Statistical Analysis Plan

Study Code	D822BC00001
Edition Number	2.0
Date	19 December 2023

A Randomized, Multicenter, Open-Label, Phase 3 Study to Evaluate the Efficacy and Safety of Acalabrutinib versus Chlorambucil plus Rituximab in Subjects with Previously Untreated Chronic Lymphocytic Leukemia

TABLE OF CONTENTS PAGE

LIST OF ABBREVIATIONS

Abbreviation or special term	Explanation
AE	Adverse event
AESI	Adverse event of special interest
ALC	Absolute lymphocyte count
ALT	Alanine aminotransferase
AST	Aspartate aminotransferase
ATC	Anatomical Therapeutic Chemical
BICR	Blinded independent central review
BOR	Best overall response
CI	Confidence interval
CLL	Chronic Lymphocytic Leukemia
CMH	Cochran Mantel Haenszel
CR	Complete response
CRi	Complete response with incomplete bone marrow recovery
CSP	Clinical Study Protocol
CSR	Clinical Study Report
CT	Computed tomography
CTCAE	Common Terminology Criteria for Adverse Events
DAE	Discontinuation of investigational product due to adverse events
DBL	Database lock
DCO	Data cut-off
DoR	Duration of response
eCRF	Electronic case report form
ECI	Adverse Events of Clinical Interest
EDR	Early discrepancy rate
CCI [redacted]	[redacted]
EoT	End of treatment
CCI [redacted]	[redacted]
CCI [redacted]	[redacted]
FAS	Full Analysis Set
HR	Hazard ratio

Abbreviation or special term	Explanation
HRQoL	Health related quality of life
ICH	International Conference on Harmonisation
IDMC	Independent data monitoring committee
ITT	Intention-to-treat
IVIG	Intravenous immunoglobulins
IRT	Interactive Response Technology
IWCLL	International Workshop on Chronic Lymphocytic Leukaemia
KM	Kaplan-Meier
LD	Longest diameter
MRD	Minimal Residual Disease
NA	Not applicable
NDA	New drug application
NE	Not evaluable
NTL	Non-target lesion
OSAE	Other significant adverse events
ORR	Overall response rate
OS	Overall survival
RTSM	Randomization and Trial Supply Management
PD	Progressive disease
PFS	Progression free survival
PK	Pharmacokinetics
PR	Partial response
PRL	Partial response with lymphocytosis
CCI	[redacted]
RDI	Relative dose intensity
SAE	Serious adverse event
SAP	Statistical Analysis Plan
SAS®	A commercially available integrated system of software products, commonly used for reporting and analysis of Clinical Studies
SD	Stable disease
SoA	Schedule of assessments

Abbreviation or special term	Explanation
TBL	Total bilirubin
TEAE	Treatment emergent adverse events
TL	Target lesion
TTNT	Time to next therapy
VAS	Visual analogue scale
WHO	World Health Organization

AMENDMENT HISTORY

Section Number Change refers to:	Date	Description of change	Change from CSP?	Rationale
N/A	15Nov2019	Initial approved SAP	N/A	N/A
General considerations	V2.0(19 December 2023)	Updated throughout to align with CSP V4.0.	No	To align with CSP
List of Abbreviations		Updated to align with the updated content of the SAP. Added abbreviation of WHO.	No	N/A
Section 1.3		Updated to align with CSP V4.0.	No	To align with CSP
Section 2.1		• Added the following definition to the FAS section: "Unless otherwise specified, all analyses using the FAS will include data only prior to crossover for patients from Arm B." • Updated Safety analysis set section to remove mention of formal analysis and updated that analyses will also be performed for crossover period. • Crossover analysis set has been renamed "Crossover period."	No	To improve previous content
Section 2.2		• Updated wording of presentation of violations and deviations in CSR to listed and summarized. • Update Description of Deviation Table as per IPD handling plan and deleted last column in table.	No	To improve previous content

		Remove reference to Deviation 3.		
Section 3.2.5		Removed specification of "tumor response" and updated "Not evaluable (NE)" to "Unknown (UNK)." This is also updated throughout the SAP document.	No	To improve previous content
Section 3.2.6		Deleted reference to BICR identifying disease progression, but not identified by investigator.	No	To improve previous content
Section 3.3.1		• Rephrase PFS calculation definition to include unit of measure. • Added in table number for "PFS: Definition of 2 Missed Visits". • Updated assessment details in Table 4. Discrepancy rates updated to Concordance rates	No	To improve previous content
Section 3.3.2		Renamed "Objective response rate" to "Overall response rate." This is also updated throughout the SAP document.	No	To align with CSP
Section 3.3.3		BOR definition update.	No	To align with CSP
Section 3.3.5		Added further definitions of subsequent anticancer therapies.	No	To improve previous content
Section 3.4.2		CCI [redacted]	No	To improve previous content
Section 3.5.1		Dose Intensity calculation update.	No	To improve previous content
Section 3.5.2		• AEs of special interest, added in Appendix A reference and made subsequent updates to Appendix A.	No	To improve previous content

		• Removed OSAE section. • Updated the ECI section. Grammar update from "…the start of a subsequent therapy (including crossover treatment, whichever occurs first)." to "the start of a subsequent therapy (including crossover treatment), whichever occurs first."		
Section 3.5.4		Section 3.5.4 has been moved and combined with Section 4.2.6.2.	No	To improve previous content
Section 4.1		Added definition of how to calculate baseline of subjects who did not receive treatment.	No	To improve previous content
Section 4.2.1		Updated to align with CSP V4.0.	No	To align with CSP
Section 4.2.2.2		• PFS – "…treatment status" reference deleted. • PFS – reference to "…distribution of number of days prior to progression…" deleted. Disagreements between BICR and investigator. Paragraph rephrased and moved after "Progression Free Survival sensitivity analyses" bullet points.	No	To improve previous content
Section 4.2.2.8		Removed reference to EAP.	No	To improve previous content
Section 4.2.3		Updated to align with CSP V4.0.	No	To align with CSP
Section 4.2.6.2		• Added in TEAE definition. • SAEs reported by SOC and PT.	No	To improve previous content

		• AEs and drug-related AEs leading to treatment discontinuation. • AESI updated to be reported by PT. • ECI section has been updated. Added in AESI summary description.		
Section 4.2.6.3, 4.2.6.4 & 4.2.6.5		Updated the time windows.	No	To improve previous content
Section 4.2.6.7		Added physical examination section.	No	To align with CSP
Section 4.2.7		• Rephrased to patient recruitment region/country. • Removed child bearing potential (females only) from demographics summary. • Removed weight group from patient characteristics at baseline and added BMI, BMI group. • Added eCRF summary for stratification factors. • Removed Binet staging from disease characteristics at baseline summary and added detail about age section.	No	To align with CSP; To improve previous content
Section 4.2.10		Section added for COVID_19 and China cohort analysis section is renumbered to 4.2.11.	No	To align with CSP
Section 5		Accrual period and Interim analysis updated to 30 months and 40 months, respectively, as per CSP V4.0.	No	To align with CSP

* Pre-specified categories are
Primary or secondary endpoints; Statistical analysis method for the primary or secondary endpoints; Derivation of primary or secondary endpoints; Multiple Testing Procedure; Data presentations; Other

1. STUDY DETAILS

This statistical analysis plan (SAP) contains a more detailed description of the analyses described in the clinical study protocol (CSP) with the exception of analyses described in section 9.5.5, regarding combined PK analyses across different studies. This SAP is based on version 4.0 of the CSP.

1.1 Study Objectives

Table 1 Study Objectives

Primary objective:	**Endpoint/variable**
To compare the efficacy of acalabrutinib relative to chlorambucil plus rituximab in subjects with previously untreated chronic lymphocytic leukemia without del(17p) or TP53 mutation.	Progression free survival is defined as time from randomization until progression per the International Workshop on Chronic Lymphocytic Leukemia 2018 criteria as assessed by blinded independent central review or death due to any cause.
Secondary objectives:	**Endpoints/variables:**
To compare acalabrutinib relative to chlorambucil plus rituximab on objective response rate, duration of response, time to next treatment, and overall survival.	• Overall response rate is defined as the proportion of patients who have a complete response, complete response with incomplete bone marrow recovery, nodular partial response or partial response, as determined by blinded independent central review and investigator per International Workshop on Chronic Lymphocytic Leukemia 2018 criteria. • Duration of response is defined as the time from the date of first documented response until date of documented progression or death in the absence of disease progression, as determined by blinded independent central review and investigator. • Time to next therapy is defined as time from randomization until institution of non-protocol specified treatment for chronic lymphocytic leukemia. • Overall survival is the length of time from randomization until the date of death due to any cause.
To evaluate minimal residual disease negativity rate during treatment and at the end of treatment.	Minimal residual disease negativity rate (peripheral blood) at the start of Cycle 9.

Table 1 Study Objectives

To characterize the pharmacokinetics of acalabrutinib and its major metabolite (ACP-5862).	• Summarized plasma concentrations of acalabrutinib and ACP-5862 at specified time points. • Pharmacokinetic parameters by population analyses as appropriate.
Safety Objective:	**Endpoints/variables:**
To assess the safety and tolerability of acalabrutinib as compared to chlorambucil plus rituximab in subjects with previously untreated chronic lymphocytic leukemia without del(17p) or TP53 mutation.	Safety and tolerability will be evaluated in terms of adverse events, vital signs, clinical laboratory, physical examinations, and electrocardiogram. Assessments related to adverse events cover: • Occurrence/frequency. • Relationship to investigational product as assessed by investigator. • Common Terminology Criteria for Adverse Events severity grade. • Seriousness. • Death. • Adverse events leading to discontinuation of investigational product. • Adverse events leading to dose reduction of investigational product. • Adverse events leading to dose delay of investigational product. Vital signs parameters include systolic and diastolic blood pressure, and pulse rate, body temperature. Assessments cover: • Observed value. • Absolute change from baseline values over time.
Exploratory Objectives:	**Endpoints/variables:**
CCI [redacted]	[redacted]

Note: Sensitivity analysis of PFS will be performed based on the investigator's assessment according to International Workshop on Chronic Lymphocytic Leukemia (IWCLL) 2018. CCI [redacted]

1.2 Study Design

This randomized, regional, multicenter, open-label, Phase 3 study will evaluate the efficacy and safety of acalabrutinib monotherapy (Arm A) versus chlorambucil plus rituximab (Arm B) in patients with previously untreated Chronic Lymphocytic Leukemia (CLL) without del(17p) or TP53 mutation.

