Janssen Research & Development *

Statistical Analysis Plan

A Randomized, Double-blind, Placebo-controlled Phase 3 Study of the Bruton's Tyrosine Kinase Inhibitor, PCI-32765 (Ibrutinib), in Combination with Either Bendamustine and Rituximab (BR) or Rituximab, Cyclophosphamide, Doxorubicin, Vincristine, and Prednisone (R-CHOP) in Subjects with Previously Treated Indolent Non-Hodgkin Lymphoma (iNHL)

Protocol PCI-32765FLR3001 (SELENE); Phase 3

PCI-32765 (ibrutinib)

Status:	Approved
Date:	15 June 2022
Prepared by:	Janssen Research & Development, LLC
Document No.:	EDMS-ERI-68097552, 6.0

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

AMENDMENT HISTORY

October 20, 2020
Added descriptive summary of COVID-19 pandemic impact analysis
May 5, 2022
Estimand framework was used to present primary endpoints, sensitivity, and supplementary analyses.
Updated interim analysis boundaries based on observed OS events (information).
Reference to separate biomarker analysis report was removed and details on planned biomarker analysis added.
Reference to IWRS version of disease status was removed from subgroup analysis as IWRS and eCRF version have high concordance.
Disease status categories refined and POD24 and PRIMA-PI added in subgroup analysis as well as covariate adjusted analysis.
Language on planned interim analysis was updated and DMC recommendation at the interim analysis added.
Reference to per-protocol analysis was removed
Table 3: Estimated crossing boundary in HR and p-value scales for the planned OS analysis was updated to reflect corresponding alpha spending at the interim analysis for PFS and reflected observed number of OS events at the PFS final analysis
Reference to audit plan was included
Language about AE of interest has been updated.
Other editorial changes were added for clarity.

ABBREVIATIONS

AE	adverse event
ALC	absolute lymphocyte count
ALT	alanine aminotransferase
ANC	absolute neutrophil count
aPTT	activated partial thromboplastin time
ASCT	autologous stem cell transplant
AST	aspartate aminotransferase
BR	bendamustine, rituximab
BSA	body surface area
CI	confidence interval
COVID-19	Coronavirus Disease 2019
CR(s)	complete response(s)
CRF	case report form
CSR	clinical study report
CT	computed tomography
DMC	data monitoring committee
DOR	duration of response
ECOG	Eastern Cooperative Oncology Group
EQ-5D-5L	EuroQol questionnaire
FACT-Lym	Functional Assessment of Cancer Therapy-Lymphoma
FACT-LymS	FACT-Lym Lymphoma Symptom Subscale
FL	follicular lymphoma
IF	information fraction
Ig	Immunoglobulin
HR	hazard ratio
iNHL	indolent Non-Hodgkin Lymphoma
INR	international normalized ratio
ITT	intent-to-treat
LDH	lactic acid dehydrogenase
MedDRA	Medical Dictionary for Regulatory Activities
MRD	minimal residual disease
MZL	marginal zone lymphoma
NCI-CTCAE	National Cancer Institute Common Terminology Criteria for Adverse Events
ORR	overall response rate
OS	overall survival
PD	progressive disease
PET	positron emission tomography
PFS	progression-free survival
PFS2	progression-free survival on next-line therapy
PK	Pharmacokinetics
PP	per-protocol
PR	partial response
PRO	patient-reported outcomes
R-CHOP	rituximab, cyclophosphamide, doxorubicin, vincristine, and prednisone
SAE	serious adverse event
SAP	statistical analysis plan
SD	stable disease
SUV	standardized uptake values
TTNT	time-to-next-treatment
TTW	time to worsening
ULN	upper limit of normal
WBC	white blood cell (count)

1. INTRODUCTION

The FLR3001 study is part of a comprehensive clinical development program to evaluate the safety and efficacy of ibrutinib (JNJ-54179060: PCI-32765) in subjects with B-cell malignancies. This randomized, double-blind, placebo-controlled, multicenter, Phase 3 study was designed to evaluate the efficacy and safety of ibrutinib in combination with bendamustine and rituximab (BR) or rituximab, cyclophosphamide, doxorubicin, vincristine, and prednisone (R-CHOP) in subjects with previously treated indolent Non-Hodgkin Lymphoma (iNHL).

This statistical analysis plan (SAP) contains definitions of analysis sets, derived variables, data handling rules, and statistical methods for the analysis of efficacy, safety, biomarkers, and pharmacokinetics. Details of the study conduct, and data collection can be found in the study protocol amendment INT-2 and electronic case report forms (eCRFs). Any changes to the protocol analysis plan, including additional analyses are documented here.

1.1. Trial Objectives

The primary objective of the study is to evaluate whether the addition of ibrutinib to either BR or R-CHOP will result in prolongation of progression free survival (PFS) compared with either BR or R-CHOP alone in subjects with previously treated iNHL (follicular lymphoma [FL] or marginal zone lymphoma [MZL]).

The secondary objectives are the evaluation of treatment groups in terms of the following:

- Overall survival (OS)
- Complete response (CR) rate
- Overall response rate, ORR (CR + partial response [PR])
- Duration of response (DOR)
- Patient-reported lymphoma symptoms and concerns as measured by the Functional Assessment of Cancer Therapy-Lymphoma Symptom Subscale (FACT-LymS) of the Functional Assessment of Cancer Therapy-Lymphoma (FACT-Lym)
- Safety of ibrutinib when combined with BR or R-CHOP

The exploratory objectives include the evaluation of the following:

- Time-to-next treatment (TTNT)
- Minimal residual disease (MRD) negative rate in FL subjects
- Patient-reported outcomes (PRO), related to general health status, utilizing EuroQol (EQ-5D-5L)
- Pharmacokinetics of ibrutinib and explore the potential relationships between ibrutinib metrics of exposure with relevant clinical or biomarker information

1.2. Trial Design

FLR3001 is a randomized, double-blind, placebo-controlled, multicenter, Phase 3 study to compare the efficacy and safety of ibrutinib in combination with either BR or R-CHOP to placebo plus BR or R-CHOP in subjects with previously treated iNHL (FL or MZL).

All subjects received background therapy of either 6 cycles of BR or 6 cycles of R-CHOP. Selection of background therapy was based on prior treatment history and cardiac function and was in the judgment of the investigator. Approximately 400 eligible subjects were stratified by (1) backbone chemotherapy treatment either BR or R-CHOP, (2) refractory vs relapsed disease (stable disease [SD]/progressive disease [PD] vs CR/PR as best response to last prior chemotherapy), (3) iNHL histology (FL vs MZL), and (4) number of prior lines of therapy (1 vs >1). Subjects were randomized in a 1:1 ratio to either Treatment Arm A (background therapy + placebo) or Treatment Arm B (background therapy + 560 mg of ibrutinib). Study drug was administered orally approximately at the same time each day and on a continuous schedule until disease progression, unacceptable toxicity, or study end, whichever came first.

