	The primary trial data from ADMIRAL should be used to inform post-HSCT OS, as it has been for other parts of the model. **It is appropriate to consider external data to validate the model.**
5	**Effectiveness of gilteritinib maintenance therapy**	It is unclear whether there is an additional benefit from maintenance therapy. Note: this issue is relevant only if trial data are not used for post-HSCT (issue 4).
6	**Utilities**	Utility values from Ara and Brazier are more plausible and progression should not be double-counted. The impact on the ICER is minimal.
7	**Cost amendments**	It is reasonable to include wastage, follow up costs after 3 years and 3.3 FLT3 tests.

1.2 The technical team recognised that the following uncertainties would remain in the analyses and could not be resolved:

- High dropout rate in salvage chemotherapy arm of ADMIRAL trial.

1.3 The cost-effectiveness results include a commercial arrangement (patient access scheme) for gilteritinib.

1.4 Taking these aspects into account, the technical team's preferred assumptions result in an incremental cost-effectiveness ratio (ICER) of £98,498 per QALY gained (see table 1). This estimate does not include the commercial arrangement for azacitidine, because this is confidential and cannot be reported here. Estimates that included this commercial arrangement would be higher than those reported above.

1.5 Based on the modelling assumptions, the intervention is likely to meet the end-of-life criteria (see table 3). However the technical team noted that the available analyses do not include the possible impact of the difference in prior midostaurin use between ADMIRAL and the population in clinical practice in England (issue 2).

1.6 The technology is unlikely to be considered innovative (see table 3).

1.7 No equality issues were identified.

2. Topic background

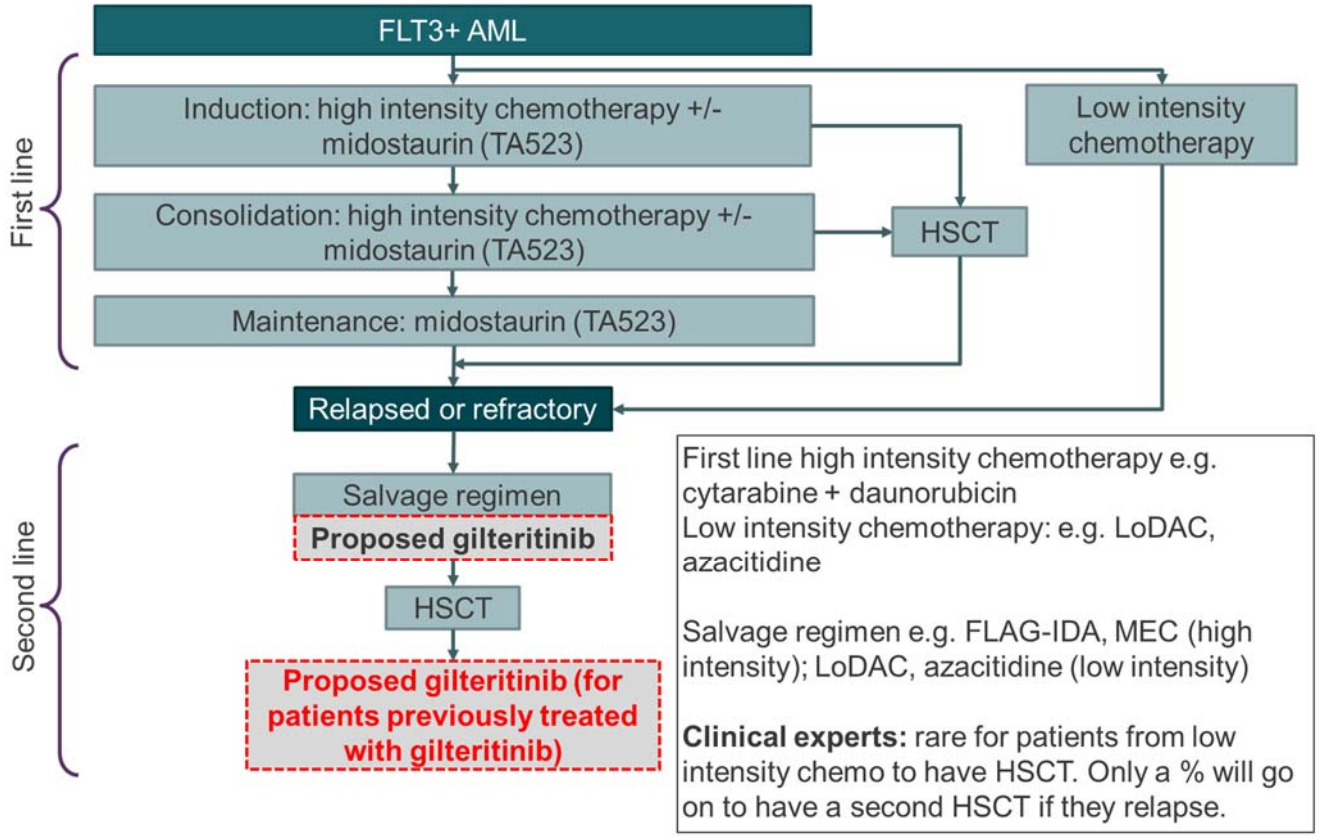
2.1 Disease background – acute myeloid leukaemia (AML)

- 2,543 new diagnoses of AML in England in 2016.
- Symptoms include fatigue, shortness of breath, weight loss, infection, fever, bruising, bleeding, bone or joint pain.
- Without treatment, progresses rapidly and is fatal within months.
- Mutations in Tyrosine kinase-3 FLT3 gene occur in around 30% of people with AML.

2.2 Gilteritinib

Marketing authorisation	Gilteritinib for the treatment of adult patients who have relapsed or refractory acute myeloid leukaemia (AML) with a FLT3 mutation (Marketing authorisation granted October 2019)
Mechanism of action	Tyrosine kinase-3 (FLT3) and AXL inhibitor. AXL is a receptor tyrosine kinase (RTK). AXL is a cell-surface receptor involved in the proliferation and survival of cells. It also mediates migration and invasiveness of cancer cells.
Administration	Oral tablet
Price	The average cost of a course of treatment of gilteritinib is anticipated to be [redacted] per patient (at list price) A patient access scheme has been agreed.

