Official Title of Study:

A PHASE 3, MULTICENTER, RANDOMIZED, DOUBLE-BLIND, PLACEBO- CONTROLLED, PARALLEL- GROUP STUDY OF THE EFFICACY AND SAFETY OF LENALIDOMIDE (REVLIMID ®) AS MAINTENANCE THERAPY FOR PATIENTS WITH B-CELL CHRONIC LYMPHOCYTIC LEUKEMIA FOLLOWING SECOND-LINE THERAPY (THE CONTINUUM TRIAL)

NCT Number: NCT00774345

Document Date (Date in which document was last revised): October 16, 2015

STATISTICAL ANALYSIS PLAN

A PHASE 3, MULTICENTER, RANDOMIZED, DOUBLE-BLIND, PLACEBO-CONTROLLED, PARALLEL-GROUP STUDY OF THE EFFICACY AND SAFETY OF LENALIDOMIDE (REVLIMID®) AS MAINTENANCE THERAPY FOR PATIENTS WITH B-CELL CHRONIC LYMPHOCYTIC LEUKEMIA FOLLOWING SECOND-LINE THERAPY

(THE CONTINUUM TRIAL)

STUDY DRUG: CC-5013

PROTOCOL NUMBER: CLL-002

DATE FINAL: October 16, 2015

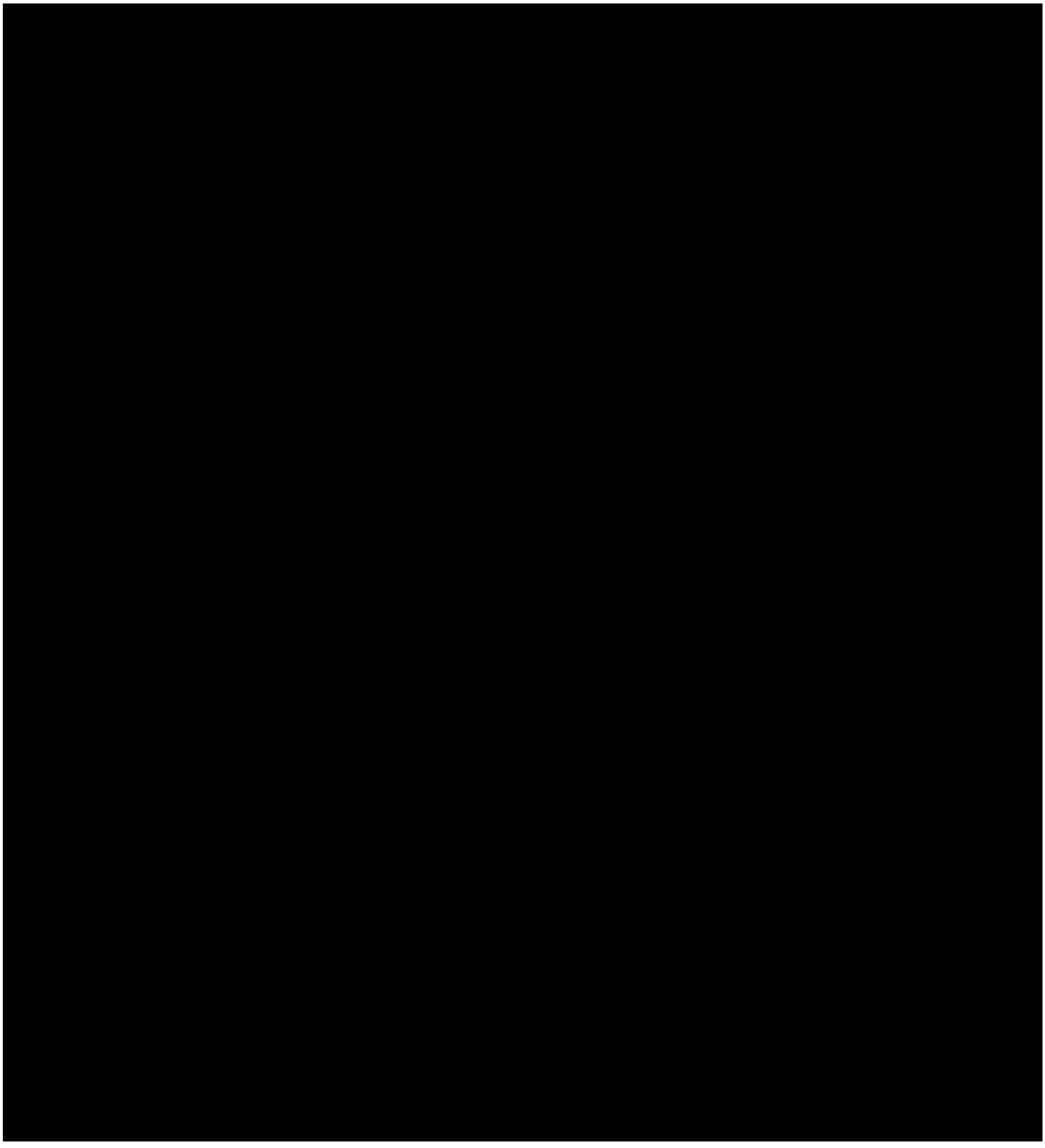
[redacted]

CONFIDENTIAL

The information contained in this document is regarded as confidential and, except to the extent necessary to obtain informed consent, may not be disclosed to another party unless such disclosure is required by law or regulations.