Three clinical cut-offs were planned. The first 2 clinical cut-offs were planned to occur when approximately 151 and 252 PFS events had been observed, respectively. The interim analysis and the primary analysis of the primary endpoint PFS were to take place at these 2 clinical cutoffs, respectively. The last cutoff will occur at the time of the end of study, when approximately 50% of the randomized subjects have died or the Sponsor terminates the study, whichever comes first.

Subject participation includes a Screening Phase, a Treatment Phase, and a Posttreatment Follow-up Phase. Subject eligibility is determined up to 30 days prior to randomization. The Treatment Phase extends from randomization until study drug discontinuation. The Posttreatment Follow-up Phase begins once a subject discontinues study drug and continues until death, loss to follow up, consent withdrawal, or study end, whichever occurs first.

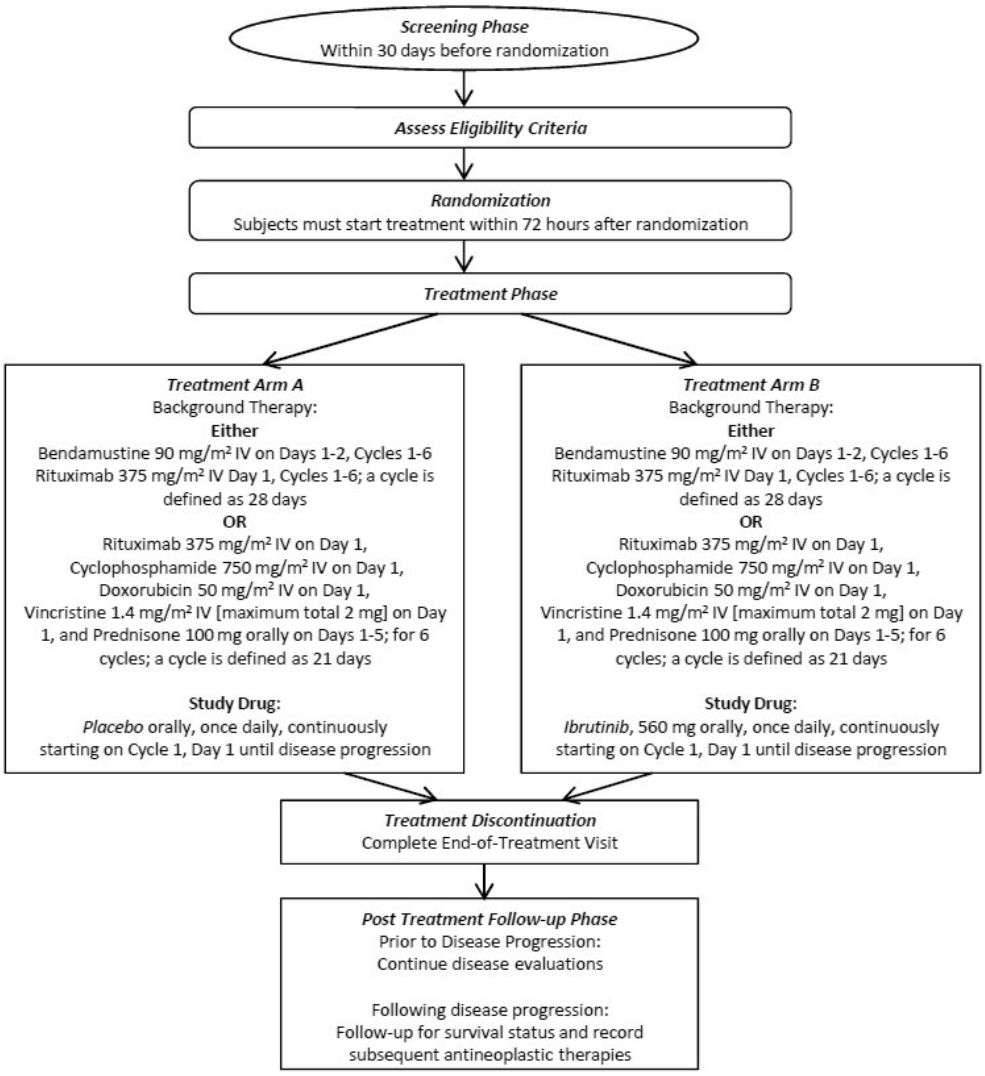
Assessment of tumor response and progression is conducted in accordance with the Revised Response Criteria for Malignant Lymphoma.[1] The investigator evaluates sites of disease by radiological imaging, physical examination, or other procedures as necessary. The primary efficacy analysis of PFS is based on investigator assessment. The assessments are performed at Screening, Week 12, Week 24, and then every 16 weeks until 3 years, and every 24 weeks thereafter until disease progression, or death, whichever comes first. Subjects who discontinue treatment for reasons other than disease progression (for reasons such as an adverse event [AE]) must continue to have regularly scheduled computed tomography (CT) scans/efficacy assessments according to the Time and Events Schedule until disease progression, death, or the clinical cutoff for the final analysis of the primary endpoint, whichever occurs first.

At each site visit, subjects are evaluated for toxicity. Safety evaluations include adverse event monitoring, physical examinations, concomitant medication usage, and clinical laboratory parameters. At some visits, blood samples are drawn for assessment of pharmacokinetic parameters and for MRD.

An independent Data Monitoring Committee (DMC) was formed and constituted according to regulatory agency guidelines. The independent DMC reviewed the unblinded safety and efficacy of the treatment combination and made recommendations as to the further conduct of the study at designated milestones. Details regarding the composition and procedures are provided in the DMC charter.

A diagram of the study design is provided below in Figure 1.

Figure 1: Schematic Overview of the Study



1.3. Statistical Hypotheses for Trial Objective

The statistical hypotheses are described as follows:

H_0: Null hypothesis, the PFS distributions of the experimental treatment group, $S_T(t)$, and the placebo group, $S_P(t)$, are the same at all time points t:

$S_T(t) = S_P(t)$, for all $t \geq 0$

versus

H_1: Alternative hypothesis, the PFS distributions of the experimental treatment group, $S_T(t)$, are stochastically different than the placebo group, $S_P(t)$:

$S_T(t) \neq S_P(t)$, for some $t \geq 0$.

1.4. Sample Size Justification

This study was designed to evaluate the effect of treatment on PFS and was powered for this endpoint. The sample size for the study was calculated based on the following considerations:

a. 1:1 randomization ratio between 2 treatment groups

b. Target hazard ratio (HR) of 0.7. Assuming the median PFS for the control arm (background therapy + placebo) is 20 months from randomization, a target HR of 0.7 corresponds to an 8.6-month increase in median PFS for the treatment group (background + ibrutinib) relative to the control (ie, 28.6 months versus 20 months, respectively)

c. Minimum 80% power

d. 2-sided overall significance level of 0.05

e. One interim analysis for efficacy at 60% of the planned total PFS events

Using the above assumptions, the study enrolled approximately 400 subjects (about 200 subjects to each arm) to observe 252 PFS events to achieve 80% power for PFS endpoints. The planned data cutoffs were based on the interim and primary analysis of PFS when approximately 151 and 252 PFS events, respectively, had occurred.