2.3 Treatment pathway



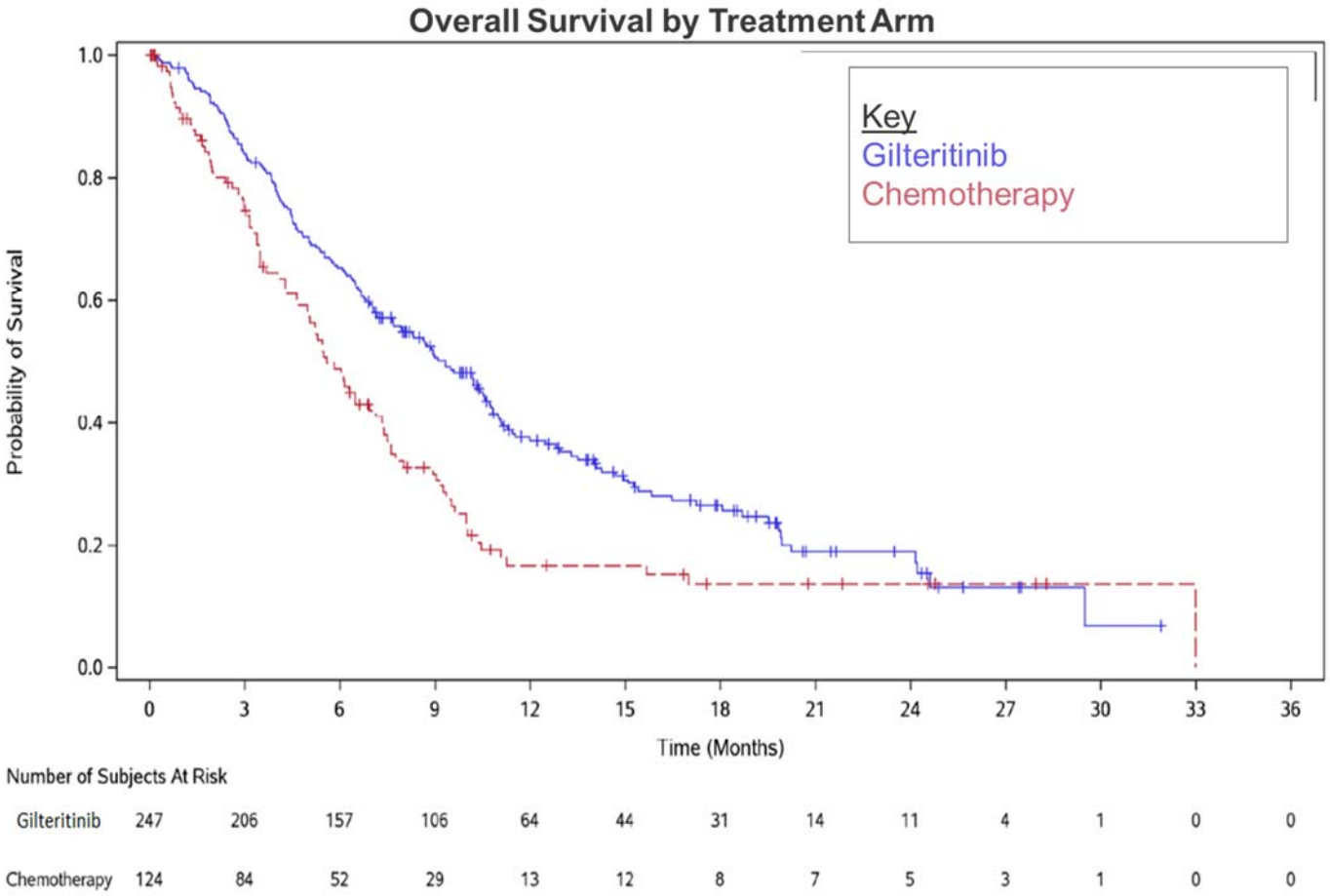
2.4 Clinical evidence

	ADMIRAL (n=371). Open-label, randomised trial
Population	Adults with r/r FLT3 mutation positive AML
Intervention	Gilteritinib 120mg/day
Comparator	Salvage chemo – investigator's choice (LoDAC, azacitidine, MEC, FLAG-Ida)
Primary outcomes	OS, CR/CRh
Secondary outcomes	EFS, LFS, duration of remission

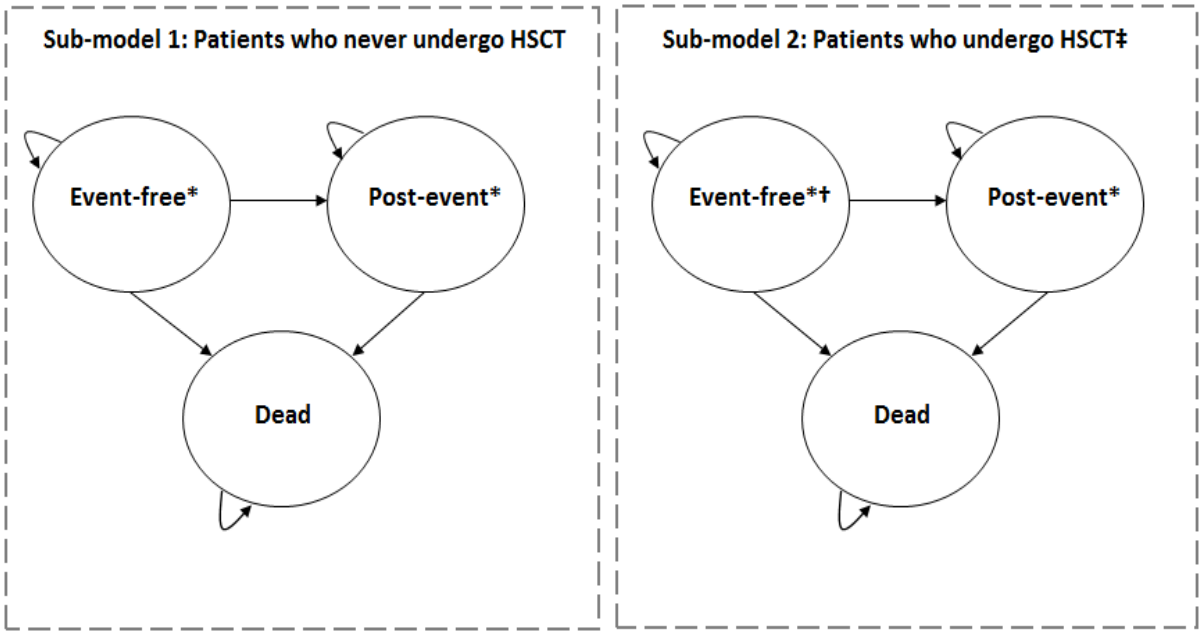
Abbreviations: CR complete remission, OS overall survival, LFS leukaemia-free survival, r/r relapsed or refractory. CR/CRh complete remission and complete remission with partial haematological recovery, EFS event-free survival

