Janssen Research & Development

Statistical Analysis Plan

A Randomized, Open-label, Safety and Efficacy Study of Ibrutinib in Pediatric and Young Adult Patients With Relapsed or Refractory Mature B-cell non-Hodgkin Lymphoma

Protocol 54179060LYM3003; Phase 3

JNJ 54179060 (Ibrutinib)

Status:	Approved
Date:	4 October 2017
Prepared by:	Janssen Research & Development, LLC
Document No.:	EDMS-ERI-144479215, 1.0

Compliance: The study described in this report was performed according to the principles of Good Clinical Practice (GCP).

TABLE OF CONTENTS

LIST OF IN-TEXT TABLES

ABBREVIATIONS

ANC	absolute neutrophil count
AUC	area under the plasma concentration-time curve
B-AL	Burkitt leukemia
B-ALL	B-cell acute lymphoblastic leukemia
BCR	B-cell receptor
β-hCG	beta-human chorionic gonadotropin
BL	Burkitt lymphoma
BLL	Burkitt-like lymphoma
BSA	body surface area
BTK	Bruton's tyrosine kinase
CIT	chemoimmunotherapy
CL/F	apparent (oral) plasma clearance
CLL	chronic lymphocytic leukemia
C_{max}	maximum observed plasma concentration
CNS	central nervous system
CR	complete response
CR_b	complete response biopsy-negative
CRF	case report form
CR_u	unconfirmed complete response
CT	computed tomography
CTCAE	Common Terminology Criteria for Adverse Events
CYP	cytochrome P450
DLBCL	diffuse large B-cell lymphoma
DMC	Data Monitoring committee
ECG	electrocardiogram
eDC	electronic data capture
EFS	event-free survival
EICNHL	European Intergroup Collaboration for Childhood non-Hodgkin Lymphoma
EMA	European Medicines Agency
GCB	germinal center B-cell like
GCP	Good Clinical Practice
HIV	human immunodeficiency virus
ICF	informed consent form
ICH	International Conference on Harmonisation
IEC	Independent Ethics Committee
ILD	interstitial lung disease
IRB	Institutional Review Board
IRC	Independent Review Committee
ITT	intent-to-treat
IWRS	interactive web response system
LVEF	left ventricular ejection fraction
MCL	mantle cell lymphoma
MRI	magnetic resonance imaging
MUGA	multiple-gated acquisition scan
NCI	National Cancer Institute
NHL	non-Hodgkin lymphoma
ORR	overall response rate
PBPK	physiologically-based pharmacokinetic
PD	progressive disease
PET	positron emission tomography
P-gp	P-glycoprotein
PQC	Product Quality Complaint
PR	partial response
RICE	rituximab, ifosfamide, carboplatin, etoposide, and dexamethasone
RVICI	rituximab, vincristine, ifosfamide, carboplatin, idarubicin, and dexamethasone
SET	Study Evaluation Team

SF	shortening fraction
SJS	Stevens-Johnson Syndrome
SLL	small lymphocytic lymphoma
SUSAR	suspected unexpected serious adverse reaction
ULN	upper limit of normal
Vd/F	apparent (oral) volume of distribution

1. INTRODUCTION

This clinical study is a part of a comprehensive ibrutinib clinical development plan to evaluate the safety and efficacy of ibrutinib for subjects with B-cell malignancies. This is a 2-part, multicenter study. A safety and pharmacokinetic run-in part (Part 1) will be conducted before starting the randomized part (Part 2) of the study. Part 2 is a randomized, open-label, Phase 3 study to compare the safety and efficacy of ibrutinib in combination with chemoimmunotherapy (CIT [RICE or RVICI]) versus CIT alone in children and young adult subjects with relapsed or refractory mature B-cell non-Hodgkin's lymphoma (NHL).

The purpose of the statistical analysis plan (SAP) is to lay out key elements including definition and statistical methods for the planned analyses for the primary, secondary and safety endpoints.

1.1. Trial Objectives

Primary Objective

- Run-in Part (Part 1): Confirm that the pharmacokinetics in pediatric subjects is consistent with that in adults.
- Randomized Part (Part 2): Assess efficacy (event-free survival [EFS]) of ibrutinib in combination with RICE or RVICI background therapy compared to RICE or RVICI background therapy alone.

Secondary Objectives

The secondary objectives for Run-in Part (Part 1) are:

- Evaluate the safety and tolerability of ibrutinib in combination with RICE or RVICI background therapy in pediatric subjects with B-cell malignancies
- Assess anti-tumor activity of ibrutinib as add-on to RICE or RVICI regimens
- Assess disease-specific biomarkers
- Assess the pharmacodynamic response
- Acceptability and palatability assessment of all ibrutinib formulations

The secondary objectives for Randomized Part (Part 2) are:

- Evaluate the safety and tolerability of ibrutinib in combination with RICE or RVICI background therapy in pediatric subjects and young adults with B-cell malignancies
- Determine the overall response rate (ORR)
- Evaluate tumor volume reduction at Day 14
- Determine the number and proportion of subjects who proceed to stem cell transplantation
- Evaluate the time to response

- Measure the duration of response
- Evaluate long-term survival (EFS at 2 and 3 years)
- Evaluate overall survival
- Assess disease-specific biomarkers
- Assess the pharmacodynamic response, if deemed appropriate based on Part 1 results
- Assess the population pharmacokinetics of ibrutinib in pediatric subjects and young adults
- Acceptability and palatability assessment of all ibrutinib formulations