1.5. Randomization and Blinding

Central randomization was implemented in this study. Subjects were randomly assigned to 1 of the 2 treatment groups based on a computer-generated randomization schedule prepared before the study and under the supervision of the sponsor. The randomization was balanced by using randomly permuted blocks and stratified by (1) background therapy treatment (BR vs R-CHOP), (2) refractory vs relapsed disease, (3) iNHL histology (FL vs MZL), and (4) number of prior lines of therapy (1 vs >1). The sponsor strived towards an adequate number of subjects randomized into either of the background therapy regimens. Subjects were randomized in a 1:1 ratio to either

Treatment Arm A (background therapy + placebo) or Treatment Arm B (background therapy + 560 mg of ibrutinib).

In this double-blind placebo-controlled study, subjects, investigators, and the sponsor's study team members will remain blinded to treatment assignment until the database has been locked for writing the clinical study report (CSR). Personnel who may be unblinded during the study are:

- The independent DMC, the independent unblinded biostatisticians and statistical programmers from an independent Statistical Support Group who are responsible for preparing interim tables, listings, and graphs for DMC review. Unblinding procedures and the control of the unblinded data are described in the DMC charter
- Sponsor's representative responsible for pharmacokinetics testing and analysis
- Sponsor safety representative to fulfill regulatory reporting requirements for suspected unexpected serious adverse events (SAE)
- In case of an urgent safety concern, site personnel and the sponsor may be unblinded if treatment assignment information is needed to determine further actions to address the urgent safety concern (eg, life-threatening event, medication error, such as an accidental overdose)

2. GENERAL ANALYSIS DEFINITIONS

2.1. Study Treatment and Study Drug

For the purpose of analysis, the term "study treatment" refers to the background therapy (BR or R-CHOP) and ibrutinib/placebo. The term "study drug" refers to ibrutinib/placebo. Outputs will be provided by treatment group (ie, ibrutinib + background therapy, placebo + background therapy).

2.2. Baseline Definitions

Unless otherwise specified, the baseline value is defined as the last available (non-missing) value collected on or prior to the administration of the first study treatment, except for efficacy variables, for which the randomization date is used to define baseline. For subjects who have been randomized but not treated with any dose, the randomization date is used as the reference date for baseline value calculation.

2.3. Study Day and Cycle Day

For efficacy data, the randomization date is considered as the reference date (Day 1). For safety data, date of first dose of study treatment is used as the reference date (Day 1).

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

Study day is calculated as follows:

Study Day (if on or after reference date) = assessment date – reference date +1,

Study Day (if before reference date) = assessment date – reference date,

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

2.4. Pooling Algorithm for Analysis Centers

The data from all investigative sites is pooled for all analyses.

2.5. Imputation of Missing Dates

In general, imputation of missing dates is made for AE onset date, AE resolution date, date of death, start and end dates of prior, concomitant, and subsequent therapies, date of progression/relapse on the last prior therapy, date of second progressive disease, and date of initial diagnosis according to the following rules. Start date is imputed before end date.

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

The above imputations are modified by the following rules:

- For initial diagnosis if such imputed date is on or after the randomization date, then randomization date - 1 day is used.
- If such imputed date for prior therapies or initial diagnosis is on or after the randomization date, then randomization date - 1 day is used. If such imputed date for subsequent therapies is before date of last dose, then date of last dose +1 day is used.
- The imputed start date for subsequent therapies is 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 is used.
    - If subsequent therapy end date is not missing and is before the imputed subsequent therapy start date, then the subsequent therapy end date is 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 is used.

- The imputed AE start date is adjusted sequentially using the following steps:
    - If the imputed date is in the same year and month but before the first dose date, then the first dose date is used, or if it is in the same year and month but after the last dose date + 30 days, then the last dose date + 30 days is 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 is used.
    - If the imputed AE start date and is after date of death, then date of death is 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 is used.
- If the imputed date is for an AE end date and is after the death date, then the death date is used, or if the imputed AE end date is before the AE start date, then the AE start date is used.
- The AE imputation rule is used for concomitant medication.

2.6. Analysis Sets

2.6.1. Efficacy Analysis Set(s)

The Intent-to-Treat (ITT) analysis set is defined as all randomized subjects. Subjects in this analysis set are analyzed according to the treatment to which they are randomized. The ITT analysis set is used to summarize the study population and characteristics, efficacy, and PRO data.

2.6.2. Safety Analysis Set

The Safety analysis set is defined as all randomized subjects who received at least 1 dose of study drug. Safety data are analyzed according to the actual treatment received.

2.6.3. Pharmacokinetics (PK) Analysis Set

This analysis set is defined as all randomized subjects who received at least 1 dose of ibrutinib and had at least 1 pharmacokinetic sample obtained.

2.7. Definition of Subgroups

The subgroup variables and the cutoff values are summarized in Table 1.

Table 1: Subgroups

Subgroup	Definition of Group	Analysis Type
Sex	Male, Female	E, S, D
Race	White, Non-White	E, S, D
Region	Asia, Other	E, S, D
Age	<65, ≥65	E, S, D
Baseline ECOG	0, 1 or 2	E, S, D
Background therapy treatment	BR, R-CHOP	E, S, D
Number of lines of prior therapy	1, >1	E
Disease status	Refractory, relapsed	E
iNHL histology	FL, MZL	E,S,D
Baseline extra nodal site involvement	Yes, no	E
Prior autologous stem cell transplant (ASCT)	Yes, no	E
Baseline LDH	Normal, elevated	E
Baseline PET SUV	Mild to moderate SUV, high SUV	E
Tumor bulk (largest diameter)	≤6 cm, >6 cm	E
PRIMA PI[2]	Low, Medium, High risk	E
POD24[3]	>24, <= 24	E
Hepatic impairment (based on NCI criteria)	Normal vs Mild vs Moderate/ Severe liver dysfunction	S
Renal impairment	Creatinine clearance <60, ≥60	S
Concomitant use of CYP3A inhibitor	Strong/Moderate, all remaining subjects	S
Concomitant use of CYP3A inhibitor	Strong, all remaining subjects	S

BR=bendamustine, rituximab; CIT=chemoimmunotherapy; CYP=cytochrome P450; E=efficacy; S=safety (adverse events); D=demographic and baseline disease characteristics, ECOG=Eastern Cooperative Oncology Group; FL=follicular lymphoma; LDH=lactic acid dehydrogenase; PET= positron emission tomography; MZL=marginal zone lymphoma; NCI=National Cancer Institute; PRIMA PI=simplified index based on beta 2 microglobulin and bone marrow involvement for patients treated with immunochemotherapy; POD24=progression of disease (POD) within 2 years of first line of CIT therapy; R-CHOP=rituximab, cyclophosphamide, doxorubicin, vincristine, and prednisone SUV=standardized uptake values

2.8. Other General Definitions

2.8.1. Year and Month

For the analyses purpose it's assumed 1-year equals to 365.25 days and 1 month equals to 30.4375 days.

