NOVARTIS

Hematology Global Development Unit

PKC412/Midostaurin/Rydapt®

CPKC412E2301/ NCT03512197

A phase III, randomized, double-blind study of chemotherapy with daunorubicin or idarubicin and cytarabine for induction and intermediate dose cytarabine for consolidation plus midostaurin (PKC412) or chemotherapy plus placebo in newly diagnosed patients with acute myeloid leukemia (AML) without FLT3 mutations

Statistical Analysis Plan (SAP) for final Clinical Study Report (CSR)

Author:	[redacted]
Document type:	SAP Documentation
Document status:	Final
Release date:	11-Jan-2021
Number of pages:	40

Document History – Changes compared to previous final version of SAP

Date	Time point	Reason for update	Outcome for update	Section and title impacted (Current)
02-Sep-2019	First version	NA	NA	NA
11-Jan-2021	Second version – SAP for final CSR after study termination due to futility	Futility outcome	Adjusting analyses based on futility outcome of the first interim analysis. The censoring rules for DFS and CIR analyses have been adjusted to consider all events rather than censoring after two or more missing assessments.	All

Table of contents

List of tables

List of figures

List of abbreviations

AE	Adverse Event
AESI	Adverse Event of Special Interest
ALT	Alanine Transaminase
AML	Acute Myeloid Leukemia
AST	Apartate Transaminase
ATC	Anatomical Therapeutic Classification
BP	Blood Pressure
CHMP	Committee for Medicinal Products for Human Use
CI	Confidence Interval
CID	Cumulative Incidence of Death
CIR	Cumulative Incidence of Relapse
CR	Complete Remission
CRi	Morphologic Complete Remission with incomplete hematological recovery
CSR	Clinical Study report
CTC	Common Toxicity Criteria
CTCAE	Common Terminology Criteria for Adverse Events
CV	Coefficient of Variation
DAR	Dosage Administration Record
DI	Dose Intensity
ECG	Electrocardiogram
ECOG	Eastern Cooperative Oncology Group
EFS	Event-Free Survival
ELN	European Leukemia Network
EOT	End of Treatment
FAS	Full Analysis Set
FPFV	First Patient First Visit
eCRF	Electronic Case Report Form
IWG	International Working Group
HR	Hazard Ratio
HSCT	Hematopotic Stem Cell Transplantation
IRT	Interactive Response Technology
ITD	internal tandem duplication
LAIP	Leukemia-associated immunophenotypes
LLOQ	Lower Limit of Quantitation
LPLV	Last Patient Last Visit
MedDRA	Medical Dictionary for Regulatory Activities
mg	milligram(s)
MN	Mutation Negative
MRD	Minimal Residual Disease
NCI	National Cancer Institute
NMQ	Novartis MedDRA queries
OS	Overall Survival
PBO	Placebo
PDI	Planned Dose Intensity
PK	Pharmacokinetics
PR	Partial Remission

PRO	Patient-reported Outcomes
PT	Preferred Term
QOL	Quality of Life
QUE	Questionnaire compliance part of eCRF
RAP	Report and Analysis Process
RDI	Relative Dose Intensity
SAE	Serious Adverse Event
SAP	Statistical Analysis Plan
SD	Standard Deviation
SI	System international
SMQ	Standardized MedDRA Queries
SOC	System Organ Class
SR	Signal Ratio
TKD	Tyrosine Kinase Domain
ULN	Upper limit of normal
WBC	White Blood Cell
WHO	World Health Organization

1 Introduction

This statistical analysis plan (SAP) describes all planned analyses for the clinical study report (CSR) of study CPKC412E2301: a phase III, randomized, double-blind clinical trial of chemotherapy with daunorubicin or idarubicin and cytarabine for induction and intermediate dose cytarabine for consolidation plus midostaurin (PKC412) or placebo in newly diagnosed patients with FLT3-MN (SR<0.05) AML (FLT3 mutant to wild type signal ratio below the 0.05 clinical cut-off).

The content of this SAP is based on protocol CPKC412E2301 amendment 1. All decisions regarding final analysis, have been made prior to final database lock but after unblinding of the study data.

The first interim analysis of the primary endpoint (Event Free Survival (EFS)) for futility has been performed and the study was stopped due to lack of efficacy on 23-Sep-2019 based on DMC recommendation. Subsequently, investigators were notified on the futility outcome and study was unblinded not only for Novartis but also the investigators. While majority of patients discontinued from the study within subsequent weeks, some patients decided to continue protocol planned study treatment. This SAP now describes the analyses for the final CSR which will be generated when all patients discontinued study treatment.

The originally planned efficacy and safety analyses are reduced based on the futility outcome. However, some subgroup analyses have been added (e.g. safety and efficacy by anthracyclin for AML induction therapy (Daunorubicin vs Idarubicin)) and some analyses are only performed now for the final CSR (e.g. QoL, PK, [redacted]).

2 Study design, objectives and endpoints

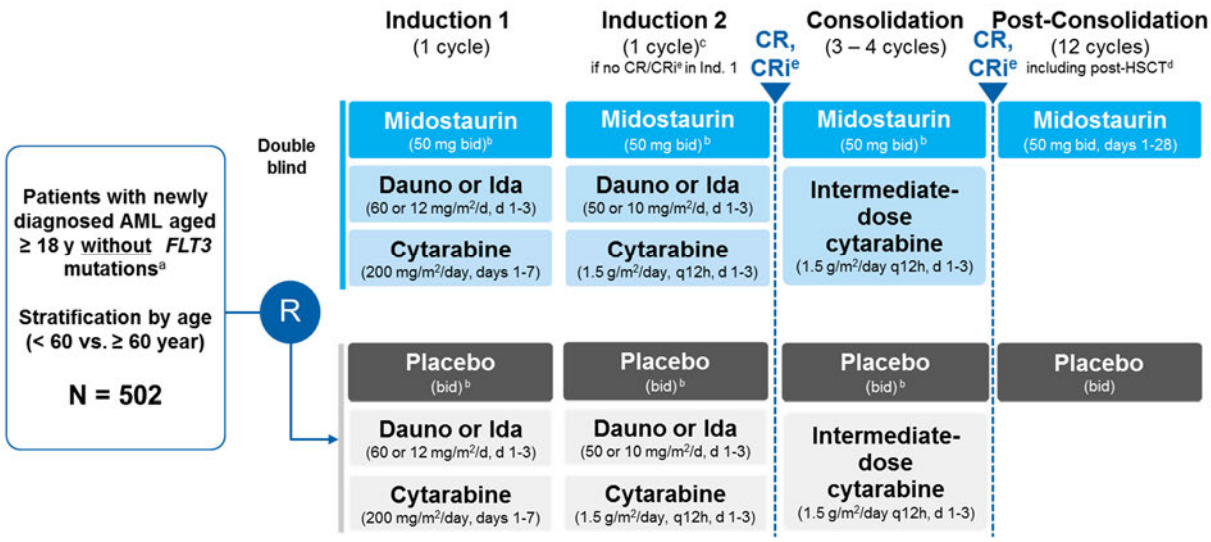
This is a randomized, double-blind, multi-center, placebo-controlled phase III study using a group sequential design with two interim analyses for the primary endpoint event-free survival (EFS) and regular safety looks. The study population comprises 502 adult patients with newly diagnosed FLT3-MN (SR<0.05) AML. Patients will be stratified according to age (<60 vs. ≥ 60 years) and will be randomized in a 1:1 ratio into one of two treatment arms:

- Midostaurin + chemotherapy

Or

- Placebo + chemotherapy

Figure 2-1 Study design



bid: twice daily; 1 cycle = 28 days, d: day(s); HSCT: Hematopoietic Stem Cell Transplantation; dauno: daunorubicin; ida: idarubicin; [a]Hydroxyurea therapy allowed for ≤ 7 days; [b]Midostaurin/Placebo treatment from day 8(Induction 1) and from day 4 (induction 2 and consolidation) until 48 hours prior to start of next treatment cycle; [c]Age ≥60: no daunorubicin / idarubicin in IND 2, cytarabine 1000mg/m^2 ; [d] Patients in CR/CRi may proceed to HSCT at any time and will enter the post-consolidation after HSCT-recovery; [e] Patients in CRi must have adequate blood count recovery to move to next treatment phase

The study contains three treatment phases:

- Induction treatment phase
- Consolidation treatment phase
- Post-consolidation phase

A patient has to meet the following criteria to continue the next treatment phase:

- A patient achieves CR or CRi **with adequate blood count recovery** defined as the following:
  - Complete Recovery in neutrophil count (≥1 x 10E9/L)
  - Minimal recovery in platelet count (≥50 x 10E9/L)

Of note, a patient not meeting the above criteria up to 93 days after start of induction phase is considered as "*induction failure*".

The **primary objective** is to determine if the addition of midostaurin to standard induction and consolidation therapy, followed by single agent post-consolidation therapy improves event-free survival (EFS) in patients with newly diagnosed FLT3-MN (SR<0.05) AML.

The **primary endpoint** is EFS and is determined according to the IWG criteria for AML [Cheson et al 2003, ELN 2017 / Döhner et al 2017] as per investigator assessment. The final analysis was planned to be when a total of 285 EFS events are observed. The first and second interim analyses were planned after observing 40% and 75% of 285 EFS events, respectively.