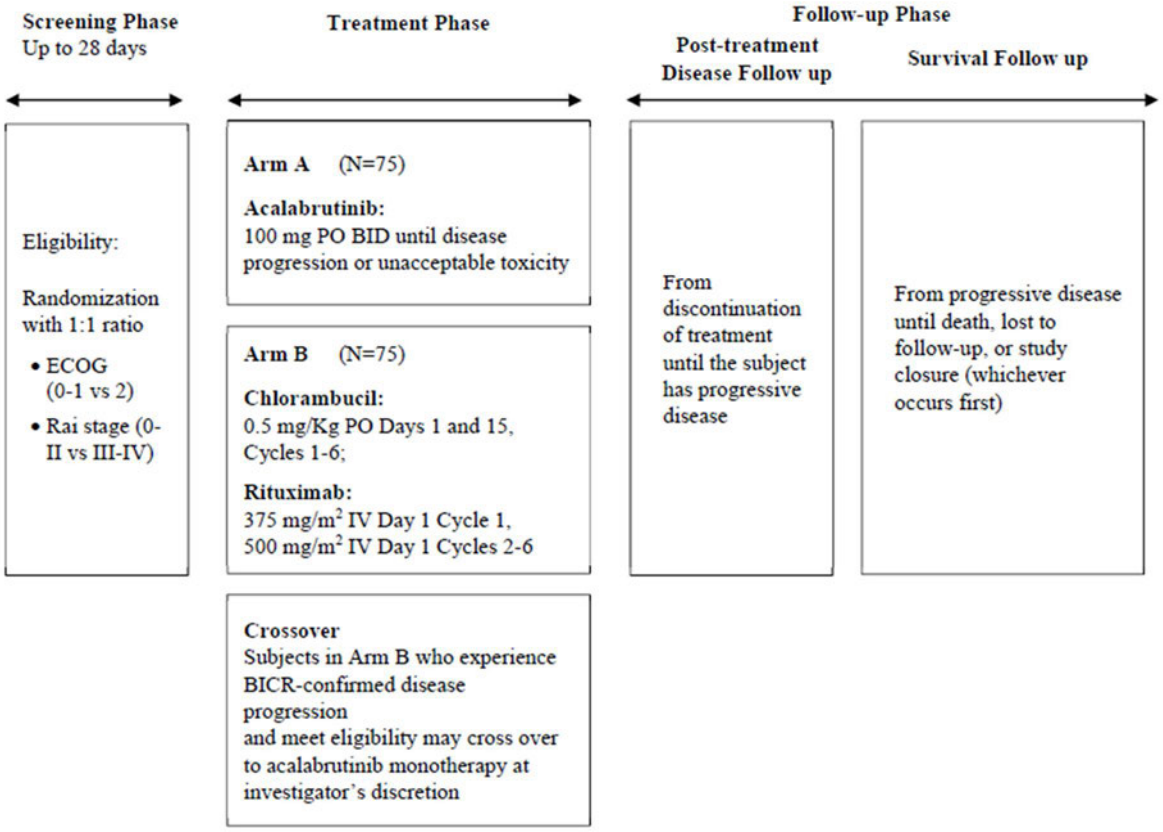
Study assessments will be conducted in accordance with the schedule of assessments, which can be found in the CSP, section 1.1. See Figure 1 below for a graphical representation of the study design.

Patient participation will include a screening phase, treatment phase, post-treatment disease follow-up phase, and survival follow-up phase. The screening phase will last up to 28 days before the first dose of study drug, during which the patient's eligibility and baseline characteristics will be determined. The treatment phase will last from randomization until study drug(s) discontinuation for the purpose of efficacy analyses, and from first dose to study drug(s) discontinuation for safety analyses. Treatment with acalabrutinib may be continued until disease progression or any other treatment discontinuation criterion is met. Treatment with chlorambucil/rituximab includes maximum 6 cycles, or until there is disease progression or any other treatment discontinuation criterion is met. In both treatment arms a cycle is defined as 28 days (±3 days for Arm B).

Patients randomized to Arm B who have blinded independent central review (BICR) confirmed disease progression may be eligible to crossover and receive single agent acalabrutinib 100 mg BID at investigator discretion provided they meet the crossover eligibility criteria, as defined in the CSP. An end of treatment (EoT) visit is required for safety assessments for any patients who permanently discontinue study drug for any reason (except for death, lost to follow-up, or withdrawal of consent), including disease progression. The EoT visit should be performed within 7 days of the last dose of all study drugs, if possible, and is not required for patients who discontinue from study drug within 10 days of a scheduled study visit, or if the EoT visit would be performed within 14 days of the safety follow-up (SFU) visit.

The survival follow-up will begin after BICR or investigator-confirmed progressive disease. During this phase the following information will be recorded: subsequent anti-CLL therapy including best response, IWCLL indication for treatment initiation, additional malignancy occurrence, and survival status. The survival follow-up will continue until death, lost to follow-up, consent withdrawal, or study closure, whichever occurs first.

Figure 1 Study design



1.3 Number of Patients

Under the exponential model assumptions, the study is expected to randomize approximately 75 patients per arm (150 patients in total from both arms) with 1:1 randomization ratio. It is planned to randomize approximately 75 to 120 patients in total (50% to 80% of total sample size) from China. Randomization will be stratified by ECOG [0-1 versus 2] and Rai stage [0-II versus III-IV]. The study is sized to achieve more than 95% power to detect a hazard ratio of 0.333 in PFS (which translates into an improvement in median PFS from 30 months to 90 months) at the 2-sided significance level of 0.05, allowing for one interim analysis conducted at approximately 76% of the target events. The sample size calculation assumes a median PFS of 30 months in Arm B (Michallet et al 2018).

The final analysis of BICR-assessed PFS is event-driven and will be conducted when enrollment is completed and there are approximately 50 BICR-assessed PFS events in total from Arm A and Arm B. The accrual period is assumed to be approximately 30 months. The dropout rate is assumed to be 5% by 14 months for both arms. The interim and final analysis are anticipated to occur approximately 40 months and 53 months, respectively, after the first patient is randomized.

2. ANALYSIS SETS

2.1 Definition of Analysis Sets

Full analysis set (FAS)

Includes all randomized patients. Patients will be analyzed by treatment group assigned in accordance with the randomization, regardless of the treatment actually received. Patients who were randomized but did not subsequently receive treatment are included in the FAS. The analysis of data using the FAS therefore follows the principles of intent to treat (ITT).

Unless otherwise specified, all analyses using the FAS will include data only prior to crossover for patients from Arm B. Overall FAS is defined as all randomized patients in the study.

Safety analysis set

Consists of all patients who received at least 1 dose of study drug. Safety data will be summarized using the safety analysis set, according to the actual treatment received. If a patient receives study drug(s) other than what was assigned, the patients will be summarized in the treatment arm of the study drug first received.

PK analysis set

All randomized patients who receive at least one dose of study drug, for whom there is at least one reportable post-dose PK concentration available, will be included in the PK analysis set.

In case there was evident violation of Good Clinical Practice (GCP), patients might be removed from above analysis set(s). Such occasions will be documented and shared with health authority.

China cohort

The China cohort is a subset of overall population and consists of all patients recruited from sites in China.

The China full analysis set (China FAS) will include all patients randomized in the China cohort and will be used for all China only efficacy analyses.

The China safety analysis set will consist of all patients recruited in the China cohort who received at least one dose of investigational product and will be used for all China only safety analyses.

Patients randomized in the China cohort are therefore included in both the overall FAS and the China FAS per the above definitions.

In addition to the evaluation of overall data for primary, secondary, safety and exploratory objectives, evaluation of consistency in efficacy and safety in the China population is required to facilitate the benefit-risk assessment for Chinese patients. Hence the safety and efficacy data in the China cohort will be analyzed separately where the same endpoint definitions and the same analysis methods are applied.

The China PK analysis set will consist of all patients recruited in the China cohort who received at least one dose of investigational product, for whom there is at least one reportable post-dose PK concentration.

Definitions of the analysis sets for each outcome variable are provided in Table 2.

Table 2 Summary of outcome variables and analysis sets

Outcome variable	Analysis set
Efficacy data	
PFS (primary)	FAS
ORR, DoR[1], TTNT, OS, MRD rate	FAS
CCI	FAS
Study population/Demography data	
Patient disposition	FAS[2]
Demography characteristics	FAS
Baseline and disease characteristics	FAS
Important deviations	FAS
Medical/surgical history	FAS
Concomitant medications/procedures	FAS
Subsequent anti-cancer therapy[3]	FAS
PK data	
PK concentrations	PK
Safety data	
Exposure	Safety
Adverse events	Safety
Laboratory measurements	Safety
Vital signs	Safety
ECG	Safety
ECOG PS	Safety

[1] *Patients who are evaluable for the analysis of DoR are those who responded in the ORR analysis.*

[2] *All patients will be included in the summary of screen failures.*
[3] *Includes both medication and radiotherapy.*

2.2 Violations and deviations

The following table describes the deviations that will be considered important protocol deviations. Important protocol deviations will be programmatically identified within the clinical database (where feasible) or detected via site monitoring. These will be listed, summarized and discussed in the CSR as appropriate:

Deviation Number	Description of the deviation (based on Protocol Deviations Plan V3.0)
1	Key Inclusion Criteria Deviations, participants who were enrolled or randomized but did not meet critical inclusion criteria.
2	Key Exclusion Criteria Deviations, participants who were enrolled or randomized but did not meet critical exclusion criteria.
3	Discontinuation Criteria for study product met but patient not withdrawn from study treatment.
4	Discontinuation Criteria for overall study withdrawal met but participant not withdrawn from study.
5	Investigational Product (IP) Deviation.
6	Excluded Medications taken
7	Deviations to study procedure
8	Other Important Protocol Deviations (including missing PI eCRF signature [corporate IPD], other to be agreed by Study Team)

The contents and details of this list may be expanded in a separate document, if considered appropriate.

Patients who receive the wrong treatment at any time will be included in the safety analysis set as described in Section 2.1. During the study, decisions on how to handle errors in treatment dispensing (with regard to continuation/discontinuation of study drug or, if applicable, analytically) will be made on an individual basis with written instructions from the study team leader and/or statistician.

The important protocol deviations will be listed and summarized by randomized treatment arm. None of the other deviations will lead to patients being excluded from the analysis sets described in Section 2.1 (with the exception of the PK analysis set, if the deviation is considered to impact upon PK).

A per protocol analysis excluding patients with specific important protocol deviations is not planned, however, a 'deviation bias' sensitivity analysis may be performed on the progression free survival endpoint excluding patients with derivations that may affect the efficacy of trial therapy if >10% of patients in either treatment group:

- Did not have the intended disease or indication or

- Did not receive any randomized therapy.

The need for such a sensitivity analysis will be determined following review of the protocol derivations ahead of database lock and will be documented prior to the primary analysis being conducted.

In addition to the programmatic determination of the deviations above, other study deviations captured from the eCRF module for inclusion/exclusion eligibility criteria will be tabulated and listed. Any other deviations from reports will be reported in an appendix to the CSR.

3. PRIMARY AND SECONDARY VARIABLES

3.1 Handling of crossover patients

Patients in Arm B who have BICR-confirmed progression will have the option to crossover to acalabrutinib. For all efficacy analyses, except for the analysis of OS, will include data prior to crossover. OS will encompass the entire study duration.

All safety analyses, except that for death, will include data prior to crossover unless noted otherwise. Death will be summarized in the following three ways: over the randomized period, crossover period and entire study duration.

The crossover period is defined only for patients who were randomized and received at least one dose of crossover therapy.

Prior to crossover first/last dose refers to the first/last dose of chlorambucil/rituximab. Post crossover first/last dose refers to the first/last dose of acalabrutinib.

3.2 IWCLL 2018 Visit Responses

For all patients, tumor response and progression will be determined at each scheduled visit as per the schedule of assessments according to the IWCLL 2018 criteria through the evaluation of physical exams, B-symptoms, radiologic evaluations, hematologic evaluations and bone marrow exams.

For the purposes of all the analyses described, IWCLL 2018 responses will be analyzed as determined by BICR or the investigator. Responses will not be derived programmatically.

