AcertaPharma

ACE-CL-007: Statistical Analysis Plan

A Randomized, Multicenter, Open-Label, 3 Arm Phase 3 Study of Obinutuzumab in Combination with Chlorambucil, ACP-196 in Combination with Obinutuzumab, and ACP-196 Monotherapy in Subjects with Previously Untreated Chronic Lymphocytic Leukemia

Final: July 8, 2015
Amendment 1: August 6, 2015
Amendment 2: March 23, 2018
Amendment 3: March 6, 2019

PPD

Prepared by: 04/08/2019
Date

Approved by: 08 APR 2019
Date

Approved by: 01/April 2019
Date

SAP Revision History:

Version / Date	Summary of Major Changes and Rationale
Final, July 8, 2015	
Amendment 1	The amendment is per feedback from Ethics Committee in Spain. This phase 3 study consists of one control arm (A) and two experimental arms (B and C) as defined below. • Arm A: Obinutuzumab + Chlorambucil • Arm B: ACP-196 + Obinutuzumab • Arm C: ACP-196 Monotherapy The primary endpoint is a superiority comparison between Arm B and Arm A as measured by IRC-assessed PFS, and the key secondary endpoint is a superiority comparison between Arm C and Arm A as measured by IRC-assessed PFS. If both Arm B and Arm C are statistically superior to Arm A, then the next question of interest is whether Arm B, an ACP-196 containing combination therapy, is more efficacious than Arm C, an ACP-196 monotherapy. Specifically, after both the primary endpoint and the key secondary endpoint have met statistical significance, an exploratory analysis will be performed [redacted] Section 7.2.2: Remove “withdrawal from study treatment due to toxicity” from the TTNT calculation formula so it will be consistent with the text description.
Amendment 2	1. Added details for primary and key secondary analysis. Same method will be applied to both. 2. [redacted] 3. Added ORR as assessed by IRC and OS into multiplicity adjustment plan. – Section 2.7 4. Removed [redacted] analysis. [redacted] analysis will be described in a separate SAP. – Section 7.5 5. Clarified TEAE and ECI definition. – Section 8.1

	6. Minor clarifications are made to align with Acerta's standard and process.
Amendment 3	1. Added clarification language around interim analysis, hierarchical testing procedure for the key secondary endpoints – Sections 2.4 and 2.6 2. Removed crossover population – Section 3 3. Added definitions for subsequent anticancer therapy, added language to clarify definitions on crossover subjects – Section 4 4. Clarified the definition of treatment-emergent period and treatment-emergent adverse events – Section 4.1.6, Section 8.1.1 5. Add language around collapse of stratification factors – Section 7. 6. Clarified the censoring rule for primary analysis for IRC-assessed PFS, added analysis table for IRC-PFS censoring – Section 7.1 7. Added operational definition of the last adequate IRC assessment – Section 7.1 8. Updated the planned sensitivity and subgroup analyses for primary efficacy endpoint – Section 7.3, Section 7.4 9. CCI [redacted] 10. Added language around analysis of clinical laboratory and ECG – Section 8.3.2 and Section 8.6

TABLE OF CONTENTS

LIST OF ABBREVIATIONS AND DEFINITIONS

Del11q	11q deletion mutation, chromosome deletion 11q22.3
Del17p	17p deletion mutation, chromosome deletion 17p13.1
AE	adverse event
ALC	absolute lymphocyte count
BID	twice per day
CI	confidence interval
CIRS	Cumulative Illness Rating Score
CLL	chronic lymphocytic leukemia
CR	complete remission (response)
CRi	CR with incomplete bone marrow recovery
CT	computed tomography
CTCAE	Common Terminology Criteria for Adverse Events
DMC	Data Monitoring Committee
DOL	duration of lymphocytosis
ECOG	Eastern Cooperative Oncology Group
eCRF	electronic case report form
CCI	[redacted]
CCI	[redacted]
FDA	U.S. Food and Drug Administration
CCI	[redacted]
HR	hazard ratio
IGHV	immunoglobulin heavy-chain variable
INV	investigator
IRC	independent review committee

IPD	important protocol deviation
ITT	intent-to-treat
IV	intravenous
IWCLL	International Workshop on Chronic Lymphocytic Leukemia
IWRS	interactive Web response system
LPR	last subject randomized
MedDRA	Medical Dictionary for Regulatory Activities
CCI [redacted]	
CCI [redacted]	
CCI [redacted]	
NCI	National Cancer Institute
nPR	nodular partial remission (response)
ORR	objective response rate
OS	overall survival
PD	progressive disease
PFS	progression-free survival
CCI [redacted]	
PLT	platelet
PR	partial remission (response)
CCI [redacted]	
CCI [redacted]	
RBC	red blood cell
SAE	serious adverse event
SAP	statistical analysis plan
SD	stable disease

TEAE	treatment-emergent adverse event
WHO	World Health Organization

1 INTRODUCTION

This statistical analysis plan (SAP) provides details of the statistical analyses that have been outlined for study ACE-CL-007 Protocol Amendment 4, dated 06 March 2018, which is titled “A Randomized, Multicenter, Open-Label, 3 Arm Phase 3 Study of Obinutuzumab in Combination with Chlorambucil, ACP-196 in Combination with Obinutuzumab, and ACP-196 Monotherapy in Subjects with Previously Untreated Chronic Lymphocytic Leukemia.”

This study is being conducted under the sponsorship of Acerta Pharma. The sponsor is conducting statistical programming and analyses using dummy treatment code before study unblinding to the Sponsor for interim/final analyses. All safety and interim efficacy analyses for the purpose of independent Data Monitoring Committee (DMC) review will be conducted in an unblinded fashion under contract by an independent contract organization.

2 STUDY DESCRIPTION

2.1 Study Design

This randomized, multicenter (i.e., approximately 200 global centers), open-label, 3-arm Phase 3 study is designed to evaluate the safety and efficacy of acalabrutinib in combination with obinutuzumab, acalabrutinib monotherapy compared with obinutuzumab in combination with chlorambucil in subjects with previously untreated chronic lymphocytic leukemia (CLL).