1.2. Trial Design

This is a 2-part, multicenter study. A safety and pharmacokinetic run-in part (Part 1) will be conducted before starting the randomized part (Part 2) of the study. Part 2 is a randomized, open-label, Phase 3 study to compare the safety and efficacy of ibrutinib in combination with CIT (RICE or RVICI) versus CIT alone in children and young adult subjects with relapsed or refractory mature B-cell NHL. All subjects in Part 1 will receive ibrutinib in combination with CIT (investigator choice of RICE or RVICI); 6 to approximately 24 pediatric subjects (1 to <18 years) will be enrolled to allow confirmation of the dose regimen. Enrollment will begin with children in the 2 older age groups (6-11, 12-17 years) to assess pharmacokinetic and safety data before allowing enrollment of children in the youngest age group (1-5 years). The SET will meet to decide on any changes to the starting dose for each age group, and when enrollment may begin for the youngest age group. At a minimum, the first 2 subjects in each age group will enroll into Part 1 before recruitment of children in that age group will begin in Part 2. In Part 2, approximately 72 additional subjects will be randomized in a 2:1 ratio to receive ibrutinib in combination with CIT (investigator choice of RICE or RVICI) or CIT alone; at least 40 subjects are targeted to be of age 1 to <18 years and at least 10 of the 40 subjects are targeted to be age <11 years. Subjects will be stratified by histology (Burkitt lymphoma [BL]/Burkitt leukemia [B-AL] versus other) and by background therapy (RICE versus RVICI). Pharmacokinetic samples will be obtained during Part 2 of the study to characterize the pharmacokinetics in pediatric subjects.

Part 1 and Part 2 of the study will be conducted in 3 phases: a Pretreatment (Screening) Phase, a Treatment Phase, and a Posttreatment Phase. The Treatment Phase will extend from enrollment/randomization until 1 of the following: 1) completion of 3 cycles of therapy, 2) transplantation, if clinically indicated, or 3) disease progression, whichever comes first. Subjects will begin the Posttreatment Phase after completion of combination therapy. During the Posttreatment Phase, subjects on ibrutinib with a response of PR or better will continue on ibrutinib monotherapy for three 28-day cycles as described below (see Dosage and Administration). All subjects will be followed to assess disease progression as described below (see Efficacy Evaluations/Endpoints). The Posttreatment Phase will continue until death, loss to follow up, consent withdrawal, or study end, whichever occurs first. The end of study is defined as when approximately 60 EFS events have occurred in Part 2 (death, disease progression, or

lack of CR or PR after 3 cycles of treatment based on blinded independent event review), or the sponsor terminates the study, whichever comes first.

1.3. Statistical Hypotheses for Trial Objectives

The primary statistical hypothesis of this study is that the addition of ibrutinib to a salvage CIT regimen will extend EFS compared to CIT alone in pediatric and young adult subjects with relapsed or refractory mature B-cell NHL.

For the primary efficacy analysis, EFS will be compared between treatment groups using a non-stratified log-rank test. The estimate of the hazard ratio between the 2 treatment groups and its associated 90% confidence interval will be computed based on the non-stratified Cox proportional hazards model.

1.4. Sample Size Justification

Since Part 1 is to confirm that the pharmacokinetics in pediatric subjects is consistent with adults, no formal power calculations to pre-determine sample size were performed. The required number of subjects in Part 1 is based on pharmacokinetics results and clinical judgment.

In Part 2, approximately 72 subjects will be randomized. The sample size calculation for the randomized part (Part 2) is based on the assumption of 100% improvement (hazard ratio = 0.5) in median EFS in subjects receiving ibrutinib plus CIT (RICE or RVICI) compared with CIT (RICE or RVICI) (10 months versus 5 months). Utilizing a 2:1 randomization, this study will enroll approximately 72 subjects (approximately 48 subjects treated with ibrutinib [and RICE or RVICI background therapy]) during Part 2. Based on a total of 60 events for the 2 treatment groups, this study will have at least 80% power, given a 1-sided alpha of 0.05. An accrual rate of 1.44 subjects per month will result in a study duration of approximately 4.2 years.

1.5. Randomization and Blinding

Central randomization will be implemented in Part 2 of this study. Subjects will be randomly assigned to 1 of 2 treatment groups based on a computer-generated randomization schedule prepared before the study by or under the supervision of the sponsor. The randomization will be balanced by using randomly permuted blocks and will be stratified by histology and background therapy, then randomized in a 2:1 ratio to either Treatment Arm A (ibrutinib and RICE or RVICI background therapy) or Treatment Arm B (RICE or RVICI background therapy only). The interactive web response system (IWRS) will assign a unique treatment code, which will dictate the treatment assignment and matching study drug kit for the subject. The requestor must use his or her own user identification and personal identification number when contacting the IWRS, and will then provide the relevant subject details to uniquely identify the subject.

As this is an open-label study, blinding procedures are not applicable.

2. GENERAL ANALYSIS DEFINITIONS

Part 1 and Part 2 will be analyzed separately. The below specifications apply mainly to Part 2 unless otherwise noted.

2.1. Visit Windows

Visit windowing will be based on phases and cycles:

Pretreatment (Screening) Phase: The Screening Phase will occur within 14 days before administration of study drug.

Treatment Phase: The Treatment Phase will extend from enrollment/randomization until 1 of the following: 1) completion of 3 cycles of therapy, 2) transplantation, if clinically indicated, or 3) disease progression, whichever comes first.

Posttreatment Phase: Subjects will begin the Posttreatment Phase after completion of combination therapy. This is defined as starting from the last dose of study medication + 1 day until death, loss to follow up, consent withdrawal, or study end, whichever occurs first.

2.2. Pooling Algorithm for Analysis Centers

The data from all investigative centers (study sites) will be pooled for all analyses.