2.5 Key trial results

ADMIRAL	Median (95% CI)		HR vs salvage chemo	p value
	Gilteritinib monotherapy	Salvage chemotherapy		
Overall survival	9.3 months	5.6 months	0.637 (95%CI 0.490, 0.830)	p=0.0004
CR/CRh	34.0%	15.3%	-	p=0.001
CR/CRh: Complete remission or complete remission with partial haematological recovery				



2.6 Model structure



**** Surviving patients are assumed to be "cured" (SMR=2) after 3 years***
† HSCT is assumed to occur at fixed timepoint (gilteritinib - [redacted]; salvage chemotherapy - [redacted]). All patients in the HSCT sub-models remain alive and event-free until this timepoint
‡ Proportion of patients entering the HSCT sub-model dependent on treatment group

- 1-month cycle length
- Lifetime time horizon (40 years)
- Weighted comparator
- Decision tree followed by a partitioned survival model
- Gilteritinib used as maintenance following HSCT in a proportion of patients

2.7 Key model assumptions

2.7.1 Company model treatment effectiveness, no HSCT, OS

Gilteritinib

Salvage chemotherapy

Both treatment groups = gilteritinib and salvage chemotherapy. Patients in both groups who had HSCT were modelled as shown in this graph.

3. Key issues for consideration

Issue 1 – Comparators

Questions for engagement	1. Is best supportive care (BSC) a relevant comparator for patients with relapsed or refractory FLT3+ AML in NHS clinical practice? 2. What proportion of patients would receive BSC in this patient population?
Background/description of issue	The ADMIRAL trial included investigator's choice of salvage chemotherapy in the comparator arm. It did not include BSC as a comparator. The **technical team** noted that BSC was included in the scope. **The company** included a blended comparator of salvage chemotherapy based on ADMIRAL in its economic model results. It did not include BSC care in its base case results. The company did include BSC as a scenario analysis by applying a HR of 2.86 to gilteritinib OS, informed by a naïve indirect comparison. **The clinical experts** noted that BSC is a relevant option in a small proportion of patients who choose not have salvage chemotherapy due to toxicity and lack of fitness for treatment. **The ERG** noted that BSC could be a relevant option for 25% to 30% of patients in this population. The ERG had several concerns about the methods, assumptions and sources used to inform the company's indirect comparison for BSC (see ERG report section 5.3.3 issue 6). These include: • the indirect comparison assumes that LoDAC is equivalent to salvage chemotherapy • the source of the hazard ratio used is unclear • proportional hazards are assumed, which may not be appropriate.
Why this issue is important	Including BSC in the blended comparator reduces the ICER. A scenario analysis (re-run by the ERG) from the company's model showed that including BSC as 20% of the weighted comparator decreased the company's base case ICER from **£47,695** to **£43,601**.

Technical team preliminary judgement and rationale	Best supportive care is a relevant comparator and should be included in the cost-effectiveness analysis. The method the company used to model BSC is subject to uncertainty; however, the technical team note that it is plausible that including BSC could reduce the ICER.
Summary of comments	Comments from the company: • Agree that BSC is a relevant comparator for people who would choose not to receive chemotherapy • Provided updated cost-effectiveness results, taking into account BSC within the weighted comparator, at different proportions: 20%, 25% and 30%. Comments from clinical experts: • No more than 20% would receive BSC in this patient population • BSC is a relevant comparator for 10-20% of patients Professional group comments: • BSC would be an option for people who could not receive intensive therapy • Bridge to transplant would not be the aim in this group • This would probably be about 20% of patients Comments from patient expert • Patients may choose not to continue with chemotherapy due to significant side effects so BSC should be considered a relevant option • The proportion of patients receiving BSC is likely to increase with age. ERG • BSC should not be included in the weighted comparator because characteristics of people who would have BSC are likely to be different to those who would receive chemotherapy o HSCT rate is the same whereas people who would have BSC are less likely to receive a HSCT • Concerns about the methods used to estimate overall survival remain
Technical team judgement after engagement	Best supportive care is a relevant comparator but the method used to model BSC leads to uncertain results.

Issue 2 - Prior midostaurin use

Questions for engagement	3. Would gilteritinib be used after prior midostaurin? 4. What proportion of rrAML patients receive midostaurin in NHS clinical practice in England? 5. If prior midostaurin use differs between the ADMIRAL trial and the population in England, how will this affect the effectiveness results from the trial?
Background/description of issue	The **clinical experts** confirmed that gilteritinib would be given after midostaurin (another FLT-3 inhibitor) in clinical practice. The reasons are that it is a far more potent FLT-3 inhibitor and they did not believe that prior exposure to midostaurin would affect response. Additionally, at relapse AML cells with the ITD are very dependent upon that pathway for survival. The **clinical experts** also noted that the proportion having midostaurin in clinical practice (around [redacted]%) in England is higher than in the trial. In the ADMIRAL trial, 13% of the gilteritinib group and 11.3% of the salvage chemotherapy groups had prior FLT-3 inhibitors. ADMIRAL subgroup results showed that for patients with no prior FLT-3 inhibitor (n=[redacted]), gilteritinib was [redacted] (HR=[redacted], 95% CI [redacted]. For the [redacted] patients with prior use of a FLT3 inhibitor, the treatment difference was [redacted](HR [redacted] 95% CI [redacted]; *p*=[redacted].
Why this issue is important	Because the proportion of patients with prior midostaurin is higher in England than in ADMIRAL, the trial results could have a different effectiveness when applied to the population in England
Technical team preliminary judgement and rationale	The trial results might have a different effectiveness in the clinical population in England. Clinical and cost-effectiveness scenario analyses around this issue should be provided.
Summary of comments	Comments from the company: • Additional clinical benefits are seen in patients who previously received midostaurin • Gilteritinib is a more potent inhibitor and it is not expected that prior midostaurin exposure would affect the response • Discussion on the Technical Engagement call suggested [redacted] patients a month are receiving midostaurin. This equates to [redacted] patients per year, whereas it is estimated there are over 400 patients with FLT3 mutation positive relapsed or refractory AML a year. This suggests [redacted] would have received midostaurin.