Approximately 510 eligible subjects will be randomized in a 1:1:1 ratio into 3 arms (n=170 each) to receive either Arm A (obinutuzumab in combination with chlorambucil per the package inserts), Arm B (ACP-196 [acalabrutinib] 100 mg BID in combination with obinutuzumab per the package insert), or Arm C (acalabrutinib 100 mg BID).

Eligible subjects will be stratified for randomization with regard to:

- Presence or absence of 17p deletion by central laboratory;
- ECOG performance score 0 or 1 versus 2; and
- Geographic region (North America and Western Europe versus Other).

Subjects randomized to Arm A eligible to cross over to acalabrutinib monotherapy will receive acalabrutinib treatment if they meet eligibility for crossover. Eligibility for crossover includes meeting the criteria for ECOG performance status and the laboratory parameters as outlined in the inclusion criteria. Subjects may not receive any new systemic therapy after IRC confirmation of disease progression and prior to the initiation of crossover therapy with acalabrutinib.

An independent DMC will be formed and constituted according to regulatory agency guidelines. Detailed information regarding the composition of the DMC and detailed DMC procedures will be provided in a separate charter. The DMC will review the safety data periodically and the interim analysis results and provide recommendations according to the charter.

One interim efficacy analysis is planned for the study (see Section 2.4).

2.2 Study Objectives

2.2.1 Primary Objective

- To evaluate the efficacy of obinutuzumab in combination with chlorambucil (Arm A) compared with acalabrutinib in combination with obinutuzumab (Arm B) based on IRC assessment of progression-free survival (PFS) per International Workshop on Chronic Lymphocytic Leukemia criteria (IWCLL; Hallek et al. 2008) with incorporation of the clarification for treatment-related lymphocytosis (Cheson 2012)—hereafter referred to as IWCLL 2008 criteria—in subjects with previously untreated CLL.

2.2.2 Secondary Objectives

- To evaluate the efficacy of obinutuzumab in combination with chlorambucil (Arm A) versus acalabrutinib monotherapy (Arm C) based on IRC assessment of PFS per IWCLL 2008 criteria
- To compare obinutuzumab in combination with chlorambucil (Arm A) versus acalabrutinib in combination with obinutuzumab (Arm B), and obinutuzumab in combination with chlorambucil (Arm A) versus acalabrutinib monotherapy (Arm C) in terms of:
  - IRC-assessed ORR per IWCLL 2008 criteria;
  - TTNT (defined as the time from randomization to institution of non-protocol–specified treatment for CLL); and
  - OS

2.2.3 Safety Objectives

Incidence of AEs and SAEs and changes in laboratory measurements

2.2.4 Exploratory Objectives

CCI

CCI [redacted]

2.3 Power and Sample Size Justification

The null and alternative hypotheses are as follows:

H_0: $PFS_{Arm\ B} = PFS_{Arm\ A}$

H_A: $PFS_{Arm\ B} \neq PFS_{Arm\ A}$

The test will be conducted to reject the null hypothesis in favor of the alternative while showing that Arm B is superior to Arm A.

The study is expected to randomize approximately 510 subjects. Subjects will be randomized in a 1:1:1 ratio to Arm A, Arm B, or Arm C. Approximately 170 subjects will be randomized to each arm. The sample size calculation is driven by hypothesis test between Arm B and Arm A.

The study is sized to achieve approximately 90% power to detect a hazard ratio (Arm B/Arm A) of 0.60 for PFS, which, under the model assumptions, CCI [redacted] in median PFS time: CCI [redacted] Given the study assumptions, the minimum detectable treatment difference at the final analysis of PFS corresponds to a CCI [redacted]

CCI [redacted]

One interim analysis and one final analysis for PFS are planned. The final analysis of PFS is planned to occur when a total of 167 PFS events have been observed, which is anticipated to occur at CCI [redacted] after the first subject is randomized. The interim analysis of PFS will be conducted when approximately two-thirds of the final analysis

PFS event goal (i.e., 111 events across Arms A and B) have been observed, which is expected to occur approximately CCI after the first subject has been randomized.

The sample size and power calculations were based on East 6 (Version 6.3.1), with use of the design "Two-Sample Test, Parallel Design, Logrank Given Accrual Duration and Study Duration."

2.4 Interim Analysis

An interim analysis will be conducted to assess futility and superiority of Arm B versus Arm A with respect to the primary efficacy endpoint, PFS as assessed by IRC. The sponsor will monitor PFS events closely, where progression is based on IRC assessment. An event-based interim analysis will occur when approximately 111 events across Arms A and B (approximately two-thirds of the required PFS events for final analysis) have been observed. A visit cutoff date will be determined for the interim analysis. A snapshot of the database will occur after all data through the visit cutoff date have been entered and cleaned. The interim analysis will be conducted based on a snapshot of the database that consists of all available data, including all available PFS events, up to the visit cutoff date.

CCI

All tests for interim and final analyses will be performed at 2-sided significance level. The Lan-DeMets alpha-spending function based on the O'Brien-Fleming boundaries (O'Brien and Fleming 1979; Lan and DeMets 1983) for superiority and futility (nonbinding) will be used.). The table below presents a summary of the planned PFS analyses, the efficacy stopping boundaries, and the estimated timing of these analyses.

Table 1: CCI [redacted]

Analysis	No. of IRC-Assessed PFS Events	Early Stopping Boundary (α Level) [a]		Estimated Timing[b]
		Inferiority (Futility)	Superiority (Efficacy)	
Interim	CCI [redacted]			
Final				

[a] P-values will be based on two-sided log rank test.

[b] Time from enrolment of first subject to data cutoff date.

In the event that the actual number of IRC-assessed PFS events is different from the estimated 111 events, the nominal levels for the interim and final analyses will be determined according to the aforementioned alpha-spending function and based on the actual information fraction at the time of the interim analysis.