The **key secondary objective** is to determine if the addition of midostaurin to standard induction and consolidation therapy, followed by single agent post-consolidation therapy improves overall survival (OS) in patients with newly diagnosed FLT3-MN (SR<0.05) AML.

The **key secondary endpoint** is OS. OS was to be initially hierarchically tested if EFS is significant in any of the interim and final analyses following a separate group-sequential plan using Haybittle-Peto stopping boundaries. Since the study was futile at the first interim analysis and therefore stopped due to futility, OS was not tested but will be analyzed and reported in the final clinical study report (CSR).

Other **secondary** [redacted] **objectives** are:

Objective	Endpoint
To compare CR + CRi with adequate blood count recovery rate in the two treatment groups.	CR and CRi with adequate blood count recovery rate according to the International Working Group (IWG) for AML [Cheson et al 2003, ELN 2017 / Döhner et al 2017] as per investigator assessment.
To compare the percentage of patients who reached CR or CRi with adequate blood count recovery with MRD negative bone marrow in the two treatment groups.	Percentage of patients with CR or CRi with adequate blood count recovery with MRD negative bone marrow.
To compare the time to MRD- between the two treatment arms in the two treatment groups.	Number of days from date of randomization to first documented MRD-
To compare the percentage of patients with MRD negative status in the post-consolidation phase in the two treatment groups.	Percentage of patients with MRD negative status during post-consolidation Phase
To compare DFS, as well as the Cumulative Incidence of Relapse (CIR) and Cumulative Incidence of Death (CID) in the two treatment groups.	DFS, as measured from the date of first CR or CRi with adequate blood count recovery to relapse or death from any cause, whichever occurs first. CIR is defined for patients with CR or CRi with adequate blood count recovery: time from achieving CR or CRi with adequate blood count recovery until onset of relapse. Patients without relapse are censored at the last adequate response assessment. Patients who died without relapse are counted as a competing cause of failure. CID is defined for patients with CR or CRi with adequate blood count recovery: time from achieving CR or CRi with adequate blood count recovery until death. Patients who did not die are censored at the last contact date. Patients who relapsed are counted as a competing cause of failure.
To compare the time to CR or CRi with adequate blood count recovery in the two treatment groups.	Number of days from date of randomization to first documented CR or CRi with adequate blood count recovery.
To compare the time to neutrophil recovery in the two treatment groups.	Number of days from the first day of a chemotherapy cycle to first day neutrophils ≥0.5 x 10^9/L. Number of days from day 1 of commencing induction therapy to first day neutrophils ≥1.0 x 109/L.

Objective	Endpoint
To compare the time to platelet recovery in the two treatment groups.	Number of days from the first day of a chemotherapy cycle to first day platelets ≥50 x 109/L. Number of days from day 1 of commencing induction therapy to first day platelets ≥100 x 10^9/L.
To assess the safety and tolerability of midostaurin in combination with chemotherapy and as monotherapy during post-consolidation.	Frequency/severity of AEs, and laboratory abnormalities.
To further characterize the pharmacokinetics of midostaurin, CGP52421 and CGP62221.	Plasma concentrations and pharmacokinetic parameters for midostaurin, CGP52421 and CGP62221.
To assess the impact of midostaurin on health related quality of life and AML symptom reduction.	Change from baseline score for each time point for the FACT-Leu and the EQ5D-5L (visual analogue scale (VAS)) by treatment arm.
[redacted]	
[redacted]	[redacted]
[redacted]	
[redacted]	[redacted]
[redacted]	[redacted]
[redacted]	[redacted]
[redacted]	[redacted]

[redacted]

Data analysis general information

The statistical analysis of this study will be performed by Novartis. SAS® version 9.4 or later (SAS Institute Inc., Cary, NC, USA) will be used for all analyses. All data are part of the Clinical Database and will be exported to SAS® files for analysis. EAST version 6.4 or later (Cytel Inc.®) will be used to determine stopping boundaries in the group sequential plan.

The final EFS analysis was supposed to be performed by Novartis. However, since the study was stopped at first interim analysis only an updated EFS analysis will be provided in the final CSR. The first interim analysis of EFS was performed by an independent external statistician and an independent external programmer (CRO not involved with the conduct of the study). The study was stopped at this first interim and no further analyses are to be performed (i.e. study stopped).

The study will end when the last ongoing patient will have completed or discontinued treatment.

All events with start date before or on the cut-off date and end date after the cut-off date will be reported as 'ongoing'. The same rule will be applied to events starting before or on the cut-off date and not having documented end date. This approach applies, in particular, to adverse event and concomitant medication reports. For these events, the end date will not be imputed and therefore will not appear in the listings.

General analysis conventions

General analysis conventions

Pooling of centers: Unless specified otherwise, data from all study centers will be pooled for the analysis. Due to expected small number of patients enrolled at centers, no center effect will be assessed.

Qualitative data (e.g., gender, race) will be summarized by means of contingency tables by treatment group; a missing category will be included as applicable.

Quantitative data (e.g., age, weight) will be summarized by means of summary statistics (mean, standard deviation, median, 25^{th} – 75^{th} percentiles, minimum, maximum) by treatment group.

3 General definitions

3.1 Cut-off date used for the analysis

The data base lock date will be the cut-off date used for the final clinical study report (CSR).

3.2 Study drug and study treatment

Study drug refers to Midostaurin (PKC412) or Placebo (PBO).

Study treatment refers to

- Midostaurin + Chemotherapy (idarubicin/daunorubicin, cytarabine) or
- Placebo + Chemotherapy (idarubicin/daunorubicin, cytarabine) (control group).

Study treatment components refer to

- Midostaurin/Placebo or
- Idarubicin/Daunorubicin or
- Cytarabine

Study drug combination partner refers to Idarubicin/Daunorubicin or Cytarabine

3.3 Date of first administration of study drug

The date of first administration of study drug is derived as the first date when a non-zero dose of study drug was administered and recorded on the exposure electronic case report form (eCRF). The date of first administration of study drug will also be referred as ***start of study drug***.

3.4 Date of last administration of study drug

The date of last administration of study drug is defined as the last date when a non-zero dose of study drug was administered and recorded on the exposure electronic case report form (eCRF).

3.5 Date of first administration of study treatment

The date of first administration of study treatment is derived as the first date when a non-zero dose of any component of study treatment was administered and recorded on the exposure eCRF. For example, if the 1st dose of study drug A is administered on 04Jan2019, and the 1st dose of its combination partner, drug B, is administered on 03Jan2019, the date of the first administration of study treatment is on 03Jan2019). The date of the first administration of study treatment will also be referred as the ***start of study treatment***.

3.6 Date of last administration of study treatment

The date of last administration of study treatment is derived as the last date when a non-zero dose of any component of study treatment was administered and recorded on exposure eCRF of any study treatment component. For example, if the last dose of Midostaurin is administered on 15Apr2019, and the last dose of a combination partner, e.g., Cytarabine, is administered on 17May2019, the date of last administration of study treatment is then on 17May2019.

3.6.1 Baseline

Baseline is the result of an investigation describing the "true" uninfluenced state of the subject.

For ***efficacy evaluations***, the last assessment *before or at the date of randomization* is taken as the "baseline" value or assessment, if available.

In the context of a baseline definition, ECOG PS assessment as well as all PRO variables will follow the efficacy rule, irrespective of addressing quality of life or disease symptoms.

For PRO completion summary and PROs, baseline is defined by instrument and scale for QOL and by instrument for QUE. For QUE, the last "partially" completed (or better) answer will be used for baseline in case of more than 1 baseline assessment. In case of >1 non-completed answers (whatever reason), the last answer will be used. For QOL, the last assessment with non-missing score within eligible baseline visit per scale/category will be used as baseline for that

scale. For Global Severity questionnaire the baseline is defined by instrument; the last available assessment per baseline will be used.

For ***safety evaluations*** (e.g. laboratory and ECG), the last available assessment *before or at the date the study treatment* administration will be used as the "baseline" assessment.

If patients have no value as defined above, the baseline result will be considered to be missing.

3.6.2 On-treatment assessment/event

All on-treatment assessments/events are any assessments/events obtained in the time interval:

> [Date of first administration of study treatment; date of last administration of study treatment +30 days], i.e., inclusive lower and upper limit.

The calculation of study treatment duration may use different rules as specified in Section 6.1.

3.6.3 Study day

For the study day calculation, if the event or assessment date is after the randomization date or after the treatment start date, the following calculations listed below apply:

The study day ***for safety assessments*** (e.g. adverse event [AE] onset, laboratory abnormality occurrence, vital sign measurement, dose interruption etc.) will be calculated as:

The date of the assessment / event - start date of study treatment + 1 day.

If an AE starts before the start of study treatment the study day displayed on the listing will be negative. In this case the study day will be calculated as:

The date of onset of the event - start date of study treatment.

The study day ***for efficacy assessments*** (e.g., efficacy endpoints EFS and OS, response assessment, quality of life [QoL] etc.) will be calculated as:

The date of the assessment / event - randomization date + 1 day.

For any assessment or events that happened **prior to the start of the study drug or randomization**, e.g., time since diagnosis of disease, the study days will be calculated for safety assessments as:

Study Day = Event date - Date of first study treatment.

Note that the day of first dose of study treatment is Day 1 and the day before the date of first study treatment is Day – 1, not Day 0.