2.3. Analysis Sets

2.3.1. Efficacy Analysis Set

The intent-to-treat (ITT) Population will consist of all randomized subjects; subjects will be analyzed based on randomization, regardless of study drug received. The ITT population will be used for all efficacy analyses, subject disposition and biomarker analyses. The primary efficacy analysis is based on the ITT population for data collected in the Part 2. For Part 1, efficacy will be provided as a secondary endpoint.

2.3.2. Safety Analysis Set

The Safety population will consist of all subjects who received at least 1 dose of treatment. The safety population will be used for all safety analyses and subjects will be analyzed based on actual study drug received.

2.4. Definition of Subgroups

Subgroup analysis will be performed for the selected variables to assess the internal consistency of the treatment benefit and/or safety (Table 1). The subgroup variables and the cutoff values are subject to change if warranted to better represent the data.

Table 1: Subgroup Definitions

Subgroup	Definition of Subgroup	Analysis Type
Age	1-5 years, 6-11 years, 12-17 years	D, E, S
Sex	Male, Female	D, E, S
Race	White, Non-White	D, E, S
Histology/Diagnosis	Burkitt lymphoma [BL]/Burkitt leukemia [B-AL] versus DLBCL vs. other	E
Background therapy	RICE versus RVICI	E
Geographic region	United States, Europe including Eastern Europe, and Other	D, E, S
Elevated LDH at Baseline	No (<350 U/L), Yes (≥350 U/L)	E
Phosphor-BTK, as well as SYK, STAT3, caspase-3, BCL-xL, and cIAP1 expression	B-cell Proteomics	E
B-cell receptor (BCR)/CD79B, CARD11, and MYD mutations	BCR pathway mutations	E
c-MYC, immunoglobulin, and T-cell receptor gene rearrangements at baseline	Baseline genetics	E
Concomitant use of any CYP3A inhibitor	Yes, No	S

Insert table abbreviations
Cytopenia is defined as yes if platelet count ≤ 100,000/uL, Hgb ≤ 11g/dL, or ANC ≤ 1500/uL is observed.
analysis type D= demographic and baseline disease characteristics
analysis type E= efficacy (EFS, OS, ORR)
analysis type S= safety (adverse events)

2.5. Study Day and Cycle Day

Assessments will be presented chronologically by study day or cycle day as described below.

For efficacy data, the randomization date is considered as the reference date (Day 1). For safety data, date of first dose of study medication (randomization date for subjects who have been randomized but not treated) will be used as the reference date (Day 1).

Reference date (Day 1) = randomization date (for efficacy data), or first dose date of study medication (for safety data).

Study Day = assessment date – reference date + 1 for assessment performed on or after the reference date; assessment date – reference date for assessment performed before the reference date.

Cycle Day = assessment date - date of the first dose for the cycle + 1.

Cycles

In defining treatment cycles in a combination therapy setting, the entire combination must be taken into account. Therefore, the nominal cycles as defined in protocol and recorded on the Case Report Form (CRF) will be used in the statistical analyses.

Cycles will be 28 days long; however, if count recovery occurs quickly and the investigator deems a subject ready to proceed with the next cycle of therapy, cycles can be shortened to as few as 21 days. As long as there is no disease progression after Cycle 1 or 2, a second and third cycle of therapy will be given, respectively.

2.6. Baseline and Endpoint

Unless specified otherwise, the baseline value is defined as the last non-missing value collected on or before the administration of the first dose of study medication. For subjects who have been randomized but not treated with any dose, randomization date will be used as the reference date for baseline value calculation.

2.7. Imputation Rules for Missing AE Date/Time of Onset/Resolution

In general, imputation of missing dates will be made for AE onset date, AE resolution date, date of death, start and end dates of prior, concomitant and subsequent therapies, and date of initial diagnosis according to the following rules. Start date will be imputed before end date.

- If date is completely missing, no imputation will be made.
- If year is missing, no imputation will be made.
- If only year is present but month and day are missing, then June 30th will be used.
- If only day is missing but year and month are available, then the 15th of the month will be used.

However, the above imputations will be modified by the following rules:

- For initial diagnosis if such imputed date is on or after the randomization date, then randomization date - 1 will be used.
- If such imputed date for prior therapies or initial diagnosis is on or after the randomization date, then randomization date - 1 will be used. If such imputed date for subsequent therapies is before date of last dose, then date of last dose +1 will be used.
- The imputed start date for subsequent therapies will be adjusted sequentially using the following steps:
  - If the imputed start date is before the treatment discontinuation date or last dose date if no treatment discontinuation date but in the same year and month, then the treatment discontinuation date or last dose date if no treatment discontinuation date will be used.
  - If subsequent therapy end date is not missing and is before the imputed subsequent therapy start date, then the subsequent therapy end date will be used as the start date.
- If the imputed date is for a date of death and is before the last date that the subject is known to be alive, the latter date will be used.
- The imputed AE start date will be adjusted sequentially using the following steps:
  - If the imputed date is in the same year and month as but day before the first dose date, then the first dose date will be used, or if it is in the same year and month as but day after the last dose date + 30 days, then the last dose date + 30 days will be used.

  - If AE end date is not missing and the imputed AE start date is after the AE end date, then the AE end date will be used.
  - If the imputed AE start date and is after date of death, then date of death will be used
  - If the imputed AE start date is in the same month and year but after the 1st subsequent therapy start date, then 1st subsequent therapy start date will be used.
- If the imputed date is for an AE end date and is after the death date, then the death date will be used, or if the imputed AE end date is before the AE start date, then the AE start date will be used.
- The AE imputation rule will be used for concomitant medication.