An independent statistician will conduct the unblinded interim analysis for review by the DMC. The committee will operate independent of the sponsor and the clinical investigators. Acerta Pharma will remain blinded with respect to the interim analysis results unless the DMC recommends trial stopping and Acerta Pharma agrees with the recommendation.

If the criterion for early efficacy or futility is met at the time of the interim analysis, the DMC may recommend stopping the study in accordance with the terms of the DMC charter. If the DMC recommends stopping early, the interim analysis will be the final efficacy analysis.

The DMC members will use their expertise, experience, and judgment to evaluate the safety data from the trial on a regular basis and to recommend to the sponsor whether the trial should continue or be stopped early for safety. No formal statistical rules recommending early stopping for safety are planned.

2.5 Final Analysis

If the study fails to cross the boundary at interim analysis, the study will proceed to final analysis. A visit cutoff date will be determined after approximately 167 IRC-assessed events have been observed. Database lock will occur after all data through the visit cutoff date have been entered and cleaned. The final analysis will be conducted based on final locked database.

2.6 Multiplicity Adjustments

To control the overall Type I error at 0.05 level, the Lan-DeMets alpha-spending function based on the O'Brien-Fleming boundary will be used to split α into α_1 and α_2 for interim and final analyses, respectively. The nominal α_1 and α_2 levels will be determined based on the actual information fraction at the time of the interim analysis. Within each analysis, the fixed sequence procedure will be utilized to adjust for multiple comparisons.

If the primary efficacy endpoint, PFS as assessed by IRC in Arm B versus Arm A, achieves statistical significance, the following secondary efficacy endpoints will be tested in a fixed sequential hierarchical manner for interim and final analyses:

1 PFS as assessed by IRC between Arms C and A
2 ORR as assessed by IRC between Arms B and A
3 ORR as assessed by IRC between Arms C and A
4 OS between Arms B and A
5 OS between Arms C and A

The fixed sequence procedure will perform the testing of PFS between Arms B and A first. If the p-value is ≤ α_i (i=1 for interim and 2 for final), the procedure will proceed to test PFS between Arms C and A at the same α_i (i=1 for interim and 2 for final) level (testing 1).

If the testing of IRC-assessed PFS achieve statistical significance at interim analysis, the IRC-assessed ORR will be tested at an alpha level of 0.05 (testing 2 and 3), given that almost all responses will have been observed at that time.

The OS will be tested at the same alpha level as the primary endpoint (testing 4 and 5), thus nominal α_1 and α_2 for interim and final analyses, respectively.

Following the fixed sequence testing procedure, if a p-value is not statistically significant, the p-value for subsequent tests will be presented as descriptive.

2.7 Randomization and Blinding

2.7.1 Randomization

Randomization will be used to minimize bias in the assignment of subjects to treatment groups, to increase the likelihood that known and unknown subject attributes (e.g., demographic and baseline characteristics) are evenly balanced, and to enhance the validity of statistical comparisons across treatment groups.

Central randomization will be implemented through use of Interactive Web Response System (IWRS). Approximately 510 eligible subjects will be randomized in a 1:1:1 ratio into one of the 3 arms (n=170 each) to receive either obinutuzumab in combination with chlorambucil per the package insert (Arm A), acalabrutinib 100 mg BID in combination with obinutuzumab per the package insert (Arm B), or acalabrutinib 100 mg BID (Arm C). Subjects will be randomized based on the following stratification factors:

- Presence versus absence of 17p deletion by central laboratory (yes versus no)
- ECOG Performance Status 0 or 1 versus 2
- Geographic region (North America and Western Europe versus Other)

The randomization code will be controlled through a centralized procedure and will not be known to study and site personnel or the subject before the subject is randomized into the study.

2.7.2 Blinding

In this open-label study, neither the subjects nor the investigators will be blinded to treatment. However, access to treatment randomization by IWRS will be controlled so that the sponsor's staff overseeing the conduct of the study or analyzing/summarizing data will not have aggregated efficacy or safety summary by randomized treatment arm before “unblinding.” Dummy treatment codes will be used to set up statistical programming for interim and final analysis.

Response assessment will be performed centrally by the IRC. An Independent Review Charter for IRC will be created to describe details of data review, data flow, and work flow.

3 ANALYSIS POPULATIONS

3.1 Intent-to-Treat Population

The Intent-to-Treat (ITT) population is defined as all subjects who were randomized, to be analyzed according to the arm to which they were randomly assigned, following "intent-to-treat" principle.

Unless otherwise specified, all analyses using the ITT population will be analyzed as randomized and include data only prior to crossover for Arm A subjects who crossed over to Arm C.

3.2 Safety Population

The safety population consists of all subjects who received any amount of study drug.

4 GENERAL CONVENTIONS

Continuous data will be summarized using descriptive statistics (number of observations, mean, standard deviation, median, 25th and 75th percentiles, minimum, and maximum). Frequencies and percentages will be used for summarizing categorical (discrete) data.

Confidence intervals, when presented, will generally be constructed at the 2-sided 95% level. For binomial variables, the normal approximation methods will be employed unless otherwise specified.

Calculation of time to event or duration of event endpoints will be based on the study day of the event or censoring date rather than visit number or visit label. Missing efficacy or safety data will not be imputed unless otherwise specified.

The following rules will be used for the days to cycle/month/year conversion:

- 1 cycle=28 days=4 weeks
- 1 month=30.4375 days
- 1 year=365.25 days.

All summaries will be presented by treatment arm unless otherwise specified. Data will be presented in data listings by subject number.

4.1 Definition

4.1.1 Definition of Baseline and Post-Baseline Value

For safety parameters, baseline is defined as the last measurement taken prior to the first dose of study drug administration. The post-baseline value is defined as a measurement taken after the first dose of study drug administration.

For efficacy parameters, the baseline value is defined as the last measurement taken prior to or on the date of randomization; the post-baseline value is defined as a measurement taken after the date of randomization.

The baseline and post-baseline definitions for laboratory parameters are similarly defined as those for safety parameters. For laboratory parameters used for response assessment and PROs, baseline values will be defined in reference to first dose date of study drug.

