Official Title: A Prospective, Open-Label, Multicenter Randomized Phase III Study to Compare the Efficacy and Safety of a Combined Regimen of Venetoclax and Obinutuzumab Versus Fludarabine, Cyclophosphamide, and Rituximab (FCR)/ Bendamustine and Rituximab (BR) in FIT Patients with Previously Untreated Chronic Lymphocytic Leukemia (CLL) Without Del(17P) or TP53 Mutation

NCT Number: NCT04285567

Document Date: Statistical Analysis Plan Version 2: 13-Mar-2024

STATISTICAL ANALYSIS PLAN

STUDY TITLE: A PROSPECTIVE, OPEN-LABEL, MULTICENTER RANDOMIZED PHASE III STUDY TO COMPARE THE EFFICACY AND SAFETY OF A COMBINED REGIMEN OF VENETOCLAX AND OBINUTUZUMAB VERSUS FLUDARABINE, CYCLOPHOSPHAMIDE, AND RITUXIMAB (FCR)/ BENDAMUSTINE AND RITUXIMAB (BR) IN FIT PATIENTS WITH PREVIOUSLY UNTREATED CHRONIC LYMPHOCYTIC LEUKEMIA (CLL) WITHOUT DEL(17P) OR TP53 MUTATION

STUDY NUMBER:	CO41685
STUDY NAME:	CRISTALLO
VERSION NUMBER:	2
ROCHE COMPOUND(S):	Venetoclax (GDC-0199 [ABT-199]; RO5537382), Obinutuzumab (GA101, RO5072759)
EUDRACT NUMBER:	2019-003327-37
IND NUMBER:	115,045
NCT NUMBER:	NCT04285567
PLAN PREPARED BY:	[redacted], M. Sc.

STATISTICAL ANALYSIS PLAN APPROVAL

SPONSOR:	F. Hoffmann-La Roche Ltd
LEGAL REGISTERED ADDRESS:	Grenzacherstrasse 124 4070 Basel, Switzerland
DATE FINAL:	See electronic date stamp below

STATISTICAL ANALYSIS PLAN VERSION HISTORY

This Statistical Analysis Plan (SAP) was developed based on the Roche SAP model document updated on 28 February 2022.

SAP Version	Approval Date	Based on Protocol (Version, Approval Date)
2	See electronic date stamp on the last page of this document	V 5.0, 7 June 2023
1	04-Sep-2023	V 5.0, 7 June 2023

STATISTICAL ANALYSIS PLAN AMENDMENT RATIONALE

Key changes to the SAP from version 1 to Version 2, along with the rationale(s) for each change, are summarized below.

Section	Description of Change	Rationale for Change
Table 1 & Section 4.3.1	Censoring Rule for PFS has been extended to all patients (irrespective of the event occurrence)	Health authorities comment
Section 4.2.3	Additional sensitivity analysis for the primary endpoint has been added for patients whose absence of Chromosome 17P Del has been confirmed at the local evaluation.	Additional sensitivity analysis is added to evaluate the potential impact of non-confirmed Chromosome 17P Del absence on the study.
Section 4.7	Details on estimation of expected events & timelines for each PFS analysis have been added	Health authorities comment

Additional minor changes have been made throughout to improve clarity and consistency.

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

LIST OF ABBREVIATIONS AND DEFINITIONS OF TERMS

Abbreviation or Term	Description
AE	adverse event
AESI	adverse event of special interest
BMI	body mass index
BR	bendamustine and rituximab
CLL	Chronic Lymphocytic Leukemia
CMH	Cochran-Mantel-Haenszel
COVID-19	Coronavirus Disease 2019 (COVID-19)
CR	complete response
CRi	Complete Remission with incomplete blood count recovery
DOR	Duration of Response
eCRF	electronic Case Report Form
EDC	Electronic Data Capture
EFS	Event-free survival
EORTC QLQ-C30	European Organisation for Research and Treatment of Cancer Quality of Life Questionnaire Core 30
EQ-5D-5L	EuroQol 5 Dimension questionnaire
FAS	full analysis set
FCR	fludarabine, cyclophosphamide and rituximab
FIT	patients with a cumulative illness rating scale [CIRS] score ≤ 6 and a normal creatinine clearance of ≥ 70 mL/min
HR	hazard ratio
IA	interim analysis
ICH	International Council on Harmonization
iDMC	independent Data Monitoring Committee
IRF	Independent Review Facility
ITT	intent to treat
IxRS	interactive voice/web-based system
iwCLL	International Workshop on Chronic Lymphocytic Leukemia

Abbreviation or Term	Description
LLoQ	lower limit of quantification
MDD	minimally detectable difference
MDASI-CLL	M.D. Anderson Symptom Inventory
MedDRA	Medical Dictionary for Regulatory Activities
MRD	Minimal Residual Disease
NCI CTCAE	National Cancer Institute Common Terminology Criteria for Adverse Events
NGS	Next Generation Sequencing
OS	overall survival
PD	progressive disease
PFS	Investigator-assessed progression-free survival
PR	partial response
PRO	patient-reported outcomes
PK	pharmacokinetic
RMST	restricted mean survival time
SAE	serious adverse event
SAP	Statistical Analysis Plan
SMART	sequential multiple assignment randomized study
SMQs	standardized MedDRA queries
TLS	Tumor Lysis Syndrome

1. INTRODUCTION

This SAP is based on Protocol CO41685 (CRISTALLO) version 5, a prospective, open-label, multicenter randomized Phase III study designed to evaluate the efficacy and safety of venetoclax (Venclexta® or Encyst®) in combination with obinutuzumab (Gazyva® or Gazyvaro®) (VEN + G) versus fludarabine, cyclophosphamide, and rituximab (FCR), and bendamustine and rituximab (BR) in FIT (patients with a cumulative illness rating scale [CIRS] score ≤ 6 and a normal creatinine clearance of ≥ 70 mL/min) patients with previously untreated chronic lymphocytic leukemia (CLL) without del(17p) or TP53 mutations. This SAP provides details of the planned analyses and statistical methods for Study CO41685. The background of this study can be found in the study protocol.

There are no changes to the planned analyses described in the protocol.

1.1 OBJECTIVES AND ENDPOINTS AND ESTIMANDS

This study will evaluate the efficacy and safety of venetoclax (Venclexa® or Venclyxto®) in combination with obinutuzumab (Gazyva® or Gazyvaro®) (VEN+G) compared with FCR and BR in FIT patients with previously untreated CLL without del(17p) or *TP53* mutations.

The primary study objective and corresponding endpoint, as well as the secondary efficacy objectives and the subset of corresponding secondary efficacy endpoints which are not derived from patient-reported outcome (PRO) measures, are reported using the estimand framework as in Table 1, in accordance with the International Conference on Harmonization E9 (R1) statistical principles for clinical trials (ICH 2020).

Table 1 Primary and Secondary Objectives and Corresponding Estimands

Primary Objective(s)	Estimand Definition
• To evaluate the efficacy of VEN+G compared with FCR/BR	• Population: Participants with previously untreated CLL without del(17p) or *TP53* mutations in FIT patients as defined by the study inclusion and exclusion criteria (see Sections 4.1.1 and Section 4.1.2 of the protocol, respectively) • Endpoint: MRD response rate, determined as the proportion of patients with MRD-negativity (defined as <1 CLL cell in 10,000 leukocytes), measured in peripheral blood (PB) using next-generation sequencing (NGS) at Month 15 • Treatment (see Section 3.1.4.2 of the Protocol): – Experimental arm A: (VEN+G) Venetoclax (Cycle 1-12)+Obinutuzumab (Cycle 1-6)

	– Control arm B: (FCR/BR) fludarabine, cyclophosphamide, and rituximab (Cycle 1-6) OR bendamustine and rituximab (Cycle 1-6) ● Intercurrent events and handling strategies: – Discontinuation of treatment for any reason (except PD) prior to Month 15: treatment policy strategy – Start of non-protocol anti-CLL therapy prior to Month 15: non-responder – Death prior to 15 months: non-responder – Missing/not evaluable MRD sample at Month 15: non responder ● Population-level summary: Rates for MRD
Secondary Objective(s)	**Estimand Definition**
● To evaluate the efficacy of VEN + G compared with FCR/BR	● Population: as defined above in the primary objective ● Endpoint: PFS, defined as the time from randomization to the first occurrence of disease progression, or death from any cause; disease progression assessed by the investigators using the International Workshop on CLL (iwCLL) criteria (Hallek et al. 2018) ● Treatment: as defined above in the primary objective ● Intercurrent events and handling strategies: – Start of a new anti-CLL treatment prior to an PFS event: censored – Discontinuation from study treatment for any reason prior to an PFS event: treatment policy – Missing two or more consecutive response assessments: censored ● Population-level summary: hazard ratio for PFS
● To evaluate the efficacy of VEN + G compared with FCR/BR	● Population: as defined above in the primary objective ● Endpoint: MRD response rate in PB at the end of treatment response visit (9 months for FCR/BR vs 15 months for VEN + G) ● Treatment: as defined above in the primary objective ● Intercurrent events and handling strategies – Discontinuation of treatment for any reason (except PD) prior to Month 9/15: treatment policy strategy – Start of non-protocol anti-CLL therapy prior to Month 9/15: non-responder – Death prior to 9/15 months: non-responder – Missing/not evaluable MRD sample at Month 9/15: non responder ● Population-level summary: Rates for MRD
● To evaluate the efficacy of VEN + G compared with FCR/BR	● Population: as defined above in the primary objective

	• Endpoint: MRD response rate in BM at the end of treatment response visit (9 months for FCR/BR vs 15 months for VEN + G)
	• Treatment: as defined above in the primary objective
	• Intercurrent events and handling strategies
	  – Discontinuation of treatment for any reason (except PD) prior to Month 9/15: treatment policy strategy
	  – Start of non-protocol anti-CLL therapy prior to Month 9/15: non-responder
	  – Death prior to 9/15 months: non-responder
	  – Not evaluable MRD sample at Month 9/15: non responder
	• Population-level summary: Rates for MRD