Response evaluations will be done every 12 weeks (±14 days) with the first on-treatment assessment occurring on Cycle 4 Day 1, the second on Cycle 7 Day 1, and so on through Cycle 25. After Cycle 25 response evaluations will be done every 24 weeks (±14 days) thereafter, until disease progression or death.

3.2.1 Radiologic tumor assessments

Baseline radiologic tumor assessments will be performed no more than 28 days before the start of randomized treatment. Up to 6 measurable disease sites (only nodal lesions $\geqslant$ 1.5 cm

in the longest diameter, clearly measurable in two perpendicular dimensions), will be followed as target lesions for each patient. If additional lesions are present but are not included in the target lesion assessment, they can be added as non-target lesions followed throughout the study.

Post-baseline radiologic tumor assessments must be performed up to 7 days prior to response evaluation.

The cranial-caudal measurement of the spleen and longest diameter of the liver will also be assessed at screening and at all subsequent response evaluations.

3.2.2 Physical exams and hematologic evaluations

Physical examinations and hematologic evaluations are also required within 28 days of the start of randomized treatment. The hematologic profile includes absolute lymphocyte count, ANC, platelet count and hemoglobin for efficacy assessments, other parameters are assessed for safety purposes. Results from the central laboratory will be used for determining response. If results from the central laboratory are not available local laboratory results will be used instead. Patients who have signs and symptoms of disease progression outside of the scheduled study visits should be evaluated by the investigator with physical examination and blood counts with differential to determine if disease progression is present.

In addition to response evaluations, hematologic profiling is performed at every scheduled visit for safety evaluation and physical examinations are performed at Day 1 of every cycle for safety assessment. For reporting purposes, all hematological parameters will be summarized as safety parameters only. The hematologic blood samples for response or disease progression determination should be confirmed by a central laboratory; however if central laboratory results are not available local laboratory results will be used instead. Any suspected case of disease progression (in the absence of laboratory or histopathologic changes meeting the criteria for PD) should be confirmed with a CT scan.

3.2.3 B-symptoms assessment

B-symptoms are defined as fulfilling at least one of the following criteria:

- Unintentional weight loss of 10% or more within the previous 6 months.
- Significant fatigue (i.e., ECOG performance status 2 or worse; inability to work or perform usual activities).
- Fever > 100.5°F or 38.0°C for ≥2 weeks without other evidence of infection.
- Night sweats for ≥ 1 month without evidence of infection.

B-symptom assessments are required within 28 days of the start of randomized treatment. In addition to response evaluation assessments B-symptoms are assessed on Day 1 of every cycle, starting at Cycle 2 for safety purposes.

3.2.4 Bone marrow aspiration and biopsy

Bone marrow aspiration and biopsy are required within 60 days before first dose of randomized treatment.

Bone marrow aspiration and biopsy are not conducted as part of routine response evaluation, however if radiologic evaluations, physical examinations, B-symptom assessments and hematologic evaluations suggest that complete response (CR) or complete response with incomplete bone marrow recovery (CRi) has been achieved then bone marrow aspiration and biopsy are performed to confirm the CR. In cases where cytopenic progression is suspected, a bone marrow aspirate or biopsy must be performed to distinguish autoimmune and drug-related cytopenias.

3.2.5 Determination of response according to IWCLL 2018 criteria

From review of the response evaluation assessments the investigator will determine disease response according to the IWCLL 2018 criteria. At each scheduled response assessment visit patients will be assigned a response of CR, CRi, nPR, PR, PRL, SD or PD. If a patient has had a response assessment that cannot be evaluated, then the patient will be assigned a response of unknown (UNK) at that visit.

The following table describes the required criteria for each response level:

Table 3 Response Assessment Criteria for CLL (modified from Hallek 2018) – IWCLL Criteria

Group	Parameter	CR[a,b]	PR[c]	PD	SD
A	Lymph nodes (target lesions only[d])	None ≥1.5 cm	Decrease ≥50% (from baseline)[e]	Increase ≥50% from baseline or from response	Change of -49% to +49%
	Liver and/or spleen size[f]	Spleen size <13 cm; liver size normal	Decrease ≥50% (from baseline)	Increase ≥50% from baseline or from response	Change of -49% to +49%
	Constitutional symptoms	None	Any	Any	Any
	Circulating lymphocyte count	Normal	Decrease ≥50% from baseline	Increase ≥50% over baseline	Change of -49% to +49%
B	Platelet count	≥100,000/μL	≥100,000/μL or increase ≥50% over baseline	Decrease of ≥50% from baseline secondary to CL	Change of -49% to +49%
	Hemoglobin	≥11.0 g/dL (untransfused and without erythropoietin)	≥11 g/dL or increase ≥50% over baseline	Decrease of ≥2 g/dL from baseline secondary to CLL	Increase < 11.0 g/dL or < 50% over baseline, or decrease > 2 g/dL
	Marrow	Normocellular, no CLL cells, no B-lymphoid nodules	Presence of CLL cells, or of B-lymphoid nodules, or not done	Increase of CLL cells by ≥50% on successive biopsies	No change in marrow infiltrate

Abbreviations: CLL=chronic lymphocytic leukemia; CR=complete response; CT=computed tomography; PD=progressive disease; PR=partial response; SD=stable disease.

Note: CR, complete response: all of the criteria have to be met; PR, partial response: for a PR at least 1 of the parameters of group A and 1 parameter of group B need to improve if previously abnormal. If only one parameter of both groups A and B is abnormal prior to therapy, only 1 needs to improve. PD, progressive disease: at least one of the above criteria of group A or group B has to be met; SD, stable disease: all of the above criteria have to be met. Constitutional symptoms alone do not define PD.

[a] CRi (CR with incomplete bone marrow recovery) refers to patients who fulfill all the criteria for a CR (including the bone marrow examinations), but have a persistent anemia, thrombocytopenia, or neutropenia apparently unrelated to CLL, but related to drug toxicity.

[b] nPR (Nodular partial response): CR with the presence of B-lymphoid nodules in the bone marrow which reflect residual disease.

[c] PRL (partial response with lymphocytosis): presence of lymphocytosis, plus ≥ 50% reduction in lymphadenopathy and/or in spleen or liver enlargement, plus one of the PR criteria for platelets or hemoglobin have to be met.

[d] Non-target lymph nodes are taken into account for response evaluation but are not measured. Therefore response/progression of non-target lesions are assessed qualitatively as CR, PR/SD or PD.

[e] Sum of the products of 6 or less lymph nodes (as evaluated by CT scans and physical examination in clinical trials, or by physical examination in general practice).

[f] Spleen size is considered normal if <13 cm. There is not firmly established, international consensus of the size of a normal liver; therefore, liver size should be evaluated by imaging and manual palpation in clinical trials and be recorded according to the definition used in a study protocol.

For a detailed description of the response parameters see Hallek 2018.

3.2.6 Blinded independent central review (BICR)

A central imaging service (BICR) will be used to provide independent assessments for the purposes of determining the primary endpoint. All imaging assessments (including unscheduled visit scan) will be sent to the sponsor-appointed central reader on an ongoing basis. Results of the review will not be reported back to site.

Further details of the BICR will be documented in the BICR Charter.

3.3 Efficacy Variables

3.3.1 Progression free survival (PFS)

PFS is defined as the time period from the date of randomization until progression or death from any cause, whichever occurs first (i.e. date of PFS event or censoring – date of randomization + 1, time period is calculated in months). Patients who withdraw from the study or are considered lost to follow-up without prior documentation of disease progression will be censored on the date of the last adequate disease assessment. However, if the patient progresses or dies immediately after two or more consecutive missed response evaluations, the patient will be censored at the time of the latest adequate disease assessment prior to the two or more missed visits. Patients who start new anti-cancer therapy before documentation of disease progressionwill be censored on the date of the last adequate disease assessment that is on or before the start date of the new anti-cancer therapy. For patients without an adequate post-baseline disease assessment or no baseline assessments, PFS will be censored on the date of randomization unless they die within 2 visits of baseline. Detailed censoring rules are described below.

Table 4 Censoring rules for PFS

Assessment	Outcome	Date of progression or censoring
Death before first response evaluation[1]	Event	Date of death
Documented PD between scheduled visits[1] (or, between a scheduled and unscheduled visit, if the unscheduled visit falls in between two scheduled visits) on or before receiving subsequent anti-cancer therapy or data cutoff, whichever occurred first	Event	Earliest date of disease assessment documenting progression
Death without documented PD[1] and not receiving subsequent anti-cancer therapy on or before data cutoff	Event	Date of death
No baseline tumor assessments and it's not known to have died in 2 consecutive visits of baseline. In cases of no measurable disease at baseline censoring will not occur.	Censored	Date of randomization
No adequate post-baseline response evaluations and is not known to have died in 2 consecutive visits of baseline	Censored	Date of randomization
Documented PD or death after subsequent anti-cancer therapy and the subsequent anti-cancer started before data cutoff date	Censored	Date of last adequate disease assessment prior to subsequent anti-cancer treatment
No documented PD or death at the time of data cutoff and subsequent anti-cancer therapy started before the data cutoff	Censored	Date of last adequate disease assessment prior to subsequent anti-cancer treatment
Documented PD or death after subsequent anticancer therapy and the subsequent anticancer started after data cutoff date	Censored	Date of last adequate disease assessment on or before data cutoff
No documented PD or death at the time of data cutoff and patient not received subsequent anticancer therapy or subsequent anticancer therapy started after the data cutoff	Censored	Date of last adequate disease assessment on or before data cutoff
Withdrew consent without documented PD or death	Censored	Date of last adequate disease assessment on or before date of withdrawn consent
Lost to follow-up without documented PD or death	Censored	Date of last adequate disease assessment on or before data cutoff
Documented PD or death after 2 or more consecutively missed response evaluations	Censored	Date of last adequate disease assessment before the consecutively missed visits

[1] Having not missed 2 consecutive response evaluations.

Note: Because several assessments contribute to response evaluation for all the scenarios described above the date of 'disease assessment' means the date of the last assessment contributing to the response evaluation.

PD = progressive disease.

An adequate disease assessment is defined by a BICR IWCLL 2018 response that is not 'UNK' (unknown) or missing.

The following rules will be used to determine if a patient has missed two consecutive response evaluations:

CCI

CCI

For the purposes of calculating the time between assessments a response evaluation of UNK would not be considered missing.

PFS will also be obtained using the same derivation rule described above but based on investigator assessment.

3.3.2 Overall response rate (ORR)

ORR is defined as the proportion of patients who have a best overall response of CR, CRi, nPR or PR assessed by BICR per IWCLL 2018 criteria at or before initiation of subsequent anti-cancer therapy. CRi refers to patients who fulfil all the criteria for a CR (including the bone marrow examinations), but have a persistent anemia, thrombocytopenia, or neutropenia apparently unrelated to CLL, but related to drug toxicity. nPR refers to patients who fulfil all the criteria for a CR but with the presence of B-lymphoid nodules in the bone marrow which reflect residual disease. See Section 3.3.3 for more information on how best overall response is determined.

As per the study inclusion criteria all patients enrolled onto the study should have active disease per IWCLL 2018 criteria. Therefore, the FAS will be used as the denominator for calculating ORR. For subjects with CLL who achieve a response (CR, CRi, nPR, PR), CT must be performed for response confirmation after 8 weeks (>56 days) after the initial response imaging assessment.

ORR plus patients with a best overall response of PRL will also be assessed, denoted as ORR+PRL.