3. INTERIM ANALYSIS AND DATA REVIEW COMMITTEE

3.1. Study Evaluation Team (SET)

A SET composed of study investigators and internal sponsor team members will convene (for all subjects in Part 1 and the first 2 subjects in each age group if not enrolled until Part 2) to evaluate the safety of ibrutinib in combination with CIT to determine if the dose is appropriate and when enrollment may start in the youngest age group. Safety review by the SET will include non-hematologic toxicities that required withholding study treatment for more than 10 days or cardiac rhythm abnormalities in those subjects receiving RVICI.

3.2. Data Monitoring Committee (DMC)

An independent DMC will be established to monitor data on an ongoing basis to ensure the continuing safety of the subjects enrolled in this study and review the interim analysis results. The DMC will consist of at least 1 medical expert in the relevant therapeutic area and at least 1 statistician. The DMC responsibilities, authorities, and procedures will be documented in its charter.

The committee will meet periodically to review interim data. After the review, the DMC will make recommendations regarding the continuation of the study. The details will be provided in a separate DMC charter. At the interim analysis, the DMC may recommend stopping the study for either futility or efficacy, if the pre-specified stopping boundary is crossed.

3.3. Interim Analysis

During Part 2, an external DMC will be used for the review of safety and will conduct an interim analysis. The pre-planned interim analysis will be conducted when approximately 30 EFS events are reached. An external DMC will determine the appropriateness for early stopping using the nonbinding stopping rules (Table 2) as well as other efficacy and safety endpoints. The 1-sided p-value required for early stopping for futility is ≥0.367 and efficacy is ≤0.006 if there are 30 EFS events at the interim analysis. This design employs the sequential testing approach as described by O'Brien and Fleming (1979)[1] to preserve the Type–I error rate.

Table 2: Stopping Boundaries (N=72 subjects, 2 Analyses)

Analysis Number	EFS events	Approximate Time	Number of Subjects	Cumulative Error		Stopping Boundary			
				Alpha	Beta	p-value (1-sided)		Observed Hazard Ratio	
				Efficacy	Futility	Efficacy	Futility	Efficacy	Futility
Interim	30 (50%)	32 months	46	0.006	0.07	≤0.006	≥0.367	<0.378	≥0.878
Final	60 (100%)	50 months	72	0.050	0.2	≤0.048	--------	<0.632	--------

EFS=event-free survival.

4. SUBJECT INFORMATION

All statistical analyses will be performed using statistical analysis system (SAS®). Analyses of disposition, demographic, baseline disease characteristics, and prior and concomitant therapy will be conducted on the ITT population. Analyses of treatment compliance and extent of exposure will be conducted on the safety population. No statistical testing is planned.

Unless otherwise specified, all continuous endpoints will be summarized using descriptive statistics, which will include the number of subjects with a valid measurement (n), mean, standard deviation (SD), median, minimum, and maximum. All categorical endpoints will be summarized using frequencies and percentages. Percentages will be calculated by dividing the number of subjects with the characteristic of interest by the number of subjects in the analysis population.

4.1. Demographics and Baseline Characteristics

Subject demographics and baseline disease characteristics will be summarized using descriptive statistics.

- Demographics and baseline characteristics: age, sex, race, ethnicity, geographic region, height (cm), weight (kg), systolic blood pressure/diastolic pressure (SBP/DBP (mmHg), body surface area (m^2)

- Baseline disease characteristics: time from initial diagnosis to randomization; time from progression/relapse since last line of treatment to randomization; diagnosis (relapsed/refractory BL, Burkitt-like lymphoma, B-AL, diffuse large B-cell lymphoma (DLBCL), DLBCL not otherwise specified, or other pediatric mature B-cell NHL.); Lansky-Karnofsky performance status; cytopenia; Phosphor-BTK, as well as SYK, STAT3, caspase-3, BCL-xL, and cIAP1 expression; B-cell receptor (BCR)/CD79B, CARD11, and MYD mutations; c-MYC; immunoglobulin; T-cell receptor gene rearrangements; and drug acceptability and palatability assessment (Visual analog scale)
- Hematology: hemoglobin, platelet count, white blood cell count, absolute neutrophil count, absolute lymphocyte count
- Chemistry: sodium, potassium, blood urea nitrogen, creatinine, aspartate aminotransferase, albumin, alkaline phosphatase, total bilirubin, lactic acid dehydrogenase, uric acid, alanine aminotransferase

4.2. Subject Disposition

Disposition information will be summarized for the ITT population and safety population. Subject enrollment will be summarized by region, country, and investigator. The number of subjects undergoing, discontinuing, and completing the study treatment as well as their reasons for treatment discontinuation will be summarized.

Descriptive statistics will be provided for time on study. Time on study is defined the same way as OS with reversed censoring, ie, subjects who died will be censored. Based on this definition, time on study is the same as length of follow up. The Kaplan-Meier method will be used to estimate the median time on study.

4.3. Exposure Related Definitions

Treatment duration is the interval between date of first dose and end-of-treatment date. If date of end of treatment is not available, use last available treatment date.

Number of treatment cycles = the last cycle number - the first cycle number + 1.

The individual anticancer drugs that make up study treatment should be summarized separately. For the treatment phase, dosing information includes total dose received (the sum of actual dose administered), average dose level per administration (the ratio of total dose and treatment duration in cycles), and relative dose intensity:

$$\frac{Sum\ of\ actual\ dose\ administered}{planned\ total\ cumulative\ dose\ for\ the\ regimen}$$

Use the highest level of dose reduction to summarize categories - ‘One dose reduction’ and ‘Two dose reductions’.

4.4. Extent of Exposure

Descriptive statistics (n, mean, standard deviation, median, and range) will be provided for total number of cycles, treatment duration, and dosing information for all study medications (ibrutinib, rituximab, ifosfamide, carboplatin, etoposide, vincristine, idarubicin, and dexamethasone).