For assessments on the day of first dose where time is not captured, a nominal predose indicator, if available, will serve as sufficient evidence that the assessment occurred prior to first dose. Assessments on the day of first dose where neither time nor a nominal predose indicator are captured will be considered prior to first dose if such procedures are required by the protocol to be conducted before first dose.

For subjects who were not treated, the baseline is defined as the last valid assessment within the study.

4.1.2 Definition of Study Day

For efficacy data summary, the study day will be calculated in reference to the date of randomization. Study Day 1 for efficacy data is defined as the date of randomization. For visits (or events) that occur on or after randomization, study day is defined as (date of visit [event] – date of randomization + 1). For visits (or events) that occur prior to randomization, study day is defined as (date of visit [event] – date of randomization).

For safety data, study day will be calculated in reference to the first dose date of study drug. Study Day 1 for safety data analysis is defined as the first dose date of study drug administration. For visits (or events) that occur on or after first dose date, dose day is defined as (date of visit [event] – date of first dose of study drug administration + 1). For visits (or events) that occur prior to first dose date, dose day is defined as (date of visit [event] – date of first dose of study drug administration).

There is no Study Day 0 for efficacy or safety analysis.

4.1.3 Definition of Prior and Concomitant Therapy

For the purpose of inclusion in prior and/or concomitant medication or therapy tables, incomplete medication or radiation start and stop dates will be imputed as detailed in Section 4.2. Based on imputed start and stop dates:

- Prior medications/procedure/radiation therapies are defined as medications with a start date occurring before the date of first dose of study treatment.
- Concomitant medications/procedures/radiation therapies are defined as medications that:
    - Had start date between the first dose date of study drug, 30 days after the last dose of study drug, or the first dose date of new anticancer therapy for CLL, whichever is earlier, or
    - Had start date before first dose date and stopped or continued after first dose date.
- In addition, medications/procedures/radiation therapies that meet the criteria for both prior and concomitant medications will be classified as both prior and concomitant medication.

4.1.4 Subsequent Anticancer Therapy for CLL

For the purpose of inclusion in subsequent anticancer therapy tables, incomplete medication or radiation start and stop dates will be imputed as detailed in Section 4.2. Based on imputed start and stop dates:

- Subsequent anticancer medications/radiation therapies are defined as medications/radiation therapies that:
  - Had an indication for the primary malignancy CLL,
  - Had start date after study treatment completion or discontinuation.

The start date of subsequent anticancer therapy is defined as the first dose date of the subsequent anticancer therapy.

4.1.5 Crossover Period

The crossover period is defined only for subjects who were randomized to Arm A and crossed over, as the period of time from the first dose date of acalabrutinib to study exit date.

4.1.6 Duration of Treatment

Duration of treatment will be calculated from the date of the first dose of study drug to the date of the last dose of study drug and prior to crossover period for Arm A subjects who crossed over, as follows:

- Duration of treatment = date of last dose of study drug - date of first dose of study drug + 1 day.
- Duration of crossover treatment will be calculated using the above formula during the crossover period.

4.1.7 Time on Study

Time on study will be calculated from the date of randomization to the study exit date, as follows:

Time on study = study exit date - date of randomization + 1 day.

Time on study for Arm A subjects will not include the crossover period. Time on study for the crossover period is defined as the following:

Time on study = study exit date – the first dose date of acalabrutinib + 1 day.

4.1.8 Treatment-Emergent Period

The treatment-emergent period is defined as the period of time from the date of the first dose of study drug through 30 days after the date of the last dose of study drug or the first date starting new anticancer therapy for CLL (including acalabrutinib for Arm A subjects who crossed over), whichever is earliest.

The treatment-emergent period for crossover period is defined as the period of time from the date of the first dose of study drug (acalabrutinib) through 30 days after the date of the last dose of acalabrutinib.

4.2 Imputation Rules for Missing and Partial Data

In general, other than for partial dates, missing data will not be imputed and will be treated as missing. The algorithms for imputation of partial dates vary depending upon the parameter.

4.2.1 Adverse Events, Concomitant Medications, Subsequent Anticancer Therapies

Imputation of partial dates will be made for AE onset and stop dates, start and end dates of concomitant medication, start date of subsequent anticancer therapy, date of initial diagnosis, and death date. If dates are completely missing, no imputation will be made. For any partial date with missing year, no imputation will be made.

The general rule for imputation is:

- If only day is missing, then the 15th of the month will be used.
- If only year is present, then June 30th will be used.

If such imputation date for initial diagnosis is on or after date of randomization, then date of randomization – 1 will be used. If such imputed date for subsequent anticancer therapies is before date of last dose, then date of last dose + 1 will be used.

If the imputed date is for an AE start date and is in the same year and month as but before the first dose date, then the first dose date will be used, or if the imputed AE start date is after the AE end date, then the AE end date will be used. If the imputed date is for an AE start date and is in the same year and month as but after the last dose date + 30 days, then the last dose date + 30 days 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.

4.2.2 Death Dates

- If death year and month are available but day is missing:
    - If mmyyyy for last contact date = mmyyyy for death date, set death date to the day after the last contact date.
    - If mmyyyy for last contact date < mmyyyy for death date, set death date to the first day of the death month.
    - If mmyyyy for last contact date > mmyyyy for death date, data error and do not impute.
- If both month and day are missing for death date or a death date is totally missing, do not impute and censor the subject survival time.

4.2.3 Date Last Known Alive

- If year and month are available but day is missing, set date to the 1st date of the month.
- If both month and day are missing, set date to January 1st of the year.

4.2.4 Laboratory Values

- Laboratory values below the lower level of quantification (Q) that are reported as "<Q" or "≤Q" in the database will be imputed by Q x 0.99 for analysis purposes. However, the original value will be reported in the Listings.
- Laboratory values above the upper level of quantification (Q) that are reported as ">Q" or "≥Q" in the database will be imputed by Q x 1.01 for analysis purposes. However, the original value will be reported in the Listings.