In this study, 'death' is an efficacy endpoint. However, it will also be included in safety analysis. For safety, the 'study day' will be calculated following the rules for safety assessments as described above. For efficacy 'time-to-event' variables, the study day will be derived following the rules for non-safety assessments as described above. The study day will be displayed in the data listings.

3.6.4 Screening failure

Screening failures are patients who have been screened and have failed the inclusion and exclusion criteria; these patients are never randomized. Patients who are randomized but never

receive study treatment are not considered as screening failures but will be included in the full analysis set.

3.7 Last contact date

The last contact date will be used for censoring of subjects in the analysis of overall survival.

The last contact date is defined as the latest complete date from the below list on or before the data cut-off date (Table 3-1). The cut-off date will not be used for last contact date, unless the subject was seen or contacted on that date. No date post cut-off date will be used.

Completely imputed dates (e.g. the analysis cut-off date programmatically imputed to replace the missing end date of a dose record) will not be used to derive the last contact date. Partial date imputation is allowed for event (death)/censoring only if coming from the 'Survival' eCRF.

The last contact date will be derived for subjects not known to have died at the analysis cut-off using the last complete date among the following:

Table 3-1 Last contact date data sources

Source data	Conditions
Date of Randomization	No condition
Last date subject was known to be alive from Survival Follow-up page	Subject status is reported to be alive or unknown
Start/End dates from further antineoplastic therapy	Non-missing medication/procedure term
Start/End* dates from drug administration record	Non-missing dose
Response assessment date	Response marked as 'done'
Laboratory/PK collection dates	Sample collection marked as 'done'
Vital signs date	At least one non-missing parameter value
ECOG performance status date	Non-missing performance status
Start/End dates of AE	Non-missing verbatim term

3.8 Analysis sets

3.8.1 Full Analysis Set (FAS)

The Full Analysis Set (FAS) comprises all patients to whom study drug has been assigned by randomization. According to the intent to treat principle, patients will be analyzed according to the treatment and stratum they have been assigned to during the randomization procedure.

3.8.2 Safety set (SS)

The Safety Set (SS) includes all patients who received at least one dose of study treatment starting at day 1 and being randomized with at least one dose of Midostaurin or placebo. Patients

will be analyzed according to the study treatment received, where treatment received is defined as the randomized treatment if the patient took at least one dose of that treatment. Patients who started therapy at day 1 but discontinued prior to randomization at day 8 will be listed separately.

3.8.3 Pharmacokinetic analysis set

Pharmacokinetic Analysis Set for all (PAS-all)

The Pharmacokinetic analysis set for all (PAS-all) includes all subjects in the safety set, and provide at least one evaluable PK concentration.

For a concentration to be evaluable:

- Dosing information must be properly documented (data and time of administration)
- The planned dose of midostaurin must be taken prior to sampling.
- For pre-dose samples: no vomiting within 4 hours after the midostaurin dose prior to sampling, the sample is collected before the next dose administration.
- For post-dose samples: no vomiting within 4 hours after midostaurin dosing

The PAS-all will be the primary population used for all pharmacokinetic analyses using trough concentration data.

Pharmacokinetic Analysis Set for full PK profile (PAS-full)

The Pharmacokinetic analysis set for full PK profiles (PAS-full) includes all subjects in the PAS-all, who provide an evaluable PK profile. A profile is considered evaluable if all of the following conditions are satisfied:

- Subject receives the planned first dose of midostaurin on C1D8 of Induction therapy
- Subject did not vomit within 4 hours of the first dose of midostaurin on C1D8 of Induction therapy
- Subject provides at least one primary PK parameter (Cmax, AUC0-t)

The PAS-full will be the primary population used for all pharmacokinetic analyses based on full PK profile data.

3.8.4 Stratum information used for analyses

For all stratified analyses based on the FAS, stratification information will be used as provided by the IRT.

Subgroup analyses by age will use the actual stratum information as derived from baseline characteristics which might be different from that used for randomization as provided by the IRT. In case no baseline characteristics for stratification factors are available, the information as provided by IRT will be used.

3.8.5 Subgroup of interest

The main efficacy endpoints EFS, OS, DFS, CIR, CID, CR/CRi and MRD as well as other key safety endpoints will be summarized by the following subgroups not only to examine the

treatment effect across demographic, baseline disease characteristics variables and anthracyclin for AML induction therapy but also to evaluate midostaurin in these different subgroups:

- Age (< 60 years / ≥ 60 years)
- Sex (male / female)
- European Leukemia Network (ELN) 2017 risk classification (favorable / Intermediate / Adverse).
- ITD or TKD signal ratios (<0.015, 0.015 - < 0.03, >= 0.03-0.05)
- Anthracyclin for AML induction therapy (Idarubicin vs. Daunorubicin).
- Following geographic regions as defined by the World Health Organization (WHO) (http://apps.who.int/ghodata/). Hong Kong and Taiwan are not members of the WHO and will be assigned to the region Western Pacific.
  - Region 1: Europe
  - Region 2: Western Pacific
  - Region 3: Americas

Other endpoints and subgroups may be added if deemed necessary. For each of subgroups, The HR and their 95% CIs be presented by means of forest plots for EFS, OS, DFS, CIR and CID.

4 Patient disposition, demographics and other baseline characteristics

The Full Analysis Set (FAS) will be used for all baseline and demographic summaries. Summaries will be reported by treatment group and for all patients. No inferential statistics will be provided.

4.1 Baseline demographics

All demographic data will be summarized by treatment group.

4.2 Disease characteristics of acute myeloid leukemia (AML)

Summary statistics and frequency counts with percentages will be tabulated for diagnosis and characteristics of AML. This analysis will include the following:

- Time since initial diagnosis to start of treatment.
- AML diagnosis, WHO classification 2016.
- Classification of AML
- Bone marrow blast count.
- Presence of Auer rods.
- Platelets, neutrophils, blasts and hemoglobin values in peripheral blood.
- Extramedullary disease assessment.
- European Leukemia Network (ELN) 2017 risk classification will be assigned based on cytogenetic abnormalities and mutations for these patients who are all FLT3 negative. In order to be assigned as intermediate or adverse risk group, patients must have

cytogenetic abnormalities and mutations assessed, however, they can be assigned as favorable when only the respective cytogenetic abnormality or mutation was assessed:

- **favorable risk group** includes any patient with any of the following cytogenetic abnormalities or mutations:
  - t(8;21)(q22;q22.1); *RUNX1-RUNX1T1*
  - inv(16)(p13.1q22) or t(16;16)(p13.1;q22); *CBFB-MYH11*
  - Biallelic mutated *CEBPA*
  - Mutated *NPM1*
- if patients are not in the favorable risk group as defined above, then they are assigned to **adverse risk group** if they have any of the following cytogenetic abnormalities or mutations:
  - t(6;9)(p23;q34.1);*DEK-NUP214*
  - t(v;11q23.3); *KMT2A rearranged*
  - t(9;22)(q34.1;q11.2); *BCR-ABL1*
  - inv(3)(q21.3q26.2) or t(3;3)(q21.3;q26.2); *GATA2,MECOM(EVI1)*
  - -5 or del(5q); -7 ; -17/abn(17p)
  - Complex karyotype
  - Mutated *RUNX1* (unless t(9;11) (p21.3;q23.3))
  - Mutated *ASXL1* (unless t(9;11) (p21.3;q23.3))
  - Mutated *TP53*
- if patients are neither in the favorable nor adverse risk group as defined above, they are in the **intermediate risk group** if there is presence of either of the following:
  - Wild type *NPM1*
  - t(9;11)(p21.3;q23.3); *MLLT3-KMT2A*
  - Any cytogenetic abnormalities not classified as favorable or adverse

4.3 Medical history

Medical history and ongoing conditions, including cancer-related conditions and symptoms entered on (e)CRF will be summarized by treatment group. Separate summaries will be presented for ongoing and historical medical conditions. The summaries will be presented by primary system organ class (SOC), preferred term (PT) and treatment group. Medical history and current medical conditions will be coded using the Medical Dictionary for Regulatory Activities (MedDRA) terminology. The MedDRA version used for reporting will be specified as a footnote in the applicable tables/listings.

4.4 Concomitant medications with specific impact on the analysis

Patients who take additional anti-neoplastic therapy before discontinuing study treatment (i.e., anti-neoplastic therapy other than study treatment) will be identified as protocol deviations.

4.5 Patient disposition

Enrollment by country and center will be summarized by treatment group using the FAS. The number of screened patients and the reason for screening failure will be displayed.

The following summaries will be provided:

- Number (%) of patients who were randomized (based on data from IRT system)
- Number (%) of patients who were randomized but not treated with Midostaurin/Placebo
- Primary reason for not being treated (based on 'End of Treatment Phase Completion' eCRF page)
- **Treatment phase: Induction, consolidation and post-consolidation**
  - Number (%) of patients who were treated *(*based on exposure eCRF pages of each study treatment component completed with non-zero dose administered*)*.
  - Number (%) of patients who are still on-treatment
  - Number (%) of patients who discontinued the study treatment
  - Number (%) of patients who entered consolidation, post-consolidation phase
  - Number (%) of patients who completed each treatment phase: induction, consolidation and post-consolidation treatment phases as per protocol.
  - Primary reason for study treatment discontinuation for each treatment phase (induction, consolidation, post-consolidation)
- **Post-treatment follow-up phase**
  - Number (%) of patients who have entered the post-treatment follow-up
  - Number (%) of patients who have discontinued from the post-treatment follow-up
  - Reasons for discontinuation from the post-treatment follow-up
- **Survival follow-up phase**
  - Number (%) of patients who have entered the survival follow-up
  - Number (%) of patients who have discontinued from the post-treatment follow-up
  - Reasons for discontinuation from the post-treatment follow-up

5 Protocol deviations

The number (%) of patients in the FAS with any protocol deviation will be tabulated by deviation category. All protocol deviations will be listed.