TABLE OF CONTENTS

LIST OF TABLES



2. LIST OF ABBREVIATIONS

Table 1: Abbreviations and Specialist Terms

Abbreviation or Specialist Term	Explanation
AE	Adverse event
ALC	Absolute lymphocyte count
ALT (SGPT)	Alanine transaminase (serum glutamate pyruvic transaminase)
ANC	Absolute neutrophil count
ASCO	American Society of Clinical Oncology
ANOVA/ANCOVA	Analysis of Variance/Covariance
AST (SGOT)	Asparate transaminase (serum glutamic oxaloacetic transaminase)
ATC	Anatomic Therapeutic Chemical
β2M	Beta-2 Microglobulin
β-HCG	Beta-human chorionic gonadotropin hormone
BUN	Blood urea nitrogen
CBC	Complete blood count
CFR	Code of Federal Regulations
CI	Confidence Interval
CLL	Chronic lymphocytic leukemia
CNS	Central nervous system
CR	Complete response
CRi	Complete response with incomplete bone marrow recovery
CRF	Case report form
CT	Computed tomography
CTCAE	Common Terminology Criteria for Adverse Events
DCF	Data Clarification Form
DLT	Dose-limiting toxicity
DMC	Data Monitoring Committee
ECG	Electrocardiogram
ECOG	Eastern Cooperative Oncology Group

Abbreviation or Specialist Term	Explanation
EMA	European Medicines Agency
F	Fludarabine
FACT-Leu	Functional Assessment of Cancer Therapy-Leukemia
FCBP	Female of childbearing potential
FDA	Food and Drug Administration
FISH	Fluorescence in Situ Hybridization
GCP	Good Clinical Practice
HGB	Hemoglobin
HCT	Hematocrit
HIV	Human immunodeficiency virus
HRQL	Human Related Quality of Life
ICH	International Conference on Harmonization
IEC	Independent Ethics Committee
IgVh	Immunoglobulin Heavy-chain Variable-region
IL-6	Interleukin 6
IL-10	Interleukin 10
IMiD	Immunomodulatory drug
IND	Investigational New Drug
IRB	Institutional Review Board
ITT	Intent to Treat
IVRS	Interactive Voice Response System
LDH	Lactate dehydrogenase
MDS	Myelodysplastic syndrome
MedDRA	Medical Dictionary for Regulatory Activities
MRD	Minimal residual disease
MTD	Maximum-tolerated dose
NCI CTCAE	National Cancer Institute Common Terminology Criteria for Adverse Events
NCI-WG	National Cancer Institute Working Group criteria for chronic lymphocytic leukemia

Abbreviation or Specialist Term	Explanation
NK cells	Natural killer cells
NOS	Not otherwise specified
nPR	Nodular partial response
NSAID	Non-steroidal anti-inflammatory drug
OS	Overall Survival
PD	Progressive disease
PFS	Progression free survival
PFS2	Progression free survival 2
PR	Partial Response
PS	Performance Status
RBC	Red blood cell
RIC	Reduced intensity conditioning
SAE	Serious adverse event
SCT	Stem cell transplantation
SD	Stable disease
SGOT	Serum-Glutamic-Oxaloacetic Transaminase
SOP	Standard Operating Procedure
SPM	Second Primary Malignancy
TNF α	Tumor necrosis factor alpha
TNM	Tumor-nodes-metastasis
TSH	Thyroid stimulating hormone
TTF	Time to treatment failure
TTP	Time to progression
ULN	Upper limit of normal
VEGF	Vascular endothelial growth factor
VEGR-1	Vascular endothelial growth factor receptor 1
VEGR-2	Vascular endothelial growth factor receptor 2
VTE	Venous thromboembolism
WBC	White blood cell count
WG	Working Group
WHO	World Health Organization

Abbreviation or Specialist Term	Explanation
ZAP 70	Zeta-Chain-Associated Protein Kinase

[redacted]

4. OBJECTIVES

Primary:

To compare the efficacy of lenalidomide versus placebo as maintenance therapy.

Secondary:

To evaluate the safety of lenalidomide *versus* placebo as maintenance therapy.

5. INVESTIGATIONAL PLAN

5.1. Overall Study Design and Plan

CC-5013-CLL-002 is a phase 3, multicenter, randomized, double-blind, placebo-controlled, parallel-group study that will compare the efficacy and safety of oral lenalidomide maintenance therapy to that of placebo maintenance therapy in subjects with B-cell CLL who achieved at least partial response (PR) with second line therapy.

The schedule of assessments for this study is included in Table 3 of the protocol. The study consists of 2 phases for each study subject: maintenance therapy and after study discontinuation.

Maintenance Phase

Subjects who meet all eligibility criteria will be randomized (1:1) by a double-blind procedure as described in Section 3.

Study treatment for each subject begins on the same day the subject undergoes randomization. Lenalidomide and placebo are administered orally (identically-matched capsules). Subjects will receive lenalidomide 2.5 mg once daily on Days 1 through 28 of the first 28-day cycle or matching placebo. If the 2.5 mg dose is well tolerated (subject does not experience any Grade 3 or 4 study drug-related toxicities or any other toxicity [Grade 1 or 2] found to be unacceptable by the investigator or the subject), subjects should be escalated starting at the second cycle to 5 mg once daily on Days 1 through 28 of each 28-day cycle or matching placebo up to disease progression. If a subject develops a Grade 1 or 2 adverse event, the drug may be interrupted until resolution, however the subject must complete one full cycle at the 2.5 mg dose level prior to escalation to the 5 mg dose level.