4.3 Software

Sample size calculation was performed using East 6 (Version 6.3.1). Statistical analyses and data summary will be conducted using SAS Version 9.4 or higher.

5 STUDY POPULATION SUMMARIES

5.1 Study Disposition

Enrollment for subjects randomized will be summarized by region, country, and study center. Subject disposition will be summarized categorically and will include the number and percentage of subjects in the ITT and Safety populations. Study treatment completion status and study follow-up completion status will be summarized by treatment arms. Subject disposition during crossover period will be summarized in a separate table.

5.2 Protocol Deviations

Number and percentage of subjects with important protocol deviation (IPD) after randomization will be summarized by treatment arm. A listing of subjects with IPDs will be produced. IPD categories, subcategory codes, and descriptions will be defined by sponsor IPD guidance and used during the course of the study. The sponsor will review IPDs throughout the study prior to database lock. The final IPD list will be used to produce the IPD summary table and listing.

5.3 Demographics and Baseline

5.3.1 Demographics and Baseline Characteristics

The following demographic and baseline characteristics will be summarized:

- Age (continuous)
- Age groups:
  - o < 65, ≥ 65
  - o < 75, ≥ 75
- Sex (male, female)
- Race

- Ethnicity
- Height (cm)
- Weight (kg)
- Body surface area (BSA) (m^2)
- ECOG performance status (0 or 1, 2)
- Region (North America, South America, Western Europe, Central and Eastern Europe, Australia and New Zealand)

Reported age at informed consent signed will be used for analyses.

5.3.2 Baseline Disease Characteristics

The following baseline disease characteristics will be summarized:

- Time (month) from initial diagnosis to randomization
- Cumulative Illness Rating Scale-Geriatric (CIRS-G)
- CIRS-G > 6 by age groups (see above)
- Bulky disease (<5 cm, ≥5 cm)
- Rai stage (0, I, II, III, and IV)
- Cytopenia at baseline
    - Neutropenia - ANC ≤l.5 x 10^9/L (yes, no)
    - Anemia - Hgb ≤11 g/dL (yes, no)
    - Thrombocytopenia - platelet counts ≤100 x 10^9/L (yes, no)
    - Any of the above (yes, no)
    - All of the above (yes, no)
- 17p deletion mutation (yes, no)
- 11q deletion mutation (yes, no)
- TP53 mutation (mutated, unmutated)
- IGHV mutation (mutated, unmutated)
- Complex karyotype (yes, no)
- High risk features:
    - Del17p, TP53 mutation, del11q, or unmutated IGHV (yes, no)
    - Del17p, TP53 mutation, or del11q (yes, no)
    - Del17p and TP53 mutation (yes, no)

- β_2 Microglobulin (mg/L) group at Baseline (≤3.5 mg/L, >3.5 mg/L)
- Prior RBC transfusion within 28 days before randomization
- Prior platelet transfusion within 28 days before randomization
- Constitutional symptoms (weight loss, fever, night sweats, fatigue)
- Creatinine clearance <60 ml/min (yes, no)
- Absolute lymphocyte counts (10^9/L)
- Absolute neutrophil count (10^9/L)
- Platelet counts (10^9/L)
- Hemoglobin level (g/dL)

5.3.3 Randomization Stratification Factors

Randomization stratification factors per IWRS recording will be summarized individually and as cross-tabulation:

- Presence or absence of 17p deletion per IWRS recording (yes, no)
- ECOG performance status (0 or 1, 2)
- Geographic region (North America and Western Europe versus Other)

5.4 Medical History

General medical history data will be coded per Medical Dictionary for Regulatory Activities (MedDRA), summarized by system organ class and preferred term, and presented as a data listing.

6 TREATMENTS AND MEDICATIONS

Medications recorded on the eCRFs will be coded using the World Health Organization (WHO) Drug Dictionary.

6.1 Prior and Concomitant Medications

Prior and concomitant medications will be summarized by Anatomical Therapeutic Chemical (ATC) Class Text and WHO Drug base substance preferred name. Each subject will be counted only once for each therapeutic class, each generic name, and overall.

6.2 Subsequent Anticancer Therapies

Subsequent anticancer therapy used to treat CLL will be coded and summarized based on line of therapy and type of therapy. Each subject will be counted only once for each line and type of therapy and overall.

6.3 Study Treatment Exposure

Study treatment exposure will be summarized by as follows:

- Duration of exposure:
  - Acalabrutinib (months): (last dose date – first dose date + 1) / 30.4375
  - Obinutuzumab (days): (last dose date – first dose date + 28)
  - Chlorambucil (days): (last dose date – first dose date + 14)
- Average daily dose (mg) (for acalabrutinib only): calculated as (total dose received [mg] / duration of exposure [days])
- Relative dose intensity for acalabrutinib: calculated as (total cumulative dose received [mg] / (duration of exposure [days] * 100 [mg] * 2) *100)
- Relative dose intensity for obinutuzumab: calculated as (total volume infused (mg) / 8000 mg * 100).
- Relative dose intensity for chlorambucil: calculated as (total cumulative dose received (mg/kg) / total dose prescribed (mg/kg) from Cycles 1 to 6 * 100).

Study treatment exposure to acalabrutinib for subjects randomized to Treatment Arm A who crossed over will be summarized as follows:

- Duration of exposure: (last dose date of acalabrutinib – first dose date of acalabrutinib + 1) / 30.4375
- Average daily dose (mg): calculated as (total dose received (mg) / duration of exposure (days))
- Relative dose intensity: calculated as (total cumulative dose received (mg) / (duration of exposure (days) * 100 (mg) * 2) *100)

6.4 Study Treatment Modifications

For subjects randomized to Arm B or C, acalabrutinib dose withholding is defined as missing dose for ≥7 consecutive days, acalabrutinib dose reduction is defined as taking lower dose level (100 mg QD) for ≥3 consecutive days. Subjects with dose withholding and reduction will be summarized by percent of subjects as well as descriptive statistics.

For subjects randomized to Arm A or B, any dose withholdings with obinutuzumab, dose withholdings and reductions with chlorambucil will be summarized by percent of subjects as well as descriptive statistics.