• To evaluate the efficacy of VEN + G compared with FCR/BR	• Population: as defined above in the primary objective
	• Endpoint: ORR, which includes CR, CRi, and PR at Month 15 assessment
	• Treatment: as defined above in the primary objective
	• Intercurrent events and handling strategies
	  – Discontinuation of treatment for any reason (except PD) prior to Month 9/15: treatment policy strategy
	  – Start of non-protocol anti-CLL therapy prior to Month 15: non-responder
	  – Death prior to Month 15: non-responder
	  – Missing CR/CRi/PR at Month 15: non responder
	• Population-level summary: ORR rate

• To evaluate the efficacy of VEN + G compared with FCR/BR	• Population: as defined above in the primary objective
	• Endpoint: CR rate, which includes CR and CRi at Month 15 assessment
	• Treatment: as defined above in the primary objective
	• Intercurrent events and handling strategies
	  – Discontinuation of treatment for any reason (except PD) prior to Month 9/15: treatment policy strategy
	  – Start of non-protocol anti-CLL therapy prior to Month 15: non-responder
	  – Death prior to Month 15: non-responder
	  – Missing CR/CRi at Month 15: non responder
	• Population-level summary: CRR rate

• To evaluate the efficacy of VEN + G compared with FCR/BR	• Population: as defined above in the primary objective • Endpoint: MRD response rate in PB of patients with a CR/CRi at Month 15 • Treatment: as defined above in the primary objective • Intercurrent events and handling strategies

	– Discontinuation of treatment for any reason (except PD) prior to Month 15: treatment policy strategy – Start of non-protocol anti-CLL therapy prior to Month 15: non-responder – Death prior to Month 15: non-responder – Missing/Non-evaluable MRD sample at Month 15: non responder ● Population-level summary: MRD rate
● To evaluate the efficacy of VEN + G compared with FCR/BR	● Population: as defined above in the primary objective ● Endpoint: MRD response rate in BM of patients with a CR/CRi at the end of treatment visit (9 months for FCR/BR vs 15 months for VEN + G) ● Treatment: as defined above in the primary objective ● Intercurrent events and handling strategies – Discontinuation of treatment for any reason (except PD) prior to Month 9/15: treatment policy strategy – Start of non-protocol anti-CLL therapy prior to Month 9/15: non-responder – Death prior to Month 9/15: non-responder – Missing/Non-evaluable MRD sample at Month 9/15: non responder ● Population-level summary: MRD rate
● To evaluate the efficacy of VEN + G compared with FCR/BR	● Population: as defined above in the primary objective ● Endpoint: Duration of objective response (DOR), defined as the time from the first occurrence of a documented objective response (OR) to the time of PD or death from any cause ● Treatment: as defined above in the primary objective ● Intercurrent events and handling strategies – Discontinuation from study treatment for any reason after OR and prior to a PFS event: treatment policy strategy – Start of a new anti-CLL treatment after OR and prior to a PFS event: Date of last disease assessment prior to start of new anti-CLL after OR will be considered as censoring date ● Population-level summary: hazard ratio for DOR
● To evaluate the efficacy of VEN + G compared with FCR/BR	● Population: as defined above in the primary objective ● Endpoint: Best response achieved (CR, CRi, PR, stable disease [SD], or PD) up to and including the assessment at Month 15 ● Treatment: as defined above in the primary objective ● Intercurrent events and handling strategies – Discontinuation of treatment for any reason (except PD): treatment policy strategy

	– Start of non-protocol anti-CLL therapy prior to an assessment: non-responder – Death prior to an assessment: non-responder ● Population-level summary: MRD rate
● To evaluate the efficacy of VEN + G compared with FCR/BR	● Population: as defined above in the primary objective ● Endpoint: Event-free survival (EFS), defined as the time between the date of randomization and the date of PD/relapse, death, or start of a new anti-CLL therapy ● Treatment: as defined above in the primary objective ● Intercurrent events and handling strategies – Discontinuation from study treatment for any reason prior to a PFS event: treatment policy strategy ● Population-level summary: hazard ratio for EFS
● To evaluate the efficacy of VEN + G compared with FCR/BR	● Population: as defined above in the primary objective ● Endpoint: Overall Survival (OS), defined as the time from randomization to the date of death from any cause ● Treatment: as defined above in the primary objective ● Intercurrent events and handling strategies – Discontinuation from study treatment for any reason: treatment policy strategy – Start of a new anti-CLL treatment: “treatment policy strategy" ● Population-level summary: hazard ratio for OS
● To evaluate the TLS rate in VEN + G arm only	● Population: as defined above in the primary objective ● Endpoint: TLS risk reduction rate in Arm A, defined as the reduction in the proportion of patients who were TLS high-risk after 3 doses of obinutuzumab compared to the proportion of patients who were TLS high-risk at baseline ● Treatment: only Arm A ● No estimand and intercurrent events have been defined ● Population level summary: rates for TLS
● To evaluate the number of hospitalizations in VEN + G arm only	● Population: as defined above in the primary objective ● Endpoint: Reduction in mandatory hospitalizations during venetoclax ramp-up in Arm A patients ● Treatment: only Arm A ● No estimand and intercurrent events have been defined ● Population level summary: number of hospitalizations

BM = bone marrow; CLL = chronic lymphocytic leukemia; CR = complete response; CRi = complete remission with incomplete blood count recovery; DOR = Duration of objective response; EFS = event-free survival; HRQol = health-related quality of life; MRD = minimal residual disease; PB = peripheral blood; PD = progressive disease; PFS = investigator assessed progression-free survival; PR = partial response; ORR = objective response rate; OS = overall survival; TLS= tumor lysis syndrome.

Table 2 Supportive and Exploratory Objectives and Endpoints

• To evaluate the quality of life of participants treated with VEN + G compared with FCR/BR	• Population: Participants with previously untreated CLL without del(17p) or TP53 mutations in FIT patients as defined by the study inclusion and exclusion criteria (see Sections 4.1.1 and Section 4.1.2 of the protocol, respectively) • Endpoint: – disease and treatment-related symptoms following treatment with the combination of VEN + G compared with FCR/BR in patients with previously untreated CLL without del(17p) or TP53 mutation as measured by the M.D. Anderson Symptom Inventory – changes in physical functioning, role functioning, and global health status/quality of life (QoL) following treatment with the combination of VEN + G compared with FCR/BR in patients with previously untreated CLL without del(17p) or TP53 mutation as measured by the European Organisation for Research and Treatment of Cancer Quality of Life Questionnaire Core 30
• To evaluate the health status utility of VEN + G compared with FCR/BR	• Population: Participants with previously untreated CLL without del(17p) or TP53 mutations in FIT patients as defined by the study inclusion and exclusion criteria (see Sections 4.1.1 and Section 4.1.2 of the protocol, respectively) • Endpoint: – Health utility effects of treatment with the combination of VEN + G compared with FCR/BR in patients with previously untreated CLL without del(17p) or TP53 mutation as measured by the EuroQol 5 Dimension (EQ-5D-5L) questionnaire
• To evaluate the efficacy of VEN + G compared with FCR/BR	• To investigate the relationship between various baseline markers and clinical outcome parameters in both arms of the study – Baseline markers will include, but are not limited to, CLL fluorescence in situ hybridization (FISH) (17p-, 11q-, 13p-, -12q), immunoglobulin heavy chain variable region (IgHV) mutation status, p53 mutation status, serum parameters, and other CLL disease markers.

• To evaluate the biomarker activities of VEN + G compared with FCR/BR	• Relationship between efficacy outcomes and potential biomarkers, including Bcl-2 expression, for patients treated with VEN + G compared with those treated with FCR/BR • To evaluate potential biomarkers, including CLL driver mutations and genomic complexity that are prognostic and/or predictive of response and resistance to treatment with VEN + G or with FCR/BR • To assess new technologies, such as the circulating tumor DNA (ctDNA) assessment, in measuring MRD and monitoring disease clonal evolution

BR = bendamustine and rituximab; CLL = chronic lymphocytic leukemia; FCR = fludarabine, cyclophosphamide, and rituximab; FISH = fluorescence in situ hybridization

1.2 STUDY DESIGN

This is an open-label, international, multicenter, randomized, Phase III study to investigate the efficacy and safety of VEN + G compared with FCR/BR in FIT patients (FIT defined by a CIRS/score ≤ 6 and a normal creatinine clearance of ≥ 70 mL/min) with previously untreated CLL without del(17p) or *TP53* mutation requiring treatment.

Approximately 165 patients will be recruited from approximately 40 centers globally and randomly assigned in a 1:1 ratio to receive either VEN + G (Arm A) or FCR/BR (Arm B). Randomization will be stratified according to the following factors:

- IGHV mutation: yes or no
- Binet stage (3 levels): A, B, or C
- Age: ≤ 65 or > 65 years

Patients randomized to Arm A (VEN + G) will receive 12 cycles of treatment, each with a precise duration of 28 days. During the first cycle, obinutuzumab will be administered intravenously (IV) on Days 1 (and 2), 8, and 15 as well as on Day 1 of Cycles 2−6. Continuous daily administration, with a 5-week ramp-up period, of venetoclax starts on Cycle 1, Day 22. Venetoclax will be administered until the end of Cycle 12.

Because it is possible that debulking from obinutuzumab therapy before starting venetoclax dosing already significantly reduces the risk of tumor lysis syndrome (TLS), TLS prophylactic measures will be re-assessed at the time of starting the first dose of venetoclax (see Protocol Section 5.1.6.1).