	• The EPAR for gilteritinib states: “in this subpopulation (people who had prior FLT3 inhibitors) results were in favour of gilteritinib in terms of CR rate (18% vs 0%) and the HR for OS 0.705 (95%CI: 0.346, 1.438). Thus, exclusion of patients with prior FLT3 inhibitors from the indication was not considered necessary.” • Gilteritinib has different mechanisms of action and is a more potent and targeted FLT3 inhibitor than midostaurin, for example gilteritinib is effective in monotherapy, unlike midostaurin • The below Kaplan-Meier curve presents the survival data from the patient subgroup from ADMIRAL previously treated with midostaurin or sorafenib (n=45) and the survival data from the trial population as a whole  Comments from clinical experts: • Gilteritinib would be used after prior midostaurin • There is no convincing evidence that prior midostaurin affects response to gilteritinib – subgroups in ADMIRAL are in small patient numbers

	• In discussion with a US clinician he stated there was at least a 50% overall response rate in patients who have had prior midostaurin therapy • In first line treatment, FLT3 ITD status would be known for around 50% of people under 70 and most of these would receive midostaurin. FLT3 status would only be known in minority of people over 70 and few of these would receive midostaurin Professional group comments • Gilteritinib would be used after prior midostaurin • About 60% of people with FLT3 positive AML would receive midostaurin first line • Reasons for not treating with midostaurin include: o lack of access to timely FLT3 testing o use of gemtuzumab with induction chemotherapy (recommended where there are not known to be adverse cytogenetics at diagnosis) • Not aware of any clinical data to suggest prior exposure to a FLT3 inhibitor would affect the effectiveness of gilteritinib in this setting – a few patients may develop tyrosine kinase domain (TKD) point mutations, but gilteritinib is active against most FLT3 TKD mutations Comments from patient expert: • Clinical experts agree gilteritinib is a more potent inhibitor so would have activity in patients despite prior targeted treatment
Technical team judgement after engagement	It is unknown whether gilteritinib would have a different effectiveness in the clinical population in England compared with in ADMIRAL.