2.8.2. Age

Age in years is calculated at the date the informed consent is signed.

2.8.3. Time from Initial Diagnosis to Randomization

Time from initial diagnosis to randomization in months is calculated as (date of randomization - date of initial diagnosis)/30.4375 and the result is rounded to the first decimal place. Partially missing initial diagnosis date is imputed based on the rules provided in Section 2.5 of the SAP.

2.8.4. Treatment-emergent Adverse Event 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 antineoplastic therapy, whichever occurs first.

3. INTERIM ANALYSIS AND DATA MONITORING COMMITTEE REVIEW

3.1. Data Monitoring Committee

An independent DMC was established to monitor data on an ongoing basis to ensure the safety of subjects enrolled in this study and to evaluate the efficacy of the treatment at the time of interim analysis. The committee met periodically to review study data, to assess the evidence of benefit or adverse effects of ibrutinib, and to monitor the conduct of the study. After the review, the DMC was to make recommendations regarding the conduct of the study including stopping the study for efficacy, if the pre-specified stopping boundary is crossed at the interim analysis. The details regarding the DMC responsibilities, authorities, and procedures are provided in a separate DMC charter.

3.2. Interim Analysis

An interim analysis using the classical O'Brien-Fleming boundary for efficacy was planned to be conducted after observing approximately 60% (151) PFS events (PD or death).[4] The stopping boundaries were implemented using Lan-DeMets spending function with the parameter resembling the conservative O'Brien-Fleming boundary using East® software (version 5.3) to control the 2-sided Type I error of 0.05 for the comparison of the PFS endpoint. The interim analysis was conducted after 163 PFS events were observed. Details of the efficacy monitoring plan are summarized in Table 2.

Table 2: Two-stage Group Sequential Design for Progression-free Survival

Analysis	Information Fraction	Number of PFS Events	Monitoring Boundary Based on Observed HR	2-sided Nominal Significance Level	2-sided Cumulative Alpha
Planned Interim	60%	151	0.646	0.0076	0.0076
Actual Interim	65%	163	0.674	0.0116	0.0116
Final	100%	252	0.778	0.0476	0.05

PFS=progression-free survival, HR=hazard ratio
The 2-sided nominal significance level and cumulative alpha in this table are based on a 2-sided log-rank test and are updated using the software package East® v6.5 (Cytel Software Corp., Cambridge, MA). The significance level for the final analysis is based on the actual information fraction for the interim analysis.

At the preplanned interim analysis, the independent DMC made the recommendation of continuing the study as the pre-specified boundary for efficacy was not crossed.

4. SUBJECT INFORMATION

4.1. Disposition Information

Disposition information will be summarized for the ITT analysis set, including the number and percentage of subjects who are randomized, treated, and discontinued treatment as well as their reason for discontinuation. The tabulation of the above information will be provided by treatment group and combined.

Descriptive statistics will be provided for the time on study. Time on study will be defined the same way as OS with reversed censoring, ie, subject who died is 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.

4.2. Demographics and Baseline Characteristics

All demographics and baseline characteristic variables will be summarized for the ITT analysis set by treatment group and combined, unless otherwise specified. No formal statistical hypothesis testing is planned.

- Demographic and baseline characteristics data: age (continuous and grouped as <65, or ≥65 years), sex, race, ethnicity, weight, height, blood pressure, temperature, heart rate, and body surface area (BSA)
- Disease characteristics at baseline: iNHL histology (FL or MZL), time from initial diagnosis to randomization, lymphoma-related B-symptoms, lactic acid dehydrogenase (LDH), number of lines of prior therapies (continuous and grouped as 1, >1), background therapy received in the study (BR or R-CHOP), refractory or relapse to last prior therapy, hepatic impairment, creatinine clearance rate (<60 mL/min, ≥60 mL/min), extra nodal disease, Eastern Cooperative Oncology Group (ECOG) performance status, tumor bulk (≤6 cm, >6 cm)
- Hematology: hemoglobin, white blood cell (WBC) count, absolute neutrophil count (ANC), absolute lymphocyte count (ALC), and platelet count
- Chemistry: creatinine, aspartate aminotransferase (AST), alanine aminotransferase (ALT), total bilirubin
- Coagulation: Prothrombin international normalized ratio (INR), Activated partial thromboplastin time (aPTT)

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 and range. 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.

4.3. Medical History

Abnormal medical history findings reported by the investigator will be summarized by body system by treatment group and combined, based on the ITT analysis set.

4.4. Extent of Exposure

Extent of exposure will be summarized for the safety analysis set. Descriptive statistics (n, mean, standard deviation, median and range) will be provided for the treatment duration (the interval between date of first dose and date of last dose of study drug), and study drug dosing information including total dose received (the sum of total dose), dose intensity (the ratio of total dose and treatment duration) and relative dose intensity (the ratio of dose intensity and 560 mg).

For each component of the background therapy, dosing information will include total number of cycles received, total dose received (the sum of actual dose administered), dose intensity (the ratio of total dose and number of cycles), and relative dose intensity (the ratio of dose intensity and planned total cumulative dose for the specific drug per cycle). For vincristine, relative dose intensity will be considered as 100% for those subjects who received 2 mg since dosing is independent of basal surface area if dose >2 mg.

For each of the study drug and background therapy components, dose reduction information will be summarized for the number of subjects with any dose reduction (at least 1 reported dose reduction), as well as the frequency and reason of dose reduction. Number and percentage of subjects with cycle delays will also be presented.

AEs leading to dose reduction will be summarized for each component of the study treatment.

4.5. Protocol Deviations

Subjects with major protocol deviations will be summarized and listed, as appropriate, by treatment group. Protocol deviations are based on clinical review primarily on the following aspects (but not limited to): (1) eligibility criteria, (2) treatment compliance, (3) subject safety, and (4) efficacy assessment deviations. Protocol deviations are closely monitored during the execution of the study and the final set of protocol deviation criteria is finalized before database lock. The major protocol deviations are reviewed and/or classified based on clinical review of the protocol deviations.

4.6. Prior and Concomitant Medications

Prior and concomitant medications will be summarized for the ITT analysis set by the World Health Organization Drug Dictionary therapeutic class, pharmacological class, and preferred term. Medications administered prior to the first dose of study drug, or any component of the background therapy treatment is considered prior medications. Concomitant therapies include those taken on or after the first dose of study drug or any component of the background therapy treatment through 30 days after the last dose.

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

- Anti-coagulation and anti-platelets.
- Cytochrome P450 (CYP)3A inhibitors/inducers

- Growth factors/ cytokines: Referred to medications coded to ATC level 3 Text = "other anti-anemic preparations" or ATC level 4 Text = "colony stimulating factors"
- Transfusions

5. EFFICACY

Unless specified otherwise all efficacy analyses are based on the ITT analysis set.

