Document Type:	Statistical Analysis Plan
Official Title:	A Phase III, randomized, double-blind, controlled, multicenter study of intravenous PI3K inhibitor copanlisib in combination with standard immunochemotherapy versus standard immunochemotherapy in patients with relapsed indolent non-Hodgkin's lymphoma (iNHL) - CHRONOS-4
NCT Number:	NCT02626455
Document Date:	22 SEP 2023

Title page

A Phase III, randomized, double-blind, controlled, multicenter study of intravenous PI3K inhibitor copanlisib in combination with standard immunochemotherapy versus standard immunochemotherapy in patients with relapsed indolent non-Hodgkin's lymphoma (iNHL) - CHRONOS-4 (Safety run-in)

Copanlisib and standard immunochemotherapy in relapsed iNHL

Bayer study drug:	BAY 80-6946 / Copanlisib		
Study purpose:	Dose-finding, efficacy and safety of copanlisib		
Clinical study phase:	III	**Date:**	22 SEP 2023
Study No.:	17833	**Version:**	3.0
Author:	PPD		

Table of Contents

Table of Tables

Table of Figures

List of Abbreviations

AE	Adverse event
ATC	Anatomical therapeutic chemical
BMI	Body mass index
BOR	Best overall response
BRM	Blind review meeting
CD20	B lymphocyte antigen CD20
CHOP	Cyclophosphamide, hydroxydoxorubicin, vincristine, prednisone
CMR	Complete metabolic response
CR	Complete response
CRR	Complete response rate
CS	Clinically significant
CSP	Clinical study protocol
CT	Computed tomography
CTC	Common Toxicity Criteria
CTCAE	Common Terminology Criteria for Adverse Events
d	Days
DCR	Disease control rate
DLT	Dose-limiting toxicity
DMC	Data monitoring committee
ECG	Electrocardiogram
ECOG	Eastern Cooperative Oncology Group
e.g.	For example (exempli gratia)
eGFR	Estimated glomerular filtration rate
EOT	End of treatment
FAS	Full analysis set
FL	Follicular lymphoma
FU	Follow-up
HbA1c	Glycated hemoglobin
ID	Identifier
i.e.	That is (id est)
IgM	Immunoglobulin M
iNHL	Indolent non-Hodgkin's lymphoma
kg	Kilogram
LKAD	Last known alive date
LPL	Lymphoplasmacytic lymphoma
MCL	Mantle cell lymphoma
MedDRA	Medical Dictionary for Regulatory Activities
mg	Milligram
min	Minute
mL	Milliliter
mmHg	Millimeter of mercury
MR	Minor response
MRI	Magnetic resonance imaging
MUGA	Multiple gated acquisition
MZL	Marginal-zone lymphoma
NCI	National Cancer Institute
NHL	Non-Hodgkin's lymphoma
OEE	Overall extent of exposure
ORR	Objective response rate
OS	Overall survival
PBMQ	Product specific Bayer MedDRA queries
PD	Progressive disease
PET-CT	Positron emission tomography-computed tomography
PFS	Progression-free survival
PI3K	Phosphatidylinositol-3-kinase
PK	Pharmacokinetic(s)

PR	Partial response
PT	Preferred term
R	Rituximab
R-B	Rituximab and bendamustine
R-CHOP	Rituximab, cyclophosphamide, doxorubicin, vincristine, and prednisone
RP3D	Recommended phase III dose
SAE	Serious adverse event
SAF	Safety analysis set
SAP	Statistical analysis plan
SAS	Statistical analysis system
SD	Stable disease
SFU	Safety follow-up
SLL	Small lymphocytic lymphoma
SMQ	Standardized MedDRA query
SOC	System Organ Class
TEAE	Treatment-emergent adverse event
US	United States (of America)
VGPR	Very Good Partial Response
WM	Waldenström macroglobulinemia

1. Introduction

Rationale of the study

The usual front-line treatment for indolent non-Hodgkin's lymphoma (iNHL) consists of the anti-CD20-Monoclonal antibody rituximab (R) given together with an alkylating agent (bendamustine; R-B), or a chemotherapy combination containing an alkylating agent (cyclophosphamide, hydroxydoxorubicin, vincristine, and prednisone; R-CHOP]). While the R-B combination is preferred in the United States (US), R-CHOP remains the first option in the European Union. In 2nd line, patients with R-sensitive disease (the large majority) would usually receive R-B following R-CHOP, and R-CHOP following R-B.

Copanlisib has activity as monotherapy in patients with relapsed or refractory iNHL (shown in the results of studies 12781- and 16349-Part A). In addition, based on the current metabolism knowledge of each compound there is low potential for drug-drug interactions between copanlisib and the combination partners. It can therefore be expected that copanlisib would increase the activity of standard 1st and 2nd line immunochemotherapy (R-B/R-CHOP) in patients with relapsed, R-sensitive iNHL warranting re-treatment with a rituximab based regimen. The clinical benefit should be expressed in a prolonged progression-free survival (PFS), and an improved quality of life in comparison to the standard immunochemotherapy. For further information regarding the background and rationale of the study, refer to the study protocol and Phase III statistical analysis plan (SAP), version 4.0, 22SEP 2023.

Protocol Version and Amendments

This analysis plan is based on an integration of the following protocol versions and integrated global amendments:

- Original protocol, Version 1.0, dated 23 JUN 2015
- Amendment 01 forming integrated protocol Version 2.0, dated 01 SEP 2015
- Amendment 02 forming integrated protocol Version 3.0, dated 18 JUL 2016
- Amendment 03 forming integrated protocol Version 4.0, dated 30 MAR 2017
- Amendment 05 forming integrated protocol Version 5.0, dated 14 AUG 2017
- Amendment 06 forming integrated protocol Version 6.0, dated 25 SEP 2018
- Amendment 08 forming integrated protocol Version 7.0, dated 14 AUG 2019
- Amendment 09 forming integrated protocol Version 8.0, dated 21 MAR 2023

and also on local amendments, not forming part of the integrated global protocol:

- Amendment 04, dated 11 MAY 2017 (local amendment valid for Japan only)
- Amendment 07, dated 12 MAR 2019 (local amendment valid for Germany only)

2. Study Objectives

The primary objective of the Safety run-in part is to determine the recommended phase III dose (RP3D) of copanlisib in combination with standard immunochemotherapy (R-B or R-CHOP) to be used in the subsequent Phase III part of the study.

The secondary objectives of the Safety run-in part are to evaluate:

- Radiological and clinical indicators of treatment efficacy.
- Safety and tolerability of copanlisib in combination with R-B/R-CHOP.

Further objectives are to evaluate:

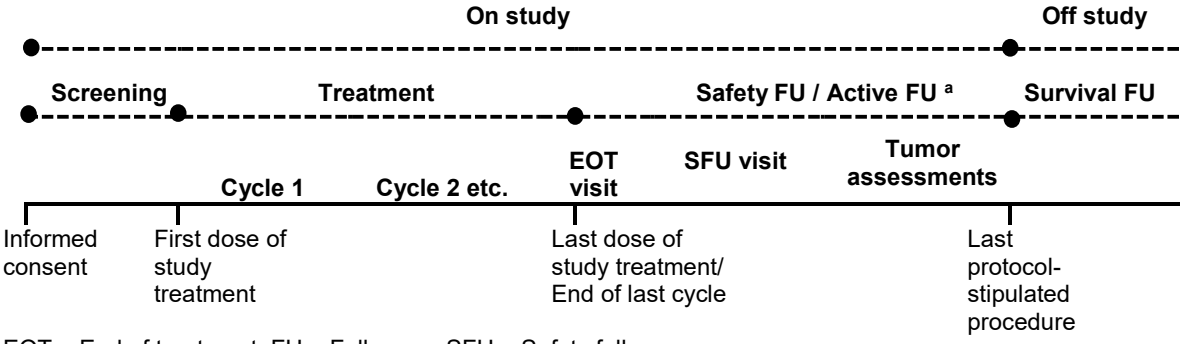
- Pharmacokinetics (PK) of copanlisib.
- Biomarkers of efficacy, mode-of-action-related effect, safety, and / or the pathomechanism of the disease.