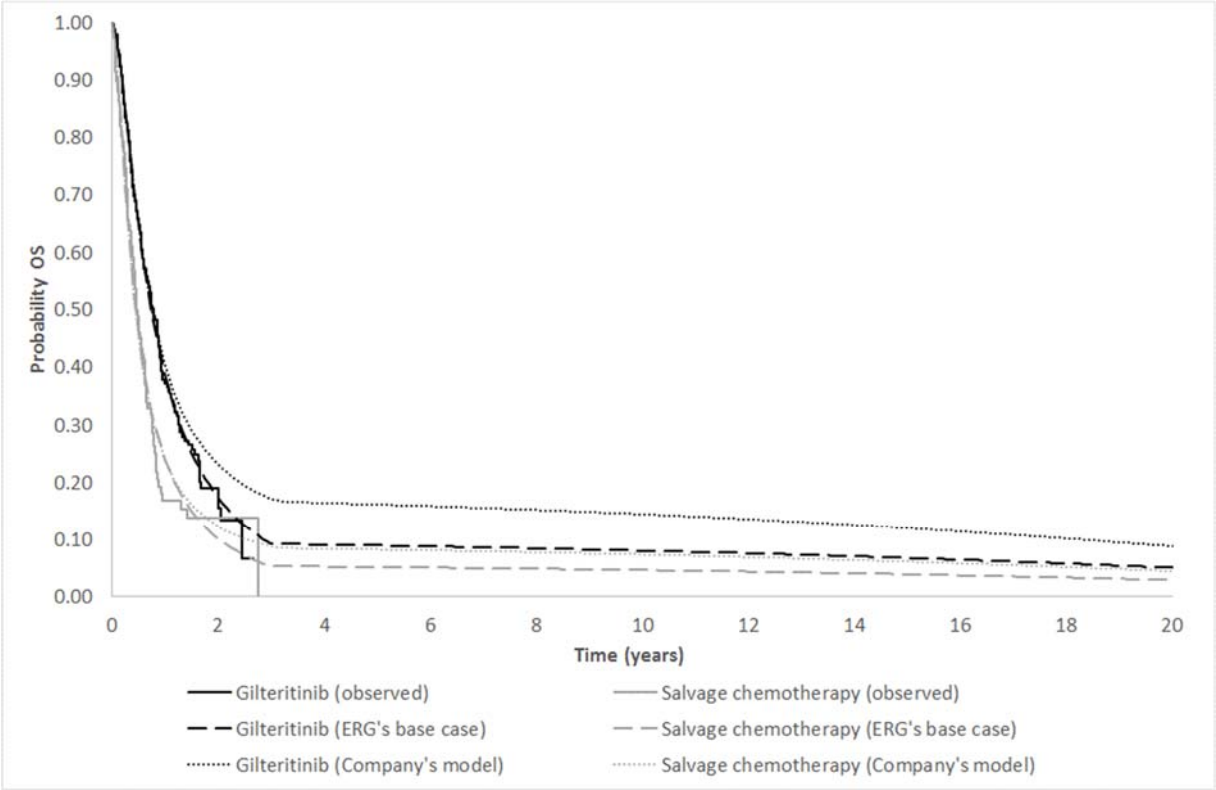
Issue 3 – Cure assumptions

Questions for engagement	6. Is it plausible to assume that all patients who remain alive at 3 years are 'cured' regardless of whether they have progressed or have had HSCT?
Background/description of issue	**The company** assumed that all patients who were alive at 3 years were 'cured', regardless of whether they had progressed or had had HSCT. After 3 years, patients' survival is modelled using an uplifted general population mortality rate (standardised mortality ratio of 2.0). The 3-year cure assumption was based on the NICE midostaurin appraisal in untreated FLT3+ AML (TA523), published literature and clinical advice. Evidence to support the cure assumption from ADMIRAL was not presented. The company have confirmed that no further data cuts are expected from ADMIRAL. **The clinical experts** agreed that 3 years was a plausible cure point. **The ERG** considers that there is uncertainty around the potential for gilteritinib to provide a 'cure'. It notes that ADMIRAL is not used to inform the assumptions of 'cure'. The ERG notes that the Kaplan-Meier curves from ADMIRAL do not show a plateau (see section 1.5 of technical report background for the OS KM curve). The ERG had some concerns over the structural constraints of the method of modelling cure but found that irrespective of how cure is modelled (using cure fractions or fixed cure timepoints) the mean OS for patients having HSCT were similar. A model based on the ADMIRAL with-HSCT OS data (see issue 4) will produce ICERs which are considerably higher than the company's base case using both methods of modelling cure (see figure 24 and table 45 ERG report).
Why this issue is important	The cure point affects the ICER. The ERG did sensitivity analysis for different cure points, the ERGs preferred base case changes from **£102,085** with a 3-year cure point to **£66,123** with a 2-year cure point and **£133,111** with a 4-year cure point. The importance and impact of the cure point on the ICER depends on the source of post-HSCT data used (see issue 4).
Technical team preliminary judgement and rationale	It is plausible that patients alive at 3 years could be considered 'cured'.
Summary of comments	Comments from the company:

	• Provided scenario analyses to explore cure points at 1 and 2 years as well as the 3 years in the original base case • Updated base case includes a 2 year cure point • Survival rate estimates for patients not receiving HSCT indicate that most do not survive to the cure point and patients who have progressed comprise a relatively small proportion of the patients alive who had a transplant. Therefore, the company considers that applying the cure point to everyone in the model does not impact on model results. Comments from clinical experts: • Plausible to assume that all patients who remain alive at 3 years are 'cured' • Patients who have progressed shouldn't be considered cured but this would be very rare if alive at 3 years Professional group comments: • Most relapses in FLT3 positive AML occur in the first 6-12 months, especially in second line therapy • If stable at 3 years most can be assumed to have been 'cured' • It would be very rare that anyone would survive to 3 years without HSCT Comments from patient expert: • It is plausible that all patients who remain alive at 3 years are 'cured' ERG • Company's updated model (2 year cure point) now suggests implausible outcomes: o 11% of gilteritinib-treated patients without HSCT and 5% of chemotherapy-treated patients without HSCT will remain alive at 2-years and will therefore achieve cure. o Around 4% of patients who receive HSCT but have progressed by 2-years are also assumed to be cured in each treatment group. o The company's submission states that HSCT is the only treatment with curative intent.
Technical team judgement after engagement	It is plausible that patients alive at 3 years can be considered 'cured'.

Issue 4 – Gilteritinib effectiveness after HSCT

Questions for engagement	7. Is it more appropriate to use external data or ADMIRAL trial data to estimate the relative effectiveness of gilteritinib after HSCT? a. Which of the extrapolated survival models (see figure 1) appears to be more clinically plausible?
Background/description of issue	In **the company's** model, post-HSCT OS is based on a Gompertz curve fitted to data from Evers *et al.* The company excluded ADMIRAL data for this group of patients from the company submission and the model because there is limited follow up (median follow up post-HSCT [redacted] months) and a small sample size. **The ERG** considers that ADMIRAL trial data is the most relevant data source. The ERG highlights that patients in Evers *et al.* were in second complete remission and did not all have FLT3+ mutations. The company's model suggests that around [redacted]% of gilteritinib-treated patients who undergo HSCT will be cured. At the final data cut-off of ADMIRAL, [redacted]% of all patients treated with gilteritinib and who received HSCT had died. In order to meet the 3-year cure rate from the company's model, the majority of surviving (censored) patients in the ADMIRAL gilteritinib-treated HSCT group would need to be considered 'cured'. The ERG did an analysis using ADMIRAL data to inform OS for people who have HSCT, which it included in its base case. A lognormal parametric curve was fitted to ADMIRAL data (both treatments pooled) until the 3-year cure point. The data in the ERG analysis fit the observed data more closely (see figure 1). **The technical team** notes that the impact of the 3-year cure point (issue 3) is reduced when using the ADMIRAL data to inform post-HSCT OS because fewer patients are estimated to survive to this timepoint. Additionally, issue 5 (maintenance treatment benefit) is no longer relevant if the ADMIRAL trial data is used to inform post-HSCT OS.

	Figure 1: OS, model-predicted versus observed, overall ADMIRAL population – ERG's base case model versus company's model 
Why this issue is important	Post-HSCT OS is a key driver of the cost-effectiveness results. Using ADMIRAL data to inform post-HSCT OS increased the company ICER (with ERG corrections) from **£54,844** to **£95,642** (**note**: the gilteritinib maintenance HR is not included in this analysis- see issue 5)
Technical team preliminary judgement and rationale	The primary trial data from ADMIRAL should be used to inform post-HSCT OS, as it has been for other parts of the model, such as for the patients who did not have HSCT.

Summary of comments	Comments from the company: • A similar approach using overall survival data from an external database was used in NICE's appraisal of olaparib for maintenance treatment of BRCA mutation-positive advanced ovarian, fallopian tube or peritoneal cancer after response to first-line platinum-based chemotherapy (TA598). • Presented post-HSCT overall survival rates in both datasets at months 1-12. Considers that the figures are comparable. • In ADMIRAL, [redacted] patients had follow-up data beyond year 1, and [redacted] patients had follow-up data beyond year 2. • Company also considered a study by Ustun et al, which is in a population with FLT3 mutation positive AML and provided a scenario analysis using this data to inform post-HSCT survival. • Company considers extrapolated data from Evers et al. 2018 to be the most appropriate data source for post-HSCT survival. Comments from clinical experts • The trial data is the best available • Both sources of data should be considered as both have limitations Professional group comments • ADMIRAL data appear robust ERG • If external information is considered relevant to include in the model, it should supplement evidence from ADMIRAL, not replace it. • With the updated 2 year cure point, the model suggests that around 60% of people in the gilteritinib group who receive HSCT will survive up to the cure point o At the final data cut-off in ADMIRAL, [redacted] of [redacted] ([redacted]) people in the gilteritinib group who received HSCT had already died
Technical team judgement after engagement	The technical team notes regarding the external data referred to in TA598 that: • Overall survival data from the main trial was not mature

	• The committee considered the population in the external study to have similar characteristics as the main trial • The external data was not used in the modelling but only used to validate the model • The intervention was not recommended for routine commissioning due to uncertainty in the long-term overall survival data The technical team considers that the primary trial data from ADMIRAL should be used to inform post-HSCT OS. It is appropriate to consider external data to validate the model.