6 Treatments (study treatment, rescue medication, concomitant therapies, compliance)

6.1 Study treatment / compliance

Duration of exposure, actual cumulative dose, dose intensity (DI) and relative dose intensity (RDI) will be summarized by treatment group, separately for each component of study treatment. The duration of exposure will also be presented for the study treatment. Duration of exposure will be categorized into time intervals; frequency counts and percentages will be presented for the number of patients in each interval (0 to $\leq$ 4 weeks, 4 to $\leq$ 8 weeks, 8 to $\leq$ 12 weeks, 12 to $\leq$ 24 weeks, 24 to $\leq$ 48 weeks, $>$ 48 weeks). The number (%) of patients, who have dose reductions or interruptions, and the reasons, will be summarized by treatment group.

Patient level listings of all doses administered on treatment along with dose change reasons will be produced. The safety set will be used for all summaries and listings of study treatment.

6.2 Cycle definition

Cycle length will be calculated from the cycles recorded in the visit panel. Each cycle ends the day before the next cycle starts.

Cycle length= (date of Day 1 of the next cycle – date of Day 1 of the current cycle).

The length of last cycle (last date of last intake to the study component) – (date of Day 1 of the last cycle.

6.3 Duration of exposure to Midostaurin/placebo

Duration of exposure to a study component (days) = (last date of last intake to the study component) – (date of first administration of the study component) + 1;

6.4 Duration of exposure to cytarabine, daunorubicin or Idarubicin in induction and consolidation treatment phases

For each treatment cycle, duration of exposure to the study treatment components (days) = (last date of last intake to the study component) – (date of first administration of the study component) + 1;

Therefore, the total duration of exposure will be the sum of exposures for each treatment cycles in induction and consolidation.

6.5 Duration of exposure to study treatment

Duration of exposure to study treatment is considered by taking into account the duration of exposure to Midostaurin/Placebo, Daunorubicin (or Idarubicin) and Cytarabine:

Duration of exposure to study treatment (days) = (last date of exposure to study treatment) – (date of first administration of study treatment) + 1;

The last date of exposure to study treatment is the latest of the last dates of exposure to Midostaurin/Placebo, Daunorubicin (or Idarubicin) and Cytarabine.

6.6 Cumulative dose

Cumulative dose of a study treatment component is defined as the total dose given during the study treatment exposure and will be summarized for each of the study treatment components. The **actual cumulative dose** refers to the total actual dose administered, over the duration for which the patient is on the study treatment as documented in the Dose Administration eCRF. Cumulative dose for Midostaurin/Placebo is presented in mg. Daunorubicin or (Idarubicin) and Cytarabine is subcutaneously administered and dose administered is collected in mg/m2.

For patients who did not take any drug the actual cumulative dose is by definition equal to zero.

The **planned cumulative dose** for a study treatment component refers to the total planned dose as per the protocol in the duration of exposure to study treatment.

6.7 Planned dose intensity (PDI), Dose intensity (DI) and relative dose intensity (RDI) by cycle

6.7.1 Planned cycles lengths

A cycle length is planned to be 28 days for all patients who reached CR, CRi with adequate blood count recovery, PR or No response at end of each cycle. If the patient has CRi without adequate blood count recovery, the planned cycle length might be prolonged to:

- Up 42 days for induction cycle 1
- Up 49 days for induction cycle 2
- Up 49 days for all consolidation and post-consolidation treatment cycles.

6.7.2 Dose Intensity (DI) by cycle analysis

6.7.2.1 Midostaurin/Placebo

For cycles except last cycle or last cycle with length as scheduled or longer

- DI _c [mg/day] = (Total actual dose in an actual cycle [mg]) / actual cycle length [days]

For all cycles

- RDI_c [%] = DI_c [mg/day] / Planned Dose Intensity [mg/day] * 100

6.7.2.2 Cytarabine, daunorubicin or Idarubicin

The same rules will be applicable as above with the unit of mg/m2.

6.7.3 Dose Intensity (DI) over study period

6.7.3.1 Midostaurin / Placebo

Dose intensity (DI) for patients with non-zero duration of exposure is defined as follows:

$$\text{DI (mg/day)} = \frac{cumulative\ dose}{\sum_{i=1}^{k}(actual\ cycle\ legnths)_i}$$

and i =1,2,3,…, k are the indices for the cycle.

For patients who did not take any drug the DI is by definition equal to zero.

Relative dose intensity (RDI) is defined as follows:

RDI [%] = DI [mg/day] / Planned Dose Intensity [mg/day] * 100

6.7.3.2 Cytarabine, daunorubicin or Idarubicin

The same rules will be applicable as above with the unit of mg/m2.

Frequencies for RDI will be produced by treatment group, study treatment component and treatment phase for the following categories:

- 0 to < 50%
- 50 to < 70%

- 70 to < 90%
- 90 to < 110%
- ≥ 110%

6.8 Dose reductions, interruptions or permanent discontinuations

'Dose changed', 'Dose interrupted', and 'Dose permanently discontinued' fields from the exposure eCRF pages will be used to determine the dose reductions, dose interruptions, and permanent discontinuations, respectively.

The corresponding fields 'Reason for dose change/dose interrupted' and 'Reason for permanent discontinuation' will be used to summarize the reasons.

A dose change is either 'change in prescribed dose level' or 'dosing error' where actual dose administered is different from the prescribed dose.

For the purpose of summarizing interruptions and reasons, in case multiple entries for interruption that are entered on consecutive days with different reasons will be counted as separate interruptions. However, if the reason is the same in this block of entries, then it will be counted as one interruption.

Reduction: A dose reduction is, where the prescribed dose level is lower than the previous prescribed dose level or where the actual dose administered is lower than the calculated dose amount based on the prescribed dose. Therefore any dose change to correct a dosing error will not be considered a dose reduction. Only dose change is collected in the CRF, number of reductions will be derived programmatically based on the change and the direction of the change. In addition, if a patient moves to a less frequent regimen, such changes will then be counted as dose reductions.

7 Analysis of the primary objective

The primary objective of this study was to determine if the addition of midostaurin to standard induction and consolidation therapy, followed by single agent post-consolidation therapy improves event-free survival (EFS) in patients with newly diagnosed FLT3-WT AML.

As the study concluded at the first efficacy interim analysis due to futility, the endpoints will be derived and summarized based on the final data, but no testing (p-values) will be provided in the final CSR.

7.1 Primary endpoint - EFS

The primary endpoint is EFS as per the International Working Group (IWG) for AML [Cheson et al., 2003, ELN 2017] criteria per investigator assessment. EFS is defined as the time from the date of randomization to failure to obtain a CR or CRi with adequate blood count recovery in induction (i.e. induction failure), relapse from CR or CRi with adequate blood count recovery, or death due to any cause, whichever occurs first.

7.2 Statistical hypothesis, model, and method of analysis

The study is designed to test the following statistical hypothesis for EFS using a stratified log-rank test (stratified according to randomization stratification factor of age (<60 vs. ≥ 60 years)) at the one-sided 2.5% level of significance:

$$H01: \theta 1 \geq 1 \text{ vs. } Ha1: \theta 1 < 1$$

Where θ1 is the hazard ratio (Midostaurin treatment arm vs. placebo arm) of EFS.

In accordance with the stratified randomization scheme used in the study, a stratified two-sided 3-look group sequential log-rank test with a cumulative type I error rate of α=0.025, one-sided, will be used to test the null hypothesis (H_{01}). The stratification factor is:

- Age : <60 vs. ≥ 60 years

For EFS, the hazard ratio for treatment effect will be estimated and its two-sided 95% confidence intervals will be reported. The estimation will be based on a cox proportional hazards model with treatment and the stratification factor in it.

In addition, survivorship functions will be estimated by using the Kaplan-Meier product-limit method and displayed as graphs. The estimations are performed by treatment group. Median EFS time and its two-sided 95% confidence intervals will be reported for both treatments.

The primary analysis will be based on the FAS. Response status will be based on investigator's assessments as per the International Working Group (IWG) criteria for AML [Cheson et al 2003, ELN 2017].

7.2.1 Handling of missing values/censoring/discontinuations

In the primary analysis, a patient who had not an EFS event at the date of the analysis cut-off will have his/her EFS censored at the time of the last adequate response assessment before the cut-off date.

Induction failure is defined as failure to achieve CR or CRi with adequate blood count recovery up to 93 days after start of induction phase.

A patient who failed to achieve CR or CRi without adequate blood count recovery in induction phase based on available response assessments, would have his/her EFS event documented at randomization date.

A patient who discontinued induction treatment phase without an efficacy assessment will also be considered to have failed to achieve CR or CRi with adequate blood count recovery and would have his/her EFS event at randomization date as well.