3. Study Design

The study includes two parts. The Safety run-in part determined the RP3D of copanlisib, in combination with standard immunochemotherapy (R-B or R-CHOP), to be used in the subsequent Phase III part, treating 21 global patients (excluding Japan) and 6 patients from Japan. The Japan cohort was added to confirm that the RP3D of copanlisib in combination with R-B or R-CHOP is safe and tolerable in Japanese patients (Amendment 04, dated 11 MAY 2017). The primary objective of the phase III part is to evaluate whether copanlisib, in combination with standard immunochemotherapy, is superior to standard immunochemotherapy in prolonging PFS in patients with relapsed, rituximab-sensitive iNHL. The study will accrue patients with follicular lymphoma (FL), marginal-zone lymphoma (MZL), small lymphocytic lymphoma (SLL), and lymphoplasmacytic lymphoma (LPL) / Waldenström macroglobulinemia (WM). In view of the prevalence of various histological types, it is expected that the large majority of study participants will be patients with FL.

The overview of study periods is presented in Figure 3–1.

Figure 3–1: Study periods



EOT = End of treatment; FU = Follow up; SFU = Safety follow-up
a: SFU for patients who discontinue study treatment due to PD; Active FU for patients who complete 12 months' study treatment without PD or discontinue study treatment for any other reason than PD (includes Safety FU).

The following design applies to both Safety run-in part and the Phase III part unless otherwise specified.

The study is composed of the following periods: Screening, Treatment, Safety follow-up, Active follow-up (if applicable) and Survival follow-up.

Screening period

The start of the screening period is defined by signing of the informed consent form. After the screening period of up to 28 days with a minimum of 7 days, patients who meet the entry criteria will be randomized to one of the two arms (only in the Phase III part) and start treatment. The start of the treatment period is defined by the first administration of study treatment.

Re-screening of patients who have failed screening will only be allowed once, following discussion with the sponsor's designated medical representative and after approval by the sponsor. Re-screened patients must re-sign the informed consent form, even if it was not changed after the patient's previous screening.

Treatment period

The start of the treatment period is defined by the first administration of study treatment. Combination therapy (copanlisib/placebo with R-B or R-CHOP) will be administered for a maximum of 6 cycles (Cycle 1- Cycle 6). Copanlisib/placebo (study drug) monotherapy will be administered from Cycle 7 onwards. Tumor assessments with the same modality will be performed every 12 weeks (± 7 days) from Cycle 1 Day 1.

All patients will be treated until the occurrence of progressive disease (PD) (per central independent blinded radiology review) as defined by the Lugano Classification [1] (or Owen criteria for patients with LPL/WM [2]) or clinical progression (e.g., Eastern Cooperative Oncology Group [ECOG] performance status of ≥ 3), unacceptable toxicity occurs, until another criterion is met for withdrawal from study or until completion of monotherapy, whichever comes first. The maximum duration of treatment with copanlisib/placebo is 12 months (including combination therapy and monotherapy).

An end-of -treatment (EOT) visit will be performed no later than 7 days after the decision is made to discontinue study treatment.

Safety follow-up period

All patients with the exception of those who died, withdrew consent or were lost to follow-up will have a Safety follow-up visit that will take place 30 days (a time window of +5 days is allowed) after the last administration of study treatment.

Active follow-up period

Only patients who complete 12 months of study treatment without PD or discontinue study drug for reasons other than PD will enter the Active follow-up period (which also serves as a Safety follow-up), except for patients who specifically object to entering the Active follow-up and patients who object to all further data collection after study drug discontinuation. Patients in the Active follow-up are to have tumor assessments by central independent blinded review as outlined in the protocol (every 12 weeks [± 7 days] from Cycle 1 Day 1 during Years 1 and 2; every 24 weeks [± 7 days] during Years 3, 4, and 5; and every 24 weeks [± 14 days] beyond Year 5, until either PD is centrally evaluated and documented, up to 2 years after primary completion date, end of study or a new anti-tumor treatment is administered, whichever occurs first).

Date of last visit

A patient's last visit date is defined as their last visit/contact prior to any Survival follow-up visits/contact.

Survival follow-up period

Except for patients who object to follow-up data collection in the Survival follow-up period, all patients are to be followed off-study for overall survival (OS) at 3-month (± 14 days) intervals, until death or until the end of study, whichever occurs first.

End of Study

The study will end when one of the following conditions is met:

- final contact with the last remaining patient has occurred,
- 10 years after the last patient started study treatment.

Safety run-in part

The Safety run-in part is based on an open-label, two dose level, 3+3 design.

During the Safety run-in part, two treatment combinations (copanlisib with R-B and copanlisib with R-CHOP) will be tested at two dose levels of copanlisib for safety and tolerability (first dose level of 45 mg and, second dose level of 60 mg). Assignment of R-B or R-CHOP will be made according to the following rules.

- Patients who received R-CHOP or Rituximab, cyclophosphamide, vincristine, and prednisone as a previous line of therapy will be assigned to the copanlisib + R-B treatment group

- Patients who received R-B as a previous line of therapy will be assigned to the copanlisib + R-CHOP treatment group.

The two dose levels will have a minimum of 3 and maximum of 6 patients evaluable for dose-limiting toxicity (DLT). See Section 4 for the definition of a DLT.

The dose levels of copanlisib in the Safety run-in part will be as follows:

Dose level 1:	45 mg of copanlisib in combination with either R-B or R-CHOP
Dose level 2:	60 mg of copanlisib in combination with either R-B or R-CHOP

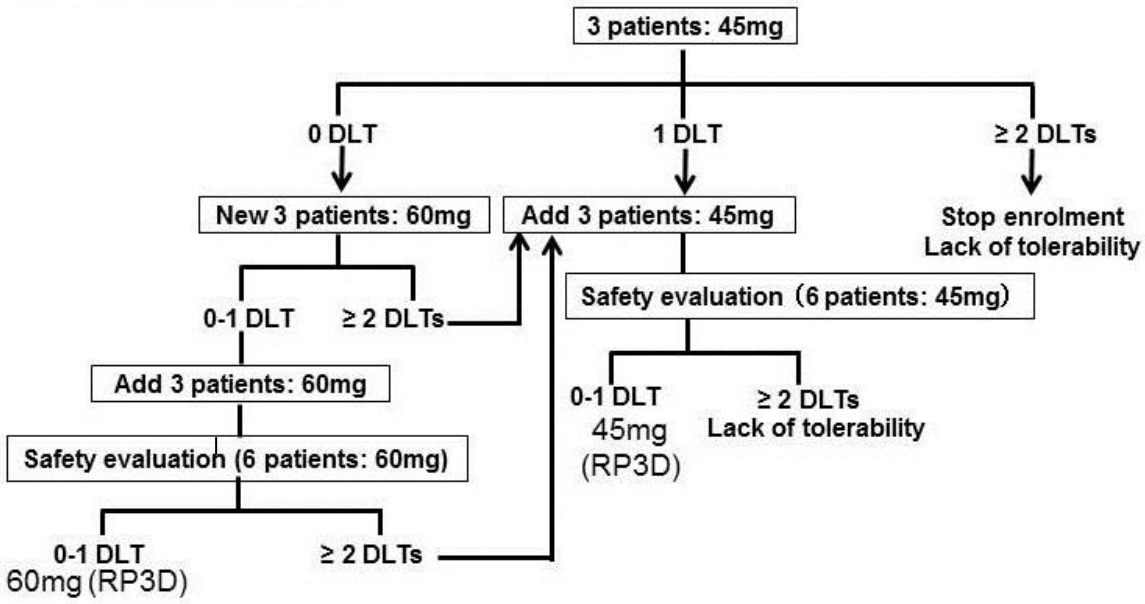
The dose-finding steps are presented in **Figure 3–2**.