5.1. Analysis Specifications

5.1.1. Level of Significance

In general, all tests are performed at a 2-sided significance level of 0.05, unless otherwise specified. All interval estimations are reported using 2-sided 95% confidence intervals (CIs). Statistical inference on the primary endpoint, PFS, is conducted at 2-sided overall significance level of 0.05, under group sequential testing design per O'Brien-Fleming boundaries, as specified in Section 3. PFS will be updated at the time of protocol-specified final analysis of OS.

If the primary endpoint achieves statistical significance at a 2-sided significance level of 0.0476 at the final analysis, testing of the secondary endpoints will be performed at the 2-sided significance level of 0.05 in a sequential hierarchical manner based on a closed testing procedure. All key secondary endpoints are ranked in sequence according to the hierarchical order specified below:

1. CR rate
2. OS
3. ORR (CR + PR)
4. Time to worsening (TTW) in the FACT-LymS

The significance level at the interim and the primary analyses of PFS is determined by the alpha-spending function specific to endpoints:

- For CR and ORR, the data are expected to be matured at the interim analysis of PFS and are not expected to change at the final analysis of PFS. Consequently, testing of CR rate and ORR endpoints is conducted at 2-sided significance level of 0.05
- For OS, testing of OS is conducted under group sequential design per O'Brien-Fleming boundaries, as specified in Section 5.3
- For TTW in the FACT-LymS, the information fraction is expected to be the same as for PFS at the interim analysis. The O'Brien-Fleming alpha-spending function as implemented by the Lan-DeMets method is used for alpha spending. The actual alpha is the same as for PFS at the interim and at the final analyses of PFS

If the null hypothesis for any endpoint fails to be rejected at the interim analysis, then any subsequent endpoint(s) listed in the hierarchical order above will not be tested until the next analysis time point, eg, primary analysis of PFS), if applicable. If the null hypothesis for an

endpoint is rejected at the interim analysis, it will remain being rejected and will not be re-tested at any subsequent time points, if any.

5.1.2. Data Handling Rules

Unless specified otherwise, missing values will not be imputed.

5.1.3. General Analysis Considerations

Disease progression and disease response is based on the modified Revised Response Criteria for Malignant Lymphoma.[1]

If the number of subjects in a stratum is small, 1 or more stratification factors will be removed from the stratified analysis. In that case, the stratified analysis with fewer stratification factors will be considered as the primary analysis. The order of removal is determined by clinical importance as follows:

1. Background therapy treatment (BR vs R-CHOP)
2. Refractory vs relapsed disease
3. Number of prior lines of therapy (1 vs >1)
4. iNHL histology (FL vs MZL)

5.2. Primary Efficacy Endpoint(s)

5.2.1. Definition

The primary endpoint is the investigator assessed PFS, defined as duration from the date of randomization to the date of disease progression or death, whichever is first reported. Subjects who are progression-free and alive or have unknown status will be censored at the last adequate disease assessment. Subjects with no postbaseline disease assessment will be censored at the randomization date. Adequate disease assessment is defined as having sufficient evidence to correctly indicate that progression has or has not occurred.

5.2.2. Primary Estimand

Primary Trial Objective: To demonstrate the superiority of ibrutinib + background therapy (BR or RCHOP) to that of placebo + background therapy in terms of PFS in patients with relapsed refractory FL or MZL.

Estimand Scientific Question of Interest: What is the effect on PFS of assigning subjects to ibrutinib + background therapy vs placebo + background therapy?

This primary estimand is the main clinical quantity of interest to be estimated in this study, which is defined by the following five attributes.[5]

Population: Patients with relapsed or refractory FL or MZL.

Treatment: Ibrutinib+ background therapy vs placebo + background therapy.

Variable: PFS (PD, as assessed by the investigator or death).

Population-level summary: Kaplan-Meier estimates of PFS, HR of ibrutinib + background therapy vs placebo+ background therapy.

Intercurrent events and handling strategies: Treatment discontinuation, use of subsequent anticancer therapy, death due to Coronavirus Disease 2019 (COVID-19).

Intercurrent Events	Name of Strategy for Addressing Intercurrent Events and its Description
Treatment discontinuation	**Treatment policy strategy:** Use time to PD or death, regardless of whether treatment discontinuation had occurred.
Use of subsequent anticancer therapy	**Treatment policy strategy:** Use time to PD or death, regardless of whether used subsequent anti-cancer therapy. **Hypothetical strategy:** Subjects are censored at the last disease assessment showing no evidence of PD before the use of subsequent anti-cancer therapy.
Death due to COVID-19	**Composite variable strategy:** Consider (pre-PD) death as a PFS event. **Hypothetical strategy: S**ubjects are censored at the last disease assessment before (pre-PD) death due to COVID-19.

5.2.3. Analysis Methods

5.2.3.1. Primary Analysis

Assuming a non-informative censoring, distinct baseline hazard for each stratum and a common proportional HR across strata, the primary estimator for the primary endpoint is the HR and its 95% CIs estimated using a stratified Cox regression model with study intervention as the sole explanatory variable with (1) background therapy treatment (BR vs R-CHOP), (2) refractory vs relapsed disease, (3) iNHL histology (FL vs MZL), and (4) number of prior lines of therapy (1 vs >1) as stratification factors.

The treatment policy strategy will be adopted for handling the intercurrent events of treatment discontinuation and use of subsequent anti-cancer therapy. The composite variable strategy will be adopted for handling the intercurrent events of pre-PD death (PFS event) due to COVID-19. However, this analysis would only be conducted if pre-PD COVID-19 related deaths account for more than 5% of the total PFS events.

The primary analysis of PFS will be based on the ITT analysis set. The Kaplan-Meier product limit method will be used to estimate the distribution and median PFS with 95% CI for each treatment group. The stratified log-rank test will be used to compare PFS distributions of the 2 treatment groups. The stratification factors at randomization to be used in the analysis are background

therapy treatment (BR vs. R-CHOP), refractory vs. relapsed disease, number of prior lines of therapy (1 vs. >1), and iNHL histology (FL vs MZL).

To further explore the potential effect of covariates in the primary analysis, a Multivariate Cox Regression analysis may be performed using a selected set of potential prognostic variables (obtained at or before baseline) as covariates. Each factor will be assessed individually for prognostic values using a univariate Cox model. Factors that are deemed to have prognostic value will be included as covariates in a Cox model to assess their significance in the presence of other factors. Selection methods will be used to identify the final set of prognostic factors. Treatment will then be added to this final model to assess the effect of treatment when adjusted for these selected prognostic factors. Potential prognostic factors include, but are not limited to, the following.