7 EFFICACY ANALYSES

All efficacy analyses, except that for OS, will be performed in ITT population and will be analyzed as randomized and include data prior to crossover for Arm A subjects who crossed over to Arm C. OS will be analyzed based on ITT population during the entire study including the crossover period.

All efficacy analyses will be performed at the 2-sided significance level.

The following three randomization stratification factors (collected via IWRS) will be used for all stratified analyses:

- Presence or absence of 17p deletion by central laboratory;
- ECOG performance score 0 or 1 versus 2; and
- Geographic region (North America and Western Europe versus Other).

For the primary efficacy analysis of IRC-assessed PFS, if there is at least one stratum that has fewer than 2 events (where a stratum is defined as stratification factor 1 * stratification factor 2 * stratification factor 3), stratification factors will be collapsed until all strata have at minimum 2 events for the primary endpoint. The stratification factors will be collapsed in the following order:

1 Geographic region (North America and Western Europe versus Other);
2 ECOG performance status (0 or 1 versus 2).

If there is still at least one stratum that has fewer than 2 events after collapsing the 2 stratification factors above, an unstratified analysis (equivalently as to collapse all 3 stratification factors) will be performed as the primary analysis.

All stratified analyses will be conducted in accordance with this prespecified pooling strategy as above for the primary efficacy analysis of IRC-assessed PFS.

7.1 Primary Efficacy Endpoint and Analysis

7.1.1 Primary Efficacy Endpoint

The primary efficacy endpoint is progression-free survival (PFS) as assessed by IRC PFS, defined as the time from date of randomization to the date of first IRC-assessed disease progression or death due to any cause.

Subjects who do not have a confirmed disease progression or death at or prior to data analysis cutoff date will be censored as follows:

Table 2. Primary Efficacy Analysis of PFS-IRC Outcome

Situation	**Date of Progression or Censoring**	**Outcome**
PFS events include death or first IRC-confirmed disease progression that occurred at or prior to the data analysis cutoff date.		
Death before first disease assessment	Date of death	Event
IRC-confirmed PD or death between scheduled assessments	Earliest date of IRC confirmed PD or death	Event
All other cases will be censored as follows:		
No baseline tumor assessments	Randomization	Censored
No adequate post-baseline assessment	Randomization	Censored
No IRC-confirmed PD or death at the time of data cutoff (including subjects who had PD or died after cutoff)	Date of last adequate IRC assessment before data cutoff	Censored
No IRC-confirmed PD or death before study exit	Date of last adequate IRC assessment before data cutoff	Censored
No IRC-assessed PD or death before start of subsequent anticancer therapy	Date of the last adequate IRC assessment before start of subsequent anticancer therapy Therapy	Censored
IRC-assessed PD or death after start of	Date of the last adequate IRC assessment	Censored

Situation	Date of Progression or Censoring	Outcome
subsequent anticancer therapy	before start of subsequent anticancer therapy	
IRC-confirmed PD or death after 2 or more consecutively missed visits	Date of last adequate IRC assessment before the consecutively missed visits	Censored

PD=progressive disease; IRC=independent review committee.

PFS will be calculated as date of disease progression or death (censoring date for censored subjects) – randomization date + 1.

The last adequate IRC assessment is defined as the last IRC-assessed overall response that is not 'UNK' (unknown), as defined in the IRC charter.

Number and percent of censored subjects for IRC-assessed PFS will be summarized by treatment arm and follow-up time.

7.1.2 Primary Efficacy Analysis

The primary efficacy analysis is to compare PFS as assessed by IRC between Arms A and B in ITT population.

The primary efficacy analysis will be performed using a stratified log-rank test adjusting for randomization stratification factors. The estimate of the hazard ratio (Arm B/Arm A) and its corresponding 95% CI will be computed using a Cox Proportional Hazards model stratified by randomization stratification factors. Randomization stratification factors will be based IWRS randomization assignment.

A Kaplan-Meier (KM) curve will be used to estimate the distribution of PFS. The proportion of subjects who are progression free will be estimated based on KM method and its corresponding 95% CI will be calculated at selected timepoints for each treatment arm. A summary of PFS events will be provided by treatment arm.

7.2 Secondary Efficacy Endpoints and Analyses

7.2.1 Progression-Free Survival by IRC (Arms C versus A)

The key (first) secondary efficacy endpoint is PFS as assessed by IRC comparing between Arms A and C. The definition and censoring rule is the same as that described in Section 7.1.

The key (first) secondary efficacy analysis, PFS as assessed by IRC comparing between Arms A and C in ITT population, will be performed in the same fashion as that for primary efficacy analysis as described in Section 7.1.

Other secondary efficacy endpoints are listed below, and each will be compared between Arms B versus A and Arms C versus A in the ITT population.

7.2.2 Overall Response Rate by IRC

Best overall assessment is defined as the best disease outcome over all visit outcomes as assessed by IRC on or before the initiation of subsequent anticancer therapy. Best overall assessment categories include complete remission (CR), complete remission with incomplete bone marrow recovery (CRi), nodular partial remission (nPR), partial remission (PR), stable disease (SD), and progressive disease (PD). If a subject takes any subsequent anticancer therapy, best overall assessment will be re-derived based on all IRC timepoint by timepoint assessments prior to the use of subsequent anticancer therapy. ORR is defined as the proportion of subjects achieving a best overall assessment of CR, CRi, nPR, or PR at or before initiation of subsequent anticancer therapy.

Best overall assessment will be summarized by number and percentage of subjects for each response category. ORR will be summarized by number and percentage of subjects, and its corresponding 95% CI will be calculated based on normal approximation (with use of Wilson's score). ORR will be analyzed using the Cochran-Mantel-Haenzel (CMH) test, with adjustment for randomization stratification factors.

Descriptive statistics will be provided for best overall response for each treatment arm. The number and proportion of subjects within each category of response will be

presented. The proportion will be estimated by dividing the number of subjects within each category of response by the number of subjects. Each subject will be counted within only one response group, with the best response during the study as the classification group.