Three patients will initially receive 45 mg of copanlisib. If 2 or more of the first 3 patients experience a DLT at this dose level, the Safety run-in would be closed and the combination considered to be not feasible. If 1 out of the first 3 patients experiences a DLT at the 45 mg dose level, 3 additional patients will need to be enrolled at this dose level. If none of the first 3 patients experiences a DLT at the 45 mg dose level, 3 new patients will start at 60 mg. If 0-1 patient out of the first 3 experiences a DLT at the 60 mg dose level, the cohort will be expanded and 3 additional patients will receive the dose of 60 mg. If 0-1 patient out of 6 experiences a DLT, 60 mg will be the recommended dose. If 2 or more out of 6 patients experience a DLT at 60 mg, 3 more patients will be assigned to receive the lower dose (45 mg) if there are only 3 patients thus far treated with 45 mg of copanlisib. If 0-1/6 patients display a DLT at this dose, then 45 mg will be the recommended dose. If 2 or more of 6 patients experience a DLT at 45 mg, then the Safety run-in would be closed and the combination considered to be not feasible with the tested schedule.

All patients in the Japan cohort were assigned to 60 mg of copanlisib, in order to confirm that the RP3D of copanlisib in combination with R-B or R-CHOP is safe and tolerable in Japanese patients (Protocol Amendment 04, dated 11 MAY 2017). Although planned in Protocol Amendment 04, all Japanese patients received the RP3D of copanlisib in combination with R-B.

Patients treated within the Safety run-in part and who in the investigator's opinion benefit from treatment can continue combination therapy (copanlisib plus R-B or R-CHOP) for a maximum of 6 cycles followed by copanlisib monotherapy with the same dose of copanlisib administered with immunochemotherapy. The maximum duration of treatment with copanlisib is 12 months (including combination therapy and monotherapy).

Figure 3–2: Overall study design of the Safety run-in part



DLT = dose-limiting toxicity; RP3D = recommended phase III dose

All patients enrolled in the Safety run-in part need to be evaluated for occurrence of DLTs during the first cycle. Patients not fully evaluable for occurrence of DLTs during the first cycle of therapy will be replaced unless the reason for dropout is a DLT (replacement criteria are detailed in Section 6.4.2 of the study protocol). Any patient who experiences a DLT should not receive further administration of copanlisib or any other study treatment. The investigator must notify the sponsor immediately of any Common Terminology Criteria for Adverse Events (CTCAE) Grade 3 or 4 AEs or clinically relevant laboratory abnormalities.

3.1 Definition of dose-limiting toxicities

Dose-limiting toxicity will be defined as any of the following occurring during Cycle 1 at a given dose level and regarded by the investigator and/or the sponsor to be possibly, probably, or definitely related to copanlisib given in combination with R-B or R-CHOP. National Cancer Institute (NCI)-CTCAE Version 4.03 will be used to assess toxicities/adverse events (AEs). The data monitoring committee (DMC) opinion will decide in case of persisting disagreement between sponsor's and investigator's assessments.

General

- Any Grade 5 hematologic or non-hematologic toxicity
- Any delay of >2 weeks of Cycle 2 due to study treatment-related toxicity

Non-hematologic DLT:

- **Any** non-hematologic toxicity Grade ≥ 3, **except**:

- Grade 3 nausea/vomiting/diarrhea of < 3 days duration or responsive to medical therapy within ≤ 3 days of occurrence
- Increased glucose levels that returned to ≤ 200 mg/dL with or without the use of insulin or oral glucose lowering medications within ≤ 3 days of occurrence
- Increased blood pressure that returned to <150 mmHg (systolic blood pressure) and <90 mmHg (diastolic blood pressure) with or without the use of anti-hypertensive medication within ≤ 3 days of occurrence

Hematologic DLT:

- Grade 4 absolute neutrophil count decrease (< 500/mm3) lasting >7 days
- Grade 4 febrile neutropenia (ANC < 1000/mm3 with fever >38.5°C)
- Grade 4 platelet count decreased (platelet count <25,000/mm3) or Grade 3 platelet count decreased (platelet count 25,000/mm3 to > 50,000/mm3) with serious bleeding
- Signs of serious bleeding and/or international normalized ratio increased or partial thromboplastin time prolonged of Grade 3 (> 2.5 upper limit of normal)

4. General statistical considerations

4.1 General principles

The statistical evaluation will be performed using the Statistical Analysis System (SAS) software package release 9.2 or higher (SAS Institute Inc., Cary, NC, US). The version used in analyses will be presented in the clinical study report. All data will be listed and all variables will be summarized according to their type. Variables measured on metrical scales will be summarized by use of descriptive statistics (number of patients, mean, standard deviation, median, minimum and maximum). Variables measured on ordinal or nominal scales will be summarized by use of frequency tables showing the number and percentage of patients falling within a particular category.

Definition of efficacy and safety endpoints, analysis strategies, structure of analysis datasets and layout of analysis data displays are following Bayer standards as documented in the Bayer standard system: Oncology Therapeutic Area Standards, Global Medical Standards and will comply with the Oncology Common Statistical Language document. The given ordering reflects the priority of the different standards, where specifications of the latter standards are to be followed only if not specified in the earlier standards. Study-specific specifications may be included in addition to the project standards, if needed.

The Blind Review Report (BRR) may contain decisions which are relevant for statistical evaluation, for details see Section 4.8.

4.2 Handling of Dropouts

A patient who discontinues study participation prematurely (i.e., prior to disease progression confirmed by independent central review or death) for any reason is defined as a "dropout" if the patient has already been assigned to treatment (Safety run-in part), even if no study drug has been taken.

Patients who discontinue during the first cycle due to reasons other than toxicity will be replaced in order to ensure that an adequate number of patients are evaluable for safety.

A patient who signed the informed consent but for any reason terminated the study before treatment is assigned is regarded a "screening failure".

Patients who discontinue study drug due to reasons other than death, PD confirmed by independent central review, or biomedical progression in patients with LPL/WM without measurable lesion in the baseline radiological assessment, will enter the Active follow-up period and will not be defined as a "dropout", except for those who specifically object to entering the Active follow-up or object to all further data collection after study drug discontinuation. The patients in the Active follow-up will have follow-up tumor assessments until disease progression is documented or new anti-tumor treatment is administered, whichever occurs first.