Subjects who complete 5 continuous cycles at the 5 mg dose level without achieving MRD-negative CR and without experiencing any Grade 3 or 4 study drug-related toxicities and without experiencing any other toxicity (Grade 1 or 2) found to be unacceptable by the investigator may be escalated to 10 mg once daily on Days 1 through 28 of each 28-day cycle or matching placebo up to disease progression.

Dose de-escalation to 7.5 mg from 10 mg and to 5 mg from 7.5 mg and to 2.5 mg from 5 mg and to 2.5 mg every other day from 2.5 mg will be permitted for those subjects who experience a Grade 3 or 4 study drug-related toxicity or any other Grade 1 or 2 toxicity found to be unacceptable by the investigator or subject.

Dose interruptions are also permitted for Grade 1 or 2 adverse events at the investigator's discretion. A guideline for the reduction of the dose of lenalidomide/placebo for DLT is provided (see Section 10.2.1 of the protocol).

Dose re-escalation is permitted if the subject is able to complete 2 full cycles at the reduced dose level without experiencing a DLT or other toxicity deemed to be unacceptable by the investigator or subject and MRD-negative CR has not been achieved yet. If that dose level is again not tolerated, the subject should be de-escalated to the dose below and remain at that tolerated dose level for the rest of the study.

Study visits and serial measurements of safety and efficacy will be performed as outlined in Table 3 of the protocol.

Subjects who discontinue maintenance therapy early (due to reasons other than PD, including unacceptable toxicity) will be followed on study until PD occurs. Subjects will be assessed for response, subsequent CLL therapies, second primary malignancies, AEs, recovery from SAEs and any new SAEs related to study procedures or prior maintenance therapy. Subjects will also be asked to continue to complete the FACT-Leu and EQ-5D questionnaires every 8 weeks.

After Study Discontinuation:

Upon experiencing disease progression, subjects will be followed for survival, information on other CLL treatments, second primary malignancies (reported as SAEs), hospitalizations, and will complete FACT-Leu and EQ-5D questionnaires at each of these visits. Patients who discontinue from progression follow up for reasons other than progression/disease transformation (i.e. withdraw consent from monthly visits, etc) should still be followed for survival, and if possible, date of progression and documentation of progression should be collected on these patients who enter the survival follow up phase prior to progression or disease transformation. Subjects will be followed until all subjects in the study have been followed for at least 5 years from randomization (or died/become lost to follow up before 5 years) or until at least 160 deaths have occurred, whichever comes later.

5.2. Study Endpoints

For all the efficacy endpoints, the determination of responses (including progression of disease) is based on the central adjudication committee (CAC) review of the CLL response data using IWCLL guidelines for the diagnosis and treatment of chronic lymphocytic leukemia [redacted] [redacted] [Appendix 22.4 of the protocol]

Co-primary Endpoints

- Progression-Free Survival (calculated as the time from randomization to the first documented progression confirmed by a blinded, independent Response Adjudication Committee or death due to any cause during or after the treatment period, whichever occurs first)
- Overall Survival (calculated as the time from randomization to death from any cause).

Secondary Endpoints

- Safety [type, frequency, and severity of adverse events (AEs) and relationship of AEs to lenalidomide]
- Tumor Response, including evaluation of minimal residual disease (MRD) by flow cytometry (improvement of the quality of response)
- Duration of response
- Health-Related Quality of Life (HRQL) by Functional Assessment of Cancer Therapy-Leukemia (FACT-Leu) and EQ-5D

- Progression Free Survival 2 (PFS2)

5.3. Data and Analysis Quality Assurance

This protocol is conducted under the sponsorship of Celgene Corporation. The data collection and management will be performed according to the Celgene Standard Operating Procedures (SOP). The quality of the analyses will be verified through a programming validation process. At a minimum, all tables of key efficacy and safety variables will be validated based on an independent programming of these results.

6. GENERAL STATISTICAL CONSIDERATIONS

6.1. Sample Size

For OS, a 61.3% improvement in median survival from randomization, from 3 years for placebo to 4.84 years on lenalidomide is considered clinically relevant. For PFS, a 61.3% improvement in median time to progression from randomization, from 20 months for placebo to 32.3 months on lenalidomide is considered clinically relevant. The OS and PFS distributions are assumed to be exponential with a constant failure (hazard) rate. When the total number of events is approximately 160 over both treatment arms, then a two-sided log-rank test with an overall significance level of 0.025 (allowing for the co-primary endpoints) would have 80% power to detect a hazard rate ratio of 0.62. To ensure timely completion of the study, 320 subjects will be enrolled, 160 in each treatment arm.

6.2. Randomization and Stratification

Stratification and randomization will be performed by IVRS, an Interactive Voice Response System, to ensure a 1:1 central randomization. Subjects who are eligible for the study will be randomized (1:1) to receive lenalidomide or placebo. Subjects will be stratified at randomization by: 1) their response to second-line induction therapy (PR, nPR, CRi or CR versus MRD-negative CR); 2) age ($\leq$ 70 versus > 70 years); and 3) presence of at least one of the following poor prognostic factors: 11q deletion, 17p deletion, unmutated IgV_H or $\beta 2M > 4.0$ mg/L (Yes versus No versus Unknown).