Issue 5 - Gilteritinib maintenance therapy

Questions for engagement	8. Would gilteritinib be used as maintenance therapy after HSCT in clinical practice? 9. Is it plausible that there is an *additional* effect of maintenance therapy on OS? 10. Is the current method of deriving and applying an additional benefit of maintenance therapy appropriate?
Background/description of issue	**The company** applied a hazard ratio (HR) to the post-HSCT OS Gompertz model to reflect additional survival benefit associated with gilteritinib maintenance therapy post-HSCT ([redacted]% of patients are assumed to have maintenance therapy). The company state that the post-HSCT OS data are immature and so derived the HR from an indirect comparison using data from Evers et al. The company acknowledge that the KM from HSCT to death does not show a favourable effect of gilteritinib post-HSCT but note that there are small patient numbers, high levels of censoring and believe that if the patients with salvage chemotherapy were followed up for longer, a steeper curve would be seen which would separate from the gilteritinib curve (see figure 2). The **clinical experts** confirmed that gilteritinib would be used as maintenance therapy after HSCT in clinical practice. As previously mentioned (issue 4), Evers *et al.* included patients in second complete remission and not necessarily with FLT3+ mutation. **The ERG** had several concerns with this approach: • The available evidence suggests the proportional hazard assumption is violated (so applying a HR is inappropriate) • There was no adjustment for differences in patient characteristics between studies • It is inconsistent to use a subset of ADMIRAL post-HSCT OS data to estimate an additional treatment effect given that the company point out the data are immature (which is why external data were used – see issue 4) • It is unclear if the data from ADMIRAL show conclusive evidence of a difference in treatment effect in post-HSCT OS. The ERG did an analysis using a HR of 1 to indicate no additional benefit of maintenance therapy, which it included in its base case.

	Figure 2: Time from HSCT to death in patients who received HSCT in ADMIRAL  **Note:** this point is relevant only if the Evers *et al.* data is used for post-HSCT gilteritinib effectiveness (issue 4)
Why this issue is important	Using a HR of 1 for gilteritinib maintenance therapy increased the company ICER (with ERG corrections) from **£54,844** to **£70,515**
Technical team preliminary judgement and rationale	It is unclear whether there is an additional benefit from maintenance therapy. Nevertheless, there are issues with how this treatment effect was applied. Given that the technical team's preferred assumption is to use gilteritinib trial data post-HSCT (issue 4), this point does not apply.
Summary of comments	Comments from the company:

	• The summary of product characteristics permits maintenance treatment "30 days after HSCT if engraftment was successful, the patient did not have grade ≥2 acute graft versus host disease and was in CRc", which the company suggests is likely to be a small population • In the ERG base case using a hazard ratio of 1, company considers that the costs of gilteritinib should also be removed and has provided such a scenario analysis Comments from clinical experts: • Gilteritinib would be used as maintenance therapy after HSCT in clinical practice • It is plausible that there is an additional effect of maintenance therapy on overall survival – some evidence showing improved leukaemia-free survival and overall survival with sorafenib (a FLT3 inhibitor) maintenance after HSCT in first line therapy • The current method is appropriate • It is unclear whether the current method is appropriate Professional group comments: • Gilteritinib would be used as maintenance post-HSCT if approved • There are no clear randomised data to support using FLT3 inhibitors in this setting – there is an ongoing study, the MORPHO study • Most clinicians believe that ongoing FLT3 inhibition is likely to suppress low level FLT3+ AML subclones that are likely to drive relapse • It is plausible that there is an additional effect of maintenance therapy on overall survival ERG • If the company expects use of maintenance therapy to be lower in clinical practice then in ADMIRAL and if maintenance therapy is associated with additional OS gain, this will affect the ICER • Although the company did not use the data from ADMIRAL to estimate OS outcomes for patients receiving HSCT, the data is still used in the company's method to estimate the additional OS treatment effect associated with gilteritinib maintenance therapy
Technical team judgement after engagement	The technical team's preferred assumption is to use ADMIRAL trial data after HSCT, which already includes the effect of gilteritinib in maintenance therapy. Therefore this point would not be relevant.

Issue 6 - Utilities

Questions for engagement	11. Are utility values from Janssen *et al.* or Ara and Brazier *et al.* more clinically plausible after the 3-year cure point? See figure 3.
Background/description of issue	**The company model** applies an age adjustment to health state utilities. After the 3-year timepoint, health utilities for all states are based on age-adjusted values estimated using Janssen *et al.* At clarification, the company explained that the impact on the results was limited but did not explain why Janssen was used over a more granular source (e.g. Ara and Brazier). **The ERG** preferred to use estimates of general population utility by Ara and Brazier. The ERG did an exploratory analysis in which general population utilities after month 36 from Janssen *et al* were replaced with model-based estimates derived from Ara and Brazier. The graph shows the company's utility values used in the model (red line) and the ERG-preferred smoothed values. **Figure 3: Company's modelled utility profile versus general population EQ-5D-3L estimates**

	[redacted] In addition, **the ERG** noted that progressive AML is double-counted because it is counted in both the health state and as an adverse event. The ERG removed the double counting in an exploratory analysis.
Why this issue is important	Applying the utilities from Ara and Brazier decreased the company ICER (with ERG corrections) from **£54,844** to **£54,532** Removing the double counting of progression decreased the company ICER (with ERG corrections) from **£54,844** to **£54,760**.

Technical team preliminary judgement and rationale	The technical team consider that the utility values from Ara and Brazier are more plausible and that progression should not be double-counted. The technical team notes that the impact of both utility changes do not have a major impact in the ICER.
Summary of comments	Comments from the company: • Agrees that utility values from Ara and Brazier are plausible and included these values in its revised base case. Comments from clinical experts: • There is not much in it but possibly Ara and Brazier is more plausible
Technical team judgement after engagement	The utility values from Ara and Brazier are more plausible and progression should not be double-counted. The technical team notes that the impact of both utility changes do not have a large impact on the ICER.