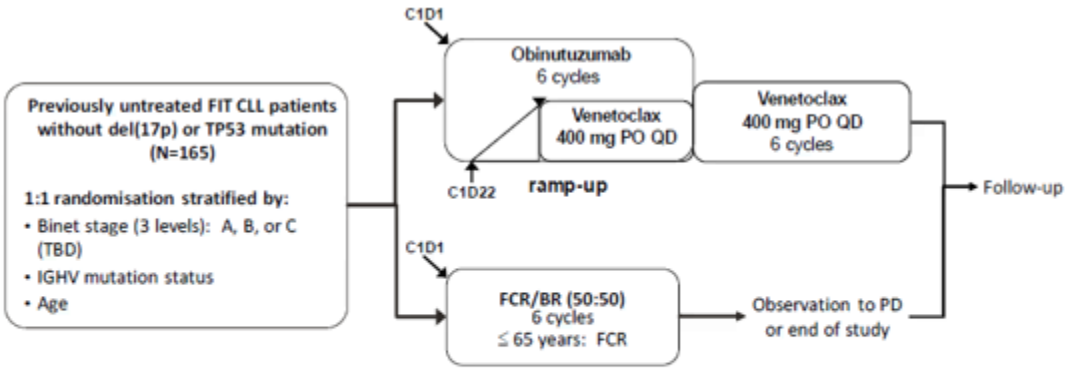
Patients randomized to Arm B (FCR/BR expected 50:50) (see Figure 1) will receive 6 cycles of FCR consisting of a single cycle of a single infusion of rituximab on Day 1 and fludarabine and cyclophosphamide infusions on Days 1−3 of each 28-day cycle or bendamustine as infusions on Days 1 and 2 and a single cycle of rituximab on Day 1 of each 28-day cycle.

All patients in Arm B can be considered for treatment with BR, whereas only patients ≤ 65 years of age will be eligible for FCR. The choice between FCR or BR in Arm B will be at the investigator's discretion.

During this study, adverse events (AE) will be graded according to the National Cancer Institute Common Terminology Criteria for Adverse Events (NCI CTCAE) (v5).

Figure 1 presents an overview of the study design. The schedules of activities for Arm A and Arm B are provided in Appendix 1 and Appendix 2 of the protocol, respectively.

Figure 1 Study Schema



BR=bendamustine and rituximab; C=Cycle; CIRS=cumulative illness rating scale; CLL=chronic lymphocytic leukemia; D=Day; FCR=fludarabine, cyclophosphamide, and rituximab; FIT=defined by a cumulative illness rating scale [CIRS] score ≤6 and a normal creatinine clearance of ≥70 mL/min; IGHV= immunoglobulin heavy chain variable region; PD=progressive disease; PO=orally (by mouth); QD=once daily; TBD=to be determined.
Note: Each cycle will consist of 28 days.

1.2.1 Treatment Assignment

This is a randomized, open-label study. After initial written informed consent has been obtained, all screening procedures and assessments have been completed, and eligibility has been established for a patient, the study site will obtain the patient's identification number and treatment assignment will be performed by an interactive voice/web-based system (IxRS).

Patients will be randomly assigned to one of two treatment arms: Arm A (VEN + G) or Arm B (FCR/BR). Randomization will occur in a 1:1 ratio using a block-stratified randomization method to ensure a balanced assignment to each treatment arm.

Randomization will be stratified according to the following criteria:

- IGHV mutation: yes or no
- Binet stage (3 levels): A, B, or C
- Age: ≤ 65 or > 65

Screening tests will be used to assess the Binet stage allowing for stratification and completion of the IxRS screening call. If a patient's Binet stage progresses between screening and randomization, then the Binet stage at the screening time point should be used for stratification.

In Arm B, the investigator at each site will determine whether the FCR or BR regimen will be used for the individual patient, based on their local practice. All patients in Arm B can be considered for treatment with BR, whereas only patients ≤65 years of age will be eligible for FCR.

1.2.1.1 Treatment Assignment

Treatment Arm A: Venetoclax and Obinutuzumab

The doses of venetoclax and obinutuzumab to be administered at each study visit is presented in Table 3.

Table 3 Treatment Arm A: VEN + G Doses per Study Visit

Cycle, Day	Dose
Obinutuzumab [a]	
Cycle 1, Day 1 [b]	100 mg or 1000 mg (follow splitting rules)
Cycle 1, Day 2 [b]	900 mg (if 100 mg on Cycle 1 Day 1)
Cycle 1, Day 8	1000 mg
Cycle 1, Day 15	1000 mg
Cycles 2−6, Day 1	1000 mg
Venetoclax [c, d, e]	
Cycle 1, Day 22−28	20 mg daily
Cycle 2, Day 1−Day 7	50 mg daily
Cycle 2, Day 8−Day 14	100 mg daily
Cycle 2, Day 15−Day 21	200 mg daily
Cycle 2, Day 22−Day 28	400 mg daily
Cycle 3, Day 1−end of Cycle 12	400 mg daily

TLS = tumor lysis syndrome; VEN + G = venetoclax + obinutuzumab.

Note: Cycles will comprise 28 days. Treatment with obinutuzumab will continue for 6 cycles, and venetoclax will end at Cycle 12.

[a] IV infusion; overnight hospitalization may be required on Day 1 following the first infusion of obinutuzumab (100 mg).

[b] Only the first dose (1000 mg) of obinutuzumab administration can be split over 2 days. Two infusion bags should be prepared for the infusion on Days 1 and 2 (100 mg for Day 1 and 900 mg for Day 2). If the first bag is completed without modifications of the infusion rate or interruptions, the second bag may be administered on the same day (no dose delay necessary, no repetition of premedication), provided that appropriate time, conditions, and medical supervision are available throughout the infusion. If there are any modifications of the infusion rate or interruptions during the first 100 mg, the second bag should be administered the following day. If patients require a delay of greater than 24 hours between the 100-mg and 900-mg infusions of obinutuzumab, please consult Section 4.3.2.2 and the Obinutuzumab Investigator's Brochure regarding drug stability.

[c] All patients will be assessed at screening and categorized in a TLS-risk category. Because of possible tumor de-bulking after obinutuzumab therapy, investigators will reassess TLS risk after obinutuzumab induction and prior to starting the venetoclax ramp-up (Days 16−21) and may assign patients to a lower risk group (see Section 5.1.7.1 for more details).

[d] The 20-mg and 50-mg doses of venetoclax will be administered in the hospital for patients who are at high risk of TLS, or if indicated to hospitalize, and thereafter at home daily for 7 days. The dose will increase every 7 days to the target dose of 400 mg, and venetoclax will be administered at home unless a patient is indicated to hospitalize.

[e] Oral tablets.

Treatment Arm B: FCR/BR

Fludarabine, Cyclophosphamide, and Rituximab

Only patients ≤ 65 years of age will be eligible for FCR due to the associated high risk of severe neutropenia and infections. Patients will receive 6 cycles of fludarabine, cyclophosphamide, and rituximab each cycle with a duration of 28 days:

- Fludarabine IV will be given on Days 1−3 (Cycles 1−6) at a dosage of 25 mg/m²
- Cyclophosphamide IV will be given on Days 1−3 (Cycles 1−6) at a dosage of 250 mg/m²
- Rituximab IV will be given before the administration of chemotherapy at a dosage of 375 mg/m$_2$ in the first cycle and at a dosage of 500 mg/m² in Cycles 2−6 with premedication according to the clinical practice of the participating sites (Eichorst et al. 2016)

Bendamustine and Rituximab

All patients in Arm B, regardless of age, can be considered for treatment with BR. Patients will receive 6 cycles of bendamustine and rituximab, each cycle with a duration of 28 days:

- Bendamustine IV will be given on Days 1 and 2 (Cycles 1−6) at a dosage of 90 mg/m²
- Rituximab IV will be given before the administration of chemotherapy at a dosage of 375 mg/m² in the first cycle and at a dosage of 500 mg/m² in Cycles 2−6 with premedication according to the clinical practice of the participating sites (Eichorst et al. 2016).

1.2.2 Independent Review Facility

No Independent Review Facility (IRF) is planned for this study.

1.2.3 Data Monitoring

The Sponsor will be responsible for data management of this study, including quality checking the data. Data entered manually will be collected via electronic data capture (EDC) through use of electronic Case Report Form (eCRF). Sites will be responsible for data entry into the EDC system. In the event of discrepant data, the Sponsor will request data clarification from the sites, which the sites will resolve electronically in the EDC system.

The Sponsor will produce an EDC Study Specification document that describes the quality checking to be performed on the data. Central laboratory data will be sent directly to the Sponsor, using the Sponsor's standard procedures to handle and process the electronic transfer of these data.

eCRFs and correction documentation will be maintained in the EDC system's audit trail. System backups for data stored by the Sponsor and records retention for the study data will be consistent with the Sponsor's standard procedures.

PRO data will be collected through the use of an electronic device provided by a vendor (see Protocol Section 7.3 for details).

Back up PRO data can be collected on a web-based system. The use of this electronic back up allows for completion of all questionnaires on a web portal. This web back-up is to be used only when the electronic device (tablet) is not functioning as expected.

For information on how to activate the web back up, please refer to the electronic Certificate of Analysis (eCOA) quick reference guide provided by the vendor.

1.2.4 Analysis Timing

There will be one analysis for the primary endpoint: Minimal Residual Disease (MRD) occurring at around 46 months after the first patient is randomized. The primary efficacy analysis is triggered when the last patient randomized has reached the Month 15 MRD assessment.

For the secondary endpoint: Investigator-assessed Progression Free Survival (PFS), two interim efficacy analyses and one final analysis with alpha spending methodologies will be applied. The first interim analysis which is conducted at the MRD primary analysis is expected to occur when 25 PFS events will be reached (when 57% of the events have been documented). In the event the expected number of PFS events are reached before or after the time of MRD analysis, the interim PFS analysis will be conducted at the time of MRD analysis regardless of the achievement of the planned number of events. In this scenario, the information fraction along with the stopping boundaries will be re-evaluated based on the number of events reached at that time.

The second interim analysis will be conducted when 33 events will be reached (i.e., 33 events will trigger the interim analysis) when 75% of the PFS events have been documented and expected to occur at around 56 months after the first patient is randomized.

The final PFS analysis will take place when 44 events will have occurred, expected to occur at around 74 months after the first randomized patient.

Upon the readout of the primary analysis, the Sponsor will report the MRD data, along with data from the first interim PFS analysis and key secondary endpoints at the primary CSR and for publication purposes.