ORR and ORR+PRL will also be obtained using the same derivation rule described above based on investigator assessment.

3.3.3 Best overall response (BOR)

Best overall response is the best response a patient has had following randomization, but prior to starting any subsequent anticancer therapy up to and including progression or the last evaluable assessment in the absence of progression. To determine the BOR, the following order is used: CR > CRi > nPR > PR > PRL > SD > PD > UNK.

Patients will be assigned a BOR of UNK (Unknown) if all post-baseline responses are either missing (not documented) or documented as NA (Not Applicable) or Unknown.

BOR will also be obtained using the same derivation rule described above based on investigator assessment.

The overall response concordance rate will be calculated as the proportion of patients in the FAS who have the same CR+CRi+nPR+PR according to the BICR and the investigator. The concordance rate of CR+CRi will be calculated similarly.

3.3.4 Duration of response (DoR)

DoR (assessed by BICR per IWCLL 2018 criteria) is defined as the time from the date of first documented response until date of documented progression or death in the absence of disease progression (i.e. date of PFS event or censoring – date of first response + 1). DoR will therefore only be calculated for patients with overall response, as documented in Section 3.3.2. Date of response will be defined as the date of the last assessment contributing to the response evaluation where an overall response is first recorded. DoR will use the same event and censoring rules as PFS (see Section 3.3.1 for further details).

DoR will also be obtained using the same derivation rule described above based on investigator assessment.

3.3.5 Time to next therapy (TTNT)

TTNT is defined as the time from the date of randomization until institution of non-protocol specified treatment for CLL (either medication or radiotherapy for CLL) or death due to any cause, whichever comes first (i.e. date of first subsequent cancer therapy/death or censoring – date of randomization + 1). Any patient not known to have had a first subsequent anti-cancer therapy will be censored at the last date that the patient was known not to have received a first subsequent anti-cancer therapy (obtained from the CAPRX1 and CAPRXRCL1 forms). If the follow-up cancer therapy eCRF has not been completed then the last known alive date will be used for censoring, considering the following eCRFs:

- AE start and stop dates
- Admission and discharge dates of hospitalization
- Study treatment date
- End of treatment date
- Laboratory sample dates
- Date of vital signs
- Date of ECGs

- Concomitant medications start and stop dates
- Disease assessment dates on IWCLL 2018 eCRF
- Start and stop dates of alternative anti-cancer treatment
- Date last known alive on survival status eCRF
- End of study date.

Patients who are ongoing randomized treatment will be censored at the date of last study visit, using the same criteria defined above. If a patient terminated the study for reason other than death without subsequent therapy, these patients will be censored at the earliest of their last known to be alive and termination dates. For the Interim Analysis, DCO date will be used.

Subsequent anticancer 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.

3.3.6 Overall survival (OS)

Overall survival is defined as the time from the date of randomization until death due to any cause regardless of whether the patient withdraws from randomized therapy or receives another anti-cancer therapy (i.e. date of death or censoring – date of randomization + 1). Any patient not known to have died at the time of analysis will be censored based on the last recorded date on which the patient was known to be alive (SUR_DAT, recorded within the SURVIVE module of the eCRF). For any patients who have not known to have died and are ongoing treatment or do not have a completed survival follow-up eCRF, at the time of OS analysis the last known alive date for each individual patient is defined as the latest among the following dates recorded on the eCRFs:

- AE start and stop dates
- Admission and discharge dates of hospitalization
- Study treatment date
- End of treatment date
- Laboratory sample dates
- Date of vital signs

- Date of ECGs
- Concomitant medications start and stop dates
- Disease assessment dates on IWCLL 2018 eCRF
- Start and stop dates of alternative anti-cancer treatment
- Date last known alive on survival status eCRF
- End of study date.

At the time of the planned final analysis, a survival sweep may be conducted. All subjects who are on study and not known to have died before the survival sweep may be contacted at that time.

If a patient is known to have died where only a partial death date is available then the date of death will be imputed as the latest of the last date known to be alive +1 from the database and the death date using the available information provided:

a. For Missing day only – using the 1st of the month

b. For Missing day and Month – using the 1st of January

If death has been recorded but the date is entirely missing, then date of death will be imputed as the date the patient was last known to be alive + 1.

3.3.7 Minimal residual disease (MRD) negative rate

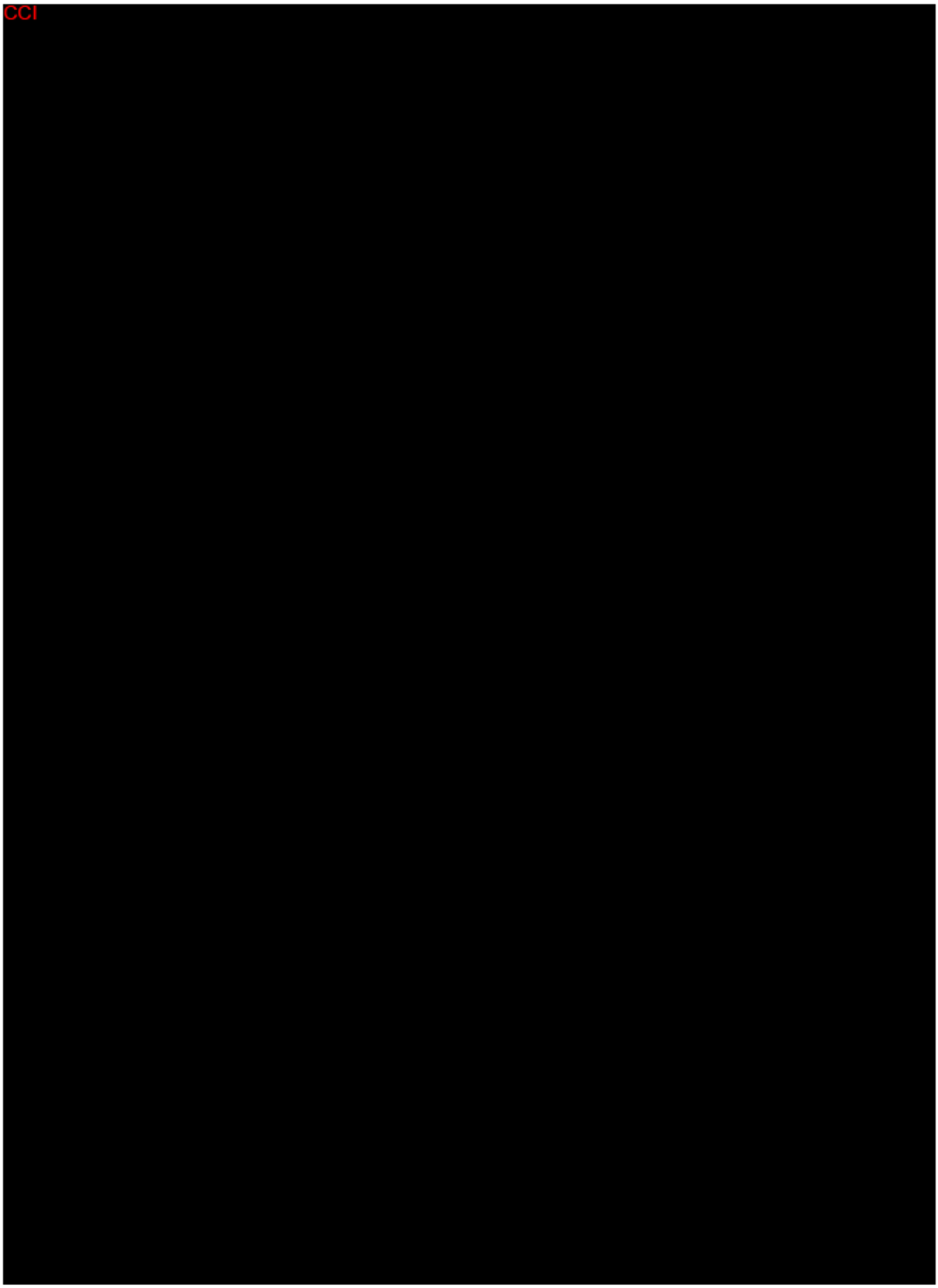
If the patient's physical examination findings, laboratory evaluations, and radiographic evaluations suggest that PR, CRi, or CR has been achieved, a peripheral blood and bone marrow sample will be assessed between 8-12 weeks from the time of supportive clinical assessments to determine MRD negativity. MRD negative rate (peripheral blood) at the start of Cycle 9 will be tested. MRD negative is defined as < 1 CLL cell per 10,000 leukocytes. MRD negative rate is defined as the proportion of subjects with MRD-negativity.

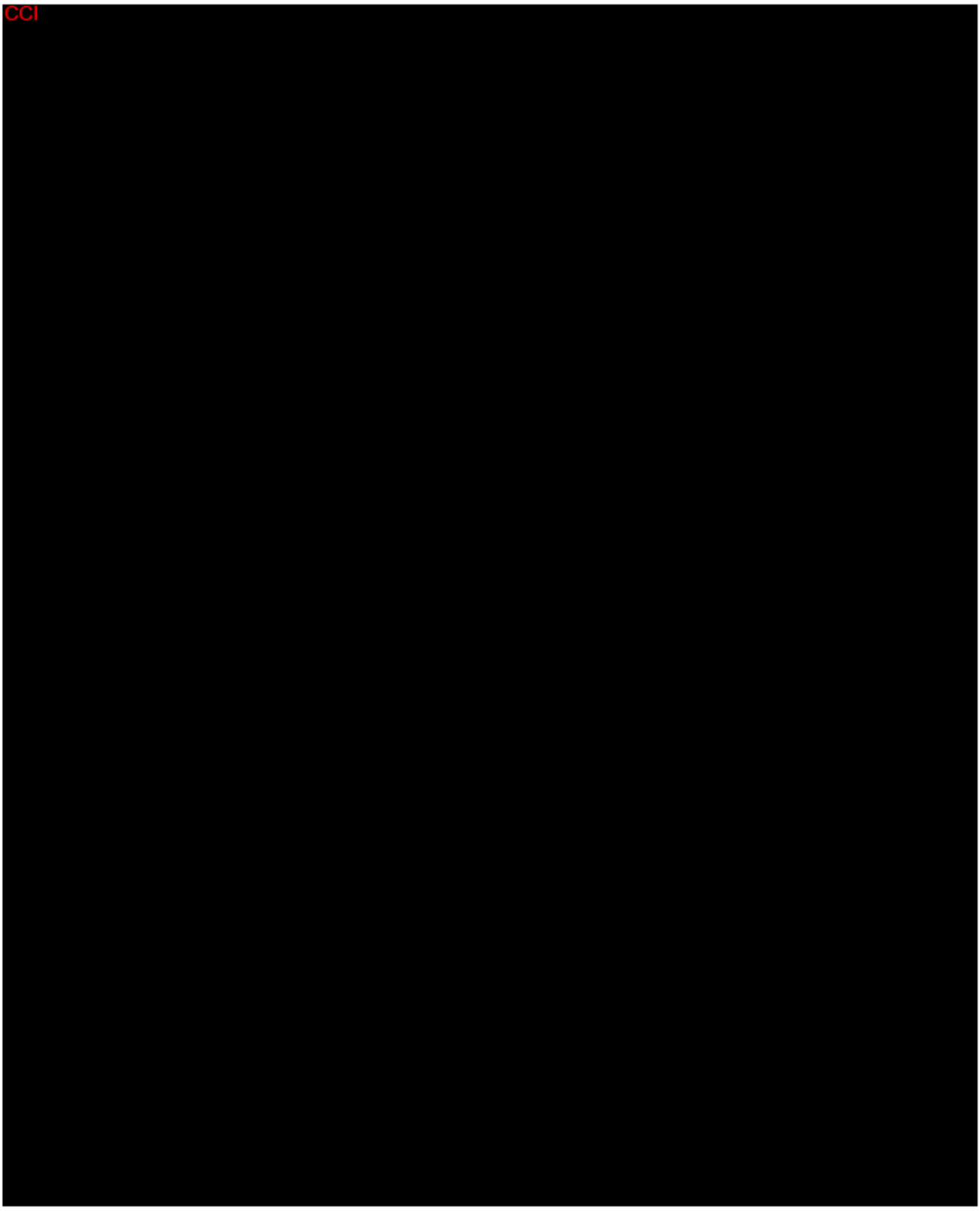
3.4 CCI

3.4.1 CCI

[redacted]