- Age (<65 vs ≥65)
- Gender (male vs female)
- Race (White vs non-White)
- Background therapy treatment (BR vs R-CHOP)
- Number of lines of prior therapy (1 vs >1)
- Disease status (refractory vs. relapsed)
- iNHL histology (FL vs MZL)
- POD24 (>24 vs ≤ 24)
- PRIMA- PI (Low vs Medium vs High risk)

Concordance between investigator determined PFS and IRC determined PFS will be primarily assessed using the Dodd 2-stage test procedure.[6] In addition, the AMIT audit method[7]will be used as sensitivity analysis for assessing concordance between IRC and investigator data. Details on the study's imaging audit plan is provided in the study audit plan.

5.2.3.2. Sensitivity and Supplementary Analyses

5.2.3.2.1. Sensitivity Analysis of Disease Assessment Follow up

Assumptions: Non-informative censoring assumed for all types of censoring.

Sensitivity Estimator: A stratified Cox regression model with study intervention as the sole explanatory variable will be performed, subjects will be censored at the last disease assessment if they progress or die after missing ≥2 consecutive planned disease assessment visits.

5.2.3.2.2. Supplementary Analysis of Estimand 2:

Estimand 2 is defined to support the primary estimand. The only attribute that changes from the definition of the primary estimand is the handling strategy for the use of subsequent anti-cancer therapies.

- Hypothetical strategy: All subjects had continued treatment as planned and had not used any subsequent anti-cancer therapies.

Under the estimand 2, time to progression or death is censored at the last disease assessment showing no evidence of progression before the start of subsequent antineoplastic therapy. The same analyses described in primary estimator will be applied.

5.2.3.2.3. Supplementary Analysis of Estimand 3

Estimand 3 is defined to support the primary estimand. The only attribute that changes from the definition of the primary estimand the handling strategy for the death due to COVID-19:

- Hypothetical strategy: all subjects had continued treatment as planned and had not died from COVID-19.

Under the estimand 3, time to PD or death is censored at the last disease assessment before pre-PD death due to COVID-19. The same analyses described in primary estimator will be applied, however, this supplementary analysis would only be conducted if pre-PD COVID-19 deaths account for more than 5% of total PFS events.

5.2.3.3. Subgroup Analyses of PFS

Subgroup analysis will be performed for the subgroups (Section 2.7) to assess the internal consistency of the treatment benefit. A forest plot, with HR between the treatment groups within each subgroup and its associated 95% CI estimated using non-stratified Cox regression model will be presented.

In addition, interaction between treatment and background therapy treatment will be tested at a 2-sided 0.2 level using a stratified Cox proportional hazards model. If the parameter for the interaction term is significant, the inferences on the treatment benefit will be made for each background therapy treatment separately. Secondary Endpoints

5.3. Secondary Endpoints

5.3.1. Overall Survival

Definition

Overall survival is defined as the interval between the date of randomization and the date of the subject's death from any cause. If the date of death is unknown, OS will be censored at the date the subject was last known to have been alive. Subjects who are known to be alive as of their last known status will be censored at their date of last contact. Subjects who are lost in follow-up will be censored at the date the subject is last known to have been alive.

Analysis method

Overall survival will be analyzed using the log-rank test for treatment comparison at the interim analyses of OS and the stratified log-rank test at the final analysis of OS. The OS distribution and median OS with its 95% CI will be estimated using the Kaplan-Meier product limit method. The HR for background therapy + ibrutinib relative to background therapy + placebo and its associated 95% CI will be calculated based on the Cox proportional hazards model at the interim analyses of OS and the stratified Cox proportional hazards model by the stratification factors at randomization at the final analysis of OS.

Multiplicity adjustment for OS

Inferential hypothesis testing for OS was planned to conducted at 2-sided overall significance level of 0.05 according to the hierarchy specified in Section 5.1.1, at different time points under the group sequential testing design per O'Brien-Fleming boundaries (Table 3). The first analysis of OS was planned to occur at the time of the planned interim analysis for PFS. The second analysis will occur at the time of the planned primary analysis of PFS when 252 PFS events have been observed. The final analysis of OS will be performed at the study end, which is defined as when 50% of the subjects have died or sponsor terminates the study, whichever comes first.

Assuming OS follows an exponential distribution and the median OS for the control arm (background therapy + placebo) is 6 years from randomization and the HR for the treatment group (background therapy+ibrutinib) relative to the control is 0.78, it was expected that approximately 54 and 104 overall survival events will be observed at the time of the first and second analysis of OS, respectively. At the interim analysis of PFS, 55 OS events were observed.

The efficacy of OS will not be claimed unless PFS reaches statistical significance. For the 3 planned analyses of OS, the type I error rate of the OS will be rigorously controlled using the Lan-DeMets spending function with the parameter resembling the conservative O'Brien-Fleming boundary which conservatively allocates alpha according to the information fraction (IF) at the time of each planned analysis. Information on the estimated boundary for OS is provided in Table 3.

Table 3: Estimated Crossing Boundary in Hazard Ratio and p-value Scales for the Planned OS Analysis

OS Analysis	Cumulative Number of OS Events	Timing (Month)	Calculated Boundary HR	Calculated Boundary p-value (2-sided)	Incremental Boundary Crossing Probability
1st analysis	55 (IF = 27.5%)	25	0.33	0.00004	0.0007
2nd analysis	120 (IF = 60%)	49	0.61	0.0076	0.094
Final analysis	200	89	0.76	0.0476	0.319

OS=overall survival, HR=hazard ratio, IF=information fraction

5.3.2. Complete Response Rate and Overall Response Rate

Definition

Complete response rate is defined as the proportion of subjects who achieve CR. Overall response rate is defined as the proportion of subjects who achieve CR or PR.

Analysis method

Subjects with no postbaseline data will be considered as non-responders. Partial response or CR after the start of subsequent therapy will not be considered. CR rate will be summarized using descriptive statistics for categorical data by treatment group.

The relative risk (background therapy +ibrutinib vs background therapy + placebo) adjusted for the stratification factors at randomization will be reported along with the associated 95% CI. Statistical inference will be evaluated using the Cochran-Mantel-Haenszel (CMH) Chi-square statistic, adjusted for the stratification factors at randomization: background therapy treatment (BR vs R-CHOP), refractory vs relapsed disease, number of prior lines of therapy (1 vs >1), and iNHL histology (FL vs MZL). Logistic regression analysis will also be performed to estimate the odds ratio and its associated 95% CI between the 2 treatment groups, adjusted for the stratification factors at randomization.

Overall response rate will be analyzed in a similar fashion as CR rate. A listing of tumor response by assessment visit will be provided by treatment group.

5.3.3. Time to Worsening in the Lym Subscale

Definition

Time to worsening (TTW) in the FACT-LymS is defined from the date of randomization to the start date of the worsening of patient symptoms. Worsening is defined by a 5-point decrease from baseline in patient symptoms. Death and missing data due to very ill as noted on the case report form (CRF) will also be considered as worsening. Subjects who have not met the definition of worsening will be censored at the last PRO assessment. Subjects with no baseline assessment or on-study assessment will be censored at date of randomization.