ORR including PR with lymphocytosis CCI assessed by IRC will also be analyzed with the same analysis method used for ORR.

7.2.3 Time to Next Treatment

Time to Next Treatment (TTNT) is defined as the time from date of randomization to the start date of non-protocol–specified subsequent anticancer therapy for CLL (or first dose date of acalabrutinib for Arm A subjects who crossed over to receive acalabrutinib) or death due to any cause, whichever comes first. Subjects who do not have the above-specified events prior to the data cutoff date will be censored at the date of last visit. TTNT will be calculated as:

(Earlier date of the start date of subsequent anticancer therapy for CLL or date of death due to any cause) – date of randomization + 1. For censored subjects, date of last visit will replace earlier date of use of subsequent anticancer therapy for CLL or date of death due to any cause in the calculation.

TTNT will be analyzed in the same fashion as that for primary efficacy endpoint as described in Section 7.1.

7.2.4 Overall Survival

OS is defined as the time from date of randomization to date of death due to any cause. Subjects who were not known to have died prior to the analysis data cutoff date will be right-censored as follows:

Table 3. Overall Survival Censoring

Situation	Censoring Date
Lost to follow-up immediately after randomization	Randomization date
Not known to have died at or prior to analysis data cutoff date	Date last known alive before analysis data cutoff date

OS will be calculated as death date (or censoring date) - randomization date + 1.

OS will be analyzed in the same fashion as that for primary efficacy endpoint as described in Section 7.1.

7.3 Sensitivity Analyses

The following sensitivity analyses will be performed for PFS as assessed by IRC between Arms B versus A and Arms A versus C in support of primary and key secondary efficacy analyses:

- Unstratified analyses;
- Subjects with the use of any subsequent anticancer therapy prior to the first IRC confirmed PD or death due to any cause will be censored at the last adequate assessment prior to the start date of the subsequent anticancer therapy;
- Subjects with PFS events after 2 or more consecutively missed visits will not be censored at the last adequate assessment. In particular, IRC-confirmed PD or death after 2 or more consecutively missed visits will be included as a PFS event;
- Excluding subjects with IPD.

7.4 Subgroup Analyses

Subgroup analyses will be performed using potential prognostic variables at screening or baseline listed below to investigate the consistency and robustness of PFS and ORR as assessed by IRC between Arms B versus A and Arms C versus A:

- Randomization stratification factors per IWRS recording:
  - Presence of 17p deletion mutation (yes vs. no)

  - ECOG at randomization (0, 1 vs. 2)
  - Geographic region (North America and Western Europe vs. Other)
- Region (North America, South America, Western Europe, Central and Eastern Europe, Australia and New Zealand)
- Age group:
  - Age <65 vs. ≥65 years
  - Age <75 vs. ≥75 years
- Sex (male vs. female)
- Race (White vs. Non-White)
- Rai Stage at screening (Stage 0-II vs. III-IV)
- Bulky disease (longest diameter of lymph node <5 cm vs. ≥5 cm at baseline)
- β2-microglobulin at baseline (≤3.5 mg/L vs. >3.5 mg/L)
- Presence of 11q deletion mutation (yes vs. no)
- TP53 mutation (mutated vs. unmutated)
- 17p del or TP53 mutation (yes vs. no)
- 17p del and TP53 mutation (yes vs. no)
- IGHV (mutated vs. unmutated)
- Del17p, TP53 mutation, del11q, or unmutated IGHV
- Del17p, TP53 mutation, or del11q
- Complex karyotype (yes vs. no)

The hazard ratio (Arm B/Arm A) and its corresponding 95% CI for each subgroup will be calculated based on an unstratified Cox regression model and displayed graphically in a forest plot.

Additional subgroups may be considered when appropriate.

7.5 Exploratory Efficacy Analyses

CCI [redacted]

7.5.1 CCI

CCI

7.5.2 CCI

CCI

7.5.3 CCI

CCI

7.5.4 CCI

CCI

CCI

7.5.5 CCI

CCI

7.5.6 PPD

CCI

8 SAFETY SUMMARIES

All safety analyses will be performed using the Safety population and will be analyzed as treated. For Arm A subjects who crossed over to Arm C, all safety analyses, except that for death, will include data only prior to crossover period.

Safety and tolerability will be assessed by the incidence of treatment emergent-adverse events (TEAEs), changes in laboratory parameters and vital signs from baseline, and ECOG performance score.

8.1 Adverse Events

Adverse event terms recorded on the eCRF will be mapped to preferred terms and system organ class through use of the MedDRA. The severity of AEs will be graded using National Cancer Institute (NCI) CTCAE v4.03 or higher for hematologic and nonhematologic AEs. The investigator will judge each event to be “not related” or “related” to study treatment.

8.1.1 Treatment Emergent Adverse Events

The TEAE is defined as adverse event reported during the treatment-emergent period (defined in Section 4.1.8). Any ongoing event that worsens in severity after the first dose date of study treatment is also defined as a TEAE. TEAEs will be tabulated in summary tables. For the purpose of determining a TEAE, incomplete onset dates will be imputed as detailed in Section 4.2.

All TEAEs will be summarized by treatment arm as treated. For each treatment arm, adverse event incidence will be summarized with frequency and percentage by system organ class and preferred term, and the denominator for the adverse event incidence will be based on the number of subjects treated in that treatment arm, unless otherwise specified. In addition, adverse event incidence rates will also be summarized by severity and relationship to study drug. Relationship to study drug is recorded on the eCRF per the investigator's judgment.

Subjects with multiple occurrences of events for a given preferred term, system organ class, or overall will be counted only once at the maximum severity and strongest relationship to study drug, respectively, for each preferred term and system organ class.

In addition, Grade 3 or Grade 4 TEAEs, TEAEs that led to permanent study drug treatment discontinuation, TEAEs that led to dose reduction, serious TEAEs, TEAEs of clinical interest, and TEAEs that resulted in death will be summarized by treatment arm as treated.