4.3 Handling of Missing Data

In order to achieve a well conducted clinical trial in accordance with Good Clinical Practice, every effort will be made to collect all data. However, despite best efforts, it may be inevitable that missing or incomplete data are reported. If a patient indicates they do not want to "continue", investigators must determine whether this refers to discontinuation of study treatment, unwillingness to attend follow-up visits, unwillingness to either have telephone contact, contact with study personnel, or to allow contact with a third party (e.g., family member, doctor). In all cases, every effort should be made to continue to follow the patient for tumor assessment until patients develop radiological progression by independent central review, or biomedical progression in patients with LPL/WM without baseline lesions evaluable by imaging. Survival information including survival status, death date and/or contact date should be determined for all patients during the Survival follow-up. All missing or partial data will be presented in the patient data listing as they are recorded on the electronic Case Report Form. Unless specified, missing data will not be carried forward or otherwise imputed in any statistical analysis.

Missing or non-evaluable tumor observations (including a scheduled assessment that was not done and an incomplete assessment that does not result in an unambiguous tumor response or detection of new lesions) will not be imputed or used in the calculation of derived efficacy variables related to tumor assessments, e.g., response. For example, if a patient misses a scan visit and PD is documented at the next available scan visit, the actual visit date of the first documented PD will be used to calculate PD-related endpoints. For complete definitions of efficacy parameters refer to Section 6.2

Partial dates may occur. To impute values, available data from other sources should be utilized. The last known alive date (LKAD) is derived from the main data sources and may be used in statistical analyses to impute death or tumor assessment dates which are partially or completely missing. Except for death, with a date of death on or prior to the data cut-off, data reported during cleaning, after the data cut-off date, will not be considered in the derivation of LKAD. For such cases the data cut-off date will be used as the LKAD. The last available date across all key data panels (laboratory [LB], tumor response [RS], demography [DM], subject visits [SV], exposure [EX], disposition [DS], procedures [XP], bone marrow assessments [BM], ECG assessments [EG], microbiology [MB], pharmacokinetics [PC], tumor lesion measurements [TR] and tumor/lesion identification[TU]) will be used to determine survival status and the LKAD for each patient.. If the day of the LKAD is missing, then the earliest plausible day of the month should be imputed. If both the day and month of the LKAD are missing, then the earliest plausible day of the year should be imputed.

When appropriate, the following rules will be implemented to avoid excluding patients from statistical analyses due to missing or incomplete data:

Date of death

If a death date is partially or completely missing, it will be imputed following the imputation rules below:

1. If there is an AE with "Fatal" outcome, the date of death will be imputed by the stop date of the AE.

2. If there is no AE with outcome "Fatal" and only the day of death is missing, it will be imputed by the first day of the month unless the LKAD is later than this date; in which case the LKAD will be used for imputation. If both day and month and/or year are missing, then the LKAD will be used for imputation.

Date of tumor scan

If a tumor assessment date is partially missing, it will be imputed following the imputation rules below:

1. If only the day is missing, it will be imputed by the first day of the month.

2. If day and month are missing, it will be imputed by the date of the last previous tumor assessment plus 28 days unless the LKAD is earlier than this date; in which case the LKAD will be used for imputation[1].

Response rates

If a patient has no post-baseline tumor assessment available (or no post-baseline laboratory/clinical tests available for LPL/WM patients without radiologically measurable lesion[s]), i.e., the overall best response assessment is missing, the patient will be non-

[1] All tumor assessments, with complete and partial dates, will be ordered sequentially by visit number. If both day and month are missing from the final tumor assessment and the penultimate tumor assessment is the patients LKAD then the date of the penultimate assessment plus 28 days will be imputed for the final tumor assessment.

evaluable, but will be included into denominator for calculation of objective response rate (ORR), disease control rate (DCR) and complete response rate (CRR).

Missing or partially missing start or end dates for AEs and concomitant medications will be imputed based on the conventions described in Appendix 9. The purpose of the imputation for AEs is to determine whether the event is treatment-emergent, as defined in Section 6.3.1. Similarly, the purpose of the imputation for concomitant medications is to determine if the medication was given concomitantly with study drug and, in the case of new anti-lymphoma therapies, to avoid excluding patients with partial anti-lymphoma therapy start dates from statistical analyses due to missing or incomplete data.

4.4 Handling of investigator assessments when PD is not confirmed by independent central review

The analysis of efficacy, based on investigator assessments, will exclude response assessments collected after the first investigator reported PD, regardless of whether or not the PD was confirmed by independent central review.

4.5 Handling of clinical laboratory values with qualifiers

Clinical laboratory values which have qualifiers (for example, > or <) will be included in clinical safety analyses and summaries with the qualifier removed.

4.6 Interim Analyses and Data Monitoring

No interim analysis is planned for this study.

When the required number of events are reached for the Phase III primary completion analysis, the combined database of Safety run-in and Phase III patients will be cut and cleaned for analysis. This will take place when approximately 280 centrally evaluated PFS events are observed in Phase III.

A DMC was established for an independent review of ongoing data from this trial in accordance with a separate DMC Charter. The DMC operated independently of the sponsor and investigators.

4.7 Data Rules

4.7.1 Time intervals

If time intervals are to be displayed other than days in statistical evaluations, then one year is considered to have 365.25 days (average length of a year, including leap years), one month is considered to have 30.44 days (average length of a month, including leap years), one week is considered to have 7 days, and one cycle is considered to have 28 days (i.e., 4 weeks). All time conversions from days will be displayed to 1 decimal place.

4.7.2 Baseline

Baseline is defined as the last measurements performed prior to the first study drug administration in Cycle 1. If the actual time is not available but date is available for certain measurements, the baseline value is defined as the last non-missing value collected on or before the date of the first dose of study treatment. For patients who have been assigned treatment but not treated with any dose, treatment assignment date will be used as the reference date for baseline value calculation.

Also consider:

- The time part for baseline flagging. Study data tabulation model/analysis data model datasets for exposure and laboratory data both capture times. If the time in either one is missing, use the dates in that case.
- Baseline can be either a scheduled or unscheduled visit.
- If the patient has a measurement on Cycle 1 Day 1 (scheduled), then this measurement will be considered as the baseline.
- If the patient has no measurements on Cycle 1 Day 1, but has a measurement at screening visit (scheduled or unscheduled), then the screening visit measurement will be considered as the baseline.
- If multiple measurements were taken at the same Cycle 1 Day 1 visit or same screening visit and all with non-missing assessment dates, then the measurement with the latest assessment date/time will be considered as baseline, whether scheduled or unscheduled.
- If both scheduled and unscheduled measurements exist for the same visit and cannot be decided by the above rules, then the measurement taken at scheduled visit will be considered as the baseline.

4.7.3 Repeated measures

If there are repeated measurements per time point [e.g., laboratory values, vital signs, electrocardiogram (ECG)], the following rules will be used (unless otherwise specified):

- Before the start of the study drug administration (i.e., for screening and baseline value), the latest non-missing measurement at scheduled visits will be used. Unscheduled visits will be used if there are no measurements at scheduled visits. If the latter is the case, the last unscheduled visit will be used.
- In case of repeated measurements at any post baseline time point, the first measurement at scheduled visits will be used.
- When two records have the same patient identifier (ID), results, date, time and visit/cycle number, with the exception of page number or record ID, the two records

will be considered duplicates and only one record will appear in the corresponding analysis dataset.