In the event of missing item(s) in a FACT-Lym assessment, the subscale scores are computed according to the FACT scoring and administration guidelines by multiplying the sum of the subscale by the number of items in the subscale then dividing by the number of items answered.

Analysis method

The distribution of TTW in the Lym subscale will be estimated using Kaplan-Meir product limit method. A stratified log-rank test will be used to compare the distributions of the 2 treatment groups. The HR for background therapy + ibrutinib relative to background therapy + placebo and its associated 95% CI will be calculated based on the stratified Cox proportional hazards model by

the stratification factors at randomization. Further details on patient reported outcomes are provided in Section 7.

5.3.4. Duration of Response

Definition

Duration of response is defined as the interval between the date of initial documentation of a response (CR or PR) and the date of first documented evidence of PD (or relapse for subjects who experience CR during the study) or death, whichever occurs first. The censoring rule for DOR is the same as PFS.

Analysis method

Duration of response will be analyzed for subjects with CR/PR. The distribution of duration of response will be estimated using the Kaplan-Meier product limit method. No formal hypothesis testing is conducted to compare the treatment groups.

5.4. Exploratory Endpoints

5.4.1. Time-to-Next Treatment

Time-to-next treatment is defined as the duration from the date of randomization to the start date of any anti-lymphoma treatment subsequent to the study treatment. Subjects without subsequent treatment will be censored at the date of the last site visit.

The distribution of TTNT will estimated using the Kaplan-Meier product limit method. No formal hypothesis testing will be conducted to compare the treatment groups.

5.4.2. Minimal Residual Disease (MRD) Negative Rate

The rate of MRD-negative response among FL subjects will be presented. Rate of MRD-negative response is defined as the proportion of FL subjects whose response is CR and who reach MRD-negative disease status, (ie, <1 FL cell per 10,000 leukocytes for detection using the MRD assay). Subjects with missing MRD data and non-CR subjects are considered non-responders. In addition, durability of MRD-negative status defined will be presented. No formal hypothesis testing will be conducted to compare the treatment groups.

5.4.3. Progression-free Survival on Next-Line Therapy (PFS2)

Progression-free survival on next-line therapy is defined as the time interval between the date of randomization and date of event, which is defined as progressive disease as assessed by investigator that starts after the next line of subsequent antineoplastic therapy, or death from any cause, whichever first. Those who are alive and for whom a second disease progression has not been observed are censored at the last disease assessment without second disease progression. It is analyzed using the same analysis methods as used for PFS.

6. SAFETY

Safety will be analyzed using the incidence, intensity, and type of AE, laboratory tests. All safety analyses will be based on the safety analysis set and performed by the treatment actually received.

6.1. Adverse Events

The verbatim terms used in the CRF by investigators to identify AE will be coded using the current Medical Dictionary for Regulatory Activities (MedDRA). The severity of AE is assessed in National Cancer Institute (NCI) common toxicity criteria for adverse events (CTCAE) Version 4.03.

Treatment-emergent AE are (1) AE that occur after the first dose of study drug or any component of the background therapy treatment, and within 30 days following the last dose of study drug or any component of the background therapy treatment or the start date of a subsequent systemic anticancer therapy, whichever is earlier (including both dates), (2) any AE that is considered study treatment-related (ie, possibly, probably, or very likely related to study treatment) regardless of the start date of the event, (3) any AE that is present at baseline but worsens in severity or is subsequently considered study treatment-related by the investigator, or (4) any AE with missing start date and its end date is during the treatment.

Treatment-emergent AE will be summarized by system organ class and preferred term, by grade, by relationship to study drug, and by action taken. For each AE, 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 (from the highest to lowest incidence in the Total column). The same summary will be provided for study drug-related AE, SAE, and study drug-related SAE, as well as AE leading to treatment discontinuation and death. Ocular events collected on the CRF will be summarized similarly to the treatment-emergent AE.

6.2. Adverse Events of Special Interest

Major hemorrhage events: a subset of hemorrhagic events which are

- Grade 3 or higher, or
- SAE of bleeding of any grade, or
- central nervous system hemorrhage of any grade

Adverse events of special interest will be summarized similarly to treatment-emergent AE.

6.3. Other Safety Observations

Other malignancies: other malignancies are defined as new malignant tumors including solid tumors, skin malignancies and hematologic malignancies, and are to be reported by investigators for the duration of study treatment and during any protocol-specified follow-up periods including post-progression follow-up for OS. Other malignancies will be summarized by preferred term for the entire study period and for the treatment-emergent AE period. In addition, AEs of clinical interest and other special AEs will be summarized by treatment group.

6.4. Death

A summary of the number of deaths since first administration of study drug or any component of the background therapy until 30 days after last study drug administration will be provided, along with the primary cause of death. A death is a study drug-related death if the primary cause is a study drug-related AE.

6.5. Clinical Laboratory Tests

Laboratory data for hematology and serum chemistry tests will be reported in SI units. Applicable laboratory results will be graded according to NCI-CTCAE Version 4.03. Note, toxicity grading for creatinine increase will be based on the NCI CTC v4.03 criteria but limited only to the part based on the ULN, the other part that is based on change from baseline will not be used for toxicity grading. Toxicity grading for all other lab parameters will be based on the NCI CTC v4.03 criteria as is. Generic normal ranges will be applied whenever reference ranges are not available. Generic normal ranges will be applied to hematology using published values.[8,9] Local ranges will be used for chemistry parameters.

The following laboratory tests will be analyzed:

- Hematology: hemoglobin, WBC count, ANC, ALC, and platelet count
- Chemistry: sodium, potassium, creatinine, AST, ALT, alkaline phosphatase, LDH, total bilirubin, and albumin.
- Beta2-microglobulin, serum immunoglobulin levels (IgG, IgM, IgA)

Descriptive statistics (mean, standard deviation, median, and range) will be calculated for the raw data and for change from baseline at each time point of assessment, as well as for change from baseline to the last value. Toxicity grade will be summarized for all hematology and chemistry parameters except for LDH, which will be summarized by abnormal/normal using standard normal ranges provided in the dataset. Change from baseline to the worst grade during the treatment will be provided as shift tables for selected parameters.

6.5.1. Creatinine Clearance

Creatinine clearance (CrCl) is calculated using the Cockroft-Gault formula:

$$CrCl_{(est)} = \frac{(140 - age[yr])(lean\ body\ wt[kg])}{(72)(serum\ creatinine[mg/dL])} \times 0.85 (if\ female)$$

For males, the factor is 1 instead of 0.85

6.5.2. Analysis of Lymphocytosis

A descriptive summary of ALC with and without lymphocytosis, time to lymphocytosis, and duration of lymphocytosis will be provided by treatment. Lymphocytosis is defined as absolute lymphocyte counts increasing ≥50% from baseline and achieving level $\geq 5\times10^9$/L For subjects with lymphocytosis, resolution of lymphocytosis is defined as (1) a decrease of ALC value to the baseline level or lower, or (2) or an achievement of ALC value that is $<5\times10^9$/L, whichever occurs

first. Time to lymphocytosis is defined as the interval between the date of first dose and the date subject had lymphocytosis first time and presented in weeks. Duration of lymphocytosis will be derived from first time subject had lymphocytosis until it recovered for first time and presented in weeks. Subjects who have not recovered will be censored at last lab assessment.