The number and percentage of subjects with dose reduction and dose interruption will be summarized. In addition, subjects with dose modifications and reasons for dose modifications will be summarized.

Details of exposure related definition is specified in Section 4.3.

4.5. Protocol Deviations

Subjects with eligibility and major protocol deviations will be listed by treatment group.

Protocol deviations will be based on clinical review, but not limited to, the following aspects: (1) eligibility criteria, (2) patient safety, (3) efficacy assessment deviation, and (4) treatment compliance. Protocol deviations will be closely monitored during the execution of the study and the final set of protocol deviation criteria will be finalized before database lock.

4.6. Prior and Concomitant Medications

For summarization purposes, medications will be coded to a generic term based on the World Health Organization (WHO) dictionary. Medications administered prior to the first dose of study medication will be considered prior medications. Concomitant therapies include those taken on or after first dose date through the date of last dose of study drug. Using this definition, a medication can be classified as both prior and concomitant.

Incidence of prior and concomitant medications will be summarized by Anatomical Therapeutic Chemical (ATC) class and drug generic term. Prior anticancer therapy will be summarized by type (eg, radiotherapy, surgery, chemotherapy). Best response to last line of prior therapy will also be summarized.

The following concomitant medications of special interest will also be summarized:

- Anti-coagulation and anti-platelets.
- CYP3A inhibitors/inducers
- Growth factors/ cytokines by background therapy
- Transfusions

5. EFFICACY

Analysis of efficacy endpoints will be conducted on the ITT population. Table 3 summarizes the efficacy endpoints and analysis methods to be performed.

Table 3: Summary of Efficacy Analyses to be Performed

Endpoint	Analysis	Analysis Method	Population
Primary			
Event-free Survival (EFS) as assessed by IRC	Primary	Non-stratified log-rank test, non-stratified Cox regression model, all PD and death considered as events regardless of subsequent antilymphoma therapy or a conditioning regimen for stem cell transplantation. EFS is defined as the time interval from randomization to death, disease progression, or lack of CR or PR after 3 cycles of treatment.	ITT
	Sensitivity	1) Alternative censoring rule 1: subjects who received subsequent antilymphoma therapy or a conditioning regimen for cell transplantation are considered to have had an EFS event at the initiation of subsequent therapy 2) Alternative censoring rule 2: subjects who received subsequent antilymphoma therapy or a conditioning regimen for cell transplantation are censored at the last disease assessment showing no evidence of PD before the use of subsequent therapy 3) Alternative censoring rule 3: subjects will be censored at the last disease assessment if they progress, die or have not responded after missing ≥2 planned disease assessment visits. 4) Investigator assessed EFS by non-stratified log-rank test and non-stratified Cox regression model; concordance rate between the IRC-determined EFS and investigator-determined EFS	ITT
	Subgroup	Non-stratified Cox regression model within each subgroup	ITT
Secondary			
Long-term survival (EFS) at 2 and 3 years		IRC assessed chi-square test for 2 and 3 year EFS rate.	ITT
Overall Survival	Primary	Non-stratified log-rank test, non-stratified Cox regression model, as the time interval from enrollment (Part 1) or randomization (Part 2) to all deaths considered as events regardless of subsequent antilymphoma therapy or a conditioning regimen for stem cell transplantation.	ITT
	Sensitivity	Censored at subsequent antilymphoma therapy or a conditioning regimen for stem cell transplantation.	
Overall Response in Part 1	Subgroup	Non-stratified Cox regression model within each subgroup Listing of IRC and investigator assessments for those subjects who respond.	ITT Enrolled subjects
Overall Response Rate in Part 2	Primary	IRC assessed ORR chi-square test, Logistic regression model	ITT subjects with measurable disease at baseline
	Sensitivity	Investigator assessed ORR chi-square test, Logistic regression model	
	Subgroup	Logistic regression model	
Tumor Volume Reduction Rate		Analysis of variance (t-test) for percent decrease in the sum of the products of the lesion diameters at Day 14	ITT subjects with measurable disease at baseline
Time to Response		Non-stratified log-rank test, as the time interval from randomization to response for those subjects who respond as	ITT subjects who responded

Table 3: Summary of Efficacy Analyses to be Performed

Endpoint	Analysis	Analysis Method	Population
		assessed by IRC.	
Duration of Response		Non-stratified log-rank test, as the time interval from date of initial documentation of a response to the date of first documented evidence of progressive disease (PD) or death as assessed by IRC.	ITT subjects who responded
Rate of Subjects Who Proceed to Stem Cell Transplantation		Chi-square test for percent of subjects who proceed to stem cell transportation, Logistic regression model.	ITT
Other			
Subsequent Therapy		Descriptive summary Non-stratified log-rank test for the effect of subsequent antilymphoma therapy or a conditioning regimen for stem cell transplantation on overall survival, will be performed, as appropriate.	ITT subjects
Risk factors		Non-stratified Cox proportional hazards model using covariates to analyze the effect on EFS and survival will be performed, as appropriate.	ITT

5.1. Analysis Specifications

5.2. Level of Significance

All statistical testing will be performed using a 2-sided test at the 10% level of significance, unless otherwise noted. Confidence intervals will be presented as 2-sided 90% confidence intervals.

Testing of Secondary Endpoints

The secondary endpoints are to be tested sequentially at the nominal 0.10 significance level (2-sided). The order of these endpoints is as follows:

1. ORR
2. Rate of subjects who proceed to stem cell transplantation
3. Duration of Response
4. OS
5. Tumor volume reduction rate
6. Time to Response

A secondary hypothesis will be tested if and only if the primary hypothesis is rejected along with all the secondary hypotheses that precede it.