CCI




CCI

CCI

[redacted]

3.5 Safety Variables

3.5.1 Exposure and dose interruptions

Exposure (i.e. duration of treatment) will be calculated as follows for acalabrutinib, rituximab and chlorambucil:

Total (or intended) exposure of study drug for acalabrutinib:

- Total (or intended) exposure = min(last dose date where dose > 0 [units], date of death, date of DCO) – first dose date +1

Total (or intended) exposure of study drug for rituximab:

- Total (or intended) exposure = min(last dose date where dose > 0 [units] + 27, date of death, date of DCO) – first dose date +1

Total (or intended) exposure of study drug for chlorambucil:

- Total (or intended) exposure = min(last dose date where dose > 0 [units] + 13, date of death, date of DCO) – first dose date +1

Actual exposure of study drug:

- Actual exposure = intended exposure – any length of time where the patient has not taken any of the planned dose. For rituximab, actual exposure will be calculated as Total exposure - sum of positive values of (date of subsequent dose – [date of previous dose + 28]). For chlorambucil, actual exposure will be calculated as Total exposure - sum of positive values of (date of subsequent dose – (date of previous dose + 14)).

The actual exposure calculation makes no adjustment for any dose reductions that may have occurred.

Missed or forgotten doses

For acalabrutinib missed and forgotten doses should be recorded on the EX eCRF page as a dose interruption with the reason recorded as "Subject forgot to take dose". These missed or forgotten doses will not be included as dose interruptions in the summary tables but the information will appear in the listing for dosing. However, these missed and forgotten doses will be considered in the derivation of actual exposure.

Patients who permanently discontinue during a dose interruption

If a patient permanently discontinues study drug during a dose interruption, then the date of last administration of study medication recorded on the DOSDISC eCRF page will be used in the programming.

Dose intensity

Relative dose intensity (RDI) is the percentage of the actual dose delivered relative to the intended dose through to treatment discontinuation. CCI

[redacted]

3.5.2 Adverse events

AEs will be collected throughout the study, from date of informed consent until 30 days after the last dose of study drug. SAEs will be collected from the date of informed consent. The following events will be considered treatment emergent:

- Adverse events with an onset date on or after first dose of study drug, and within 30 days after last dose of study drug or the initiation of a subsequent therapy (including crossover treatment), whichever occurs first.
- Worsening of pre-existing events on or after first dose of study drug, and within 30 days after last dose of study drug or the initiation of a subsequent therapy (including crossover treatment), whichever occurs first.

The Medical Dictionary for Regulatory Activities (MedDRA) (using the latest or current MedDRA version) will be used to code the AEs. AEs will be graded according to the

National Cancer Institute Common Terminology Criteria for AEs (using the NCI CTCAE version 5.0).

AEs of special interest

The following events are adverse events of special interest (AESIs) for subjects randomized to the acalabrutinib treatment arm and must be reported to the sponsors expeditiously irrespective of regulatory seriousness criteria or causality:

Ventricular arrhythmias (e.g., ventricular extrasystoles, ventricular tachycardia, ventricular arrhythmia, ventricular fibrillation, etc.) as defined in Appendix A.

Adverse Events of Clinical Interest

Some clinical concepts (including some selected individual preferred terms and higher-level terms) have been considered "Events of Clinical Interest" (ECI) to the acalabrutinib program. ECI include cardiac events, cytopenias (anemia, leukopenia and thrombocytopenia), hemorrhage, hepatotoxicity, hypertension, infection, intestinal lung disease/pneumonitis, tumor lysis syndrome, and secondary primary malignancy.

These ECIs have been identified as a list of categories provided by the patient safety team. ECIs are defined in Appendix A.

Other categories may be added as necessary or existing terms may be merged. An AstraZeneca medically qualified expert after consultation with the Global Patient Safety Physician has reviewed the events of clinical interest and identified which higher-level terms and which preferred terms contribute to each ECI. Further reviews may take place prior to database lock (DBL) to ensure any further terms not already included are captured within the categories.

3.5.3 Clinical laboratory data

Clinical safety laboratory assessment is included in Table 6 and specified into the CSP SoA for the timing and frequency.

Table 6 Laboratory Variables

Clinical Chemistry [a]	
Calcium	Urea or blood urea nitrogen
Chloride	Uric acid
Magnesium	Creatinine
Phosphate/Phosphorus	Total bilirubin
Potassium	Glucose
Sodium	Albumin

Table 6 Laboratory Variables

AST	Total protein
ALT	Triglycerides
Alkaline phosphatase (ALP)	Cholesterol
Gamma glutamyl transferase (GGT)	Lactate dehydrogenase (LDH)
Hematology [b]	
White blood cell (WBC) count with differential	Platelet count
Red blood cell (RBC) count	Absolute neutrophil count (ANC)
Hematocrit	Absolute lymphocyte count (ALC)
Hemoglobin	
Pregnancy Test (females of childbearing potential only)	
Urine human chorionic gonadotropin (hCG) or Serum βhCG	
Hepatitis B and C Testing [c]	
HBsAg	Hepatitis B surface antibody (HBsAb)
Anti-HBc	Hepatitis C (hepatitis C virus [HCV]) antibody
Hepatitis B PCR (clinically indicated)	Hepatitis C PCR (clinically indicated)
Other Tests	
T/B/NK Cell Count	Serum immunoglobulin levels
β2-microglobulin [b]	
Cytogenetics and FISH Panel [b,d]	Genetic molecular prognostic molecules [b,e]

[a] In case a patient shows an AST or ALT ≥3×ULN together with total bilirubin ≥2×ULN please refer to CSP Appendix D 'Actions required in cases of increases in liver biochemistry and evaluation of Hy's Law,' for further instructions.

[b] These tests will be performed at the central laboratory.

[c] Hepatitis serology testing must include HBsAg, HBsAb, anti-HBc and hepatitis C (HCV) antibody. In addition, any patients testing positive for anti-HBc must have quantitative PCR testing for HBV DNA during screening and monthly thereafter. Monthly monitoring should continue until 12 months after last dose of study drug(s). Any patient with a rising viral load (above lower limit of detection) should discontinue study drug and have antiviral therapy instituted and a consultation with a physician with expertise in managing hepatitis B. Since IVIG may cause false positive hepatitis serology, monthly PCR testing is not required in patients who are currently receiving or received prophylactic IVIG within 3 months before study enrollment and have a documented negative anti-HBc test before the initiation of IVIG therapy. PCR testing should be performed when clinically indicated (eg, in the setting of rising transaminase levels). Patients with a known history of hepatitis C or who are hepatitis C-antibody positive should have quantitative PCR testing for HCV RNA performed during screening. No further testing is necessary if the PCR results are negative at screening.

[d] Cytogenetics and FISH Panel include 17p del, 13q del, trisomy 12, 11q del by FISH and stimulated karyotyping.

[e] Genetic molecular prognostic molecules panel include, but is not limited to, sequencing of p53 mutations, immunoglobulin heavy-chain variable (IGHV) mutational status.

NCI CTCAE grading

NCI CTCAE grades will be derived programmatically for applicable laboratory parameters. The programmatically derived grades will be used for summary tables.

As applicable, values will be converted to standard units and will be graded using NCI CTCAE V5.0.

Lymphocytosis

Lymphocytosis is defined as an ALC > 5000 cells/μL and an increase 50% or more from baseline.

Duration of lymphocytosis is defined as the duration of time from the earliest date on which the ALC value met the lymphocytosis criteria at a post-baseline assessment to the earliest date on which a subsequent ALC value met the resolution criteria.

For patients with lymphocytosis, 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.

4. ANALYSIS METHODS

4.1 General Principles

The primary objective of the study is to test the following hypothesis:

H0: BICR-assessed PFS of Arm A is the same as Arm B

H1: BICR-assessed PFS of Arm A is not the same as Arm B

The study will be considered positive (a success) if the primary PFS analysis result for Arm A versus Arm B is statistically significant.

The below mentioned general principles will be followed throughout the study:

- Descriptive statistics will be used for all variables, as appropriate. Continuous variables will be summarized by the number of observations, mean, standard deviation, median, upper and lower quartiles minimum, and maximum. For log-transformed data it is more appropriate to present geometric mean, coefficient of variation (CV), median, minimum and maximum. Categorical variables will be summarized by frequency counts (n) and percentages (%) for each category.
- Unless otherwise stated, percentages will be calculated out of the total number of patients in the analysis set for the corresponding treatment group. Overall totals (i.e. data from Arm A and Arm B combined) will be calculated for baseline summaries only.
- For continuous data, the mean and median will be rounded to 1 additional decimal place compared to the original data. The standard deviation will be rounded to 2 additional

decimal places compared to the original data. Minimum and maximum will be displayed with the same accuracy as the original data.

- For categorical data, percentages will be rounded to 1 decimal place.
- SAS® version 9.4 will be used for all analyses.

In general, for efficacy and CCI [redacted] (except for laboratory parameters) the last observed measurement prior to randomization will be considered the baseline measurement. However, if an evaluable assessment is only available after randomization but before the first dose of randomized treatment then this assessment will be used as baseline. For laboratory parameters used for efficacy parameters, the last observation before the first dose of study drug will be used. However, if the randomized patient never received any study treatment, the last assessment prior to randomization will be used. For safety endpoints the last observation before the first dose of study drug will be considered the baseline measurement unless otherwise specified. For demography and disease characteristics of subjects who did not receive treatment, the last observed measurement prior to randomization will be used as the baseline measurement. For analyses of subjects who crossover to acalabrutinib baseline for safety endpoints will be defined as the last assessment prior to the first dose of crossover therapy.

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

If two visits are equally eligible to assess patient status at baseline (e.g., screening and baseline assessments both on the same date prior to first dose with no washout or other intervention in the screening period), the average can be taken as a baseline value. For non-numeric assessments where taking an average is not possible then the best value would be taken as baseline as this is the most conservative. In the scenario where there are two assessments on day 1, one with time recorded and the other without time recorded, the one with time recorded would be selected as baseline. Where safety data are summarized over time, study day will be calculated in relation to date of first treatment.

In all summaries change from baseline variables will be calculated as the post-treatment value minus the value at baseline. The percentage change from baseline will be calculated as (post-baseline value - baseline value) / baseline value × 100.

Endpoints will be summarized and analyzed using the analysis sets indicated in Table 2.

4.2 Analysis Methods

The following table (Table 7) details which endpoints are to be subjected to formal statistical analysis, together with pre-planned sensitivity analyses, making it clear which analysis is regarded as primary for that endpoint.