However, a patient who discontinued due to lack of efficacy in the consolidation and post-consolidation treatment phase without any supportive response assessment will be censored at the last adequate response assessment date.

EFS events documented after the initiation of HSCT or new anti-neoplastic therapy will be considered as EFS event for the primary EFS analysis.

If an EFS event is observed after two or more missing or non-adequate response assessments, then EFS will be censored at the last adequate response assessment before the EFS event. If an

EFS event is observed after a single missing or non-adequate response assessment, the actual date of event will be used. An adequate response assessment is considered any disease assessment indicating response status apart from 'unknown' or 'not done'. The event and censoring times for EFS are summarized as follow:

Table 7-1 Primary EFS analysis (event date / outcome)

Situation		Options for event date	Outcome
A	No baseline assessment	Randomization date	Censored
B1	Relapse at / or before next scheduled assessment date or death	Date of relapse/death	Event
B2	Relapse/death after exactly one missing assessment	Date of relapse/death	Event
B3	Relapse/death after two or more missing assessments	Date of last adequate assessment	Censored
B4	Relapse/death after HSCT or anti-neoplastic therapy	Date of relapse/death	Event
C	Induction failure (incl. lack of post-baseline efficacy assessment)	Randomization date	Event
D	No relapse/ no induction failure / no death	Date of last adequate response assessment	Censored
E	Treatment discontinuation due to lack of efficacy without documented relapse in the consolidation and post-consolidation treatment phases.	Date of last adequate response assessment	Censored

Of note, a response assessment (especially relapse) can also be done based on one component of response, e.g. peripheral blood.

Missing adequate response assessment is defined as an assessment not done or classified as 'unknown'. For simplicity, the 'missing adequate response assessment' will also be referred as 'missing assessment'.

As detailed in Table 7-1, the EFS censoring and event date options depend on the presence and the number of missing assessments.

An exact rule to determine whether there are any missing assessments is needed after end of the induction phase. This rule will be based on the distance between the last adequate response assessment date and the event date.

The threshold D1 is defined as the interval between the response assessments plus the protocol-allowed window around the assessments. Similarly, the threshold D2 is defined as 2 times the protocol-specified interval between the response assessments plus the protocol allowed window around the assessments.

During consolidation and post-consolidation treatment phases:

Response assessment are performed **between day 21 and day 28** of each cycle:

- Therefore, for patients with a previous CR or CRi with adequate blood count recovery, the next assessment will be done on the next cycle between day 21 and day 28. The threshold D1 is defined as 28 days (maximum cycle length for patient with CR or CRi

with adequate blood count recovery) + 7 (in case the last assessment was done on day 21) = 35 days and the threshold D2 as 70 days.

- For patients with a previous CRi without adequate blood count recovery, the cycle will be prolonged up to 21 days. Therefore, the threshold D1 is defined as 7 (in case the last assessment was done on day 21) + 21 = 28 days and the threshold D2 is 56 days.

During follow-up phase:

- The threshold D1 is defined as 90 days and the threshold D2 as 2 * 90 days = 180 days

The analysis will assume one missing assessments if the distance between the last adequate response assessment and the event date is within (D1, D2] and 2 or more missing assessment if the distance is larger than D2.

EFS will be summarized between the 2 treatment groups within each of the following subgroups as defined in section 3.8.5.by forest plots

HRs and 2-sided 95% CIs will be provided based on the stratified Cox proportional hazards model. This model will use the BY statement to derive estimates by subgroup. Point estimates of HR and 2-sided 95% CI will be provided by subgroup in a forest plot.

8 Analysis of the key secondary objective: overall survival (OS)

The key secondary objective is to determine if the addition of midostaurin to standard induction and consolidation therapy, followed by single agent post-consolidation therapy improves overall survival (OS) in patients with newly diagnosed FLT3-MN AML.

8.1 Key secondary endpoint - overall survival (OS)

OS is defined as the time from date of randomization to date of death due to any cause. If a patient is not known to have died, survival will be censored at the date of last contact.

Table 8-1 Primary OS analysis (event date / outcome)

Situation	Options for event date	Outcome
A Death	Death date	Event
B Death after New anticancer therapy / HSCT	Death date	Event
C Withdrawal of consent	Last contact date prior to withdrawal of consent	Censored
D No death	Last contact date	Censored

8.2 Statistical hypothesis, model, and method of analysis

Assuming proportional hazards model for OS, the following statistical hypothesis for OS will be tested using a stratified log-rank test (stratified according to randomization stratification factor of age) at the one-sided 2.5% level of significance:

$$H02: \theta 2 \geq 1 \text{ vs. } Ha2: \theta 2 < 1$$

Where θ2 is the hazard ratio (Midostaurin treatment arm vs. placebo arm) of OS.

The analyses for OS will be based on the FAS population according to the treatment group patients were randomized and the strata they were assigned at randomization. For the final CSR, OS will be summarized only descriptively. No testing will be performed. For the original testing plan, please refer to study protocol. OS will be estimated using the Kaplan-Meier method. The median OS along with 95% confidence intervals will be presented by treatment group. The stratified Cox regression will be used to estimate the HR of OS, along with 95% confidence interval. All OS analyses will be based on the FAS.

OS will be summarized between the 2 treatment groups within each of the subgroups as defined in Section 3.8.5.

HRs and 2-sided 95% CIs will be provided based on the stratified Cox proportional hazards model. This model will use the BY statement to derive estimates by subgroup. Point estimates of HR and 2-sided 95% CI will be provided by subgroup in a forest plot.

8.3 Analysis of secondary endpoints

The secondary efficacy variables include CR or CRi with adequate blood count recovery rate, MRD, DFS, CIR and CID. The assessment of these endpoints, except MRD, will be based on the IWG criteria for AML [Cheson et al 2003, ELN 2017 / Döhner et al 2017] as per investigator assessment. All these endpoints will only be descriptively summarized with 95% CIs but without any p-values.

The rate of CR/CRi with adequate blood count recovery will be analyzed based on the FAS. However, DFS, Cumulative Incidence of Relapse (CIR) and Cumulative Incidence of Death (CID) will be analyzed based on data from responders (CR or CRi with adequate blood count recovery within 93 days after start of treatment) in the FAS. Assessment of relapse from CR or CRi with adequate blood count recovery, DFS, CIR and CID will not consider whether a patient received HSCT. All secondary analyses will include all data observed up-to the cut-off date, unless otherwise specified.

8.3.1 CR/CRi rate

Number and percentage of patients achieving CR or CRi with adequate blood count recovery rate will be reported by cycle of induction therapy (i.e. cycle 1 or cycle 2) and by treatment phases. Patients with CR /CRi with adequate blood count recovery after 93 days after start of chemotherapy are considered as induction failure and thus will not be reported as responders. Those patients will be listed separately.

Response assessment as per investigator evaluation will be listed for all patients by treatment group. Bone marrow, platelet, neutrophils, auer rods, extramedullary disease and transfusion assessment will be listed as well.

8.3.2 Time to CR or CRi with adequate blood count recovery

Time to CR or CRi with adequate blood count recovery is defined as the time from randomization to CR or CRi with adequate blood count recovery within 93 days after start of chemotherapy whichever occurs first. Patients without experiencing CR, CRi with adequate blood count recovery will be censored according to the following events:

- Patients experiencing induction failure will be censored at maximum follow-up (i.e. date of FPFV to date of LPLV used for the analysis).

The **date of last adequate response assessment** is the date of the last response assessment other than unknown before an event or a censoring reason occurred. In this case the last response evaluation date at that assessment will be used.

Table 8-2 Censoring rules for time to CR or CRi with adequate blood count recovery

Situation		Start-date	End-date	Outcome
A	CR or CRi with adequate blood count recovery within 93 days after start of chemotherapy	Date of randomization	Date of first onset of CR or CRi with adequate blood count recovery within 93 after start of treatment (before cut-off date)	Event
B	Induction failure	Date of randomization	At maximum follow-up (i.e. FPFV to LPLV used for the analysis) (before cut-off date)	Censor

8.3.3 Disease Free survival (DFS)

DFS is defined as the time from CR or CRi with adequate blood count recovery to relapse or death due to any cause. Patient who did not relapse nor died will be censored at the last adequate response assessment before cut-off date. DFS will be analyzed for patients who achieved CR or CRi with adequate blood count recovery within 93 days after start of chemotherapy.

Discontinuation due to lack of efficacy or failure to meet continuation criteria (collected on the 'End of treatment' and 'End of post treatment follow up' disposition pages without supporting documentation of relapse as per IWG criteria for AML [Cheson et al 2003, ELN 2017 / Döhner et al 2017] will not be considered relapse for DFS derivation.

DFS will be estimated using the Kaplan-Meier method. The median DFS along with 95% confidence intervals will be presented by treatment group. The stratified Cox regression will be used to estimate the HR of DFS, along with 95% confidence interval.