6.6. Other Safety Parameters

Frequencies of ECOG score will be reported over time.

7. PATIENT REPORTED OUTCOMES

Two PRO instruments, the FACT-LymS and EQ-5D-5L will be administered in this study:

The FACT-LymS includes 15 items, and scores range from 0 to 60, with each item scored on a 5-point scale from 0 (not at all) to 4 (very much). Higher scores represent better health status with respect to fewer disease related symptoms and fewer disease related concerns of patients.

The EQ-5D-5L is a standardized instrument for use as a measure of health outcome. For purposes of this study, the EQ-5D-5L will be used to generate weighted utility scores for use in cost effective analyses. The EQ-5D-5L is a 5-item questionnaire and a visual analogue scale ranging from 0 (worst imaginable health state) to 100 (best imaginable health state). The scores for the 5 separate questionnaires are categorical and should not be analyzed as cardinal numbers. However, the scores for the 5 dimensions are used to compute a single utility score ranging from zero (0.0) to 1 (1.0) representing the general health status of the individual.

For each of the PRO scales, compliance, descriptive statistics (number of observations, mean, standard deviation, median, minimum, maximum) of scores at baseline and postbaseline assessments, as well as change from baseline to postbaseline assessments will be reported by treatment groups. Time to worsening (TTW) in the Lym subscale analysis is specified in Section 5.3.3. Exploratory analyses may be conducted to examine time to worsening by clinical response status (ie, responder vs. non-responder) or by patient baseline characteristics of interest (ie, ECOG performance status).

In the event of missing item(s) in a FACT-Lym assessment, the subscale scores will be computed according to the FACT scoring and administration guidelines by multiplying the sum of the subscale by the number of items in the subscale then dividing by the number of items actually answered.

Time to clinically meaningful worsening and improvement in the FACT-LymS is defined as the interval from the date of randomization to the start date of the worsening and improvement separately. Instrument validation has found the minimum important difference (MID) for FACT-LymS to be in the range of 3 to 5 points, with 5 being the more conservative.[10,11] Worsening and improvement is defined by a 5 or greater point reduction and increase from baseline in the Lym lymphoma subscale, respectively.[11,12] Death will also be considered as worsening. Subjects who have not met the definition of worsening will be censored at the last FACT-LymS assessment. Subjects with no worsening event and with no baseline assessment or on-study assessment will be

censored at date of randomization. Time to worsening in the FACT-LymS will be estimated using Kaplan-Meier methods. The HR for ibrutinib relative to placebo and its associated 95% CI will be calculated based on the stratified Cox proportional hazards model by the stratification factors at randomization. The best and worst change of the lymphoma subscale from baseline to post-baseline may be explored using waterfall plot.

For FACT-LymS and EQ-5D-5L score, a mixed effects model with repeated measures analysis may be conducted estimating change from baseline at each time point between 2 treatments. ITT subjects who have a baseline value and at least 1 post-randomization value are included in the analysis. Change from baseline will be fitted to a mixed effects model including subjects as a random effect, and baseline value, treatment indicator, time in week as a continuous variable, treatment-by-time interaction, and stratification factors at randomization as fixed effects.

8. PHARMACOKINETICS

The plasma concentration data of ibrutinib and ibrutinib metabolite PCI-45227 at each time point will be summarized using descriptive statistics and will be listed for all subjects per treatment group, based on the PK analysis set.

All concentrations below the lowest quantifiable concentration or missing data will be labeled as such in the concentration data presentation. Concentrations below the lowest quantifiable concentration (ie, <0.500 ng/mL) will be treated as zero in calculating the summary statistics.

APPENDIX 1 CHANGE TO PROTOCOL PRE-PLANNED ANALYSIS

1. COVID-19 IMPACT

Impact of the COVID-19 pandemic is analyzed and presented, including COVID-19 related treatment and study disposition, protocol deviations, doses modifications, disease evaluation compliance and adverse events is summarized.

REFERENCES

1. Cheson BD, Pfistner B, Juweid ME, et al. Revised response criteria for malignant lymphoma. J Clin Oncol. 2007;25(5):579-586.
2. Bachy E et al. A Simplified scoring system in de novo follicular lymphoma treated initially with immunochemotherapy. Blood, Am Soc Hematology. 2018;132:49-58.
3. Casullo Cet al. Early Relapse of Follicular Lymphoma after Rituximab plus Cyclophosphamide, Doxorubicin, Vincristine, and Prednisone Defines Patients at High Risk for Death: An Analysis from the National LymphoCare Study. Journal of Clinical Oncology, American Society of Clinical Oncology. 2015; 33:2516-2522.
4. O'Brien PC, Fleming TR. A multiple testing procedure for clinical trials. Biometrics. 1979;35(3):549-556.
5. ICH E9 (R1) Addendum on estimands and sensitivity analysis in clinical trials to the guideline on statistical principles for clinical trials, November 2019.
6. Dodd LE, Korn EL, Freidlin B, et al. An audit strategy for progression-free survival. Biometrics. 2011;671092–671099.
7. Amit O, Mannino F, Stone AM, et al. Blinded independent central review of progression in cancer. EU J Cancer. 2011;47:1772-1778.
8. Kratz A, Ferraro, M, Sluss, PM, Lewandrowski, KB. Laboratory Reference Values; NEJM. 2004; 351: 1548-1563.
9. Provan D, Singer CRJ., Baglin T, Lilleyman J. Oxford Handbook of Clinical Hematology. Second Edition, Oxford University Press. 2004.
10. Hlubocky FJ, Webster K, Cashy J, et al. The Development and Validation of a Measure of Health-Related Quality of Life for Non-Hodgkin's Lymphoma: The Functional Assessment of Cancer Therapy—Lymphoma (FACT-Lym). Lymphoma, vol. 2013, Article ID 147176, 9 pages, 2013.
11. Carter GC, Liepa A, Zimmerman AH, Morschhauser F. Validation of the Functional Assessment of Therapy - Lymphoma (FACT-LYM) in patients with relapsed/refractory mantle cell lymphoma. Poster presented at: Annual meeting of Am Soc Hematology; 7 December 2008; San Francisco, California.
12. Cella D, Webster K, Cashy J, et al. Development of a measure of health-related quality of life for non-Hodgkin's lymphoma clinical research: The Functional Assessment of Cancer Therapy – Lymphoma (FACT-Lym) [abstract]. Am Soc Hematology. 2005;106:Abstract 750.