Issue 7 - Costs

Questions for engagement	12. In NHS clinical practice in England, would gilteritinib tablets be wasted if patients stopped taking it unexpectedly, for example because of death? 13. Should all drug costs be applied as a one-off cost in the first cycle of the model? 14. Is it more plausible to assume that for patients alive after 3 years (after the assumed 'cure') a. patients who have had HSCT have 1 outpatient visit every year indefinitely, **or** have no follow-up costs? b. patients who have not had HSCT would require 1 outpatient visit every 6 months **or** have no follow-up costs? 15. Is it reasonable to remove progression costs from the model after 3 years (after the assumed 'cure')? 16. Is it reasonable to assume 3.3 or 2.0 FLT3 tests will be required to identify 1 patient (in other words, does FLT3 occur in around 30% of patients which would result in 3.3 tests per patient)?
Background/description of issue	**The company:** • did not include any wastage costs in their model • included gilteritinib and chemotherapy costs as one-off costs in the first cycle of the model

<table>
<tr><td></td><td>

assumed there would be no follow-up costs for all patients surviving beyond the 3 year ‘cure’ point, whether they had HSCT or not
applied the cost of relapse and progression to patients considered ‘cured’
assumed 2 FLT3 tests are needed to identify one patient with FLT3 positive mutation

The clinical experts said:

patients would be followed up indefinitely but that patients are only likely to have one outpatient appointment every 3-4 months after the 3 year ‘cure’ point.

The ERG did exploratory analyses which assumed that:

half a pack (14 days’ supply) of wastage was applied for all patients who die before the 3-year cure point
patients who are alive after 3 years have:

1 outpatient visit per year (for those who had HSCT) and
1 outpatient appointment every 6 months (for patients who did not have HSCT)

no cost of relapse or progression were applied after 3 years due to this being inconsistent with the assumption that patients are ‘cured’ at this point
on average, 3.3 FLT3 tests are needed to identify 1 patient with FLT3 positive mutation

The ERG also noted the unconventional approach of applying drug costs as a one-off cost in the first cycle of the model but was unable to amend this in its analysis. They noted that discounting cannot be applied properly, gilteritinib treatment duration is underestimated (because some patients were still having gilteritinib at data cut off and this is not accounted for) and treatment duration is not linked to progression. The ERG believes the joint impact of using a one-off cost for treatment and excluding wastage will result in an underestimation of gilteritinib costs but the exact impact of this on the ICER is unclear.
</td></tr>
<tr><td>Why this issue is important</td><td>
Applying the ERG’s cost assumptions increased the company ICER (with ERG corrections) from £54,844 to:
<table>
<tr><th>Exploratory analysis</th><th>ICER</th></tr>
<tr><td>Include wastage</td><td>£58,355</td></tr>
</table>
</td></tr>
</table>

	All other resource use cost amendments: **£54,999** The impact of the company including drug costs as a one-off cost could not be explored.
Technical team preliminary judgement and rationale	The technical team consider that: • it is reasonable to include wastage • drug costs should be applied in the conventional way, by applying them in each cycle, taking into account the treatment discontinuation data from ADMIRAL and discounting appropriately • it is reasonable to include follow up costs after 3 years, but the number of outpatient appointments may be lower. This has a minor impact on the ICER • it is not reasonable to include relapse and progression costs to 'cured' patients • it is reasonable to include 3.3 FLT3 tests per patient
Summary of comments	Comments from the company • Drug wastage o Highlighted NICE's appraisal of sorafenib for advanced hepatocellular carcinoma (TA474) where the committee accepted drug wastage for up to 7 days of treatment. o Presented revised base case to include wastage for 7 days of gilteritinib treatment (0.25 packs) • Application of drug costs o Considers this method has negligible effect on model results • Average patient will receive [redacted] cycles of therapy • Resource use after cure point o The updated company base case includes 1 outpatient visit every 12 months for people who have had HSCT o For people who have not had HSCT, agrees that 1 outpatient visit per 6 months is reasonable • Progression costs after 3 years o Reasonable to remove costs of progression and relapse after 3 years • FLT3 testing o Reasonable to assume 3.3 tests are required to identify 1 patient

	o Explored the impact of the testing rate in the revised company base case Comments from clinical experts: • There would be some wastage but very little in practice • Drug costs per cycle seems more appropriate • People alive after 3 years who have had HSCT would have 1 outpatient visit every 2-3 months (of the options, 1.5 visits every month is more plausible than no follow-up costs) • People alive after 3 years who have had HSCT would have 1 outpatient visit every 3-4 months • People alive after 3 years who have not had HSCT would have 1 outpatient visit every 6 months. • It is reasonable to remove progression costs from the model after 3 years • FLT3 testing is already considered standard of care so it may not need to be included in the model • It is reasonable to assume FLT3 occurs in around 30% of patients so 3.3 tests per patient identified would be needed Professional group comments: • Patients with relapsed or refractory AML have poor prognosis and may rapidly develop complications and need to stop treatment unexpectedly. It may be better to issue prescriptions initially as half courses (e.g. 14 days) to avoid wastage costs. • By 3 years, patients will remain under active haematology follow-up but will generally be seen in clinic roughly every 3-4 months. People who have ongoing complications from HSCT may be seen more frequently but unlikely to be more than monthly. • It is reasonable to remove progression costs from the model after 3 years • Most patients being considered for gilteritinib therapy will already be known to have a FLT3 mutation from screen performed at first diagnosis. FLT3 mutation status is rechecked at the time of relapsed/refractory disease but 80-90% will still have the mutation so suggest 1.1-1.2 tests to identify each patient. Comments from patient expert: • Tablets would not be wasted if patients stopped taking gilteritinib unexpectedly

	• Patients who had been cured (after 3 years) would have follow up appointments as described regardless of the intervention received that led to the cure • All patients should receive the FLT3 test to ensure they get the most appropriate and effective treatment for them, so this is a reasonable assumption ERG • Treatment costs should not be continued beyond the cure point
Technical team judgement after engagement	• It is reasonable to include wastage • The technical team would prefer drug costs to be applied in the conventional way, by applying them in each cycle, so that the treatment discontinuation data from ADMIRAL and discounting can be appropriately taken into account • Follow up costs should be included after 3 years. This has a small impact on the ICER. • Relapse and progression costs should not be included for 'cured' patients • It is appropriate to assume 3.3 FLT3 tests are required to identify one person in the model