6.3. Analysis Populations

6.3.1. Intent-to-Treat (ITT) Population

The Intent-to-Treat (ITT) population is defined as all subjects who are randomized, independent of whether they received study treatment or not. The ITT population will be used for the primary efficacy analysis.

Subject disposition, demographics, baseline characteristics, and all efficacy analyses will be based on the ITT population unless specified otherwise. All subjects in the ITT population will be analyzed according to the treatment they were randomized to receive and not according to what they actually received, if different.

6.3.2. Safety Population

The safety population is defined as all randomized subjects who receive at least 1 dose of the study treatment (either lenalidomide or placebo).

Drug exposure and all safety analyses (including adverse events, labs, and deaths) will be based on the safety population. Subjects will be analyzed according to the initial treatment actually received.

6.4. Reporting Conventions

The algorithms, imputations, and conventions that will generally apply to programmed manipulations of the data for summary tabulations and individual subject data listings are described in this section.

Summary tables, listings, and any supportive SAS output will include a "footer" of explanatory notes that will indicate, at a minimum, the:

- Program source (e.g, SAS program name, including the path, that generates the output) and
- Data extraction date (e.g, the database lock date).

The purpose of the data extraction date is to link the output to a final database, either active or archived, that is write-protected for replication and future reference. An output date will also appear on each output page and will indicate the date the output was generated by the analysis program.

6.4.1. Tables

Data in tables will be summarized for each of the treatment groups and overall subjects when appropriate. Tables will contain the number of subjects in the analysis population for each column and the total number. Presentations will be summarized by cycle if applicable. In selected presentations the event or assessment indicating the most severe event by NCI CTCAE toxicity grade or most abnormal assessment (analysis of the "worst" case) for each subject over the course of the study will be summarized.

Time points in the summaries will reflect the time periods described in the protocol and CRF, such as 'Baseline', 'Cycle 1', 'Cycle 2', etc.

Age will be displayed in years and relative days (e.g., study days) will be displayed in days. Dosing exposure and treatment durations will also be displayed in days; all other time intervals will be presented in weeks.

The following descriptive statistics will be displayed for each quantitative variable: n, mean, median, standard deviation (SDev), minimum, and maximum.

Qualitative variables will be presented as category frequencies and percentages. The denominator for calculating percentages will be either the number of subjects in the analysis population or the number of non-missing observations in the treatment group for the particular variable presented. Visit windows will not be used in the data analyses. The cycle number recorded on the CRF, if available, will be used to mark the data collection time points. Actual visit or assessment dates will be used to calculate intervals of time.

If a given table template specifies the summarization of a given qualitative CRF variable and no such response appears on a CRF for any of the subjects in the study population appropriate to that table, the entire row may be deleted from the summary rather than present a row of zeroes for that response.

6.4.2. Listings

Data listings will be sorted by treatment group (for unblinded analysis), subject number, and cycle number, if applicable.

6.4.3. Dates

Dates will be displayed in ddmmmyyyy format (e.g., 30JUN2002).

Calculations using dates will adhere to the following conventions:

- Study days after the start day of study drug will be calculated as the difference between the date of interest and the first date of dosing of study medication (e.g, lenalidomide) plus 1 day. The generalized calculation algorithm for relative day: STUDY DAY = [(TARGET DATE – DSTART) + 1 where DSTART = the start day of study drug]. Note that Study Day 1 is the first day of treatment of study drug. For study days before the start day of study drug the algorithm is: STUDY DAY = (TARGET DATE – DSTART). Negative study days are reflective of observations obtained during the baseline/screening period.
- Age (expressed in days) is calculated: AGE = Date of randomization – DTBIRTH + 1. In practice, age will be converted to years by dividing the difference by 365.25 days, then truncating to a whole number.
- The conversion between weeks and months can be performed by using the following conversion formula (without truncation):
  MONTHS = WEEKS *7/30.4.
- Partial dates will be reported as recorded in the data listings. There will be no imputation for missing parts of dates, unless for the purpose of
  1. determining treatment emergent AEs when the AE start date is partial, or
  2. differentiating medication taken prior to the study treatment from concomitant medication taken during the study when the medication start date is partial.

The conventions for imputing partial dates for such purposes are as follows:

1. The onset date has year but both month and day missing
   - If the year of onset = year of first study drug dose, then the onset date = First study drug dose date.
   - If the year of onset < year of first study drug dose, then the onset month = 12 and the onset day=last day of the month.
   - If the year of onset > year of first study drug dose, then the onset month = 1 and the onset day=1.

2. The onset date has year and month but day missing
   - If the year and month of onset = year and month of first study drug dose, then the onset date = first study drug dose date.

- If the year of onset < year of first study drug dose, or the year of onset = year of first study drug dose but the month of onset < the month of first study drug dose, then the onset day = last day of the month.
- If the year of onset > year of first study drug dose, or the year of onset = year of first study drug dose but the month of onset > month of first study drug dose, then the onset day = first day of the month.

7. SUBJECT DISPOSITION

For all subjects randomized, subject disposition will be summarized. Subject disposition includes the number of subjects in the following categories:

- ITT Population
- Safety Population

Reasons for treatment termination will be collected on the CRF and will be summarized for all randomized subjects with the following categories:

- AEs that, in the judgment of the Investigator, may cause severe or permanent harm or which rule out continuation of study drug.
- Disease progression with or without histologic transformation
- Subject withdraws consent
- Subject lost to follow-up
- Death
- Protocol violation

A separate listing will be provided for subjects not treated.