5.3. Data Handling Rules

Unless specified otherwise, imputation rules for missing and partial dates are described in Section 2.7.

5.3.1. General Analysis Considerations

Descriptive statistics and subject listings will be used to summarize the data. For continuous variables, the number of observations, means, standard deviations, medians, and ranges will be used. For discrete variables, frequency will be provided. For time-to-event variables, Kaplan-Meier estimates will be provided.

Comparisons between the 2 treatment arms in Part 2 will be performed as follows: for the continuous variables representing change from baseline to a particular postbaseline timepoint, analysis of variance will be used. For discrete variables, Chi-square test will be used. For time-to-event variables, non-stratified log-rank test and non-stratified Cox proportion hazard model will be used unless specified otherwise. All tests will be conducted at 2-sided test at the 10% level of significance and 90% confidence intervals will be provided, unless otherwise stated.

Unless otherwise specified, disease progression and disease response will be based on assessments from an IRC, according to principles adapted from the International Pediatric NHL Response Criteria (Sandlund 2015).[2] Disease response will be evaluated by radiological imaging, physical examination, or other procedures as necessary, including but not limited to review of hematology, clinical chemistry, and lumbar puncture results. Criteria for response categories, as well as the process and convention of the IRC are detailed in a separate IRC charter.

5.4. Primary Efficacy Endpoint

5.4.1. Event-free Survival (EFS) Based on IRC

The primary efficacy endpoint is EFS in Part 2, which is defined as the time interval from randomization to death, disease progression, or lack of CR or PR after 3 cycles of treatment, based on blinded independent event review by the IRC, whichever occurs first. The process and conventions of the IRC will be detailed in a separate IRC charter. The analysis of the primary endpoint (EFS) will be based on the ITT population.

5.4.2. Primary Analysis Methods for EFS

Subjects who continue to respond after 3 cycles of treatment, are progression free and alive will be censored at the date of last disease assessment. Subjects with no baseline or any post-baseline disease assessment will be censored at the date of randomization.

The analysis of the primary endpoint (EFS) will be based on the ITT population. For the primary efficacy analysis, EFS will be compared between treatment groups using a non-stratified log-rank test. The Kaplan-Meier method will be used to estimate the distribution functions of EFS for each treatment group. The number of events, subjects censored, the estimate of medians, and 90% confidence interval for the medians will be presented. The plot of EFS using the Kaplan-Meier method will be presented. The estimate of the hazard ratio between the 2 treatment groups and its associated 90% confidence interval will be computed based on the non-stratified Cox proportional hazards model.

5.4.3. Sensitivity Analysis of EFS

Sensitivity analysis will be performed for EFS as follows:

1. Use of subsequent antilymphoma therapy or a conditioning regimen for cell transplantation prior to documented PD or death.

   Two alternative censoring rules will be performed for subjects who start subsequent therapy prior to first documentation of disease progression or death due to any cause:

   a. Subjects will be considered to have had an EFS event at the initiation of subsequent therapy.

   b. Subjects will be censored at the last disease assessment showing no evidence of PD before the use of subsequent antineoplastic therapy.

2. Disease assessment follow up: Subjects will be censored at the last disease assessment if they progress, die or have not responded after missing ≥2 planned disease assessment visits.

3. EFS based on investigator assessments: For this sensitivity analysis, the same censoring rules and analysis methods used for primary analysis of IRC reported EFS will be used. The concordance rate between the IRC-determined EFS and investigator-determined EFS will be evaluated. The number and percentage of the following cases determined by investigator and by IRC will be cross-tabulated:

   a. PD events and censored cases

   b. CR or PR after 3 cycles of treatment

5.4.4. Long-term Survival for EFS

To evaluate long-term survival for EFS the proportion of subjects with EFS at 2 and 3 years will be analyzed using a 2-sided chi-square test at the 10% level of significance. This will be done for IRC and investigator assessed EFS.

5.4.5. Subgroup Analysis of EFS

Subgroup analysis will be performed for the selected potential prognostic variables (as listed in Section 2.4) to assess the consistency and robustness of the treatment benefit for EFS. A non-stratified log-rank test analysis method for EFS will be used for each of the subgroup analyses. Median EFS with 90% CI, and hazard ratio between the 2 treatment groups within each subgroup and their 90% CI will be calculated using non-stratified Cox regression model. Subgroup analysis will be presented graphically in a forest plot.

5.5. Secondary Endpoints

5.5.1. Overall Survival

Overall survival is defined as the duration from the date of enrollment (in Part 1) or randomization (in Part 2) to the date of the subject's death. Overall survival will be analyzed using the non-stratified log-rank test for treatment comparison in Part 2. Survival time of living subjects will be censored on the last date a subject is known to be alive or lost to follow-up.

The overall survival distribution and median overall survival with its 90% confidence interval will be estimated using the Kaplan-Meier product-limit method. The hazard ratio for ibrutinib+CIT relative to CIT alone and its associated 90% confidence interval will be calculated based on the non-stratified Cox proportional hazards model.

The same subgroup analysis used for EFS will be performed for OS if the number of events within each subgroup is sufficient.

5.5.2. Sensitivity Analysis of Overall Survival

Sensitivity analysis will be performed for OS where censoring occurs at subsequent antilymphoma therapy or conditioning regimen for stem cell transplantation.

5.5.3. Overall Response Rate

Overall response rate is defined as the proportion of subjects achieving a best overall response of either CR (including CR_b and CR_u) or PR as evaluated by IRC.

All randomized subjects with a measurable disease at baseline will be included in this analysis. Subjects with missing post-randomization data are considered non-responders.