Table 8-3 Censoring rules for DFS analysis

Situation		Start date	End Date	Outcome
A	Relapse from CR or CRi with adequate blood count recovery or death due to any reason	Date of first occurrence of CR or CRi with adequate blood count recovery	Date of Relapse from CR or CRi with adequate blood count recovery or death due to any reason	Event
B	No Relapse from CR or CRi with adequate blood count recovery nor death due to any reason	Date of first occurrence of CR or CRi with adequate blood count recovery within 93 days after start of treatment	Date of last adequate assessment	Censored

Situation		Start date	End Date	Outcome
C	**New anticancer therapy given/HSCT** prior to relapse from CR or CRi with adequate blood count recovery or death	Date of first occurrence of CR or CRi with adequate blood count recovery	Date of Relapse from CR or CRi with adequate blood count recovery or death due to any reason	Event

8.3.4 Cumulative Incidence of Relapse (CIR)

Cumulative Incidence of Relapse (CIR) is defined for patients with CR or CRi with adequate blood count recovery within 93 days after start of chemotherapy and is time from achieving the CR or CRi with adequate blood count recovery until the onset of relapse from CR or CRi with adequate blood recovery. Patients without relapse are censored at the last adequate response assessment. Patients who died without relapse are counted as a competing cause of failure.

Table 8-4 Censoring rules for CIR analysis

Situation		Start date	End Date	Outcome
A	Relapse from CR or CRi with adequate blood count recovery	Date of first occurrence of CR or CRi with adequate blood count recovery within 93 days after start of treatment	Date of Relapse from CR or CRi with adequate blood count recovery	Event of interest
B	Death without relapse	Date of first occurrence of CR or CRi with adequate blood count recovery within 93 days after start of treatment	Death date	Competing event, censored
C	No Relapse from CR or CRi with adequate blood count recovery	Date of first occurrence of CR or CRi with adequate blood count recovery within 93 days after start of treatment	Date of last adequate assessment	Censored
D	Treatment discontinuation due to 'Lack of efficacy' or 'failure to meet continuation criteria' without documented relapse, i.e. clinical relapse based on investigator claim	Date of first occurrence of CR or CRi with adequate blood count recovery within 93 days after start of treatment	Date of last adequate assessment	Censored
E	**New anticancer therapy given/HSCT** prior to relapse from CR or CRi with adequate blood count recovery	Date of first occurrence of CR or CRi with adequate blood count recovery within 93 days after start of treatment	Date of Relapse from CR or CRi with adequate blood count recovery or death due to any reason	Event of interest

Cumulative Incidence of Relapse (CIR) will be estimated using the Kaplan-Meier method. The median CIR along with 95% confidence intervals will be presented by treatment group. The stratified Cox regression will be used to estimate the HR of CIR, along with 95% confidence interval.

8.3.5 Cumulative Incidence of Death (CID)

Cumulative Incidence of Death (CID) is defined for all patients achieving CR or CRi with adequate blood count recovery within 93 days after start of chemotherapy and is measured from

the date of achievement of CR or CRi with adequate blood count recovery until the date of death due to any reason. Patients not known to have died are censored on the last contact date. Patients who experienced relapse without death are counted as a competing cause of failure.

Table 8-5 Censoring rules for CID analysis

Situation		Start date	End Date	Outcome
A	Death due to any reason	Date of first occurrence of CR or CRi with adequate blood count recovery within 93 days after start of treatment	Death date	Event of interest
B	Relapse without death	Date of first occurrence of CR or CRi with adequate blood count recovery within 93 days after start of treatment	Relapse date	Competing event, censored
C	No death	Date of first occurrence of CR or CRi with adequate blood count recovery within 93 days after start of treatment	Date of last contact date	Censored
D	New anticancer therapy given/HSCT prior to deaths	Date of first occurrence of CR or CRi with adequate blood count recovery within 93 days after start of treatment	death due to any reason	Event of interest

CID will be estimated using the Kaplan-Meier method. The median CID along with 95% confidence intervals will be presented by treatment group. The stratified Cox regression will be used to estimate the HR of CID, along with 95% confidence interval.

8.3.6 Measurable Residual Disease (MRD) by flowcytometry (LAIP approach)

8.3.6.1 MRD status

MRD status will be reported for all patients as best response of MRD assessment outcome. Thus, patients who reached at least once MRD- will have their MRD status as negative. If all MRD assessment are positive or undetermined then MRD status will be reported as such.

It is acknowledged that for patients who failed induction therapy, no MRD assessment should be available. MRD status of these patients will be reported as ‘not done due to induction failure’. If patient failed induction therapy but have MRD assessment evaluated, MRD status will be reported as is in the analysis. Of note, some patients reached CR or CRi with adequate blood count recovery after 93 days after start of treatment and are considered to have failed induction therapy.

Patients completing successfully induction therapy (i.e. CR or CRi with adequate blood count recovery within 93 days after start of treatment) without MRD assessment will be reported as ‘missing’ for their status of MRD. This will be the case also for consolidation and post-consolidation until CR or CRi with adequate blood count recovery is not maintained.

The percentage of patients with MRD negative/positive/undetermined/missing (irrespective of investigator response assessment as per ELN criteria) be summarized along with exact 95% CI by treatment group and by treatment phase (i.e. Induction (I and II), consolidation and post-consolidation). If sufficient number of patients entering post consolidation, comparisons of the MRD levels between the end of the consolidation phase during the post-consolidation phase will be performed by treatment groups. Status of MRD for patients moving to HSCT and re-entering post-consolidation will be summarized prior and post HSCT.

In addition, the percentage of patients with CR and MRD-, CR with MRD+, CR and undetermined status of MRD and CR with missing MRD data will be reported by treatment groups. Similarly for CRi with adequate blood count recovery and other response categories (CRi without adequate blood count recovery, PR, etc.). Only CR and CRi with adequate blood count recovery will be counted within 93 days after start of chemotherapy will be considered.

Moreover, the kinetics of MRD levels (quantitative) will be displayed in terms of boxplot by over time for each treatment phase.

8.3.6.2 Time to MRD negative

The time to MRD negative status is defined as the time from randomization to first occurrence of MRD negativity. Patients without reaching MRD negative status level will be censored according to the following events:

- Patients experiencing induction failure will be censored at maximum follow-up (i.e. date of FPFV to date of LPLV used for the analysis).
- Patients not experiencing induction failure will be censored at their last adequate MRD assessment. If a patient does not have any adequate MRD assessment, patients will be censored at randomization date.

An adequate MRD assessment is considered any assessment indicating MRD status apart from ‘unknown’ or ‘not done’ or ‘undetermined’. ‘Undetermined’.

8.3.6.3 Duration of MRD negative

The duration of MRD- will be derived only for patients who reached MRD- and is defined from first occurrence of MRD – to MRD +, relapse or death whichever occurs first.

Patients without MRD+, relapse from CR nor death will be censored at last adequate MRD assessment.

8.3.6.4 Landmark analysis

A landmark analysis of OS at specific timepoints (e.g. at 3 months) will be performed to assess any potential impact of MRD status on OS by treatment group and overall. For that, patients who completed induction therapy will be classified as per their MRD status at end of induction: MRD+, MRD-, MRD undetermined, MRD missing.

8.3.7 Time to neutrophil/platelet recovery

The time to neutrophil will be assessed for the following criteria:

- Number of days from start of treatment to the first day neutrophils $\geq 0.5 \times 10^9/L$.

- Number of days from start of treatment to the first day neutrophils ≥1.0 x 10^9/L.

Similarly, the time to platelet transfusion independent recovery will be assessed for the following criteria:

- Number of days from start of treatment to the first day platelets ≥50 x 10^9/L. This partial recovery will be considered only if no transfusion was provided within the last 48h preceding the partial platelet recovery (i.e. ≥50 x 10^9/L).
- Number of days from start of treatment to the first day platelets ≥100 x 10^9/L. Moreover, the time to adequate blood count recovery will be assessed as:
- Number of days from start of treatment to the first day platelets ≥50 x 10^9/L and the first day neutrophils ≥1.0 x 10^9/L. Patients with platelet values >=100 x 10^9/L and neutrophils >=1.0 x 10^9/L prior to start of treatment will be excluded for this analysis.

Patient not meeting the recovery criteria mentioned above will be censored at last laboratory assessment. In case of no laboratory assessment, patient will be censored at start of Midostaurin. Patients who died, failed induction will be censored at maximum follow-up (i.e. LPLV). If patients reached recovery after HSCT, this will be taken into account and not censored at time of HSCT.

The analyses above will be repeated for induction and consolidation treatment phases.

Time to platelet and to neutrophil recovery will be plotted by treatment group along with their corresponding number of events, HR and 95% confidence intervals.

8.3.8 Transfusions

All transfusions will be listed and summarized using the FAS. The number of transfusion per patients will be summarized by treatment phase and for each cycle.

8.3.9 Hematopotic stem cell transplantation (HSCT)

Patient moving to HSCT will be reported by treatment phase. HSCT in CR or CRi with adequate blood count recovery and relapse/induction failure will be distinguished in the analysis. HSCT analyses will be based on the FAS.

9 Safety analyses

All safety analyses will be based on the safety set. Safety summaries (tables, figures) include data from the on-treatment period see (Section 3.5.2). Safety data will be presented in listing for patients who started treatment but were not randomized.

9.1 Deaths

Death summaries for on-treatment, and all (including both on-treatment and post-treatment) fatal cases will be produced by treatment group, SOC and PT. All deaths on safety set will be listed, post treatment deaths will be flagged. Separate listings of deaths prior to randomization will be provided for all screened subjects. Death summaries at day 30 and day 60 after start of treatment will also be summarized by treatment group, age and anthracyclin.