8.1.2 Adverse Events of Clinical Interest

Adverse events of clinical interest include cardiac events, cytopenia, hemorrhage, hepatic events, hypertension, infection, intestinal lung disease/pneumonitis, tumor lysis

syndrome, and secondary malignancy. Major hemorrhage is defined as any hemorrhagic event that is serious, or Grade ≥3 in severity, or that is a central nervous system (CNS) hemorrhage (any severity grade). Secondary 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-disease progression phase for overall survival. These events will be summarized similarly to TEAEs by treatment arms.

8.2 Deaths

All reported deaths will be summarized by treatment arm. A by-subject listing that includes date of death and cause of death will be provided.

8.3 Laboratory Assessments

8.3.1 Data Processing Methods

All laboratory values will be converted to and reported as SI units and classified as normal, low, or high based on the normal ranges provided by the central laboratory. In general, only data from the central laboratory will be summarized and analyzed. Hematologic parameters, including platelet counts, hemoglobin, and neutrophils, will be assessed by the grading scale for hematologic toxicity in CLL studies with use of both the IWCLL 2008 guidelines and NCI CTCAE v4.03. All other gradable laboratory parameters will be graded using the NCI CTCAE v4.03.

Per the grading scale in the IWCLL guidelines, 1) ANC: Both baseline grade and post-baseline grade are defined based on absolute values, and (2) hemoglobin and platelet: baseline grade is not applicable (no criterion is provided to define baseline grade), and post-baseline grade is based on percentage decrease from baseline value.

Gradable parameters that have criteria available for both low and high values (e.g., hypercalcemia for a high value of calcium and hypocalcemia for a low value of calcium) based on the NCI CTCAE v4.03 will be summarized for both criteria (low and high). Subjects will be counted only once for each criterion/direction. The same subject can be counted for both criteria if the subject has laboratory values that meet each criterion. Subjects who meet the criteria for Grade 1 or higher for the high direction will be

summarized under Grade 0 for summarization of the low direction and vice versa. Change from baseline to post-baseline value will be calculated for each parameter.

8.3.2 Analysis of Clinical Laboratory

Subject incidence of hematologic toxicities for platelet counts, hemoglobin, and neutrophils, graded per IWCLL 2008 guideline and CTCAE, will be summarized by treatment arm and grade.

Treatment emergent laboratory abnormality, per NCI CTCAE v4.03 grading, will be summarized for hematology and chemistry laboratory parameters of clinical interest.

Shift from baseline grade to maximum post-baseline CTCAE grade will be summarized for hematology and chemistry laboratory parameters of clinical interest.

8.3.3 Analysis of Lymphocytosis

For all subjects with baseline and any post-baseline ALC measurements, ALC at peak summary will be provided by treatment arm. Median percentage change in ALC from baseline with its 95% CI will also be displayed graphically over time.

The number and percentage of subjects with at least once occurrence of treatment-related lymphocytosis defined as an elevation in ALC of ≥50% compared with baseline and a post-baseline assessment >5,000/μL will be summarized. ALC at peak and time to peak ALC for subjects who meet the above criteria for lymphocytosis will be summarized by descriptive statistics.

Duration of lymphocytosis (DOL) is defined as the duration of time from the earliest date on which the ALC value met the lymphocytosis criteria above at a post-baseline assessment to the earliest date on which a subsequent ALC value met the resolution criteria. Resolution of lymphocytosis is defined as 1) a decrease of ALC value to the baseline level or lower or 2) an achievement of ALC value that is below 5,000/μL, whichever occurs first.

DOL = Earliest date of meeting resolution criteria -Earliest date of meeting lymphocytosis criteria + 1.

Confirmation is not required for resolution of lymphocytosis. Subjects who developed lymphocytosis but whose lymphocytosis was not resolved prior to the analysis cutoff date will be censored at the last sample date with nonmissing ALC value at or prior to the analysis cutoff date. KM curves will be used to estimate the distribution of DOL. The 50th percentile of KM estimates along with their two-sided 95% CIs will be used to estimate the median DOL.

8.3.4 Analysis of Serum Immunoglobulins

Serum immunoglobulins (IgA, IgG, and IgM) will be summarized using descriptive statistics at each scheduled post-baseline time point. An additional IgG summary to exclude subjects who received IV immunoglobulin on the study will be provided.

8.4 ECOG Performance Status

The ECOG performance status grade ranges from 0 to 5. The ECOG performance score will be summarized by descriptive statistics. Shift to maximum post-baseline score in ECOG performance score will be summarized by treatment arms.

8.5 Vital Signs

Body temperature, pulse rate (beats/min), systolic and diastolic blood pressure (mmHg), respiratory rate (breaths/min), and weight (kg) will be summarized. Systolic and diastolic blood pressure will be graded using NCI CTCAE v4.03. Shift to maximum post-baseline score will be summarized for blood pressure parameters by treatment arms.

8.6 Electrocardiogram

Subject incidence of ECG abnormality at baseline and post-baseline visits will be summarized by treatment arms. Number and percentage of subject with QTcB or QTcF values >480 msec at screening visit will be summarized in a table.

9 REFERENCES

Cheson BD, Byrd JC, Rai KR, et al. Novel targeted agents and the need to refine clinical endpoints in chronic lymphocytic leukemia. J Clin Oncol 2012;30:2820-2.

Common Terminology Criteria for Adverse Events (CTCAE) Version 4.03, USDHHS, NIH, NCI; publish date May 28, 2009 (v4.03: June 14, 2010).

Goede V, et al. Obinutuzumab plus chlorambucil in patients with CLL and coexisting conditions. N Engl J Med 2014;370:1101-10.

Guidance for Industry Clinical Trials Endpoints for the Approval of Cancer Drugs and Biologics (May 2007).

Hallek M, et al. Guidelines for the diagnosis and treatment of chronic lymphocytic leukemia: a report from the International Workshop on Chronic Lymphocytic Leukemia updating the National Cancer Institute–Working Group 1996 guidelines. Blood 2008;111:5446-56.