4.7.4 Overall extent of exposure

As a general rule, and in accordance with the Oncology Therapeutic Area Standard, leading "0 mg" (prior to the first positive amount of drug) and trailing "0 mg" records (not followed by any positive amount of drug), will not be included in the calculation of any drug duration or amount. Similarly, trailing "drug interruptions" will not be used in statistical tables. A footnote will be included, stating that "Interruption becoming permanent study treatment discontinuation before resumption of study treatment is not counted as an interruption".

Overall extent of exposure (OEE) for copanlisib, is defined as the time from first respective study drug intake (day_{first}) until last study drug intake (day_{last}), including 7 additional days in order to consider the weekly dosing regimen, and is calculated as:

$OEE = day_{last} - day_{first} + 7$

OEE for R-CHOP, is defined as the time from first respective study drug intake (day_{first}) until last study drug intake (day_{last}), including 21 additional days to consider the every 3 weeks (Q3W) dosing regimen and prednisone/prednisolone dosing up to Day 6, and is calculated as:

$OEE = day_{last} - day_{first} + 21 - [day_{(last)} - day_{(first)}$ of last cycle of R-CHOP]

OEE for R-B, is defined as the time from first respective study drug intake (day_{first}) until last study drug intake (day_{last}), including 28 additional days to consider the Q4W dosing regimen and Bendamustine dosing up to Day 2, and is calculated as:

$OEE = day_{last} - day_{first} + 28 - [day_{(last)} - day_{(first)}$ of last cycle of R-B]

If the respective treatment ends with dose interruptions, the day of the last actual dose will be considered to be day_{last}.

4.8 Validity Review

During study conduct, important protocol deviations and major findings affecting the assignment of patients to analysis sets will be identified through ongoing data reviews based on the Safety run-in protocol deviation document. A Blind Review Meeting (BRM) will be scheduled, just prior to unblinding of Phase III and data release, in which the study team will review and finalize the list of minor/major and important/other deviations. Patient validity (based on the presence or absence of major deviations) and details relevant to the statistical evaluation will be discussed and agreed upon during the BRM and documented in the Blind Review Report (BRR). Important deviations and major findings which may affect the assignment of patients to analysis sets, data integrity or patient safety will be listed in the BRR. Any changes to the statistical analysis prompted by the results of the BRM will be documented in an amendment and, if applicable, in a supplement to this SAP.

5. Analysis Sets

The following analysis sets will be used in the Safety run-in part:

Full analysis set (FAS): All patients in the Safety run-in part who receive at least one dose of study drug.

Safety analysis set (SAF): All patients in the Safety run-in part who receive at least one dose of study drug.

Pharmacokinetic analysis set: All patients in the Safety run-in part with at least one valid PK measurement (i.e., patients who took at least one dose of study drug and had at least one PK sample collected and measured) will be included in the evaluation of PK concentration and parameters.

6. Statistical Methodology

6.1 Population Characteristics

Analyses will be performed on the SAF. Population characteristics will be summarized for global patients by background therapy (R-B vs R-CHOP) and dose levels of copanlisib (45mg and 60mg). As all patients in the Japan cohort received 60 mg of copanlisib in combination with R-B, the summaries of R-B background therapy will include a standalone summary for the Japan 60 mg cohort. A requirement of the SAS macro, used to create summaries by background therapy, is that the layout of the treatment groups, Copa 45mg, Copa 60 mg and Japan (Copa 60 mg), is the same for both background therapies. As a consequence, R-CHOP background therapy summaries will display the Japan (Copa 60 mg) treatment cohort with zero patients.

The summaries will not include overall totals as the Phase III part will provide a comprehensive review of safety.

6.1.1 Disposition

Sample sizes by background therapy and country/region will be presented and will include the number of centers enrolling patients per country/region. The number of patients enrolled, assigned to treatment, and valid for the different analysis sets FAS and SAF will be summarized by dose level, and country/region.

The number of screening failures (patients discontinuing the screening period) together with primary reasons for discontinuation will be presented overall. The number of patients assigned to treatment, in addition to the number of patients discontinuing the treatment, Safety follow-up, Active follow-up and Survival follow-up periods together with the primary reason for discontinuation will be presented by background therapy and dose level/Japan cohort.

In addition, the number of patients assigned to treatment with important protocol deviations will be presented by background therapy and by country/region for each dose level. The

frequency of major deviations and number of patients valid for each analysis set will be presented by background therapy and dose level/Japan cohort.

A separate patient listing of protocol deviations related to COVID-19 pandemic (e.g., delay of tumor assessments, delay, interruption or discontinuation of the treatment due to the impact of COVID-19 pandemic) will be provided by background therapy. These protocol deviations will be entered in RAVE as Manual protocol deviations and identified from the specification of "COVID".

Similarly, if any Safety run-in patients are affected by the Ukraine-Russia conflict, a separate patient listing of protocol deviations related to the Ukraine-Russia conflict (e.g., delay of tumor assessments, delay, interruption or discontinuation of the treatment due to the impact of the conflict) will be provided. These protocol deviations will be entered in RAVE as Manual protocol deviations and identified from the specification of "Regional Crisis".

6.1.2 Demographic and other baseline characteristics

The following demographics and baseline characteristics will be summarized by background therapy (R-B vs R-CHOP) and dose level/Japan cohort: age, sex, ethnicity, race, weight, height, body mass index (BMI), systolic and diastolic blood pressure, heart rate, temperature, New York Heart Association classification, estimated Glomerular Filtration Rate (eGFR), glycated hemoglobin (HbA1c), medical history of diabetes and medical history of hypertension. Age, BMI, eGFR and HbA1c will be analyzed as continuous variables and in addition categorized with the following categories:

- Age group 1 (years): <65, ≥65
- Age group 2 (years): <80, ≥80
- Age group 3 (years): <65, 65-74, 75-84, ≥85
- BMI group (kg/m²): <18.5, ≥18.5- <30, ≥30
- eGFR group (mL/min/1.73 m^2): Normal: ≥90 vs. Mild impairment: 60 - <90 vs. Moderate impairment: 30- <60 vs. Severe impairment: 15 - <30)
- HbA1c group: <5.7% vs ≥5.7% – <6.5% vs ≥6.5% – ≤8.5% vs >8.5%

Medical history of diabetes and hypertension, identified by Product specific Bayer MedDRA Queries (PBMQ), will be summarized as follows:

- Diabetic History: [PBMQ] Medical history of diabetes (no/yes)
- Hypertension History: [PBMQ] Medical history of hypertension (no/yes)

The following additional baseline characteristics will be summarized:

- Histology of tumor per investigator's assessment (FL, MZL, SLL, and LPL/WM)
- Staging and size (sum of products of diameters) of tumor at study entry
- FL grade (for patients with FL) at study entry
- ECOG performance status at baseline
- Time from date of initial diagnosis to the date of first study treatment (months)
- Time from most recent progression to the date of first study treatment (weeks)
- Time from first progression to the date of first study treatment (months)
- Number of lesions (number of target and number of non-target lesions) at baseline
- Serum Immunoglobulin M (IgM) level (for LPL/WM patients) at baseline

The retrospective evaluation of histopathological diagnosis at baseline will be performed centrally. These data are exploratory and will be listed for the Clinical Study Report appendix.

Quantitative data will be summarized by arithmetic mean, standard deviation, median, minimum and maximum. Frequency tables will be provided for qualitative data.