The interim analysis is described in detail in Section 4.7.

2. STATISTICAL HYPOTHESES AND SAMPLE SIZE DETERMINATION

2.1 STATISTICAL HYPOTHESES

The primary efficacy endpoint is: MRD response rate (determined as the proportion of patients with MRD-negativity defined as < 1 cell in 10,000 leukocytes) measured in peripheral blood (PB) using next-generation sequencing (NGS) at Month 15.

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

– MRD response rate of VEN + G versus FCR/BR (i.e., H_0: VEN + G – FCR/BR = pre-specified difference versus H_1: VEN + G – FCR/BR ≠ pre-specified difference)

Hypothesis tests will be two-sided, unless otherwise indicated. The type I error (α) for this study is 0.05 (two-sided).

2.2 SAMPLE SIZE DETERMINATION

The sample size for the primary MRD analysis is determined based on the hypothesis to be tested to establish clinically relevant statistical superiority in the MRD response rate. To detect an improvement in MRD response rate at a two-sided significance level of 0.05, approximately 165 patients are required at the MRD primary analysis to achieve a power of 92% assuming a target improvement in MRD response rate of 30%. The minimum detectable difference (MDD) of MRD response rate is 14.1%.

The primary MRD analysis is expected to occur approximately 46 months after first-patient in (FPI).

The calculation of sample size and estimates of the MRD analysis timelines are based on the following assumptions, which are based on the completed enrollment:

- Patients will be randomized to FCR/BR and VEN + G arms in a 1:1 ratio
- Two-sample test for proportions at the two-sided 0.05 level of significance
- MRD response rate for FCR/BR arm at Month 15 (52%)
- FCR/BR ratio in the control arm is 55/45
- > 90% power to detect difference in MRD response rate of 30% (response rate of 82% for VEN + G arm) at Month 15
- Proportion of patients with non-evaluable MRD including samples which cannot be assayed and patients who have died or discontinued the study prior to Month 15 MRD assessment (15%)
- Misclassification rate for MRD status assessed by clonoSEQ NGS assay.

The sample size calculation was performed using rpact package in R version 4.0.3.

3. ANALYSIS SETS

The participant analysis sets for the purposes of analyses are defined in Table 4.

Table 4 Participant Analysis Sets

Participant Analysis Set	Description
Full Analysis Set (FAS)	All randomized participants; participants will be included in the analyses according to the treatment they were randomized to.
Safety Population (SE)	All participants who receive at least one dose of any study treatment (i.e., obinutuzumab, venetoclax, or FCR/BR); participants will be analyzed according to the treatment that they actually received (which may be different from the treatment the patient was randomized to). Patients receiving any dose of venetoclax or obinutuzumab will be included in the venetoclax + obinutuzumab arm; all other treated patients will be included in the control arm.

FAS = full analysis set; SE = safety analysis set;
FAS will be used for the analysis of Patients Reported Outcomes (PRO).

4. STATISTICAL ANALYSES

Unless otherwise specified, the analyses described in this section are based on all randomized patients.

The timing for interim and final analyses are provided in Section 1.2.4.

Unless specified otherwise, all continuous variables will be summarized with descriptive statistics (e.g., number of non-missing values, mean, standard deviation, median, minimum, and maximum). All categorical variables will be summarized with frequency counts and percentages. Data will be presented by treatment arm.

Hypothesis tests will be two-sided, unless otherwise indicated. The type I error (alpha) for this study is 0.05 (two-sided).

4.1 GENERAL CONSIDERATIONS

For all efficacy analyses (including PROs), patients will be grouped according to the treatment assigned at randomization (full analysis set [FAS] population). For all safety analyses, patients will be grouped according to the treatment received (patients with any dose of venetoclax or obinutuzumab will be analyzed in the VEN+G arm).

Baseline for non-safety analyses is defined as the last available measurement obtained on or prior to randomization. Baseline for safety analyses is defined as the last available measurement prior to first exposure to any of the study drugs. Patients with missing baseline assessments will not be imputed.

Data collected in scheduled and unscheduled visits will be mapped to visits that appear in the schedule of assessments per the protocol using the actual study day of assessment. If there are multiple values in the same visit window, values closest to the target study day will be used in the analysis.

For OS, data from patients who did not have death documented will be censored on the last date they were known to be alive. Patients who do not have information after baseline will be censored at the date of randomization. For PFS, data for patients who do not have documented disease progression or who have not died will be treated as censored observations on the date of the last response assessment. If no assessments were performed after the baseline visit, PFS data will be censored at the date of randomization. A detailed censoring rules for PFS have been provided in Table 5. For the response endpoints, patients with no response assessments (for any reason) will be considered non-responders.

For the multi-item European Organization for Research and Treatment of Cancer (EORTC) Quality of Life-Core 30 (QLQ-C30) subscales with 50% or more of the constituent items completed, a pro-rated score will be computed consistent with the scoring manuals and validation papers (see the protocol). For subscales with less than 50% of the items completed, the subscale will be considered as missing.

All patients have been assessed for MRD response in PB samples using NGS and will also be assessed for MRD response in the bone marrow (BM) sample using NGS in responders (i.e., partial response [PR] or suspected complete response [CR]).

According to the estimands framework, intercurrent events can be handled in different ways (i.e., composite strategies or treatment policy). As per ICH-9 estimands addendum, treatment policy strategy is defined as follows: the occurrence of the intercurrent event (IE) is considered irrelevant in defining the treatment effect of interest: the value for the variable of interest is used regardless of whether or not the intercurrent event occurs.

4.2 PRIMARY ENDPOINTS ANALYSIS

The primary efficacy objective for this study is to evaluate the efficacy of VEN + G compared with FCR/BR on the basis of the following endpoint:

MRD response rate, determined as the proportion of patients with MRD-negativity (defined as <1 CLL cell in 10,000 leukocytes), measured in PB using NGS at Month 15.

4.2.1 Definition of Primary Endpoint(s)/Estimand(s)

MRD response rate measured in PB at month 15, as defined in Section 1.1 of SAP (see Table 1).

MRD response rate is determined as the proportion of patients with MRD-negativity (defined as < 1 CLL cell in 10,000 leukocytes), measured in PB at Month 15. This will be measured on the basis of assessments performed centrally using NGS with MRD negativity defined using a cutoff of 10^{-4} (less than 1 CLL cell in 10,000 leukocytes). MRD

status derivation will be provided by the vendor and determined through the following calculations:

- If CLL count (as a % of leukocytes) ≥ 0.01%; then MRD response status = non-responder (positive)
- If CLL count (as a % of leukocytes) < 0.01% AND limit of detection (LOD) < 0.01% then MRD response status = responder (negative)
- If CLL count (as a % of leukocytes) < LOD is considered Indeterminate, then MRD response status = below limit of quantification (BLQ) corresponding to Non-evaluable (non-responder)

The primary estimand is defined as follows:

Population set: FAS

Variable: MRD response rate

Intercurrent Events:

  - Exposure to new anti-CLL treatment prior to Month 15
  - Discontinuation of treatment for any reason (except progressive disease [PD]) prior to Month 15
  - Death prior to Month 15
  - Missing/Non-evaluable MRD sample at Month 15

Handling of intercurrent events: a treatment policy with regards to the discontinuation of treatment (except PD) prior to Month 15 will be applied. Intercurrent events such as death prior to month 15, exposure to new anti-CLL treatment prior to Month 15 and missing/non-evaluable MRD sample at Month 15 will be considered as non-responder.

Patients discontinuing from the study for any reasons prior to Month 15 will not have any MRD sample, therefore they will be automatically considered non-responders.

Population-level summary: MRD rate

4.2.2 Main Analytical Approach for Primary Endpoint(s)

MRD response rate in the 2 randomized treatment arms will be compared based on Cochran-Mantel-Haenszel (CMH) test according to the following stratification factors as entered in the IxRS:

- IGHV mutation: yes or no
- Binet stage (3 levels): A, B, or C
- Age: ≤ 65 or > 65 years

If the null hypothesis of no difference in MRD response rate is rejected and the observed MRD response rate is favorable for the VEN + G experimental arm, then it is concluded that VEN + G significantly increases the chance to respond compared to FCR/BR. Difference in MRD response rate will be estimated and two-sided 95% Anderson-Hauck

CI will be provided. Within each arm, 2-sided 95% Clopper-Pearson exact CIs will also be presented.

This endpoint will be tested at the significance level alpha (α) = 5%.

4.2.3 Sensitivity Analyses for the Primary Endpoint

To check the robustness of the primary estimand for MRD and underlying assumptions, the following sensitivity analyses for MRD will be performed on the FAS population:

- If more than 5% of patients experience a confirmed/suspected coronavirus disease 2019 (COVID-19) diagnosis, a sensitivity analysis will be performed for the primary endpoint by removing any patients who experience a confirmed/suspected COVID-19 diagnosis.
- To assess the impact of discordant cases for Chromosome 17pdel between central evaluation (FISH) and local evaluation, a sensitivity analysis will be performed for the primary endpoint by removing any patients who resulted to be discordant.
- to assess the impact of cases whose Chromosome 17pdel absence was not confirmed at local evaluation, a sensitivity analysis will be performed for the primary endpoint by removing any patients who resulted not to be confirmed at the local evaluation.

4.2.4 Supplementary

Two supplementary analyses taking into account the impact of missing MRD assessments at Month 15 will be performed.

Two supplementary estimands are identified:

1. To assess the impact of missing data on primary MRD analysis, for patients who discontinued the treatment prior to Month 15 and Month 15 MRD assessment is missing, their last non-missing MRD response will be carried forward and used as the response at Month 15.

The first supplementary estimand is defined as follows:

Population set: FAS

Variable: MRD response rate

Intercurrent Events:

- Exposure to new anti-CLL treatment prior to Month 15
- Discontinuation of treatment for any reason (except PD) prior to Month 15
- Death prior to Month 15
- Non-evaluable MRD sample at Month 15
- Missing MRD sample at Month 15

Handling of intercurrent events: A treatment policy with regards to the discontinuation of treatment prior to Month 15 will be applied. Intercurrent events such as death prior to

month 15, exposure to new anti-CLL treatment prior to Month 15, Non-evaluable MRD sample at Month 15 will be considered as non-responders.