Issue 8 – Quality of life and costs associated with administration

Background/description of issue	This issue was raised during the technical engagement stage and was not included in the draft technical report.
Why this issue is important	The clinical expert highlighted a potential benefit of gilteritinib is that it is an oral treatment that does not need to be administered in hospital, whereas salvage chemotherapy requires an inpatient stay. The ERG noted that the difference in costs between the 2 treatments was reflected in the administration costs included in the model. However, the ERG noted that the model did not assume any difference in quality of life between the 2 treatments to account for the different methods of administration.
Summary of comments	Comments from the company: • Patient reported outcomes were not able to be collected in ADMIRAL from people in the salvage chemotherapy group • Wehler et al reported a "disutility associated with other high intensity chemotherapy when minimum adverse events"

	• Company applied disutility of -0.044 to high intensity chemotherapy • For people receiving high intensity chemotherapy, company model assumes that patients were in hospital for 28 days in cycle 1, and from cycle 2 onwards the hospitalisation estimate from ADMIRAL was applied ERG • The disutility is applied in every model cycle for the whole time horizon. It is unclear if this is clinically appropriate • The updated hospital costs appear reasonable.
Technical team judgement after engagement	It is reasonable to include the disutility for chemotherapy treatment, and the updated hospital costs for administration. This has a small impact on the ICER.

4. Issues for information

Tables 1 to 3 are provided to stakeholders for information only and not included in the technical report comments table provided.

Table 1: Technical team preferred assumptions and impact on the cost-effectiveness estimate

Alteration	Technical team rationale	ICER (weighted comparator excl. BSC)	Change from company's base case
Company original base case	**–**	**£47,695**	**-**
Company ICER (with ERG corrections)	Technical team agreed with ERG's amendments. See section 5.4.2 of ERG report and table 3	**£54,844**	**+£7,149**
1. BSC relevant comparator	Issue 1 - not included in technical team's ICER because method produces results that are too uncertain	-	ICER likely to be lower
2. 3-year cure point	Issue 3 – no change to base case	-	-
3. Gilteritinib effectiveness after HSCT	Issue 4 – use ADMIRAL data to inform effectiveness	£95,642	+£47,947

Alteration	Technical team rationale	ICER (weighted comparator excl. BSC)	Change from company's base case
4. Gilteritinib maintenance therapy	Issue 5 – no change to base case, not relevant in technical team ICER due to issue 4 decision	-	-
5. Utilities	Issue 6 a. Ara and Brazier utilities **b.** Remove AE double counting of progression	 £54,532 £54,760	 +£6,837 +£7,065
6. Costs	Issue 7: a. Include wastage **b.** Follow up costs after 3 years and 3.3 FLT3 tests	 £58,355 £54,999	 +£10,660 +£7,304
Cumulative impact of the technical team's preferred assumptions on the cost-effectiveness estimate (draft technical report)		**£102,085**	+£54,390
Company revised base case post-engagement (includes an amendment to the dispensing fee of gilteritinib)	**–**	**£43,346**	**-**
Technical team's preferred ICER with amendment to dispensing fee		£103,066*	+£59,720
7. Quality of life and costs associated with administration	Issue 8 - Include disutility for high-intensity chemotherapy and revised hospital costs	£98,498	+£55,152
Cumulative impact of the technical team's preferred assumptions on the cost-effectiveness estimate		**£98,498**	+£55,152

**calculated by NICE technical team*

Table 2: Outstanding uncertainties in the evidence base

Area of uncertainty	Why this issue is important	Likely impact on the cost-effectiveness estimate
High dropout rate in salvage chemotherapy group of trial	The majority of salvage chemotherapy patients finished study treatment by cycle 2 of treatment. This may be due to the open-label design of the trial. This led to high censoring for duration of remission and LFS endpoints, [redacted]% of patients were censored early. The comparative effectiveness estimates are therefore uncertain.	Unknown

Table 3: Other issues for information

Issue	Comments
Implementation of ERG corrections	The ERG highlighted a number of errors in the company model which were corrected. The correction of errors related to (for full details see ERG report section 5.4.2): • General population mortality risk • Post-HSCT OS modelling. • A logical consistency constraint to EFS • A correction to health utilities applied after month 36 The technical team accepts these corrections (see table 2).
Cancer Drugs Fund	The company has not expressed an interest in gilteritinib being considered for funding through the Cancer Drugs Fund. No further data cuts are expected from ADMIRAL. Gilteritinib is unlikely to be a candidate for the Cancer Drugs Fund.
Innovation	The company considers the drug to be innovative. However, the technical team considers that all relevant benefits associated with the drug are adequately captured in the model.
Equality considerations	No equalities issues were identified by the company, consultees and their nominated clinical experts and patient experts.
End of life criteria	The technical team considers that the short life expectancy criterion is met, because median overall survival was 5.6 months in the salvage chemotherapy group of ADMIRAL; the clinical expert stated that median survival is around 2-3 months in this patient population and both the company's and ERG's base case showed that modelled survival in BSC is less than 2 years (ERG 0.33 years for BSC and 1.69 years for salvage chemotherapy). Based on the analyses currently available, the technical team considers that the extension to life criterion is likely to be met, because both the company's and the ERG's base case economic models showed that gilteritinib extended mean overall survival by over 3 months compared with salvage chemotherapy ERG 2.34 years compared to BSC and 0.98 for salvage chemotherapy). In addition, ADMIRAL showed a median overall survival gain of 3.7 months for gilteritinib compared with salvage chemotherapy. The technical team noted that neither of these analyses include the possible impact of the difference in prior midostaurin use between ADMIRAL and the population in clinical practice in England (issue 2). The

Issue	Comments
	company indicated in its response to technical engagement that the end of life criteria would be met whether or not prior midostaurin had been received.

Authors

Stephen O'Brien

Appraisal committee chair

Kirsty Pitt

Technical lead

Alex Filby

Technical adviser

Jasdeep Hayre

Associate director

With input from the lead team:

Rob Forsyth

Lead team member

Natalie Hallas

Lead team member

Stella O'Brien

Lead team member