6.1.3 Medical history

Medical history findings (as defined in protocol Section 9.3.2) will be coded by Medical Dictionary for Regulatory Activities (MedDRA) codes. Medical history will be presented for each MedDRA Primary System Organ Class (SOC) and Preferred Term (PT) by background therapy (R-B vs R-CHOP) and dose level/Japan cohort. Ongoing conditions will be presented in a similar way.

6.1.4 Prior and concomitant medication

Prior and concomitant medications will be coded by the Anatomical Therapeutic Chemical (ATC) classification system according to the World Health Organization Drug Dictionary. Summaries will be provided by ATC class and subclass. The number of patients taking at least one prior medication will be reported in addition to the number of patients who have taken medication relating to each relevant ATC class and subclass. The same will be reported for concomitant medications. Note that the same medication can appear multiple times in the table as it can have several ATC codes.

A medication that has been stopped after first administration of study treatment is considered as concomitant, while those that have been stopped before first administration are considered as prior. Medications with missing start and stop date but flagged as being ongoing at end of study will be considered to have started prior to start of study treatment and ended after stop

of study treatment. Medications starting more than 30 days after the last day of study treatment will be excluded from the summary of concomitant medications.

Prior and concurrent anti-cancer therapy

The mean, minimum, median, standard deviation and maximum number of prior systemic anti-cancer therapy lines as well as number of patients with 1, 2, ≥3 lines of prior therapy will be summarized by background therapy (R-B vs R-CHOP) and dose level/Japan cohort for SAF. Time since last systemic anti-cancer therapy[2] will be summarized using descriptive statistics by background therapy (R-B vs R-CHOP) and dose level/Japan cohort for SAF. The time between the last day of the last course of systemic anti-cancer therapy and the day of confirmation of the most recent progression will also be displayed as a classification table, showing proportions of patients with ≤ 6 months vs. > 6 to <12 months vs. ≥ 12 months, by background therapy (R-B vs R-CHOP) and dose level/Japan cohort.

In addition, the following frequency tables will be provided by background therapy (R-B vs R-CHOP) and dose level/Japan cohort for the SAF:

- Prior radiotherapies
- Prior systemic anti-cancer therapies (displayed by 1. ATC classes, 2. drug names)
- Type of prior systemic anti-cancer therapy
- Prior treatment with PI3K inhibitors (refer to the Tables, Listings and Figures (TLF) specification document for details of how prior PI3K inhibitors are identified).
- Prior diagnostic and/or therapeutic procedures
- Concurrent diagnostic and/or therapeutic procedures
- Radiotherapies during Survival follow-up
- Systemic anti-cancer therapies during Survival follow-up (displayed by 1. ATC classes, 2. Drug names)
- Type of systemic anti-cancer therapy during Survival follow-up

6.1.5 Treatment duration and exposure

Investigator-recorded cycles for study drug exposure will be mapped to a calculated cycle. The analysis of treatment duration and exposure will use the mapped, calculated cycle. A description of how investigator defined cycles are mapped to analysis cycles can be found in the study document “Analysis Data Reviewer’s Guide”.

[2] Time between last day of last systemic anti-cancer therapy and date of first treatment.

Descriptive statistical summaries will be provided separately for copanlisib, R-B and R-CHOP by background therapy (R-B vs R-CHOP) and dose level/Japan cohort for the following variables:

- Overall extent of exposure (OEE, as defined in section 4.7.4) for FL, MZL, SLL, LPL/WM and all SAF patients
- Overall extent of exposure for FL, MZL, SLL, LPL/WM and all SAF patients, by categories:
    - 0-90 days
    - 91-180 days
    - 181-270 days
    - 271-365 days
    - >365 days
- Number of cycles for FL, MZL, SLL, LPL/WM and all SAF patients (statistical summaries in addition to number of patients with each number of cycles)
    - For copanlisib: OEE divided by 28 days
    - For R-B: OEE divided by 28 days.
    - For R-CHOP: OEE divided by 21 days
- Number of infusions during treatment period for FL, MZL, SLL, LPL/WM and all SAF patients. For R-CHOP, the number of infusions will be based on rituximab, cyclophosphamide, doxorubicin, and vincristine infusions as prednisone is taken orally.

In addition, the following analyses will be provided for the SAF, by background therapy (R-B vs R-CHOP) and dose level/Japan cohort, over all cycles and by calculated cycle:

- Total amount of dose actually administered: sum of actual dose (mg) per calculated cycle for each patient
    - For copanlisib, prednisone/prednisolone, the actual dose per timepoint is defined as

      Prescribed dose [mg] × (Total amount of dose administered [mg OR mL] / Total amount of dose prior to administration [mg OR mL])
    - For rituximab, cyclophosphamide, doxorubicin, vincristine and bendamustine, the actual dose per timepoint is defined as

Prescribed dose [mg] / Body surface Area (BSA) [m^2] × (Total amount of dose administered [mL] / Total amount of dose prior to administration [mL])

- Percent of planned dose received = Actual dose [mg] / Planned dose [mg] × 100%.
    - For copanlisib, a standard dose of either 45mg or 60 mg is set for all cycles, i.e., infusion days (Day 1, Day 8 and Day 15 of each cycle), and for planned dose (might however be modified by individual dose interruptions). Completed cycles therefore have a planned dose of either 135 mg or 180 mg. For incomplete cycles, the planned dose depends on the number of days (d) in that calculated cycle:
        - 45 mg at dose level 1 or 60 mg at dose level 2 if $0< d <7$
        - 90 mg at dose level 1 or 120 mg at dose level 2 if $7\leq d <14$
        - 135 mg at dose level 1 or 180 mg at dose level 2 if $14\leq d \leq 28$
    - For rituximab, a standard of 375 mg/m² body surface is set for all cycles for planned dose.
    - For cyclophosphamide, a standard of 750 mg/m² body surface is set for all cycles for planned dose.
    - For doxorubicin, a standard of 50 mg/m² body surface is set for all cycles for planned dose.
    - For vincristine, a standard of 1.4 mg/m² body surface is set for all cycles for planned dose and a maximum of 2.0 mg/m² is also previewed.
    - For bendamustine, a standard of 90 mg/m² body surface is set for all cycles for planned dose (might however be modified by individual dose-reductions or interruptions), i.e., infusion days (Day 1 and Day 2 of each cycle). For incomplete cycles, the planned dose depends on the number of days (d) in that calculated cycle:
        - 90 [mg/ m^2] ×BSA [m^2] if $0< d <2$
        - 90 [mg/ m^2] ×BSA [m^2] ×2 if $2\leq d \leq 28$
    - For prednisone/prednisolone, a standard of 100 mg is set for all cycles, i.e., orally (PO) days (Day 2 to Day 6 of each cycle), and for planned dose (might however be modified by individual dose-reductions or interruptions). Completed cycles therefore have a planned dose of 500 mg. For incomplete

cycles, the planned dose depends on the number of days (d) in that calculated cycle:

- 0 mg if 0< d <2
- 100 mg if 2≤ d <7 and increase 100 mg per day
- 500 mg if 7≤ d ≤21

For patients with dose reduction[3], interruption or delay, the number of dose reductions, interruptions (dose administered is non-zero), or delays (zero dose is administered) per patient and their reasons will be summarized separately for copanlisib, rituximab[4], cyclophosphamide, doxorubicin, vincristine, prednisone/ prednisolone and bendamustine, by background therapy (R-B vs R-CHOP) and dose level/Japan cohort in the SAF. Interruptions with the same start date will be counted as one interruption.