8. DEMOGRAPHICS AND BASELINE CHARACTERISTICS

Subjects' age, weight, and other continuous demographic and baseline variables including disease characteristics will be summarized using descriptive statistics (e.g., mean, standard deviation, median, minimum and maximum), while the response to previous CLL treatment, race and other categorical variables will be summarized with frequency tabulations for each treatment arm separately and pooled over all treatment arms. Cytogenetic abnormalities will be identified and summarized by treatment arm.

Homogeneity of these variables will be assessed by t-test for continuous measures and by Fisher's exact test for categorical measures. If any significant differences ($p<0.10$) are found between treatment arms, additional exploratory analyses will be performed to assess the impact of this lack of baseline homogeneity on relative treatment effects.

Summaries for the demographics and baseline characteristics will be provided for the ITT population and safety population separately. Individual subject listings will be provided.

Medical history data will be summarized using frequency tabulations for each treatment arm by system organ class and preferred term separately and pooled over all arms using the safety population.

9. DRUG EXPOSURE

The following measures of drug exposure will be analyzed for lenalidomide and placebo.

Treatment Duration

The first dose date of lenalidomide or placebo is considered the start date of the overall study treatment. For each cycle, the lenalidomide/placebo start date will be considered the cycle start date and the day before that (after the 1st cycle) will be considered the end day of the previous cycle. The end date of the last cycle will be calculated as the start date of last treatment cycle plus 27 days, unless subject discontinues study or dies before the end of the last cycle, in which case the end date of the last cycle will be the treatment discontinuation date or the death date. The overall study treatment end date will be the same as the end date of the last cycle of the last study drug. The overall treatment duration (days) is defined as:

[(The overall study treatment end date) – (the first study drug start date) + 1]

For each study drug (i.e., lenalidomide or placebo), the treatment duration will also be calculated, and defined as:

[(The end of the last cycle of the study drug) – (the first dose date of the study drug)+1]

Cumulative Dose

Cumulative dose will be calculated for lenalidomide arm. The cumulative dose is defined as the sum of all doses taken across the treatment period (in milligrams).

Dose Exposure

Dose exposure will be calculated for lenalidomide arm. Dose exposure is defined as the number of days on the study drug during the treatment phase (excluding the periods of dose interruptions.

Average Daily Dose

Average dose will be calculated for lenalidomide arm. Average daily dose is defined as the cumulative dose divided by dose exposure (mg/day).

Dose Intensity

Dose intensity will be calculated for lenalidomide arm. Dose intensity is defined as the cumulative dose divided by treatment duration (mg/day).

Descriptive statistics of treatment duration, cumulative dose, dose exposure, average daily dose, and dose intensity, will be presented by treatment arm. In addition, descriptive statistics for treatment duration in weeks and associated frequency tables, as well as the number of cycles on study drug, will be presented by treatment arm, for both induction therapy period and the whole double-blind treatment period.

The dosing statistics as described above may be presented by cycle.

Dose reduction/interruption and re-escalation will be summarized. Summaries will include subjects who have at least one dose reduction/interruption due to AE, time to first dose reduction/interruption due to AE. Additional descriptive statistics will include, subjects who have a second dose reduction/interruption due to AE, duration of first and second dose

reduction/interruption due to AE and interval between first and second dose reduction/interruption due to AE.

Number of subjects who reached 5 mg at Cycle 2 and 10 mg at Cycle 7 will also be summarized.

As an exploratory analysis, the dosing statistics above might be summarized by treatment arm and study drug for different age groups (≤ 70 yrs and > 70 yrs) to examine the possible difference in dosing tolerability by age.

[redacted]

11. EFFICACY ANALYSIS

Efficacy analyses will be performed on the intent-to-treat (ITT) population that includes all subjects randomized. Primary analyses will be performed in two steps with a Bonferroni adjustment used to control the error rate.

The first step is to compare PFS after at least 160 subjects across both treatment arms have progressed or died. An interim analysis for OS will also be performed at this time. If the PFS difference between the two treatment groups is significant and OS results do not show a negative effect for the lenalidomide treatment group, then lenalidomide treatment will be deemed efficacious. Furthermore, if the OS interim reaches the pre-specified O'Brien-Fleming boundaries and PFS is not statistically significant but not negative for lenalidomide, then lenalidomide will be deemed efficacious. These analyses will be performed by an independent statistician not associated with the trial. The results will be shared only with the DMC who will decide the future course of the study. Specifically, if either (or both) of the two efficacy conditions described above are met, then the DMC could recommend that the blind be broken and analyses be performed by the sponsor.

The second step will be the comparison of OS. This analysis will be performed when at least 160 subjects across both treatment arms have died. This analysis will be performed regardless of the outcome of the analysis of PFS. If no significant differences in PFS are found during the first step analysis, but differences are found in OS during the second step in favor of lenalidomide treatment group, the treatment will be deemed efficacious. Exploratory analysis on PFS will also be performed at this time.