Table 7 Formal Statistical Analyses to be Conducted and Pre-planned Sensitivity Analyses

Endpoints Analyzed	Notes
PFS	Primary analysis: Stratified log-rank test adjusted for ECOG (0-1 versus 2) and Rai stage (Stage 0-II versus III-IV) with BICR defined response Sensitivity analyses [a] 1. Unstratified log-rank test 2. Analysis of alternative censoring rule - without censoring prior to subsequent therapies 3. Analysis of alternative censoring rule – without censoring prior to two or more consecutively missed response evaluations 4. Ascertainment bias analysis - using investigator assessments
Overall response rate	CMH chi-square test adjusted for ECOG (0-1 versus 2) and Rai stage (Stage 0-II versus III-IV) using both BICR and investigator defined responses
Duration of response	Stratified log-rank test adjusted for ECOG (0-1 versus 2) and Rai stage (Stage 0-II versus III-IV) using both BICR and investigator defined responses
Time to next therapy	Stratified log-rank test adjusted for ECOG (0-1 versus 2) and Rai stage (Stage 0-II versus III-IV)
Overall survival	Stratified log-rank test adjusted for ECOG (0-1 versus 2) and Rai stage (Stage 0-II versus III-IV) Sensitivity analysis considering censoring crossover patients prior to receiving acalabrutinib monotherapy
Minimal residual disease	CMH chi-square test adjusted for ECOG (0-1 versus 2) and Rai stage (Stage 0-II versus III-IV)

[a] See 4.2.2.2 for further details

4.2.1 Multiplicity

To control the overall type I error at 0.05 level, the Lan-Demets alpha spending function based on the O'Brien-Fleming boundary is used to split alpha for the interim and final analyses of PFS. If exactly 76% of PFS events at final is available at the time of the interim analysis, that is 38/50 PFS events based on BICR have occurred, the 2-sided alpha level to be applied for the interim and the final analysis will be 0.02 and 0.044, respectively. The nominal alpha levels for the interim and final analyses will be determined based on actual number of PFS events observed at the time of the analyses.

4.2.2 Efficacy analyses

All efficacy analyses will use the FAS unless otherwise specified.

4.2.2.1 Time-to-event endpoint considerations

Pooling Strategy

Stratified log-rank tests and Cox proportional hazards model will adjust for the following factors:

- ECOG (0-1 versus 2)
- Rai stage (Stage 0-II versus III-IV)

There are at least 2 events per treatment group in each stratum (where a stratum is defined as stratification factor 1 * stratification factor 2), stratification factors should be collapsed until all strata have at minimum two events for the primary endpoint. Stratification factors will be removed from the model in the following order, until at least 2 events per treatment group in each stratum are available:

- Rai stage (Stage 0-II versus III-IV)
- ECOG (0-1 versus 2)

All analyses will then be conducted in accordance with the pooling strategy defined for the primary analysis.

If there are secondary endpoints that still will not conform to the requirement on number of events above defined, unstratified log-rank tests and Cox proportional hazards model will be used for the analysis of the secondary endpoint. This will be supported by unstratified sensitivity analyses of the primary endpoint.

Kaplan Meier analysis

The median time-to-event will be presented for each treatment arm alongside its 95% confidence interval calculated using the Brookmeyer and Crowley technique and the stratified log-rank test for comparing the treatment arms.

Kaplan-Meier plots of time-to-event endpoints will be presented which will display the survival curves, the number of patients at risk, the median time-to-event and the 95% confidence intervals for each treatment arm.

Hazard ratio and confidence interval estimation

Time-to-event endpoints will be analyzed using a 2-sided stratified log-rank test adjusting for ECOG (0-1 versus 2) and Rai stage (Stage 0-II versus III-IV) for generation of the p-value.

The stratification variables in the statistical modelling will be based on the values entered into IRT/RTSM at randomization, even if it is subsequently discovered that these values were incorrect.

HR will be calculated from a stratified Cox proportional hazards model together with its corresponding, 95% CI. Treatment is the only covariate, ties are handled by the Efron approach and the CI is calculated using a profile likelihood approach.

Supporting summary measures

Survival rates, estimated from the KM curve, will also be presented at 6 monthly intervals, starting at 6 months and finishing at the last 6-monthly interval prior to the maximum event time.

4.2.2.2 Progression free survival (PFS)

The primary analysis will compare PFS according to the hypotheses in Section 4.1 and as assessed by BICR per IWCLL 2018 criteria using the methods described in Section 4.2.2.1.

Proportionality assumption

The assumption of proportionality of hazards will be assessed. Proportional hazards will be tested firstly by examining plots of complementary log-log (event times) versus log (time) and, if these raise concerns, by fitting a time dependent covariate (adding a treatment-by-time interaction term) to assess the extent to which this represents random variation. If a lack of proportionality is evident, the variation in treatment effect can be described by presenting piecewise HR calculated over 6-monthly time-periods. In such circumstances, the HR from the primary analysis can still be meaningfully interpreted as an average HR over time unless there is extensive crossing of the survival curves. If lack of proportionality is found this may be a result of a treatment-by-covariate interaction, which will be investigated.

PFS sensitivity analyses

The following sensitivity analyses will be performed:

- In addition to the stratified primary analysis the unstratified log rank test and unstratified hazard ratios will be calculated.
- The primary analysis will also be repeated without PFS censoring to subsequent therapies. Therefore, patients who receive a subsequent therapy prior to progression or death will be assigned an event time according to their recorded progression or death date. Patients who receive a subsequent therapy and did not progress or die will be censored according to the rules in table 4 other than the subsequent therapy rule.

- The primary analysis will be repeated except that the actual PFS event times, rather than the censored times, of patients who progressed or died in the absence of progression immediately following two, or more, tumor assessments will be included. This will be supported by a Kaplan-Meier plot of the time to censoring where the censoring indicator of the primary PFS analysis is reversed (what was originally a censored event in the primary PFS analysis becomes an actual event and what was originally a PFS event becomes a censored event) will be presented.
- Ascertainment bias will be assessed by analyzing the investigator data. The stratified log rank test will be repeated on PFS using the investigator data based upon IWCLL 2018. The HR and CI will be presented.
- Concordance between Investigator and BICR assessed progression will be presented for each treatment group.

If there is an important discrepancy between the primary analysis using the BICR data and this sensitivity analysis using investigator data then the proportion of patients with site confirmation but no central confirmation of progression will be summarized; such patients have the potential to induce bias in the central review due to informative censoring. An approach of imputing an event at the next visit in the central review analysis may help inform the most likely HR value (Fleischer et al 2011), but only if an important discrepancy exists.

The primary analysis may also be repeated using the randomization strata as entered into the database rather than those entered into the IVRS system, but only if the proportion of patients with a discrepancy is more than 10%.

Additional supportive summaries

In addition, the number of patients prematurely censored will be summarized by treatment arm together with baseline prognostic factors of the prematurely censored patients. A patient would be defined as prematurely censored if they had not progressed (or died in the absence of progression) and the latest scan prior to DCO was more than one scheduled tumor assessment interval plus 2 weeks prior to the DCO date.

Additionally, summary statistics will be given for the number of days from censoring to DCO for all censored patients.

A summary of the duration of follow-up will be provided using median time from randomization to date of censoring (date last known to have not progressed) in censored (not progressed) patients only, presented by treatment group.

Summaries of the number and percentage of patients who miss two or more consecutive tumor assessments will be presented for each treatment group.

Subgroup analyses

Subgroup analyses will be conducted comparing PFS between Arm A versus Arm B in the following subgroups of the FAS (but not limited to):

- Age at randomization (< 65 versus ≥ 65)
- Sex (male versus female)
- ECOG (0-1 versus 2)
- IGHV mutation status (mutated versus non-mutated)
- Region/Country (China versus other)
- Rai stage at screening (Stage 0-II versus III-IV)
- β2-microglobulin at baseline (≤ 3.5 mg/L versus > 3.5 mg/L)
- Bulky disease (longest diameter of lymph node < 5 cm versus ≥ 5 cm at baseline)
- Presence of 11q deletion.
- Complex karyotype (Yes versus No)

The HR and corresponding 95% CI for each subgroup will be calculated based on an unstratified Cox proportional hazards model with with treatment as the only covariate. The Cox models will be fitted using SAS® PROC PHREG with the Efron method to control for ties, using the by statement to obtain HR and 95% CI for each subgroup level separately. These hazard ratios and associated two-sided 95% CIs will be summarized and presented on a forest plot, along with the results of the overall primary analysis. Separate analyses will be run for each subgroup, i.e. subgroup levels will not be treated as strata. The subgroup analyses for the stratification factors will be based on the values entered into the IRT/RTSM, all other factors will be based on values recorded on the eCRF.

Other baseline variables may also be assessed if there is clinical or biological justification or an imbalance is observed between the treatment arms. The purpose of the subgroup analyses is to assess the consistency of treatment effect across expected prognostic factors. If a baseline imbalance is observed between treatment arms, ad-hoc subgroup analysis may be used to investigate any potential for impact on the main results.

Unless there is a marked difference between the results of the statistical analyses of the PFS from the BICR data and that of the site investigator tumor data, these subgroup analyses will only be performed on the PFS endpoint using the BICR data.

No adjustment to the significance level for testing of the subgroup and sensitivity analyses will be made since all these analyses will be considered supportive of the analysis of PFS.

KM plots of PFS by treatment and 11q deletion status/IGHV mutational status will be produced as supportive figures.

4.2.2.3 Overall response rate

ORR (assessed by BICR per IWCLL 2018 criteria) will be compared between treatment arms (Arm A versus Arm B) using the Cochran-Mantel-Haenszel (CMH) chi-square test, adjusted for randomization stratification factors ECOG (0-1 versus 2) and Rai stage (Stage 0-II versus III-IV).

Summaries will be produced that present the number and percentage of patients with a tumor response (CR, CRi, nPR, PR) based upon the number of patients in the FAS per BICR assessments alongside the CMH p-value, the odds ratio and its 95% CI.

The analysis of ORR will be repeated using investigator responses. The overall response (CR+CRi+nPR+PR) and CR/CRi concordance rate will also be summarized separately.

As supportive analysis, ORR+PRL using BICR assessments and investigator assessments will be analyzed in the same manner as ORR.

4.2.2.4 Best overall response

For each treatment arm, best overall response (BOR) will be summarized by n (%) for each category (CR, CRi, nPR, PR, PRL, SD, PD and UNK) with tables and figures as appropriate. No formal statistical analyses are planned for BOR.

4.2.2.5 Duration of response

DoR as assessed by BICR per IWCLL 2018 criteria will be analyzed using the methods described in Section 4.2.2.1. Only patients with a BOR of CR, CRi, nPR or PR in the FAS will be included in this analysis.

The DoR will be repeated using investigator derived responses.

4.2.2.6 Time to next therapy

TTNT will be analyzed using the methods described in Section 4.2.2.1.

The status of patients at the time of analysis will be summarized. This will include the number (%) of patients who are known to have received a subsequent therapy, the number (%) of patients who died prior to receiving a subsequent therapy, and the number (%) of patients who discontinued the study prior to starting a new therapy and were alive at the date of discontinuation. The subset of patients who discontinued the study prior to death will be further summarized by the number (%) of patients who were lost to follow-up prior to starting a subsequent therapy and the number (%) of patients withdrew consent prior to starting a new therapy.