The duration of dose delays in days will be summarized for copanlisib by background therapy (R-B vs R-CHOP) and dose level/Japan cohort in the SAF. The duration is defined as the difference between the date of delay and the date of next exposure information (infusion, interruption or delay). Due to the treatment regimen with one week break at the end of each cycle, delays on day 15 of R-CHOP combination therapy, or on day 22 of R-B combination therapy, are capped at 7 days, in case the calculated duration is greater than 7.

Patient listings will be provided for dose modifications.

6.2 Efficacy

The analysis of efficacy will be summarized with descriptive statistics for the FAS, by background therapy and dose level/Japan cohort. Summaries of tumor response by independent central review and investigator's assessment will be provided.

Tumor response

Efficacy will be evaluated based on radiological tumor evaluations of neck, chest, abdomen and pelvis by using contrast enhanced computed tomography/magnetic resonance imaging (CT/MRI), and positron emission tomography-computed tomography (PET-CT) as an optional imaging modality. Scans will be performed at Screening, during Treatment as well as during the Active follow-up period (every 12 weeks [± 7 days] from Cycle 1 Day 1 during Years 1 and 2; every 24 weeks [± 7 days] during Years 3, 4, and 5; and every 24 weeks [± 14 days] beyond Year 5). Response assessment will be based on the Recommendations for Initial Evaluation, Staging, and Response Assessment of Hodgkin and Non-Hodgkin Lymphoma: The Lugano Classification and for patients with LPL/WM according to the Owen Criteria and CT/MRI, if applicable. As long as the patient has not experienced PD, investigator's assessment is sufficient for case management. In the event of radiological progression, radiological real-time confirmation by central independent blinded evaluation is needed

[3] Per the clinical study protocol (CSP), there will be no dose reductions of copanlisib, CHOP or bendamustine during the first cycle of study treatment.
[4] Per the CSP, there will be no dose reductions for rituximab in any cycle.

before a final decision is made by the investigator. Evaluations from independent central blinded review will also be used for response endpoints containing radiological tumor assessments.

LPL/WM patients who do not have radiologically measurable lesions at Screening will not have further radiological assessments as per protocol. Their tumor response will be measured only by laboratory/clinical tests and evaluated by investigators. However, because these patients may on progression develop measurable disease without simultaneous increase in IgM, the imaging material should be submitted for central review and PD confirmation. LPL/WM patients who have radiologically measurable lesions at Screening will continue to have radiological assessments and, in addition, will have laboratory tests performed on the same days.

Bone marrow biopsy will be mandatory at Screening. If the baseline biopsy is positive for lymphoma infiltration, it will be mandatory to perform it again to confirm the first complete response (CR).

The CT-based response status consists of complete response (CR), unconfirmed CR, partial response (PR), stable disease (SD), and progressive disease (PD). The PET-CT response status consists of complete metabolic response (CMR), partial metabolic response, no metabolic response, equivalent with stable disease, and progressive metabolic disease, equivalent with progressive disease. PET-CT response overrides in most cases the response by CT and/or MRI alone: e.g., if PET-CT response = CMR and CT/MRI response = PR, overall response is CR.

The response status for LPL/WM patients being assessed according to the Owen Criteria consists of CR, Very Good Partial Response (VGPR), PR, Minor Response (MR), SD, and PD.

Definition of exploratory efficacy variables

Objective response rate (ORR)

Assessed in all patients, ORR is defined as the proportion of patients who have a best response rating, while on treatment or in Active follow-up (i.e., until the time of the Phase III primary completion analysis), of complete response (CR) or partial response (PR) according to the Lugano Classification and for patients with LPL/WM a response rating of CR, VGPR, PR, or MR according to the Owen Criteria. As initiation of new anti-tumor therapy ends the Active follow-up, any tumor assessments taking place after the initiation of new anti-cancer therapy will not be included in the calculation of ORR.

Disease control rate (DCR)

DCR is defined as the proportion of patients who have a BOR of CR, unconfirmed CR, VGPR, PR, unconfirmed PR, MR, or confirmed stable disease (SD). Unconfirmed SD is

defined as SD on or before Study Day 76, which is not confirmed by an additional tumor assessment of SD, or better, on Study Day 77 or later.**Complete response rate (CRR)**

Assessed in all patients, CRR is defined as the proportion of patients who have a best response rating of CR while on treatment or in Active follow-up (i.e., until the time of the Phase III primary completion analysis) according to Lugano classification and for patients with LPL/WM, a response rating of CR according to the Owen criteria.

Overall survival (OS)

OS is defined as the time (in days) from start of study treatment until death from any cause. OS of patients alive at the time of the Phase III primary completion analysis will be censored at the last date they were known to be alive (last known alive date; LKAD). Death after the data cut-off, reported during data cleaning, will be considered for establishing LKAD at data cut-off.

See Section 4.3 for details regarding derivation of LKAD and imputation methods for partial/missing death dates.

Frequency tables of best overall response (BOR), ORR, DCR and CRR by background therapy (R-B vs R-CHOP) and dose level/Japan cohort will be presented.

Due to the small number of patients a descriptive summary of survival will be provided by dose level/Japan cohort, for each background therapy and overall. It will include the number of patients with an event (death), number of patients censored, median OS, OS interquartile range and OS range with and without censored values. Survival rates, and corresponding confidence intervals, at 6 monthly time intervals from start of study treatment will also be provided. There will be no inferential analysis of OS treatment differences.

In addition the following patient listing of exploratory efficacy variables will be provided by background therapy (R-B vs R-CHOP) and dose level/Japan cohort: patient ID, histology, start and stop dates of study treatment, BOR, first response date, time from start of study treatment to PD or death from any cause, PFS, date of death, primary cause of death and OS.

6.3 Safety

Analyses will be performed on the SAF by background therapy (R-B vs R-CHOP) and dose level/Japan cohort.

6.3.1 Adverse Events

Adverse events (AEs) will be coded according to MedDRA Version 25.1 or later and using NCI-CTCAE Version 4.03 (or higher) dictionary. The final versions used will be stored in the clinical database and provided in a footnote to the respective tables. Severity of AEs and hematological/biochemical toxicities based on laboratory measurements will be graded using NCI-CTCAE dictionary. AEs will be classified as related or not related to study drug, as determined by the investigator.

A treatment-emergent AE (TEAE) is defined as any event occurring or worsening after start of study drug administration until 30 days after the last study drug intake.

Descriptive summary tables (frequency and percentage of patients, not of events) will be presented by background therapy (R-B vs R-CHOP) and dose level/Japan cohort, by MedDRA SOC/PT and by worst NCI CTCAE grade for the following:

- Overview of TEAEs
- TEAEs
- Drug-related TEAEs (copanlisib or R-B/R-CHOP)
- Treatment-emergent serious AEs (TESAEs)
- Drug-related TESAEs
- DLTs (see Section 3.1 for the definition of DLTs)
- TEAEs leading to drug interruption
- TEAEs leading to dose reduction
- TEAEs leading to permanent treatment discontinuation

As relationship to study drug is assessed separately for copanlisib, R-B, R-CHOP, drug-related adverse events will be summarized separately for each study treatment. Deaths (including all grade 5 TEAEs) will be summarized by background therapy (R-B vs R-CHOP) and dose level/Japan cohort. As “AE” is not captured as a primary cause of death in the Survival follow-up, for the overall summary of deaths, deaths with “AE” or “other” as the primary cause will be pooled together under the primary cause “other”.