Secondary analyses will be performed to characterize change in tumor response, duration of improved response, HRQL, and PFS2. If either of the co-primary endpoints is significant, a fixed sequence testing procedure will be used for the secondary endpoints with the following order: tumor response, duration of response. The Kaplan-Meier procedure will be used to characterize the time-to-event curves in these analyses when there is censoring. Stratified log ranks tests will be used to compare the time-to-event curves. Analysis of variance/covariance (ANOVA/ANCOVA) will be used to compare HRQL and EQ-5D total scores and subscores across the two treatment arms. Otherwise, summary statistics (mean, standard deviation, median, minimum and maximum) and relevant confidence intervals will be provided.

11.1. Co-primary Endpoints and Analysis

11.1.1. Progression Free Survival

Progression-free survival will be calculated as the time from randomization to the first documented progression confirmed by a blinded, independent Response Adjudication Committee or death due to any cause during or after the treatment period, whichever occurs first. The progression date will be assigned to the earliest time when any progression is observed without prior missing assessments. If withdrawal of consent or lost to follow-up occurs before documented progression or death, then these observations will be censored at the date when the last complete tumor assessments determined a lack of progression.

For subjects who do not develop progressive disease at the time of analysis, if the response from the last visit was 'not evaluable', the most prior visit date with complete evaluation indicating no progressive disease will be used as the censor date. These rules are based on FDA guidance for cancer trial endpoints [redacted] and the following table specifies the application of the guidance for various common situations for the calculation of PFS.

Detailed decision rules for the primary and a sensitivity analysis of PFS are summarized in the following tables:

Table 2: Primary Analysis of PFS

Situation	Date of Progression or Censoring	Situation Outcome
No baseline tumor assessments	Randomization	Censored

Progression documented between scheduled visits	Earliest of:
	• Date of physical examination for lymphadenopathies, spleen and liver showing:
	  ○ appearance of new lymphadenopathies or new organ infiltration or
	  ○ increase by 50% or more in the greatest determined diameters of previously existing lymphadenopathies
	  ○ increase by 50% or more in liver or spleen size
	• Date of lab test showing 50% or more increase in lymphocytes
	• Date of post-treatment lab test showing occurrence progression of cytopenia attributable to CLL
	• Date of test confirming transformation to a more aggressive histology
		Progressed

No progression	Date of last adequate assessment with evidence of no progression	Censored
Study discontinuation for reasons other than disease progression	Date of last adequate assessment with evidence of no progression	Censored
New anticancer treatment started with no claim of progression	Date of last adequate assessment with evidence of no progression prior to the start of new anticancer treatment	Censored
Death before first PD assessment while on	Date of death	Progressed

study		

Table 2: Primary Analysis of PFS (Continued)

Situation	Date of Progression or Censoring	Situation Outcome
Death or progression after an extended lost-to-follow-up time (two or more missed assessments) while on study	Date of last adequate assessment with evidence of no progression	Censored
Death between adequate assessment visits while on study	Date of death	Progressed
Death within two months after study discontinuation	Date of death	Progressed
Death beyond two months after study discontinuation	Date of last adequate assessment with evidence of no progression	Censored

Table 3: Sensitivity Analysis for PFS

Situation	Date of Progression or Censoring	Situation Outcome
No baseline tumor assessments	Randomization	Censored
Progression documented between scheduled visits	Date of next scheduled visit	Progressed
No progression	Date of last adequate assessment with no evidence of progression	Censored
Study discontinuation for reasons other than disease progression	Date of last adequate assessment with no evidence of progression during follow-up	Censored
New anticancer treatment started with no claim of progression	Date of last adequate assessment with no evidence of progression during follow-up regardless of new anticancer treatment	Censored
Death before first PD assessment while on study	Date of death	Progressed
Death between adequate assessment visits while on study	Date of death	Progressed

Table 3: Sensitivity Analysis for PFS (Continued)

Situation	Date of Progression or Censoring	Situation Outcome
Death or progression after an extended lost-to-follow-up time (two or more missed assessments) while on study	Date of last adequate assessment with no evidence of progression	Censored
Death within two months after study discontinuation	Date of death	Progressed
Death beyond two months after study discontinuation	Date of study discontinuation visit	Censored

Kaplan-Meier product limit methods will be used to estimate the survivorship functions for the time-to-event endpoints (e.g., PFS, time to response, and OS). A two-sided log-rank test stratified by the 3 strata used in the randomization will be used as the primary analytic method to compare survivorship functions for time-to-event variables in the 2 treatment groups. In terms of the survivorship functions for each treatment group, the hypotheses of interest were:

H_0: $F_A(t) = F_P(t)$ for all t

Versus

H_1: $F_P(t) \neq F_A(t)$ for all t

where F_P is the survivorship function for placebo and F_A is the survivorship function for lenalidomide.

Median PFS will be estimated using Kaplan-Meier estimates, and the 95% confidence intervals (CI) will be computed using the method of Brookmeyer and Crowley. Hazard ratio will be calculated using Cox model stratified for the 3 strata (i.e., response to second-line induction chemotherapy, age, and presence of at least one prognostic factor) to account for the stratified randomization. The Cox model will also be used to identify prognostic factors.

Cross-tabulations will be provided by treatment group to summarize improvements from the best response during induction therapy.

For the primary analysis, the comparison of PFS between the treatment arms using the stratified log rank test, the overall two-sided significance level is 2.5%.

An unstratified log-rank test will also be performed as a secondary analysis for PFS, in addition to the stratified analysis described above.

11.1.2. Overall Survival (OS)

Overall survival is calculated as the time from randomization to death from any cause. OS will be censored at the last date that the subject was known to be alive for subjects who were alive at the time of analysis and for subjects who were lost to follow-up before death was documented.