Overall response rate will be estimated according to the crude proportion of confirmed responders (PR or better) based on the best overall response and summarized by treatment arm. Overall response rate will be compared between treatment arms using a 2-sided chi-square test at the 10% level of significance and 90% confidence intervals.

The relative risk (ibrutinib + CIT vs. CIT) will be reported along with the associated 95% confidence interval. Statistical inference will be evaluated using Chi-square statistic.

For overall response rate based on investigators' assessment, same analysis methods used for IRC reported overall response will be used.

Subgroup analysis for ORR may be performed if the number of responders within each subgroup is sufficient.

For enrolled subjects in Part 1, IRC and investigator response will be listed.

5.5.4. Tumor Volume Reduction Rate

The tumor volume reduction rate will be measured as the mean change in the sum of the products of the lesion diameters (SPD) from baseline to Day 14. The means will be compared between treatment arms using a 2-sided t-test at the 10% level of significance. The mean difference and its associated 90% confidence interval will be provided.

5.5.5. Time to Response

Time to response for subjects with either CR (including CR_b and CR_u) or PR as assessed by IRC is defined as the interval between the date of randomization and the date of initial documentation

of a response. Descriptive summary for time to response will be provided for each treatment group

Duration of Response

Duration of response for subjects achieving either CR (including CR_b and CR_u) or PR is defined as the interval between date of initial documentation of a response to the date of first documented evidence of progressive disease (PD) or death as assessed by IRC. The distribution for duration of response and median survival time with its 90% confidence interval will be estimated using the Kaplan-Meier product-limit method.

5.5.6. Subjects Who Proceed to Stem Cell Transplantation

The rate of subjects who proceed to stem cell transplantation will be estimated according to a proportion summarized by treatment arm. The rate will be compared between treatment arms using a 2-sided chi-square test at the 10% level of significance and 90% confidence intervals. Logistic regression analysis will also be performed to estimate an odds ratio and the associated 90% CI between the 2 treatment groups.

5.5.7. Other Efficacy Endpoints

Exploratory efficacy endpoints are as follows:

- Non-stratified Cox proportional hazards model using covariates to analyze the effect on EFS and survival will be performed, as appropriate.
- Non-stratified log-rank test for the effect of subsequent antilymphoma therapy or a conditioning regimen for stem cell transplantation on overall survival, will be performed as appropriate.
- The number (%) of subjects with subsequent antilymphoma therapy or a conditioning regimen for stem cell transplantation will be summarized by therapy type. A listing of subjects with subsequent therapy will also be provided.

6. SAFETY

Safety will be analyzed using the incidence and severity of AEs, laboratory tests, and electrocardiogram (ECG) measurements. Unless specified otherwise, all safety analyses will be based on the safety analysis set. Descriptive statistics will be reported for all safety data. Inferential statistics are not planned to be performed on safety data.

The baseline value for safety analysis is defined as the value collected at the time closest to and prior to the start of study medication.

Unless otherwise stated, safety data will be summarized by treatment arm as treated.

6.1. Adverse Events

Table 4 provides a summary of adverse events analyses to be performed.

Table 4: Summary of Adverse Event Analyses to be Performed

Category	Analysis	Sorted By	Cut off	Drug-Related TEAE
General	Overall summary			✔
	TEAEs	SOC+ PT+ toxicity grade; PT+ toxicity grade	10%, 5%	✔
	Serious TEAEs	SOC+ PT + toxicity grade; PT+ toxicity grade	2%	✔
	Grade 3 or worse TEAE	PT+ toxicity grade	2%	✔
	TEAEs leading to treatment discontinuation (ibrutinib+CIT/CIT)	PT + toxicity grade		✔
	TEAEs leading to death	PT + toxicity grade		
	TEAEs leading to dose modification or modification (ibrutinib+CIT/CIT)	PT + toxicity grade		
	AEs of clinical interest (Hemorrhagic events)	PT + toxicity grade		
	Other safety observations (eg, other malignancies, eye disorder)	PT + toxicity grade		
	Deaths within 30 days of last dose	Reason for death		
	Atrial Fibrillation and Atrial Flutter: overall summary			
Subgroup	Overall summary			
	TEAEs	SOC+ PT+ toxicity grade		
Exposure adjusted incidence rate	Overall summary			
	TEAEs	SOC+ PT+ toxicity grade; PT+ toxicity grade	1 per 100 patient month	
	Serious TEAEs	SOC+ PT + toxicity grade; PT+ toxicity grade	0.1 per 100 patient month	

6.1.1. All Adverse Events

The verbatim terms used in the CRF by investigators to identify AEs will be coded using the Medical Dictionary for Regulatory Activities (MedDRA), version 19.0 or higher. These coded AE terms are referred to as preferred terms (PT); classification into System Organ Class (SOC) is a result of the coding process.

Treatment-emergent AEs (TEAEs) are defined as 1) those that occur in TEAE period as defined below 2**)** present before first dose, but worsened in toxicity grade during treatment; 3) had missing start date and its end date is during the treatment; 4) was a drug-related event. Drug-related AEs are those assessed by investigator as being possible, probable, or very likely

related to study drug. To determine TEAE, partially missing AE start dates will be imputed according to the rules stated in Section 2.7.

For each TEAE, the percentage of subjects who experience at least 1 occurrence of the given event will be summarized. Tables will be sorted by frequency in incidence (the highest to lowest incidence).

For subjects using CYP3A inhibitor during treatment-emergent period, all TEAEs will be summarized by SOC, PT, maximum severity, any CYP3A inhibitor (Yes vs. No), and strong CYP3A inhibitor (Yes vs. No).