Missing MRD samples at Month 15 will be replaced by their last non-missing MRD response (carried forward).

Patients discontinuing from the study for any reasons prior to Month 15 will not have any MRD sample, therefore they will be automatically considered non-responders.

Population-level summary: MRD rate

The same statistical approach adopted in Section 4.2.2 will be applied.

2. To assess the impact of missing data on primary MRD analysis, for patients who are missing the Month 15 assessment but have an available non-missing assessment thereafter, the subsequent available MRD response closest to Month 15 will be used as the response at Month 15.

The second supplementary estimand is defined as follows:

Population set: FAS

Variable: MRD response rate

Intercurrent Events:

- Exposure to new anti-CLL treatment prior to Month 15
- Discontinuation of treatment for any reason (except PD) prior to Month 15
- Death prior to Month 15
- Non-evaluable MRD sample at Month 15
- Missing MRD sample at Month 15

Handling of intercurrent events: A treatment policy with regards to the discontinuation of treatment prior to Month 15 will be applied. Intercurrent events such as death prior to Month 15, exposure to new anti-CLL treatment prior to Month 15, Non-evaluable MRD sample at Month 15 will be considered as non-responders.

Missing MRD samples at Month 15 will be replaced by the subsequent available MRD response closest to Month 15.

Patients discontinuing from the study for any reasons (except PD) prior to Month 15 will not have any MRD sample, therefore they will be automatically considered non-responders.

Population-level summary: MRD rate

The same statistical approach adopted in Section 4.2.2 will be applied.

4.2.4.1 Subgroup Analyses for Main Estimand Endpoint(s)

The generalizability of the primary estimand results when comparing VEN + G arm to FCR/BR arm will be investigated by estimating the treatment effect in subgroups based on stratification factors, key baseline demographics, disease characteristics, FCR only patients, and BR only patients.

The odds ratios of MRD response rate in PB at Month 15 and their 95% confidence intervals, as well as the sample sizes will be reported separately for each level of the following subgroups in forest plots: stratification factors and baseline characteristics as mentioned in Section 4.6.2.

Since the study was powered for the FAS population, all subgroup analyses will be exploratory only. Additional subgroup analyses might be performed as required.

4.3 SECONDARY ENDPOINTS ANALYSIS(SES)

4.3.1 Investigator-Assessed Progression-Free Survival: Main Estimand

Investigator-assessed Progression free survival (PFS) as reported in Section 1.1 of SAP (see Table 1) is defined as the time from randomization to the first occurrence of disease progression (PD) (determined using standard International Workshop on CLL (iwCLL) guidelines – Hallek et al 2018) or death from any cause, whichever occurs first.

For patients who have not progressed or died at the time of analysis, PFS will be censored on the date of the last disease assessment. Similarly, patients who discontinued from the study for any reasons prior to PFS will be censored at the date of the last disease assessment before discontinuation.

If no disease assessments were performed after the baseline visit, PFS will be censored at the time of randomization.

The primary estimand is defined as follows:

- Population set: FAS
- Variable: Time to the first occurrence of a PFS event
- Intercurrent events and handling strategies:
  - Start of a new anti-CLL treatment prior to a PFS event: Date of last disease assessment prior to start of new anti-CLL will be considered as censoring date
  - Discontinuation from study treatment for any reason prior to a PFS event: treatment policy
  - Missing two or more consecutive response assessments: Patients with two or more consecutive missing response assessments will be censored at the last response assessment prior to the first missed response assessment.

Censoring rules for PFS are also summarized in Table 5.

Table 5 Censoring Rules Analysis of PFS

Situation	Date of PFS event or censoring	Outcome
No baseline disease assessments	Date of randomization	Censored
New anti-CLL treatment started before documentation of disease progression or death	Date of last disease assessment prior to start of new anti-lymphoma therapy	Censored
Two or more missed consecutive response assessments	Date of last disease assessment prior to the first consecutive missed assessment	Censored
Alive and without disease progression documentation	Date of last disease assessment	Censored
Death or disease progression between planned disease assessments	Date of death or first disease assessment showing disease progression, whichever occurs first	Event
Death after first on-treatment disease assessment	Date of death	Event

PFS = investigator assessed progression-free survival.

- Population-level summary: hazard ratio (HR) for PFS

4.3.1.1 Main Analytical Approach for PFS

PFS comparison will be made using a two-sided log-rank test at 0.05 significance-level, adjusted for the interim analyses as reported in Section 4.7.1 stratified by:

- IGHV mutation: yes or no
- Binet stage (3 levels): A, B, or C
- Age: ≤65 or >65 years

The Kaplan-Meier method will be used to estimate the PFS distribution for each treatment arm and construct curves for the visual description of the difference between the treatment arms. Estimates of the treatment effect will be expressed as HRs using a stratified Cox proportional-hazards analysis, including 95% confidence intervals depending on the primary endpoint readout.

The generalizability of the PFS results when comparing VEN + G arm to FCR/BR arm will be investigated by estimating the treatment effect in subgroups based on stratification factors, key baseline demographics, disease characteristics, FCR only patients, and BR only patients.

The HRs and their 95% confidence intervals, as well as the sample sizes will be reported separately for each level of the following subgroups in forest plots: stratification factors and baseline characteristics as mentioned in Section 4.6.2.

The proportional hazards assumption on PFS may be examined using both graphical and analytical methods if hazards are not proportional. The log [-log] of the survival function versus time for PFS may be plotted for the comparison between VEN + G and FCR/BR arms. If the curves are not parallel, indicating that hazards are not proportional,

supportive analyses may be conducted to account for the possible non-proportional hazards effect using the restricted mean survival time (RMST) method.

The RMST will be computed for PFS using the area under the curve from baseline to several time points (6, 12, and 18 months). The RMST will be computed for each treatment arm and the difference with its 95% CI (by Greenwood method) and p-values (by Z test) will be provided for descriptive purposes.

4.3.1.2 Sensitivity Analyses for PFS

To check robustness of the PFS analysis and underlying assumptions, the following sensitivity analyses for PFS will be performed on the FAS population:

- An unstratified log-rank test for the primary PFS comparison between treatment arms will be conducted.
- If 3 or more patients die due to COVID-19: a sensitivity analysis for PFS will be conducted considering the patients who died censored to the date of study treatment discontinuation (before date of death).
- If 3 or more patients discontinue study treatment due to COVID-19 AE: the main analytic approach for the PFS will be performed with such patients censored to date of study treatment discontinuation.

4.3.1.3 Supplementary Estimand for PFS

An analysis on PFS will be performed without taking into account missing assessments on PFS.

For patients who have not progressed or died at the time of analysis, PFS will be censored on the date of the last disease assessment. If no disease assessments were performed after the baseline visit, PFS will be censored at the time of randomization.

The supplementary estimand is defined as follows:

Population set: FAS

Variable: Time to the first occurrence of a PFS event

Intercurrent events and handling strategies:

- Start of a new anti-CLL treatment prior to a PFS event: Date of last disease assessment prior to start of new anti-CLL will be considered as censoring date
- Discontinuation from study treatment for any reason prior to a PFS event: treatment policy
- Missing two or more consecutive response assessments: treatment policy

Population-level summary: hazard ratio for PFS.

4.3.2 MRD Response Rate in PB at End of Treatment Response Visit (8-12 weeks after last dose of FCR/BR [~9 months] vs. 8-12 weeks after last dose of VEN + G [~15 months]

MRD response rate is defined as the proportion of patients with MRD-negativity measured in PB at Month 9 assessment for FCR/BR arm and at Month 15 assessment for VEN + G arm (EOT assessment i.e., ~3 months after treatment completion/early termination) as defined in Section 1.1 of SAP (see Table 1).

Treatment duration will be different according to the randomized treatment arm as follows:

- FCR/BR treatment duration: 6 cycles, each cycle with a duration of 28 days
- VEN + G treatment duration: 12 cycles, each cycle with a duration of 28 days

MRD response rate is determined as the proportion of patients with MRD-negativity (defined as < 1 CLL cell in 10,000 leukocytes), measured in PB at about 3 months after EOT, considering different treatment duration. This will be measured on the basis of assessments performed centrally using NGS with MRD negativity defined using a cutoff of 10^{-4} (less than 1 CLL cell in 10,000 leukocytes). MRD status will be determined through the following calculations:

- If CLL count (as a % of leukocytes) ≥ 0.01%; then MRD response status = nonresponder (positive)
- If CLL count (as a % of leukocytes) < 0.01% AND limit of detection (LOD) < 0.01% then MRD response status = responder (negative)
- If CLL count (as a % of leukocytes) < LOD is considered Indeterminate, then MRD response status = below limit of quantification (BLQ) corresponding to Non-evaluable (non-responder)

The estimand for this secondary endpoint is defined as follows:

Population set: FAS

Variable: MRD response rate

Intercurrent Events:

- Exposure to new anti-CLL treatment prior to EOT + 3 months according to the different treatment durations (i.e., prior to Month 9 for FCR/BR and prior to Month 15 for VEN + G)
- Discontinuation of treatment for any reason (except PD) prior to Month 9 for FCR/BR and prior to Month 15 for VEN + G
- Death prior to Month 9 for FCR/BR and prior to Month 15 for VEN + G
- Missing/Non-evaluable MRD sample at Month 9 for FCR/BR and at Month 15 for VEN + G

Handling of intercurrent events: A treatment policy with regards to the discontinuation of treatment will be applied. Intercurrent events such as death, exposure to new anti-CLL

treatment and missing/non-evaluable MRD sample, all occurring prior to Month 9 for FCR/BR arm and prior to Month 15 for VEN+G arm are considered as non-responder.