The analysis of OS will include survival information for all randomized subjects. Subjects who discontinued from the treatment phase of the study and who had possibly received other anti-cancer therapies and then subsequently died will be included in the analysis as death. However, sensitivity analyses will be performed in which these subjects will be censored at the date of the first dose date of the anti-cancer therapy.

OS will be compared between treatment arms using the stratified log-rank test at 2-sided 0.025 alpha level. The same methods used for the PFS analysis will be used for the OS analysis.

11.2. Secondary Endpoints and Analysis

11.2.1. Change in Tumor Response

Tumor response, including PD, will be assessed according to the IWCLL guidelines for the diagnosis and treatment of chronic lymphocytic leukemia (Hallek, 2008). Subjects' tumor response status change based on the best response during the treatment period relative to the response at baseline (i.e., from PR, nPR, CRi, CR, MRD-negative CR to PD, nPR, CRi, CR, MRD-negative CR) will be compared in the two arms using Van Elteren's test at a two-sided 0.05 significance level stratified for the 3 strata used at randomization. Responses from subjects after they received other anti-cancer treatments will not be counted; however, these subjects will be included in the denominator. Cross tabulations will be provided to summarize response frequency change comparing to baseline.

11.2.2. Duration of Improved Response

For those subjects who show improvement from their best response achieved during induction therapy while on study treatment (e.g. PR to CR or CR to MRD-CR), the duration of the improved response will be summarized. The duration of improved response will be calculated at the time from first improvement to the time of disease progression. Duration of improved response will be censored at the date of last assessment showing no worsening from the improved response. A two-sided unstratified log rank test at 0.05 significance level will be used to test the duration of improved response between the two treatment arms.

11.2.3. Health-Related Quality of Life (HRQL)

For Health-Related Quality of Life (HRQL) by Functional Assessment of Cancer Therapy-Leukemia (FACT-Leu) and EQ-5D, changes from baseline in overall score and sub-scores will be analyzed using repeated measures of ANOVA/ANCOVA using the baseline scores as covariates where appropriate.

Various schemes will be assessed for missing data imputation if needed.

11.2.4. Progression Free Survival 2 (PFS2)

PFS2 is defined as the time from randomization to second objective disease progression, or death from any cause, whichever occurred first. Patients alive and for whom a second objective disease

progression has not been observed should be censored at the last time known to be alive and without second objective disease progression. In situations where OS and PFS2 cannot reliably be determined, it may be possible to rule out significant lack of efficacy of further treatments by looking at the outcome of the next line therapy. For this analysis, an event is defined as second objective disease progression, or death from any cause, or the start of the CLL therapy after next line treatment, whichever occurred first [redacted]

11.3. Subgroup Analysis

The effect of treatment on the efficacy endpoints of PFS and OS will be compared between treatment arms within subgroups defined by the following variables:

- Age group (≤ 70, > 70 years)
- Response to 2nd line therapy: (PR, nPR, CRi or CR *versus* MRD-negative CR)
- Presence of at least one of the following poor prognostic factors: 11q deletion, 17p deletion, unmutated IgV_H or β2M > 4.0 mg/L (Yes versus No versus Unknown)

The methods described in Sections 11.1.1 and 11.2.1 for the endpoints of PFS and OS respectively, will be used for each subgroup.

[redacted]

12. SAFETY ANALYSIS

All safety analyses will be based on all treated subjects and will be presented by treatment group. Safety measurements will include adverse events, clinical laboratory information, vital signs, and deaths. All analyses of safety data will be conducted using the safety population unless specified otherwise.

12.1. Adverse Events

Adverse events (AE) will be classified using the Medical Dictionary for Regulatory Activities (MedDRA) classification system. The severity of the toxicities will be graded according to the NCI CTCAE version 3.0 whenever possible.

The adverse event of infection will also be analyzed according to the grading system as recommended by the IWCLL guidelines for the diagnosis and treatment of CLL [redacted] (Appendix 22.4).

Treatment emergent adverse events (AEs) will be summarized by treatment group and overall. A treatment emergent adverse event is defined as any AE occurring or worsening on or after the first treatment of any study drug, and within 30 days after the last dose of the last study drug received. Adverse events will be summarized by treatment arm and overall. Adverse events are documented on the CRF together with their severity, according to the NCI CTCAE version 3.0 or higher, also referred to as NCI toxicity grading. For the categorization of the adverse events, the Medical Dictionary for Regulatory Activities (MedDRA®) dictionary (version 16.0) will be used.

AE frequency will be tabulated by body system, MedDRA preferred term for each treatment regimen during the Treatment Phase as well as for the Follow-up Phase when appropriate. In the by-subject analysis, a subject having the same event more than once will be counted only once. AEs will be summarized by worst NCI CTCAE version 3.0 grade. In the case that the AEs or event frequencies are judged to be clinically important, an exact test may be used to analyze the difference between the treatment groups. In addition, AEs of special interest (Event of Interest),

including second primary malignancies (SPM), will be defined and summarized by treatment

arm.

AEs leading to death or to discontinuation from treatment, events classified as NCI CTCAE version 3.0 Grade 3 or higher, study-drug-related events, serious adverse events (SAEs) and events of interest (including second primary malignancies) will be summarized separately.

Individual AE listings by subject will be provided.