9.2 Adverse events (AEs)

Reporting of AEs

All AEs recorded during the study will be listed and summarized. AEs will be summarized via treatment group by presenting the number and percentage of patients having at least one AE (regardless of study drug treatment in each primary system organ class) and for each preferred term using MedDRA coding (MedDRA version 22.0 or later will be used). A patient with multiple occurrences of an AE will be counted only once in the respective AE category.

AE summaries will be presented by primary system organ class, preferred term, and maximum CTC grade (CTCAE version 5.0 will be used). A patient with multiple CTC grades for an AE category will be summarized under the maximum CTC grade recorded for the event.

AE summaries will include all treatment-emergent AEs starting on or after Study Day 1 (i.e. on or after the day of the first intake of study treatment) and starting no later than 30 days after study treatment discontinuation.

AE summaries

The following incidences of AE summaries will be produced by treatment group:

- AEs regardless of study drug relationship by primary system organ class and preferred term.
- AEs regardless of study drug relationship by primary system organ class and preferred term by age (< 60 vs. => 60 years).
- AEs regardless of study drug relationship by primary system organ class, preferred term and anthracyclin.
- AEs regardless of study drug relationship by primary system organ class, preferred term, anthracyclin and age(<60 vs. >= 60 years)
- AEs with suspected relationship to study drug by primary system organ class and preferred term.
- Deaths by primary system organ class and preferred term.
- Deaths by primary system organ class and preferred term and age (<60 vs. >= 60 years)
- Deaths by primary system organ class and preferred term with Idarubicin /Daunorubicin during Induction
- Deaths by primary system organ class and preferred term and age with Idarubicin /Daunorubicin during Induction (<60 vs. >= 60 years)Deaths by primary system organ class and preferred term after Hematopoietic stem cell transplantation (HSCT)
- SAEs regardless of study drug relationship by primary system organ class and preferred term
- SAEs with suspected relationship to study drug by primary system organ class and preferred term
- Non-SAEs by primary system organ class and preferred term
- SAEs with fatal outcome by primary system organ class and preferred term

- AEs leading to study drug discontinuation, regardless of study drug relationship by primary system organ class and preferred term
- AEs requiring dose adjustment or temporarily study-drug interruption regardless of study drug relationship by primary system organ class and preferred term
- AEs requiring additional therapy regardless of study drug relationship by primary system organ class and preferred term.
- AEs requiring hospitalization regardless of study drug relationship by primary system organ class and preferred term.

9.2.1 Adverse events of special interest (AESI)

An adverse event of special interest is a grouping of adverse events that are of scientific and medical concern specific to compound midostaurin. These groupings are defined using MedDRA terms, SMQs (standardized MedDRA queries), HGLTs (high level group terms), HLT (high level terms) and PTs (preferred terms) in the Case Retrieval Sheet. For each specified AESI, number and percentage of patients with at least one event of the AESI occurring will be summarized. A listing of all grouping levels down to the MedDRA preferred terms used to define each AESI will be generated.

9.3 Laboratory data

On analyzing laboratory assessments, data from all sources (central and local laboratories) will be combined. The summaries will include on-treatment laboratory assessments.

The following summaries will be produced for the laboratory data (by laboratory parameter):

- Shift tables using CTC grades to compare baseline to the worst post-baseline value will be produced for hematology and biochemistry laboratory parameters with CTC grades.

For laboratory parameters where CTC grades are not defined.

- Shift tables to the worst post-baseline value will be produced using the low/normal/high classifications based on laboratory reference ranges.

Liver function parameters

Liver function parameters of interest are total bilirubin (TBIL), Alanine transaminase (ALT) and enzymes aspartate transaminase (AST). The number (%) of patients with worst post-baseline values as per Novartis Liver Toxicity guidelines will be summarized:

- ALT or AST > 3xULN
- ALT or AST > 5xULN
- ALT or AST > 8xULN
- ALT or AST > 10xULN
- ALT or AST > 20xULN
- TBIL > 2xULN
- TBIL > 3xULN
- ALT or AST > 3xULN & TBIL > 2xULN

9.4 ECG and cardiac imaging data

Analyses of QT prolongation will be based on QTcF measurement. In general, QTcF would be referred to as QTc throughout the document.

Baseline definition

Baseline is defined from pre-dose ECGs performed on Cycle 1 Day 1 if available. Otherwise, the pre-dose ECG measurements taken on most recent day prior to the start of any study treatment within 14 days is considered as baseline. Unscheduled visits are included.

Population for analysis

Notable ECG assessments:

Patients will be considered evaluable (included in "Total") for outlier analysis if they are at risk at baseline and have at least one post-baseline ECG measurement. For change from baseline analyses they have to have at least one baseline ECG measurement and one post-baseline measurement.

A patient is considered at risk at baseline, if the baseline of this patient is missing or:

- ≤ 450 ms for abnormality "value of > 450 ms and ≤ 480 ms"
- ≤ 480 ms for abnormality "value of > 480 ms and ≤ 500 ms"
- ≤ 500 ms for abnormality "value of > 500 ms"

QTc analysis

QTc analysis will be mainly based on QTcF and QTcB. Notable abnormalities will be summarized for the following:

- An increase of 30 to 60 ms in QTcF (and QTcB)
- An increase of > 60 ms in QTcF (and QTcB)
- Patients with any QTcF/QTcB value of > 450 and ≤ 480 ms
- Patients with any QTcF/QTcB value of > 480 and ≤ 500 ms
- Patients with any QTcF/QTcB value of > 500 ms

9.5 Vital signs

Vital sign assessments are performed in order to characterize basic body function. The variables collected in studies: height [cm], body temperature [°C], heart rate [beats per minute (bpm)], systolic and diastolic blood pressure [mmHg], and respiratory rate [breaths per minute].

Patients with clinically notable vital sign abnormalities will be summarized and listed and assessments collected later than 30 days after study treatment discontinuation will be flagged in the listings. The criteria for clinically notable abnormalities are depicted in Table 9-1 and Table 9-2.

Table 9-1 Clinically notable elevated vital sign values

Variable	Criteria
Systolic BP	≥ 180 mmHg and an increase ≥ 20 mmHg from baseline

Variable	Criteria
Diastolic BP	≥ 105 mmHg and an increase ≥ 15 mmHg from baseline.
Body temperature	≥ 39.1°C (102.3°F)
Heart rate	≥ 120 bpm with increase from baseline of ≥ 15 bpm
Respiratory rate	≥ 30 bpm

Table 9-2 Clinically notable below normal vital sign values

Variable	Criteria
Systolic BP	≤ 90 mmHg and a decrease ≥ 20 mmHg from baseline
Diastolic BP	≤ 50 mmHg and a decrease ≥ 15 mmHg from baseline
Body temperature	≤ 35°C (95°F)
Heart rate	≤ 50 bpm with decrease from baseline of ≥15 bpm
Respiratory rate	≤ 10 breaths per minute

9.6 Pharmacokinetic endpoints

9.6.1 Exposure to the sum of active moities

The exposure variable will be derived for all PK exposure endpoints (see Table 9-3; not applicable to Tmax) to consider the contribution of the active metabolites to the total exposure of the compound.

The exposure to the sum of active moieties combines the concentration of the parent-(Midostaurin) and the two active metabolites (CGP52421 and CGP62221) scaled based on their relative potencies, and parent to metabolite molecular weight ratio for AML indication as shown below:

$$Midostaurin + CGP52421 * (0.06) * \frac{\text{Midostaurin mwt}}{\text{CGP52421mwt}} + CGP62221 * (1.40) * \frac{\text{Midostaurin mwt}}{\text{CGP62221mwt}}$$

Where mwt is the molecular weight of each analyte:

- Midostaurin =570.65
- CGP52421=587.23
- CGP62221=556.63

9.6.2 PK parameters

PK parameters for midostaurin and the active metabolites (CGP52421, CGP62221) and sum of active moieties will be determined using non-compartmental method(s) using Phoenix WinNonlin (Version 6.4 or later- Certara L.P.) for the patients who had full PK sampling on Cycle 1 Day 8 of the induction therapy.

PK parameters listed in Table 9-3 will be estimated and reported, when feasible (dependent on sampling scheme, and validity of samples).

Concentration data from patients collected after C1D8 (during induction, consolidation, post-consolidation phases at trough (predose) and C3h according to the PK sampling schedule [see

Study protocol CPKC412E2301 Table 7-8] will be analyzed where possible, and relevant NCA parameters determined (Cmin, C3h).

AUC0-t (AUC0-12h at C1D8) and Cmax are defined as primary parameters (contributing to PAS-full definition).

NCA PK parameters and all concentrations will be summarized and reported. Summary statistics will include n (number of values to be reported), arithmetic and geometric mean, median, SD, CV, geometric CV, minimum and maximum.

Table 9-3 Non-compartmental PK parameters

AUC0-t	The AUC from time zero to a measurable concentration sampling time (t) (mass x time x volume-1). Note: as the last sampling time is at 12 h, AUC0-12h will be determined after the first dose
AUClast	The AUC from time zero to the last measurable concentration sampling time after the first dose (tlast) (mass x time x volume-1)
Cmax	The maximum (peak) observed plasma, blood, serum, or other body fluid drug concentration after the first dose administration (mass x volume-1)
Cmin	Minimal observed pre-dose concentration (when feasible)
C3h	Concentration at 3 hours post-dose (when feasible)
Tmax	The time to reach maximum (peak) plasma, blood, serum, or other body fluid drug concentration after single dose administration (time)
[1] For parent drug only [2] AUC%Extrap and Rsqadj will be used in the interpretation of the primary PK parameters and therefore will be included in the listings only.	