Statistical Methods:

MRD response rate of the 2 arms will be compared based on CMH according to the following stratification factors as entered in the IxRS:

- IGHV mutation: yes or no
- Binet stage (3 levels): A, B, or C
- Age: ≤65 or >65 years

If the null hypothesis of no difference in MRD response rate is rejected and the observed MRD response rate is favorable for the VEN+G experimental arm, then it is concluded that VEN+G significantly increases the chance to respond compared to FCR/BR. Difference in MRD response rate will be estimated and two-sided 95% Anderson-Hauck CI will be provided. Within each arm, 2-sided 95% Clopper-Pearson exact CIs will also be presented.

4.3.3 MRD Response Rate in BM at End of Treatment Response Visit (9 months FCR/BR vs. 15 months VEN+G)

MRD response rate (determined as the proportion of patients with MRD-negativity as defined in Section 4.2.1) measured in BM at the completion of treatment assessment (EOT assessment i.e., ~3 months after treatment completion/early termination) as defined in Section 1.1 of SAP (see Table 1).

The estimand for this secondary endpoint is defined as follows:

Population set: FAS

Variable: MRD response rate in BM at end of treatment response visit

Intercurrent Events:

- Exposure to new anti-CLL treatment prior to EOT+3 months according to the different treatment durations (i.e., prior to Month 9 for FCR/BR and prior to Month 15 for VEN+G)
- Discontinuation of treatment for any reason (except PD) prior to Month 9 for FCR/BR and prior to Month 15 for VEN+G
- Death prior to Month 9 for FCR/BR and prior to Month 15 for VEN+G
- Non-evaluable MRD sample in BM at Month 9 for FCR/BR and at Month 15 for VEN+G

Handling of intercurrent events: A treatment policy with regards to the discontinuation of treatment will be applied. Intercurrent events such as death, exposure to new anti-CLL treatment and non-evaluable MRD sample in BM, all occurring prior to Month 9 for FCR/BR arm and prior to Month 15 for VEN+G arm are considered as non-responder.

Importantly, not all patients will have a BM assessment for MRD because, in order to have a bone marrow biopsy patients must be at least in a PR or suspected CR.

Summary measure: MRD response rate

MRD response rate will be analyzed in a similar manner as the secondary endpoint described in Section 4.3.2.

4.3.4 ORR which includes CR, CRi (complete remission with incomplete blood count recovery), and PR at Month 15 Assessment

Overall Response Rate (ORR) is defined as the rate of a clinical response of CR, complete response with incomplete bone marrow recovery (CRi), or PR at Month 15, as determined by the investigator according to the iwCLL guidelines (Hallek et al. 2018).

The estimand for this secondary endpoint is defined as follows:

Population set: FAS

Variable: ORR response rate at Month 15

Intercurrent Events:

- Exposure to new anti-CLL treatment prior to Month 15
- Discontinuation of treatment for any reason (except PD) prior to Month 15
- Death prior to Month 15
- Missing CR/CRi/PR assessments at Month 15

Handling of intercurrent events: A treatment policy with regards to the discontinuation of treatment and/or study discontinuation will be applied. Intercurrent events such as death prior to month 15, exposure to new anti-CLL treatment prior to Month 15 and missing CR/CRi assessments at Month 15 will be considered as non-responder.

Summary measure: ORR rate

ORR response rate will be analyzed in a similar manner as the secondary endpoint described in Section 4.3.2.

Subgroup analysis will be conducted for ORR as described in Section 4.2.4.1.

4.3.5 Complete Response (CR) Rate which includes CR and CRi at Month 15 Assessment.

Complete response rate (CRR) defined as rate of a clinical response of CR or CRi at Month 15 as determined by the investigator according to the iwCLL guidelines (Hallek et al. 2008).

The estimand for this secondary endpoint is defined as follows:

Population set: FAS

Variable: CRR response rate at Month 15

Intercurrent Events:

- Exposure to new anti-CLL treatment prior to Month 15
- Discontinuation of treatment for any reason (except PD) prior to Month 15
- Death prior to Month 15
- Missing CR/CRi assessments at Month 15

Handling of intercurrent events: A treatment policy with regards to the discontinuation of treatment will be applied. Intercurrent events such as death prior to month 15, exposure to new anti-CLL treatment prior to Month 15 and missing CRR information at Month 15 will be considered as non-responder.

Summary measure: CRR rate

CRR response rate will be analyzed in a similar manner as the secondary endpoint described in Section 4.3.2.

4.3.6 MRD Response Rate in PB in Patients with a CR/CRi at Month 15

MRD response rate in patients with CR (determined as the proportion of patients with MRD-negativity and in complete response). MRD is measured in the PB at Month 15. Clinical Response is defined as a clinical response of CR or CRi at Month 15 as determined by the investigator according to the iwCLL guidelines (Hallek et al. 2008).

The estimand for this secondary endpoint is defined as follows:

Population set: FAS

Variable: MRD response rate in patients with CR

Intercurrent Events:

- Exposure to new anti-CLL treatment prior to Month 15
- Discontinuation of treatment for any reason (except PD) prior to Month 15
- Death prior to Month 15
- Missing/Non-evaluable MRD sample at Month 15

Handling of intercurrent events: A treatment policy with regards to the discontinuation of treatment will be applied. Intercurrent events such as death prior to month 15, exposure to new anti-CLL treatment prior to Month 15, missing/Non-evaluable MRD sample at Month 15 will be considered as non-responder.

Summary measure: MRD rate

CRR response rate will be analyzed in a similar manner as the secondary endpoint described in Section 4.3.2.

4.3.7 MRD Response Rate in BM in Patients with a CR/CRi at End of Treatment Visit (8-12 weeks after last dose of FCR/BR (~9 months) vs. 8-12 weeks after last dose of VEN + G (~15 months).

MRD response rate in patients with CR (determined as the proportion of patients with MRD-negativity and in CR or CRi). MRD is measured in BM at Month 9 for FCR/BR arm and at Month 15 for VEN + G arm (EOT assessment i.e., 3 months after treatment completion/early termination).

Clinical Response is defined as a clinical response of CR or CRi at Month 9 for FCR/BR arm and at Month 15 for VEN + G arm (as determined by the investigator according to the iwCLL guidelines (2008).

The estimand for this secondary endpoint is defined as follows:

Population set: FAS

Variable: MRD response rate in patients with CR

Intercurrent Events:

- Exposure to new anti-CLL treatment prior to EOT + 3 months (i.e., Month 9 for FCR/BR and prior to Month 15 for VEN + G) according to the different treatment durations
- Discontinuation of treatment for any reason (except PD) prior to Month 9 for FCR/BR and prior to Month 15 for VEN + G
- Death prior to Month 9 for FCR/BR and prior to Month 15 for VEN + G
- Missing/Non-evaluable MRD sample at Month 9 for FCR/BR and at Month 15 for VEN + G

Handling of intercurrent events: A treatment policy with regards to the discontinuation of treatment will be applied. Intercurrent events such as death, exposure to new anti-CLL treatment, missing/non-evaluable MRD sample are considered as non-responder.

Summary measure: MRD rate

Statistical Methods:

MRD response rate will be analyzed in a similar manner as the secondary endpoint described in Section 4.3.2.

Furthermore, concordance analysis will be performed to address potential discrepancies between PB and BM results, proportion of patients achieving MRD negativity in both PB and BM will be compared at Month 9 for FCR/BR and Month 15 for VEN + G.

4.3.8 DOR, defined as the time from the first response to the time of PD or death from any cause, whichever comes first

Duration of overall response (DOR), defined as the time interval from the date of the first occurrence of a documented overall response (CR, CRi, or PR, as assessed by the investigator) to the first date of occurrence of PD as determined by the investigator or death from any cause, whichever occurs first.

DOR will be restricted only to those patients who experienced an overall response (OR) as defined above.

Patients who have not had an event (PD or death) will be censored at the time of the last tumor assessment date. If no tumor assessments were performed after the date of the first occurrence of a response, duration of response will be censored at the date of the first occurrence of a response.

The estimand for this secondary endpoint is defined as follows:

Population set: FAS

Variable: Time to the first occurrence of an investigator assessed PD or death from any cause for patients with OR

Intercurrent Events and handling strategies:

- Start of a new anti-CLL treatment after OR and prior to a PFS event: Date of last disease assessment prior to start of new anti-CLL after OR will be considered as censoring date
- Discontinuation from study treatment for any reason after OR and prior to a PFS event: treatment policy strategy

Population-level summary: hazard ratio for DOR

Statistical Methods:

DOR comparison will be made using a two-sided log-rank test stratified by:

- IGHV mutation: yes or no
- Binet stage (3 levels): A, B, or C
- Age: ≤65 or >65 years

The Kaplan-Meier method will be used to estimate the DOR distribution for each treatment arm and construct curves for the visual description of the difference between the treatment arms. Estimates of the treatment effect will be expressed as hazard ratios using a stratified Cox proportional-hazards analysis, including 95% confidence intervals.

4.3.9 Best Response Achieved (CR, CRi, PR, SD, or PD) up to and Including the Assessment at Month 15

Best response achieved (determined as the proportion of patients in CR or CRi or PR or Stable Disease [SD], or Progressive Disease [PD] as assessed by the investigator) up to and including the assessment at Month 15.

Patients reaching their best response as CR, CRi or PR will be considered Responders while those reaching SD or PD will be Non-Responders.

The estimand for this secondary endpoint is defined as follows:

Population set: FAS

Variable: Best response rate

Intercurrent Events:

- Exposure to new anti-CLL treatment prior to an assessment
- Discontinuation of treatment for any reason (except PD)
- Death prior to an assessment

Handling of intercurrent events: A treatment policy with regards to the discontinuation of treatment will be applied. Intercurrent events such as death, exposure to new anti-CLL treatment, will be considered as non-responder only if they occur prior to the response assessment contributing to the best response.

Summary measure: Best response rate

Best response rate will be analyzed in a similar manner as the secondary endpoint described in Section 4.3.2.

4.3.10 EFS, defined as the time between date of randomization and the date of disease progression/relapse, death, or start of a new anti-leukemic therapy

Event-free survival (EFS) is defined as the time from randomization to the first occurrence of a PFS event as defined under Section 4.3.1, but including the start of a new anti-leukemic therapy as an event whichever occurs first.