In patients who received a subsequent anti-cancer therapy, a summary table of the first subsequent anti-cancer therapy by treatment arm will be provided, as well as response to first subsequent anti-cancer therapy by treatment arm (if available).

4.2.2.7 Overall survival (OS)

OS will be analyzed using the methods described in Section 4.2.2.1.

The status of patients at the time of analysis will be summarized. This will include the number (%) of patients who died, the number (%) still in survival follow-up and the number (%) of patients who discontinued the study prior to death. Patients who discontinued the study prior to death will be further broken down by presenting the number (%) of patients who discontinued voluntarily, the number (%) of patients lost to follow-up and the number (%) of patients who discontinued the study for any other reason.

A sensitivity analysis for OS will also be performed where patients who received crossover therapy will be censored one day prior to the first dose of crossover therapy.

The number of patients prematurely censored will be summarized by treatment arm. A patient would be defined as prematurely censored if their survival status was not defined at the DCO.

In addition, duration of follow-up will be summarized using medians:

- In all patients: time from randomization to the date of death (i.e. overall survival) or to the date of censoring (date last known to be alive) for censored patients regardless of treatment arm.

4.2.2.8 Minimal residual disease negative rate

Peripheral blood MRD negative rate at cycle 9 will be compared between treatment arms (Arm A versus Arm B) using the CMH chi-square test, adjusted for randomization stratification factors.

4.2.2.9 Improvement of disease-related symptoms and sustained hematologic improvement

Disease-related symptoms

Constitutional symptoms include weight loss, fever, night sweats, and fatigue. For each symptom number and percentage of subjects with symptom absent at each post-baseline timepoint will be summarized in the subset of subjects with symptoms present at baseline.

Sustained hematologic improvement

Sustained hematologic improvement is defined as hematologic improvement that persisted continuously for ≥56 days (8 weeks) without blood transfusion or growth factors. This sustained improvement requires that 1) two assessments meet the improvement criteria and are at least 56 days apart, 2) all assessment(s) with non-missing data between those two assessments must meet the improvement criteria, and 3) there is no concomitant use of blood transfusion/growth factor during this time period. The proportion of subjects achieving sustained hematologic improvement in the subset of subjects with cytopenia(s) at baseline, prior to subsequent anticancer therapy, will be summarized by treatment arm.

Cytopenia is defined as at least one of the following 3 criteria being met:

- Neutropenia: ANC $\leq 1.5 \times 10^9$/L
- Anemia: Hgb ≤ 11 g/dL
- Thrombocytopenia: platelet count $\leq 100 \times 10^9$/L.

Hematologic improvement in the subset of patients with cytopenia(s) is defined as follows for each parameter:

- Improvement to hemoglobin >11 g/dL from baseline or increase >=50% over baseline.
- Improvement to platelet counts > 100 × 10^9/L from baseline or increase >=50% over baseline.
- Improvement to ANC>1.5× 10^9/L or increase >=50% over baseline.

4.2.3 Data cut offs

There are two planned analyses for this study. The first planned analysis is an interim analysis (IA) when approximately 38 PFS events based on BICR have been observed in the study (25.3% maturity or 76% information fraction). The second planned analysis is the final analysis, which will be conducted when there are approximately 50 BICR-assessed PFS events in total from Arm A and Arm B.

The only data considered post DCO will be the survival update forms completed as part of the survival sweep. Patients who are known to have died or to be alive after the DCO due to the survival sweep will be censored at the DCO. Any other data entered post-DCO will be excluded from the analyses. AEs and medications which started before the DCO but finished after the DCO will be reported in the analysis as ongoing and the end date will be reported as missing.

4.2.4 General considerations for safety and CCI

Time windows will be defined for any presentations that summarize values by visit. The following conventions will apply:

- The time windows will be exhaustive so that data recorded at any time point has the potential to be summarized. Inclusion within the time window will be based on the actual date and not the intended date of the visit.
- All unscheduled visit data have the potential to be included in the summaries.
- The window for the visits following baseline will be constructed in such a way that the upper limit of the interval falls half way between the two visits (the lower limit of the first post-baseline visit will be Day 2). If an even number of days exists between two consecutive visits, then the upper limit will be taken as the midpoint value minus 1 day. For example, the visit windows for vital signs data are:

- Day 8, visit window 2 – 11
- Day 15, visit window 12 – 18
- Day 22, visit window 19 – 25
- Day 29, visit window 26 – 32
- Day 36, visit window 33 – 39
- Day 43, visit window 40 – 46

- For summaries showing the maximum or minimum values, the maximum/minimum value recorded on treatment will be used (regardless of where it falls in an interval).
- Listings should display all values contributing to a time point for a patient.
- For visit-based summaries
  - If there is more than one value per patient within a time window then the closest value to the scheduled visit date will be summarized, or the earlier, in the event the values are equidistant from the nominal visit date. The listings will highlight the value for the patient that contributed to the summary table, wherever feasible. Note: in summaries of extreme values all post baseline values collected are used including those collected at unscheduled visits regardless of whether or not the value is closest to the scheduled visit date
- For summaries at a patient level, all values will be included, regardless of whether they appear in a corresponding visit-based summary, when deriving a patient level statistic such as a maximum.
- For summaries showing shift from baseline (e.g. NCI CTCAE grades) all values will be included, regardless of whether they appear in a corresponding visit-based summary.

Missing safety data will generally not be imputed. However, safety assessment values of the form of "< x" (i.e. below the lower limit of quantification) or > x (i.e. above the upper limit of quantification) will be imputed as "x" in the calculation of summary statistics but displayed as "< x" or "> x" in the listings. Additionally, adverse events that have missing causality (after data querying) will be assumed to be related to study drug.

Furthermore:

- For missing diagnostic dates, if day and/or month are missing use 01 and/or Jan. If year is missing, put the complete date to missing.
- For missing AE and medication/radiotherapy start dates, the following will be applied
  a. Missing day- Impute the 1st of the month unless month is the same as month of the first dose of study drug then impute first dose date
  b. Missing day and month -Impute 1st January unless year is the same as first dose date then impute first dose date

c. Completely missing-impute first dose date unless the end date suggests it could have started prior to this in which case impute the 1st January of the same year as the end date

- For missing AE and medication/radiotherapy end dates, the following will be applied:

a. Missing day - Impute the last day of the month

b. Missing day and month – impute 31st December. Flags will be retained in the database indicating where any programmatic imputation has been applied, and in such cases, any durations would not be calculated.

4.2.5 CCI

[redacted]

4.2.5.1 CCI

[redacted]

4.2.5.2 CCI

[redacted]

4.2.5.3 CCI

[redacted]

4.2.6 Safety

All safety summaries will use the safety analysis set. Safety analyses, except for death, will include data only prior to crossover for patients from Arm B. Selected outputs will be repeated for the crossover period.

4.2.6.1 Exposure

Summary statistics will be calculated for total exposure, actual exposure and RDI for each study drug. Additionally, a categorical summary of RDI will be produced, with categories < 80%, 80 – 120% and > 120%.

For patients who received acalabrutinib monotherapy, dose withholding and dose reduction will be summarized along with the reasons. For patients who received rituximab, dose delay and infusion interruption will be summarized along with the reasons. For patients who received Chlorambucil, dose interruption (defined as skip one or more dose in one cycle), dose delay and dose reduction will be summarized along with the reasons. Acalabrutinib dose reduction is defined as taking lower dose level (100 mg QD) for ≥3 consecutive days, dose withholding is defined as missing dose for >=7 consecutive days.

Total exposure, actual exposure, RDI and dose modifications will also be analyzed for the crossover period.

4.2.6.2 Adverse events (AEs)

All TEAEs, both in terms of current MedDRA preferred term and NCI CTCAE grade, will be summarized descriptively by count (n) and percentage (%) for each treatment group. The current MedDRA dictionary will be used for coding.

An overall summary of the number and percentage of patients in each of the following categories will be presented:

- All treatment emergent adverse events (AEs)
- All related AEs
- AEs with outcome of death
- Related AEs with outcome of death
- Serious AEs (SAEs)
- Related SAEs
- AEs leading to discontinuation of IP
- Related AEs leading to discontinuation of IP
- AEs leading to IP dose interruptions/delays*
- Related AEs leading to IP dose interruptions/delays*
- AEs leading to IP dose reduction

- Related AEs leading to IP dose reduction

  *AE leading to dose delay only applies to Arm B.

The overall summary will be repeated for the crossover period.

All AEs will be listed, including AEs which are reported more than 30 days after last dose. AEs which occur after the first dose of crossover therapy will be flagged in the listings as well as AEs which occur after the initiation of any other subsequent therapies.

AEs and drug-related AEs

AEs and drug-related AEs will be summarized by system organ class (SOC) and preferred term (PT) in descending order of frequency and separately by maximum NCI CTCAE toxicity grade. In addition, AEs with NCI CTCAE grade ≥3 and drug-related AEs with NCI CTCAE grade ≥3 will be summarized by SOC and PT. An additional table will present the number and percentage of patients with AEs by PT.

AEs will also be summarized separately for the crossover period by SOC and PT. Further AE summaries may be produced for the crossover period in an exploratory fashion if sufficient AEs occur to make these summaries worthwhile.

Serious adverse events (SAEs) and drug-related SAEs

SAEs and drug-related SAEs will be summarized by SOC and PT in descending order of frequency. SAEs will also be listed separately.

AEs and drug-related AEs leading to treatment discontinuation

AEs and drug-related AEs leading to treatment discontinuation will be summarized by SOC and PT in descending order of frequency. AEs leading to treatment discontinuation will also be listed separately.

AEs leading to dose interruptions/delays and reduction

AEs leading to dose interruptions/delays and dose reduction, as indicated in the eCRF will be summarized by SOC and PT in descending order of frequency.

AEs leading to death and drug-related AEs leading to death

AEs leading to death and drug-related AEs leading to death will be summarized by SOC and PT in descending order of frequency. AEs leading to death will also be listed separately.

Adverse events of special interest (AESI)

A summary table of certain MedDRA preferred terms will be produced and will also show the individual preferred terms which constitute each AESI grouping. AESIs will also be summarized by PT.

Events of clinical interest (ECI)

ECIs are defined in section 3.5.2.A summary of the number and percentage of patients with ECIs will be provided by ECI category and ECI subcategory. In addition, for each ECI category, a separate summary table will be displayed with number and percentage of ECIs by ECI subcategory and Preferred Term. Refer to Appendix A for ECI.

Analyses for crossover period

Summary tables from the above AE section, including AEs, AEs leading to death, AEs leading to dose discontinuation, serious AEs, and overall summary of ECIs, will be repeated for patients in the crossover period, provided the crossover period includes at least 10 patients. Otherwise, crossover related safety data will be listed only.

The AE summary tables for crossover patients will include all AEs that occurred after the start of crossover treatment up until the end of the 30 day follow-up period. The 30 day follow-up period will be defined as 30 days following discontinuation of treatment.

In addition to AE summaries, selected summaries of clinical laboratory parameters, vital signs and ECG data might be repeated for crossover patients only provided at least 10 patients receive crossover therapy.

4.2.6.3 Clinical laboratory data

Laboratory data obtained until 30 days after the last dose of study treatment or the start of a subsequent therapy (including crossover treatment), whichever occurs first, will be used for reporting. Project-specific reference ranges will be applied for relevant laboratory variables (hematology, clinical chemistry).