Treatment-Emergent Period

In general, the treatment-emergent period is defined as the time from first dose date through 30 days after last dose date, or day before subsequent antilymphoma therapy or a conditioning regimen for stem cell transplantation, whichever occurs first.

6.2. Adverse Events of Clinical Interest and Other Safety Observations

Hemorrhagic events will be identified by hemorrhage Standardized MedDRA Query [SMQ] excluding laboratory terms and be tabulated. Major hemorrhage is a subset of hemorrhagic events which are Grade ≥3 or serious or belong to central nervous system (CNS) hemorrhage/hematoma (preferred terms are listed in Attachment 1).

Other malignancies: are defined as new malignant tumors including solid tumors, skin malignancies and hematologic malignancies and are to be reported for the duration of study treatment and during any protocol-specified follow-up periods including post-progression follow-up for overall survival.

Adverse events of interest and other safety observations will be summarized by treatment arms.

6.3. Deaths

A summary of the number of deaths during the treatment phase and up to 30 days after last dose will be provided, along with the primary cause of death. In particular, frequencies of deaths due to study treatment-related adverse events will also be reported. A death is study medication-related death if the primary cause is a drug related AE.

6.4. Clinical Laboratory Tests

Laboratory data of hematology and serum chemistry up to 30 days after last dose or the end of treatment visit date, whichever is later, will be reported in SI units.

Laboratory results will be graded according to NCI-CTCAE version 4.03 or later. Note that toxicity grading for creatinine increase will be based on the NCI CTC v4.03 or later criteria, but limited only to the part based on the upper limit of normal (ULN), the other part, that is based on change from baseline, will not be used for toxicity grading. Generic normal ranges will be applied whenever reference ranges are not available.

The following laboratory tests will be analyzed:

- Hematology: hemoglobin, platelet count, white blood cell count, absolute neutrophil count, absolute lymphocyte count
- Chemistry: sodium, potassium, blood urea nitrogen, creatinine, aspartate aminotransferase, albumin, alkaline phosphatase, total bilirubin, lactic acid dehydrogenase, uric acid, alanine aminotransferase
- Serum or urine pregnancy testing for young women of childbearing potential only

Descriptive statistics (mean, standard deviation, median, and range) will be calculated for the raw data and for their changes from baseline at each timepoint of assessment as well as for the changes from baseline to the last value. Parameters will be summarized by toxicity grade. Change from baseline to the worst grade during the treatment will be provided as shift tables for selected parameters. In addition, treatment-emergent worsening in toxicity grade will be summarized for selected hematology and chemistry parameters.

The frequencies of abnormal treatment emergent uric acid will be summarized by treatment arm.

6.4.1. Creatinine Clearance

Creatinine clearance (CrCl) is calculated using Schwartz's formula:

$$CrCl = \frac{(k * Ht)}{Cr_{serum}}$$

Note: K =0.55 for girls and boys ≤ 10yrs
K=0.70 for boys > 10yrs

6.5. Electrocardiogram

QT prolongation and other clinically significant ECG abnormalities will be summarized and listed. Descriptive statistics will be calculated for the ECG parameters at baseline.

The ECG parameters that will be summarized are heart rate, RR interval, QT interval, and QTc. The QTc will be computed using Bazett's correction, ie, QTcB = $QT/\sqrt{RR}$, and Fridericia's correction, ie, QTcF = $QT/\sqrt[3]{RR}$.

All treatment-emergent abnormal findings will be tabulated, displaying the number of subjects with abnormal findings after dosing up to the end (Day 28) of the last cycle. An abnormal finding is considered to be treatment-emergent if it occurred during treatment and up to 30 days after the last dose.

7. PHARMACOKINETICS AND PHARMACODYNAMICS ANALYSIS

Details of analysis plan in PK/PD and results will be presented in a separate report.

8. BIOMARKER ANALYSIS

Details of analysis plan in biomarker and results will be presented in a separate report.

9. CHANGES TO PROTOCOL SPECIFIED ANALYSES

10. REFERENCES

1. O'Brien PC, Fleming TR. A multiple testing procedure for clinical trials. Biometrics. 1979;35(3):549-556.

2. Sandlund JT, Guillerman RP, Perkins SL, et al. International Pediatric Non-Hodgkin Lymphoma Response Criteria. J Clin Oncol. 2015;33:2106-2111.

ATTACHMENT 1: ADVERSE EVENTS OF SPECIAL INTEREST

Preferred terms for CNS hemorrhage are as follows:

- Acute haemorrhagic leukoencephalitis
- Basal ganglia haemorrhage
- Brain stem haemorrhage
- Brain stem haematoma
- Central nervous system haemorrhage
- Cerebellar haematoma
- Cerebellar haemorrhage
- Cerebral arteriovenous malformation haemorrhagic
- Cerebral haematoma
- Cerebral haemorrhage
- Cerebral microhaemorrhage
- Encephalitis haemorrhagic
- Epidural haemorrhage
- Extradural haematoma
- Haemorrhage intracranial
- Haemorrhagic cerebral infarction
- Haemorrhagic stroke
- Haemorrhagic transformation stroke
- Intracerebral haematoma evacuation
- Intracranial haematoma
- Intracranial tumour haemorrhage
- Intraventricular haemorrhage
- Pituitary haemorrhage
- Putamen haemorrhage
- Ruptured cerebral aneurysm
- Spinal cord haemorrhage
- Spinal epidural haematoma
- Spinal epidural haemorrhage
- Spinal subarachnoid haemorrhage
- Spinal haematoma
- Spinal subdural haematoma
- Spinal subdural haemorrhage
- Subarachnoid haemorrhage
- Subdural haematoma
- Subdural haemorrhage
- Subgaleal haematoma
- Thalamus haemorrhage