Patients who have not had an event will be censored at the date that they are last known to be event-free on or prior to the clinical data cutoff date. Data for patients who are randomized without any post-baseline assessments will be censored at the date of randomization plus 1 day.

The estimand for this secondary endpoint is defined as follows:

Population set: FAS

Variable: Time to the first occurrence of EFS

Intercurrent Events:

- Discontinuation from study treatment for any reason prior to an EFS event

Handling of intercurrent events: A treatment policy with regards to the intercurrent events listed above will be applied

Population-level summary: hazard ratio for EFS

Statistical Methods:

EFS will be analyzed in a similar manner as PFS (Section 4.3.1.1) to estimate event rates for each treatment arm, as well as the HR between the two treatment arms with 95% CI.

4.3.11 OS, defined as the time between the date of randomization and the date of death due to any cause

Overall Survival (OS), defined as the time between the date of randomization and the date of death due to any cause.

Patients who are alive (including lost to follow-up) at the time of the analysis will be censored at the date when they were last known to be alive on or prior to the clinical data cutoff date. Data for patients who are randomized without any post-baseline information will be censored at the date of randomization.

The estimand for the key secondary endpoint is defined as follows:

Population: FAS

Variable: Time from randomization to death due to any cause

Intercurrent Events and handling strategies:

- Start of a new anti-CLL treatment: “treatment policy
- Discontinuation from study treatment for any reason: treatment policy

Population-level summary: hazard ratio for OS

Statistical Methods:

OS will be analyzed in a similar manner as the PFS in Section 4.3.1.1 to estimate landmark survival rates for each treatment arm and the HR between the two treatment arms with 95% CI.

The generalizability of the OS results when comparing VEN+G arm to FCR/BR arm will be investigated by estimating the treatment effect in subgroups based on stratification factors, key baseline demographics, disease characteristics, FCR only patients, and BR only patients.

The HR and their 95% confidence intervals, as well as the sample sizes will be reported separately for each level of the following subgroups in forest plots: stratification factors and baseline characteristics as mentioned in Section 4.6.2.

4.3.12 TLS Risk Reduction Rate in Arm A, defined as the reduction in the proportion of patients who were TLS high-risk after 3 doses of obinutuzumab compared to the proportion of patients who were TLS high-risk at baseline

TLS risk reduction rate in Arm A is defined as the reduction in the proportion of patients who were TLS high-risk after 3 doses of Obinutuzumab during the run-in period (before initiating venetoclax on Cycle 1, Day 22) compared to the proportion of patients who were TLS high-risk at baseline. High risk category is defined as per eCRF.

Proportion of TLS high-risk at baseline and after 3 doses of Obinutuzumab will be reported in a descriptive manner or will be compared by Mc Nemar test for paired data depending on the number of patients reported in each category.

4.3.13 Reduction in Mandatory Hospitalizations during Venetoclax Ramp-up in Arm A Patients

Reduction in mandatory hospitalizations during venetoclax ramp-up in Arm A is defined as the actual number of protocol-mandated hospitalizations for TLS monitoring during venetoclax ramp-up period after 3 doses of obinutuzumab compared to the number of protocol-mandated hospitalizations for TLS monitoring during venetoclax ramp-up expected at baseline.

Ramp-up period for venetoclax is defined as period from Cycle 1, Day 22-28 and Cycle 2, Day 1-Day 7 where the 20 mg and 50 mg daily doses of venetoclax, administered for patients at TLS-high risk requires mandated hospitalizations (the hospitalizations at 100, 200 and 400 are only needed if the patient has had a TLS event at one of the lower doses).

Total number of hospitalizations in high-risk TLS pts at baseline (expected to be N = 2 hospitalization) will be compared with the number of protocol mandated hospitalizations during the first 2 doses of the ramp-up (i.e., If a patient was high-risk TLS at baseline and remains high-risk TLS after debulking with obi, 2 vs 2 hospitalizations; if a patient was high-risk TLS at baseline and is low or medium risk after obi, comparison will be 2 vs 0).

Number of hospitalizations in TLS high-risk at baseline and after the 2 doses of the venetoclax ramp-up will be displayed in listings and reported in a descriptive manner.

4.3.14 Key/Confirmatory Secondary Endpoints/Estimands

The following secondary endpoints are considered key endpoints and they will be formally tested at the time of the primary efficacy analysis (MRD at month 15):

- PFS
- MRD response rate in PB at end of treatment visit
- MRD response rate in BM at end of treatment visit

The key secondary endpoints listed in this section will be evaluated in the order specified above. The overall two-sided significance level of the key secondary endpoints is controlled at alpha = 5% by use of a hierarchical testing procedure. If the primary

endpoint reaches statistical significance, the following secondary endpoints will be tested in the order specified above recycling the alpha level = 5%. Key secondary endpoints will be tested at the appropriate significance level in the order specified in the sequence. If for one endpoint, the null hypothesis cannot be rejected, then the results for this and all subsequent endpoints reported in the sequence cannot be considered statistically significant.

The hierarchical testing procedure with the boundaries determined as described above ensures that the overall type I error for the primary and key secondary endpoints will be controlled at 5% alpha (α) level (Hung et al. 2007; Glimm et al. 2010).

PFS is the first among the key secondary endpoints in the list and it is tested at different time points (i.e., two interim analyses and one final analysis) over study duration. As it's very likely that PFS will reach statistical significance at the final analysis, the statistical significance for the subsequent endpoints in the list, whose readout will occur at the time of the first PFS interim analysis, can be declared only once PFS reaches statistical significance.

4.3.15 Supportive Secondary Endpoints/Estimands

4.3.15.1 Patient-Reported Outcomes Analyses

The secondary PRO endpoints for the study are as follows:

- Change in disease and treatment-related symptoms following treatment with the combination of VEN + G compared with FCR/BR in patients with previously untreated CLL without del(17p) or *TP53* mutation as measured by M.D. Anderson Symptom Inventory (MDASI-CLL).
- Change in physical functioning, role functioning and health-related quality of life (HRQoL) following treatment with the combination of VEN + G compared with FCR/BR in patients with previously with previously untreated CLL without del(17p) or TP53 mutation as measured by the European Organisation for Research and Treatment of Cancer Quality of Life Questionnaire Core 30 (EORTC QLQ-C30).

Analyses will be conducted on the FAS population.

Scoring for the MDASI-CLL and EORTC QLQ-C30 questionnaires (Appendix 5 and Appendix 6 of the protocol) will be based on the corresponding user manual. For scales with more than 50% of the constituent items completed, a prorated score will be computed that is consistent with the scoring manuals and validation papers. For subscales with less than 50% of the items completed, the subscale will be considered as missing. Summary statistics of the MDASI-CLL and EORTC QLQ-C30 scales and their changes from baseline will be calculated at each assessment timepoint for both study arms.

Visit summary and change from baseline analyses will be performed for the MDASI-CLL severity, interference, and CLL scales, as well as the EORTC QLQ-C30 Role functioning (RF), Physical functioning (PF), and Global health status/QoL (GHS/QOL) scales. Summary statistics (number of patients, mean, standard deviation, median, minimum, maximum) of score(s) and score change(s) from baseline to each time point will be presented by treatment arms.

Repeated measures mixed-effects model will be used for comparing the MDASI-CLL and EORTC QLQ-C30 RF, PF, GHS/QOL scale scores between treatment arms. The model will include a term for intercept, a term for linear time trend, a term for treatment group, and a term for treatment-by-time interaction. Covariates will be added as appropriate. Time points with less than 20% patients who completed the MDASI-CLL and QLQ-C30 scales, where all subsequent time points also have less than 20% completion will be excluded.

The analysis of all other scales will be considered exploratory.

4.3.15.2 Health Economic Analyses

Health utility status, as assessed by the EuroQol 5 Dimension questionnaire (EQ-5D-5L), will be evaluated. Scores at baseline and change from baseline scores for each timepoint will be quantified using descriptive statistics.

Summary statistics (number of patients, mean, standard deviation, median, minimum, maximum) of score(s) and score change(s) from baseline to each time point will be presented by treatment arms on the EQ-5D-5L utility score and visual analog scale. In addition, the proportion of patients endorsing each level of the EQ-5D-5L dimensions will be provided by treatment arms at each time point. Further analyses of the results from the health economics data will be reported separately from the Clinical Study Report.

4.4 EXPLORATORY ENDPOINTS ANALYSIS

The exploratory outcome measures for this study are as follows (provided these data are available in a sufficiently timely and complete form to perform these analyses at the time of the final analysis):

- Subgroup analyses of primary and selected secondary endpoints have been detailed in the specific sections.
- Descriptive summary statistics and the change from baseline will be calculated by treatment arm at each assessment for all remaining scales of the EORTC QLQ-C30.
- The relationship between various baseline markers and clinical outcome parameters including the primary MRD outcome will be assessed in patients from both arms of the study (including but not limited to CLL FISH (17p-, 11q-, 13p-, + 12q), IGHV mutation status, p53 mutation status, serum parameters, Bcl-2 expression and other CLL disease markers. These will be compared using forest plots.
- Durability of MRD-negativity (at 10^{-4}) over time in PB measured by NGS
- Relationship between MRD at EOT assessment and PFS on the basis of PB assessed using NGS (landmark analysis).
- Relationship between MRD at EOT assessment and PFS on the basis of BM assessed using NGS (landmark analysis)
- Exploratory analyses of MRD negativity by timepoint will also be performed with MRD negativity defined using a cutoff of 10^{-4} (less than 1 cell in 10,000 leukocytes) within the limit of sensitivity of the technology.

4.5 SAFETY ANALYSES

Safety endpoints include AE, serious adverse event (SAE) and adverse events of special interest (AESI). Safety analyses will be based on the safety population which is defined as all patients who received at least one dose of any study medication.