Hematology, clinical chemistry, T/B/NK Cell Count and Serum Immunoglobulin will be listed individually by patient and suitably summarized and may be plotted as needed. Patients who received IVIG on the study will be excluded from the summary for IgG. For all laboratory variables, which are included in the version 5.0 of NCI CTCAE, the NCI CTCAE grade will be calculated.

Numerical laboratory data (absolute and change from baseline) will be summarized by scheduled visit using standard summary statistics (mean, standard deviation, minimum, median, maximum, and number of observations).

For hematology and clinical chemistry, where applicable, data shift tables will summarize the change from baseline to maximum on-treatment NCI CTCAE grade, the change from baseline to maximum value, and the change from baseline to minimum value.

Central laboratory data will be used for summaries of hematology, however local laboratory data will be used when central data is not available. All laboratory hematology data will be summarized together, regardless of whether the tests were performed for safety purposes or for evaluation of disease response. Summary of T/B/NK cell count will be based on central laboratory data. Summaries of clinical chemistry and serum immunoglobulin levels will use local laboratory data.

To identify potential Hy's Law (see CSP Appendix D for further details) cases, patients who meet any of the following criteria at any point during the study will be highlighted:

- AST ≥ 3 × ULN
- ALT ≥ 3 × ULN
- TBL ≥ 2 × ULN

For all patients who meet the biochemical criteria for Hy's Law (potential Hy's Law), the relevant laboratory parameters will be tabulated showing all visits for these patients.

Any data outside the laboratory reference ranges will be flagged and explicitly noted on the listings that are produced. To help identify potential Hy's Law cases, patients who have ALT or AST ≥ 3 × ULN or total bilirubin ≥ 2 × ULN at any time during the study (i.e. not necessarily at the same time) will be flagged in the listings.

All laboratory results will be listed.

Analysis of Lymphocytosis

For all patients with baseline and any post-baseline ALC measurements, ALC at peak summary will be provided.

The number of patients with at least one occurrence of lymphocytosis will be summarized. The following analyses will be conducted for patients with lymphocytosis by treatment arm: ALC at peak, time to lymphocytosis and duration of lymphocytosis for patients who have lymphocytosis will be summarized with descriptive statistics.

4.2.6.4 Vital signs

Summaries of vital signs data will include all data obtained until 30 days after the last dose of study treatment or the start of a subsequent therapy (including crossover treatment), whichever occurs first. Pulse rate, systolic and diastolic blood pressure, body temperature, respiratory rate and weight will be listed by patient and summarized using standard summary statistics for the absolute value and change from baseline at each visit.

4.2.6.5 ECOG performance status

Summaries of ECOG data will include all data obtained until 30 days after the last dose of study treatment or the start of a subsequent therapy (including crossover treatment), whichever

occurs first. The number (%) of patients fulfilling each ECOG performance status category will be presented for each visit.

4.2.6.6 Electrocardiogram (ECG) data

ECG details will be listed individually by patient.

4.2.6.7 Physical examination

Shift to maximum post-baseline value will be summarized by treatment arm.

4.2.7 Demographics and baseline characteristics

The following will be summarized for all patients in the FAS by treatment group:

- Patient disposition (including screening failures and reason for screening failure) both for randomized treatment and separately for crossover therapy. For crossover patients the end of study reason will be the same for both the original treatment period and the crossover treatment period, but reasons for end of treatment will be distinct for the treatment periods.
- Important protocol deviations by category, summarized overall by treatment arm
- Inclusion in analysis sets
- Demographics (age, age group [< 50, ≥ 50 - < 65, ≥ 65 - < 75 and ≥ 75 years], sex, race and ethnicity),
- Patient characteristics at baseline (height, weight, BMI, BMI group, nicotine use [never, current, former], alcohol use [never, current, former])
- Stratification factors according to IRT/RTSM and eCRF data (ECOG performance status, Rai staging)
- Patient recruitment by region/country
- Medical history (past and current)
- Disease characteristics at baseline:
  - ECOG performance status according to the clinical database
  - Rai staging according to the clinical database
  - Time since original diagnosis to randomization (in months)
  - CIRS-G
  - Age ≥ 65 years, >18 and <65 years (creatinine clearance 30 to 69 ml/min, CIRS-G>6)
  - Bulky disease [<5, ≥5, no measurable lymph nodes]
  - 11q deletion
  - 13q deletion
  - IGHV (mutated, unmutated)
  - Complex karyotype

- Beta-2 microglobulin (>3.5 mg/L, ≤3.5 mg/L)
- Cytopenia (neutropenia ALC ≤1.5 × 10^9/L, anemia: Hemoglobin ≤11 g/dL, thrombocytopenia: PLT ≤100× 10^9/L, all of the above, any of the above, none of the above)
- Prior PBC transfusion within 28 days before randomization
- Prior platelet transfusion within 28 days before randomization
- Constitutional symptoms (weight loss, fever, night sweats, fatigue, any of above)
- Creatinine clearance group (<60 mL/min, ≥60 mL/min)
- Absolute lymphocyte count
- Absolute neutrophil count
- Platelet count
- Hemoglobin level.

The medications will be coded following AZ standard drug dictionary/WHO Drug dictionary as applicable.

4.2.8 Concomitant and other treatments

Information on any treatment within the four weeks prior to initiation of study drug and all concomitant treatments given up to 30 days after discontinuation of study drug, or objective disease progression (whichever is later), with reasons for the treatment, will be recorded in the eCRF. Thereafter, only subsequent regimens of anti-cancer therapy will be recorded in eCRF.

Other anti-cancer therapies, investigational agents, and radiotherapy should not be given while the patient is on study drug.

Medications received prior to, concomitantly, or post-treatment will be coded using the WHODrug Dictionary Anatomical Therapeutic Chemical (ATC) Classification codes. Concomitant medications will be summarized for the FAS by ATC classification codes.

For the purpose of inclusion in prior and/or concomitant medication or therapy summaries, incomplete medication or radiotherapy start and stop dates will be imputed as detailed in Section 4.2.4.

Prior medications, concomitant and post-randomized treatment medications are defined based on imputed start and stop dates as follows:

- Prior medications are those taken prior to study drug with a stop date prior to the first dose of study drug.
- Concomitant medications are those with a stop date on or after the first dose date of study drug (and could have started prior to or during treatment).
- Post-treatment medications are those with a start date after the last dose date of study drug.

In addition, all post-treatment anti-cancer therapies will be summarized for the full analysis set.

The following summaries will be produced:

- Summary of concomitant medications
- Summary of Post study drug cancer therapies

All concomitant and other treatment data will be listed.

Missing coding terms should be listed and summarized as "Not coded".

4.2.9 Pharmacokinetic data

Plasma concentrations of acalabrutinib and ACP-5862 will be summarized for the PK analysis set by nominal sample time using standard summary statistics for PK concentrations (geometric mean, geometric coefficient of variation, geometric mean ± standard deviation, arithmetic mean, standard deviation, minimum, maximum and n). All plasma concentrations will be listed.

4.2.10 COVID-19

Additional analyses may be performed to explore the impact of coronavirus disease 2019 (COVID-19) on the safety and efficacy results reported for the study. The following may be explored:

- Safety data analysis separately for patients affected by COVID-19.
- Listing of protocol deviations related to COVID-19 and the reasons for the protocol deviation may be provided.

Sensitivity analysis on key efficacy endpoints may be performed if there is more than 10% of the total events are COVID-19 related. In these sensitivity analyses, any patient who had a death with primary or secondary cause as COVID-19 infection will be censored at their COVID-19 infection death date. COVID-19 deaths will be identified by primary/secondary cause of death.

4.2.11 China cohort analysis

The China cohort analysis will be performed using the same methodology as for the entire study population. The primary PFS analysis will be conducted on the China FAS using the same method outlined in Section 4.2.2.2. Efficacy analyses of the secondary endpoints ORR, DOR, TTNT, and OS will be performed for the China FAS. Key demographic and baseline characteristics and PK analysis will be repeated for the China cohort as well.

All statistical analyses for the China cohort will be considered exploratory. No adjustment for multiplicity will be made.

The majority of the safety and tolerability analyses outlined in Section 4.2.6 will be repeated for the China safety analysis set.

5. INTERIM ANALYSES

This study includes one interim analysis. The interim analysis will be performed to test for superiority of Arm A relative to Arm B.

This analysis will be performed when approximately 38 PFS events based on BICR have been observed in the study (25.3% maturity or 76% information fraction). The accrual period is assumed to be 30 months. The dropout rate is assumed to be 5% by 14 months for both arms. The interim analysis is anticipated to occur approximately 40 months after the first patient is randomized.

An independent data monitoring committee (IDMC) will be utilized for this study. The IDMC will evaluate efficacy for the interim analysis. Members of the IDMC will be external to the sponsor. An IDMC charter will be developed which will specify the Committee's responsibilities, authorities, and procedures along with details of the interim analysis planning, decision-making guidance, and dissemination of the results as well as the recommendations and decisions after the interim analysis.

6. CHANGES OF ANALYSIS FROM PROTOCOL

In Section 2 of this document China cohort analysis sets have been defined in addition to the analysis sets defined in the CSP. Nodular partial response has been clarified per IWCLL 2018 in Sections 1.1, 3.2, 3.3, 4.2.2 of this document.

7. REFERENCES

Fayers PM, Aaronson NK, Bjordal K, Groenvold M, Curran D, Bottomley A, on behalf of the EORTC Quality of Life Group. The EORTC QLQ-C30 Scoring Manual (3rd Edition). Published by: European Organisation for Research and Treatment of Cancer, Brussels 2001.

Fleischer F, Gaschler-Markefski B, Bluhmki E. How is retrospective independent review influenced by investigator-introduced informative censoring: A quantitative approach. Stat Med. 2011; 30(29):3373-86

Hallek M, Cheson BD, Catovsky D, Caligaris-Cappio F, Dighiero G, Dohner H, et al. iwCLL guidelines for diagnosis, indications for treatment, response assessment, and supportive management of CLL. Blood 2018;131(25):2745-761.

Michallet AS, et al. Rituximab plus bendamustine or chlorambucil for chronic lymphocytic leukemia: primary analysis of the randomized, open-label MABLE study. Haematologica. 2018;103(4):698-706.

van Hout, B, Janssen, MF, Feng, YS, Kohlmann, T, Busschbach, J, Golicki, D, et al. Interim scoring for the EQ-5D-5L: Mapping the EQ-5D-5L to EQ-5D-3L value sets. Value Health. 2012;15(5):708-15.

Yellen SB, Cella DF, Webster K, et al. Measuring fatigue and other anemia-related symptoms with the Functional Assessment of Cancer Therapy (FACT) measurement system. J Pain Symptom Manage. 1997;13(2):63–74.

SIGNATURE PAGE

This is a representation of an electronic record that was signed electronically and this page is the manifestation of the electronic signature

Document Name: d822bc00001-sap-ed-2		
Document Title:	Statistical Analysis Plan Edition 2	
Document ID:	CCI [redacted]	
Version Label:	2.0 CURRENT LATEST APPROVED	
Server Date (dd-MMM-yyyy HH:mm 'UTC'Z)	**Signed by**	**Meaning of Signature**
05-Jan-2024 18:09 UTC	PPD	Management Approval
09-Jan-2024 01:44 UTC	PPD	Author Approval

CCI [redacted]