Descriptive statistics (n, arithmetic mean, CV% mean, standard deviation (SD), median, geometric mean, CV% geo-mean, minimum and maximum) will be presented by treatment for PAS-full for all PK parameters defined in Table 9-3 except Tmax, where only n, median, minimum and maximum will be presented. All individual PK parameters will be listed using the PAS-all.

9.7 Patient-reported outcomes

The Functional Assessment of Cancer Therapy – Leukemia (FACT-Leu) and the 5-level EQ-5D version (EQ-5D-5L) will be collected and assessed.

In order to summarize patient-reported outcome (PRO) variables, assessments will be time slotted using time-windows as defined in study protocol. The baseline assessment is defined according to the rule for efficacy variables (Section 3.5.1).

The FAS will be used for all PRO summaries.

Frequency tables for compliance to completing all patient-reported questionnaires will be provided by treatment group and time point for FACT-Leu and EQ-5D-5L (Yes, fully completed/ Yes, partially completed/ No, patient missed scheduled assessment visit/ No, patient refused due to poor health/ No, patient refused (unrelated to health)/ No, study staff felt patient was too ill/ No, questionnaire not available in the appropriate language/ No, institutional error/

No, other). The number of patients with at least one post-baseline value of the respective questionnaire will be used as the denominator for frequency tables.

Descriptive statistics (mean, standard deviation, median, minimum, and maximum) will be used to summarize the scored scales at each scheduled assessment time point for the FACTLeu and EQ5D-5L (VAS) by treatment arm. Additionally, change from baseline in the scores at the time of each assessment will be summarized. Patients with an evaluable baseline score and at least one evaluable post baseline score during the treatment period will be included in the change from baseline analyses.

9.7.1 FACT-Leu

Assessment schedule

The questionnaire was to be completed at:

- Screening
- Induction phase : Cycle 1 Day 1 (C1D1), Cycle 2 Day 1 (C2D1) and end of induction phase
- Consolidation phase: prior to each consolidation cycle.
- Post-consolidation phase: Day 1 of each cycle and at end of post-consolidation phase
- Follow-up phase. Every 6 months.

Data derivation

For each subscales in Table 9-4 Scoring will be done according to the instructions of the authors (Cella 1997). Higher score values for Physical well-being (PWB), Emotional well-being (EWB) and Additional Concerns (AC) indicate greater symptomatology or problems whereas higher scores for Social/family well-being (SWB) and Functional well-being (FWB) indicate lower symptomatology or problems.

Table 9-4 Fact-Leu scales

(Sub)Scale	Item numbers	Score range
Physical well-being (PWB)	GP1 to GP7	0 – 28
Social/family well-being (SWB)	GS1 to GS7	0 – 28
Emotional well-being (EWB)	GE1 to GE6	0 – 24
Functional well-being (FWB)	GF1 to GF7	0 – 28
Additional Concerns (AC)	BRM3, P2, BRM2, ES3, LEU1, TH1, TH2, HI12, BMT6, C2, C6, An7, N3, LEU5, LEU6, BRM9 and LEU7	0 – 68

Analyses

- For each treatment group, calculated scores and changes from baseline will be summarized by time window.
- The change from baseline will be evaluated in the calculated subscale scores using a mixed models for repeated measures (MMRM) up to end of post-consolidation phase. The following factors and covariates are included in the repeated measurement model: time, treatment, treatment by time interaction, age (<60/≥60 years) and baseline score.

Estimates for differences between treatments and 95% CI will be provided by time point. The test for treatment differences based on the MMRM will be provided. The use of this test will depend on the amount and type of missing data. Potential bias might be assessed by the following sensitivity analyses which will be carried out:

- Missing data will be imputed by the worst value observed for Midostaurin arm and the best for Placebo arm (high score = high level of quality of life) at the respective time window
- Missing data will be imputed by the mean of all non-missing values observed at the respective time window.

A plot will be prepared for the change from baseline score by time point and treatment group based on the repeated measurement model.

[redacted]

9.13 Interim analysis of EFS

There were two interim analysis planned of the primary endpoint. The first one was performed for futility when approximately 40% of the total number of EFS events are observed and the second one was supposed to take place when 75% of the 285 events are documented.

9.13.1 First EFS interim analysis for futility

This analysis is expected to take place around 12 months from the date of first patient randomized in the study assuming an increasing recruitment rate to reach 30 patients / month in month 6. The primary intent of the first interim analysis is to allow the study to stop early for lack of efficacy (futility). There is no intent to carry out an analysis to declare superior efficacy at the time of the first interim analysis. At least, 283 patients (56%) are expected to be randomized at the time of the interim futility analysis, i.e., when approximately 114 EFS events have occurred. However the first interim analysis took place with 145 EFS events. (i.e. 50% of EFS events).

A user-defined gamma spending function ($\gamma = -1.2$) will be used as a beta-spending function to determine the non-binding futility boundary at the time of the 1st interim analysis. The futility boundary at the first interim is calculated as hazard ratio of 0.97. Nevertheless, the boundary was adjusted based on the actual number of EFS events (i.e. 0.91).

9.13.2 Second EFS interim analysis

Since the study was stopped at first interim analysis, no further analyses were performed.

10 Sample size calculation

The assumption of median EFS of 12.0 months for the control treatment arm for sample size calculations is based on available data for patients with FLT-MN (Bacher et al 2008). It is expected that treatment with test treatment arm will result in a 37.5% reduction in the hazard rate (corresponding to an increase in median EFS from 12.0 months to 17.8 months under the exponential model assumption).

If the true hazard ratio is 0.675, a total of 285 EFS events are required to have 90% power at an one-sided overall 2.5% level of significance to reject the null hypothesis (HR≥1) using a log-rank test and a 2-look group sequential design with Haybittle-Peto boundary to determine efficacy boundary and gamma spending function ($\gamma = -1.2$) to determine the non-binding futility boundary. Considering a recruitment period of approximately 20 months assuming an increasing recruitment rate to reach 30 patients / month in month 6, 502 patients will need to be randomized to the two treatment arms in a 1:1 ratio. Assuming about approximately 10% patients will be lost to follow-up for EFS, a total of 502 patients will need to be randomized. Given the above assumptions, it is estimated that the 285th EFS event will be observed at approximately 40 months from the date of first patient randomized in the study. The sample size calculation was conducted with software package [East 6.4].

11 Change to protocol specified analyses

No change from protocol specified analysis was made. [redacted]

12 Appendices

Please refer to original SAP.

13 Reference

[Döhner H, Estey E, Grimwade D, Amadori S, Appelbaum FR, Büchner T, Dombret H, Ebert BL, Fenaux P, Larson RA, Levine RL, Lo-Coco F, Naoe T, Niederwieser D, Ossenkoppele GJ, Sanz M, Sierra J, Tallman MS, Tien HF, Wei AH, Löwenberg B, Bloomfield CD (2017)] Diagnosis and management of AML in adults: 2017 ELN recommendations from an international expert panel.

[Allison PD (2000)] Survival analysis Using the SAS system. Cary, NC, SAS Institute Inc.

[Brookmeyer R. and Crowley J. (1982)] A Confidence Interval for the Median Survival Time, Biometrics 38, 29-41

[Collet D (1994)] Modelling survival data in medical research. London, Chapman & Hall

[Dmitrienko A, Molenberghs G, Chuang-Stein C, Offen W (2005)] Analysis of clinical trials using SAS. SAS Institute Inc., SAS press

[Farewell VT, Prentice RL (1980)] The approximation of partial likelihood with emphasis on case-control studies. Biometrika 67:273-278

[Glimm E, Maurer W, Bretz F (2010)] Hierarchical testing of multiple endpoints in groupsequential trials. Statistics in Medicine 29:219-228

[Hsieh FY (1995)] A cautionary note on the analysis of extreme data with Cox regression. The American Statistician 49, 226-228

[Lachin J (2000)] Biostatistical methods: The assessment of relative risk. New York; John Wiley & Sons, Inc.

[Lan KKG, DeMets DL (1983)] Discrete sequential boundaries for clinical trials. Biometrika 70(3):659-663

[O'Brien PC, Fleming TR (1979)] A multiple testing procedure for clinical trials. Biometrics 35:549-553

[Zanger UM, Schwab M.Cytochrome (2013) P450] enzymes in drug metabolism: regulation of gene expression, enzyme activities, and impact of genetic variation, Pharmacol Ther. 2013:138(1):103-41

[Ganoci L, Božina T, Mirošević Skvrce N et al (2017)] Genetic polymorphisms of cytochrome P450 enzymes: CYP2C9, CYP2C19, CYP2D6, CYP3A4, and CYP3A5 in the Croatian population. Drug Metab Pers Ther.32(1):11-21

[University of Washington - Drug Interaction data base program CYP3A4 Gene Polymorphism (2017)]

[University of Washington - Drug Interaction data base program CYP3A5 Gene Polymorphism (2017)]