[redacted]

Adverse events (AE) will be classified using the Medical Dictionary for Regulatory Activities (MedDRA) classification system. The severity of the toxicities will be graded according to the NCI CTCAE version 3.0 whenever possible.

The adverse event of infection will also be analyzed according to the grading system as recommended by the IWCLL guidelines for the diagnosis and treatment of CLL [redacted] (Appendix 22.4).

12.2. Clinical Laboratory Evaluations

Laboratory data will be graded according to NCI CTCAE version 3.0. Shift tables demonstrating the changes from baseline to the worst severity grade observed during the treatment phase will be displayed for each treatment arm for relevant lab parameters for hematology, chemistry, and urinalysis.

12.3. Vital Sign Measurements

For vital signs, cross-tabulations showing the number of subjects with values below, within and above the normal ranges pre-treatment versus post-treatment will be summarized by treatment arm for the treatment phase. For weight, means, medians, standard deviations, minimum, and maximum will be provided by cycle.

12.4. Electrocardiogram (ECG)

ECG results by visit will be summarized by treatment arm. In addition, shift tables for ECG evaluations from baseline to the worst category during the treatment will be provided for each treatment arm.

12.5. Deaths

Deaths during the treatment phase and within 30 days from the last dose of study treatment will be tabulated by the primary cause of death and treatment arms. Separate summary tables will be made for the deaths that are suspected to drug related and for those that are considered non-drug related.

The same summary tables will be provided for the deaths that occur during follow-up phase separately.

13. OTHER TOPICS

13.1. Quality of Life

The Health-Related Quality of Life (HRQL) outcomes assessment include Functional Assessment of Cancer Therapy-Leukemia (FACT-Leu) and EQ-5D.

The subscale scores at each assessment as well as the changes from baseline for post-baseline assessments will be summarized and compared between treatment groups using repeated measures of ANOVA/ANCOVA using the baseline scores as covariates where appropriate. In addition, frequency tables for categorized change from baseline scores (Improvement, No Change, and Worsening) will be presented by treatment group.

Various schemes will be assessed for missing data imputation if needed.

14. INTERIM ANALYSIS

14.1. Administration of the Interim Analyses

An independent DMC will be convened which will be composed of medical oncologists with experience in treating subjects with CLL and a statistician, all of whom are not otherwise involved in the study as investigators. An independent unblinded statistician will perform the interim analyses. The DMC will review safety data on an ongoing basis including adverse events and clinical laboratory data. During the course of the study, the DMC will review the efficacy data in accordance with the guidelines for the pre-planned interim analyses. Operational details for the DMC will be provided in the DMC charter.

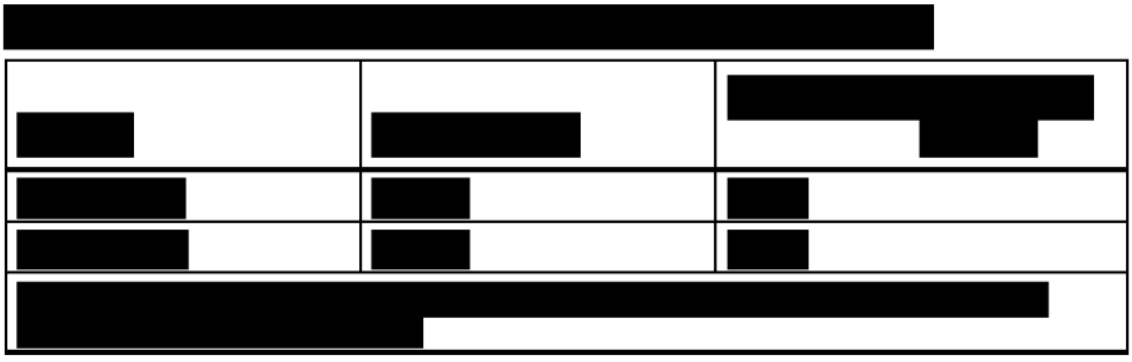
The interim analysis results will not be disseminated among investigators and those directly involved with the study conduct unless recommended by the DMC.

14.2. Statistical Approaches for Control of Alpha and Other Considerations

The co-primary efficacy variable, OS, will be compared between the 2 treatment arms using a group sequential log-rank test corresponding to 2 unequally spaced analyses: one interim analysis at approximately 60% information for OS (performed at the same time as the primary analysis for PFS, ie, after about 160 subjects progressed or died) and one final analysis at 100% information after 160 subjects die. The boundary for declaring superiority for the treatment arm is based on an alpha-spending function of the O'Brien-Fleming type with overall $\alpha = 0.025$, two-sided.

At the interim analysis for OS, a log-rank statistic will be calculated and compared with the upper (superiority) boundary. If the value of the log-rank statistic is above this upper boundary, a claim of superiority for the treatment arm could be considered.

Equivalently, the p-value for the two-sided log-rank test will be calculated and a superiority claim for the lenalidomide arm considered if the p-value is less than the nominal p-value corresponding to these boundaries. The table below gives the nominal p-values for rejecting the null hypothesis corresponding to the O'Brien-Fleming boundaries. (The interim bound is approximate and will be re-calculated based on the amount of information available when the interim analysis is actually performed.)



In addition to the log-rank test comparing the 2 arms, conditional and predictive power will be calculated at this interim analysis, together with estimation of the hazard ratio and its 95% confidence interval.

[redacted]