4.5.1 Extent of Exposure

Treatment exposure to study medication will be summarized by treatment arm for each medication administered. The following measures will be included in the summary output:

- Duration of treatment
- Dose intensity (%) – calculated as the total cumulative dose actually received divided by the planned cumulative dose
- Number of doses/cycles
- Total cumulative dose
- Missed doses
- Number of dose discontinuation and reason for discontinuation
- Number of dose modifications and reasons for modification
  - Number of dose interruptions and reasons for dose interruptions
  - Number of dose reductions and reasons for dose reductions

In addition, the number and percentage of patients withdrawing from each treatment and the primary reason for discontinuation will be summarized by treatment arm.

4.5.2 Adverse Events

Coding of AEs will be done using the most recent version of the Medical Dictionary for Regulatory Activities (MedDRA). Verbatim terms of AEs will be mapped to MedDRA thesaurus terms and graded according to the National Cancer Institute Common Terminology Criteria for Adverse Events (NCI CTCAE) grading system, version 5.0. All AEs occurring during or after the first treatment will be included in the summary tables.

AE summaries will be presented by system organ class, preferred term, and treatment arm, for the following:

- All AEs
- AEs by NCI CTCAE grade
- Grade 3-4 AEs
- SAEs
- Fatal AEs
- Grade 3 neutropenia rate and Grade 4 neutropenia rates along with dose interruptions due to Grade 3 and Grade 4
- Grade ≥3 infections
  - Opportunistic infections (Using Roche AEGT)

- Richter's transformation
- AESIs, including:
  - Cases of potential drug-induced liver injury that include an elevated ALT or AST in combination with either an elevated bilirubin or clinical jaundice, as defined by Hy's law.
  - Suspected transmission of an infectious agent by the study drug (refer to protocol for full definition)
  - TLS of any grade, irrespective of causality
  - Second primary malignancies

- AEs leading to treatment discontinuation (any study drug)
- AEs leading to study withdrawal
- AEs leading to dose interruptions
- AEs leading to dose reduction
- AEs suspected to be caused by any study drug
- AEs stratification by treatment phase (combination therapy phase and single agent treatment phase)
- All AEs by phase
- Grade 3−4 AEs by phase
- SAEs by phase
- Fatal AEs by phase
- Venetoclax dose interruptions by phase
- Venetoclax dose reductions by phase
- Venetoclax withdrawal by phase

In the case of multiple occurrences of the same AE within a patient, the most extreme severity recorded will be used in the summary tables.

Deaths reported during the study treatment period and those reported after treatment completion/discontinuation will be summarized by treatment arm.

A listing of all AEs and SAEs will be presented (separately).

An analysis of the safety profile for patients with confirmed/suspected COVID-19 will be performed, and key safety summaries for these patients will be produced separately.

Note – abnormal laboratory data that are clinically significant are reported as AEs.

4.5.3 Additional Safety Assessments

4.5.3.1 Laboratory Data

Clinical laboratory tests will be performed at local laboratories throughout the study. Laboratory tests will be grading according to the NCI CTCAE v4.0. Laboratory data outside the normal ranges will be identified.

Laboratory data including NCI CTCAE v4.0 will be presented using summary statistics of change from baseline by visit and treatment arm. Baseline will be defined as the last valid measurement before the first dose of study medication.

In addition, shift tables will be generated to cross-tabulate the number of patients grouped into each NCI CTCAE v4.0 grade at baseline versus post-baseline assessments.

4.5.3.2 Vital Signs

Vital signs data will be summarized over time for absolute values and changes from baseline without any replacement for missing data. Descriptive statistics will be provided by treatment arm.

4.5.3.3 ECGs

Electrocardiogram (ECG) data at the screening visit will be summarized by treatment arm.

4.6 OTHER ANALYSES

4.6.1 Summaries of Conduct of Study

Enrollment, major protocol deviations, study drug administration, and patient disposition will be summarized by treatment arm in all randomized patients. A summary of patient disposition will include whether treatment was completed or discontinued early and the reason for early treatment discontinuation. Descriptive statistics will be used in evaluating the conduct of the study. Median length of follow-up will be estimated overall and by-treatment in FAS population using the reverse Kaplan Meier (K-M) method.

4.6.2 Summaries of Treatment Group Comparability/Demographics and Baseline Characteristics

Descriptive summaries will be provided for the FAS population by treatment groups for the following demographic and baseline variables:

- Stratification factors
  - Binet Stage (3 levels): A, B or C
  - IGHV mutation: yes or no
  - Age: ≤ 65 or > 65 years
- Demographics
  - Gender
  - B-symptoms (no vs. yes)
  - Geographic region (US, Australia, Western Europe)
  - Age (continuous and categorical (< 40, 40−59, 60−69, ≥ 70)
  - Age group: ≤ 65, $> 65; < 75, \geq 75$.
  - Race
  - Ethnicity
  - TLS risk category (Low, Medium and High)

- Bulky disease (<5cm, ≥5 cm; <10cm, ≥10 cm)
- Total Cumulative Illness Rating Scale (CIRS) score
- Estimated creatinine clearance [according to the formula of Cockroft-Gault] (continuous and categorical with cut-off value ≥70 mL/min)
- Cytogenetic factors (11q and 13q, and trisomy 12)
- Serum beta2-microglobulin
- Eastern Cooperative Oncology Group (ECOG) Performance Status at screening.
- Time from first diagnosis to randomization

4.7 INTERIM ANALYSES

Interim analyses (IA) will be conducted on the key secondary endpoint PFS. Findings of IAs will be released regardless of the results obtained, for publication purposes. No independent Data Monitoring Committee (iDMCs) are in place for the interim analyses.

Two planned interim efficacy analyses will be conducted for PFS as detailed in Section 1.2.4. PFS will be tested at the significance level determined using the Lan-DeMets spending function with an O'Brien-Fleming boundary such that the overall two-sided type I error rate will be maintained at the 0.05 level.

Cristallo assumptions for PFS rates have been revised according to the Phase III GAIA-CLL13 study main results recently published (Eichhorst B, Niemann, CU, Kater AP et al.).

With a target HR=0.413 of the VEN+G arm over the control arm (FCR/BR), CLL13 KM-PFS rate at 3 years (75.5% in the control arm and 87.7% in the VEN+G arm) were used to estimate the number of events expected to occur at each PFS interim and final analyses over a longer follow-up time (more than 6 years).

Based on the planned numbers of PFS events, the O'Brien-Fleming boundary for statistical significance at the first and second IA will be p = 0.0055 and p = 0.0176, respectively (see Table 6). As the timing of the first IA is determined by the MRD analysis timepoint, and not driven by the number of PFS events, the actual p-value will be determined based on the number of events reached at that time. If the boundary for statistical significance is achieved at one of the IAs, results will be considered confirmatory as the boundary was crossed. In this scenario, the subsequent interim and/or final PFS analysis will be undertaken in order to have more mature data but the results will be considered descriptive.

The boundary for efficacy at the final analysis will be adjusted to incorporate the α spent at the IAs, such that the overall two-sided type I error rate will be maintained at the 0.05 level.

4.7.1 Planned Interim Analysis

The first IA for PFS will be conducted at the time of MRD primary analysis with 25 expected PFS events and the second IA will be performed at the occurrence of 33 PFS events. PFS will be tested at the significance level determined using the Lan-DeMets

spending function with an O'Brien-Fleming boundary such that the overall two-sided type I error rate will be maintained at the 0.05 level as detailed in Table 6.

Table 6 Projected Interim and Final PFS Analysis

Analysis	No of Events	% PFS Information	Projected Cutoff Date [a]	Projected MDD [b] HR	Projected Boundary (p-value)[c]	Cumulative Type I Error (two sided)	Cumulative Power
1st Interim (at MRD analysis)	25	57	Month 46	0.328	p < 0.0055	0.0055	28%
2nd Interim (at 33 events)	33	75	Month 57	0.439	p < 0.0176	0.0193	58%
Final	44	100	Month 74	0.546	p < 0.0439	0.05	83%

HR = hazard ratio; MDD = minimally detectable difference; No. = number; PFS = progression free survival

[a] Study month at which required number of events are projected to occur, where Study Month 1 is the month the first participant is enrolled. Analysis results will be available after data cleaning.

[b] The largest observed HR that is projected to be statistically significant.

[c] The projected boundary for statistical significance for the number of events shown (actual boundary to be calculated at time of analysis based on actual number of events).

5. SUPPORTING DOCUMENTATION

This section is not applicable, since there is no additional supporting document.

6. REFERENCES

Eichhorst B, Fink AM, Bahlo J, et al. First-line chemoimmunotherapy with bendamustine and rituximab versus fludarabine, cyclophosphamide, and rituximab in patients with advanced chronic lymphocytic leukaemia (CLL10): an international, open-label, randomised, phase 3, non-inferiority trial. Lancet Oncol 2016;17(7):928-42.

Eichhorst B, Niemann, CU, Kater AP et al. First-Line Venetoclax Combinations in Chronic Lymphocytic Leukemia. N Engl J Med. 2023, 388:1739-54.

Glimm E, Maurer W, Bretz F. Hierarchical testing of multiple endpoints in group-sequential trials. Stat Med 2010;29:219-28

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

Hallek M, Cheson BD, Catovsky D, et al. iwCLL guidelines for diagnosis, indications for treatment, response assessment, and supportive management of CLL. Blood. 2018;131:2745-60.

Hung HM, Wang SJ, O'Neill R. Statistical considerations for testing multiple endpoints in group sequential or adaptive clinical trials. J Biopharm Stat 2007;17: 1201-10

International Conference on Harmonization. ICH E9 (R1) addendum on estimands and sensitivity analysis in clinical trials to the guideline on statistical principles for clinical trials. 2020 [updated: 17 February 2020; cited: 12 January 2021].

Sharon Anderson & Walter W. Hauck (1983) A new procedure for testing equivalence in comparative bioavailability and other clinical trials, Communications in Statistics - Theory and Methods, 12:23, 2663-92, DOI: 10.1080/03610928308828634

Signature Page for Statistical Analysis Plan -
System identifier: RIM-CLIN-524856

Approval Task	[redacted] Company Signatory 13-Mar-2024 14:54:20 GMT+0000