AEs occurring prior to starting treatment and AEs occurring at least 30 days after the last dose of study treatment (i.e., post-treatment AEs) will also be summarized by background therapy (R-B vs R-CHOP) and dose level/Japan cohort.

The categories of TEAEs, Treatment-emergent copanlisib-related AEs, and TESAEs will be presented according to time-of-onset (new onset or worsening).

The following categories will be presented:

- Day 1 – Day 90: to include all copanlisib-treated patients
- Day 91 – Day 180: to include all copanlisib-treated patients treated for at least 3 months
- Day 181 – Day 270: to include all copanlisib-treated patients treated for at least 6 months
- Day 271 – Day 360: to include all copanlisib-treated patients treated for at least 9 months
- Day 361 or greater: to include all copanlisib-treated patients treated for at least 12 months

In addition, patient listings will be provided by background therapy (R-B vs R-CHOP) and dose level/Japan cohort, for the following AEs:

- Patients who died during study treatment up to 30 days after last dose of study treatment: patient ID, sex, age, race, histology, sponsor AE identifier, start and stop dates of study treatment, date of death, days within first and last dose date, cause of death, MedDRA SOC and PT and relationship to study drug (related to copanlisib, related to R-B/R-CHOP, related to both, and no).
- Patients who died later than 30 days after last dose of study treatment: patient ID, sex, age, race, histology, sponsor AE identifier, start and stop dates of study treatment, date of death, days within first and last dose date, cause of death, MedDRA SOC and PT, and relationship to study drug (related to copanlisib, related to R-B/R-CHOP, related to both, and no).
- Treatment-emergent SAEs: patient ID, sex, age, histology, race, sponsor AE identifier, MedDRA SOC and PT, CTCAE toxicity grade, start and stop dates of study treatment, reason for seriousness, start and stop date of AE, relationship to study drug (related to copanlisib, related to R-B/R-CHOP, related to both, and no), protocol required procedure related (yes/no), outcome, actions taken with each study drug.
- Listing of study treatment-related SAEs will be provided separately for each copanlisib dose level, the Japan cohort (all patients in this cohort received 60 mg of copanlisib), R-B (excluding the Japan cohort), R-B (Japan cohort only) and R-CHOP, with similar information.
- TEAEs leading to permanent discontinuation of study treatment, a dose reduction (only after cycle 1) or interruption will be listed separately for each copanlisib dose level, the Japan cohort (all patients in this cohort received 60 mg of copanlisib), R-B (excluding the Japan cohort), R-B (Japan cohort only) and R-CHOP: patient ID, sex, age, histology, race, sponsor AE identifier, MedDRA SOC and PT, CTCAE toxicity grade, start and stop dates of study treatment, start and stop dates of AE, relationship to study drug (related to copanlisib, related to R-B/R-CHOP, related to both, and no),

protocol required procedure related (yes/no), serious (yes/no), reason for seriousness, outcome, actions taken with each study drug.

- DLTs: patient ID, sex, age, histology, race, sponsor AE identifier, MedDRA SOC and PT, CTCAE toxicity grade, start and stop dates of study treatment, reason for seriousness, start and stop date of AE, relationship to study drug (related to copanlisib, related to R-B/R-CHOP, related to both, and no), serious (yes/no), reason for seriousness, protocol required procedure related (yes/no), outcome, actions taken with each study drug.

- TEAEs of pneumonitis (The standardized MedDRA query (SMQ); 'Interstitial lung disease' with narrow search (i.e., category 2A) will be used to select these AEs): patient ID, sex, age, race, histology, sponsor AE identifier, MedDRA SOC and PT, CTCAE toxicity grade, start and stop dates of study treatment, serious (yes/no), reason for seriousness, start and stop date of AE, relationship to study treatment (related to copanlisib, related to R-B/R-CHOP, related to both, and no), protocol required procedure related (yes/no), serious (yes/no), reason for seriousness, outcome, actions taken with each study drug.

- TEAEs of hyperglycemia (Bayer MedDRA labeling grouping; 'Hyperglycemia' will be used to select these AEs): patient ID, sex, age, race, histology, sponsor AE identifier, MedDRA SOC and PT, CTCAE toxicity grade, start and stop dates of study treatment, serious (yes/no), reason for seriousness, start and stop date of AE, relationship to study treatment (related to copanlisib, related to R-B/R-CHOP, related to both, and no), protocol required procedure related (yes/no), serious (yes/no), reason for seriousness, outcome, actions taken with each study drug

- TEAEs of arterial hypertension (Bayer MedDRA labeling grouping; 'Hypertension' will be used to select these AEs): patient ID, sex, age, race, histology, sponsor AE identifier, MedDRA SOC and PT, CTCAE toxicity grade, start and stop dates of study treatment, serious (yes/no), reason for seriousness, start and stop date of AE, relationship to study treatment (related to copanlisib, related to R-B/R-CHOP, related to both, and no), protocol required procedure related (yes/no), serious (yes/no), reason for seriousness, outcome, actions taken with each study drug

COVID-19 pandemic-relevant adverse events

The COVID-19 pandemic-relevant AEs will be identified using the SMQ “Covid-19” with a broad search. A patient listing of these AEs will be provided. The serious AEs and those leading to treatment discontinuation and/or deaths will be flagged in the listing.

6.3.2 Safety Parameters

Quantitative data [hematology, blood chemistry, coagulation, urinalysis, vital signs ECG and multiple gated acquisition (MUGA)] will summarized by background therapy and dose level/Japan cohort and will be described by the following summary statistics: arithmetic mean, standard deviation, median, minimum and maximum. These summary statistics will be presented at each post-baseline visit for the original data as well as for the change from baseline. Frequency tables will be provided for qualitative data.

Data from unscheduled visits will not be included in summary analyses over time (by time point), ; unless they are the last measurement prior to first study drug intake, but will be included in other analyses not summarized by time point, such as shift tables, incidence of abnormalities, etc. When summarizing worst-case or best-case data, data from unscheduled visits will be considered. Unscheduled laboratory assessments will be included in the patient listings of clinical laboratory parameters.

Changes in worst grade of laboratory toxicities at last pre-treatment value compared to worst grade post-baseline value and changes in worst grade of laboratory toxicities under treatment will be provided by background therapy and dose level/Japan cohort. For the change in worst grade of laboratory toxicities under treatment, observations up until the last on treatment visit will be considered. In case there is more than one observation for the same timepoint, the worst grade will be taken. In addition, treatment-emergent abnormal laboratory values (above and below the normal range) will be summarized by background therapy, dose level/Japan cohort and timepoint. For the summary of high treatment-emergent laboratory abnormalities, the denominator represents the number of distinct patients at baseline with a normal or lower than normal laboratory assessment who also had at least one valid laboratory value (including unscheduled visits) after start of study treatment but ≤ 30 days after end of study treatment. Patients with missing or high abnormal values at baseline are not included in the denominator. The numerator represents the number of distinct patients with at least one high laboratory assessment (including unscheduled visits) after the start of treatment who had a normal or lower than normal laboratory assessment at baseline. A similar approach is used for low treatment-emergent abnormalities except that patients with missing or low abnormal values at baseline are not included in the denominator. Urinalysis data will be summarized by frequency tables.

ECOG performance status values and change from baseline values will be summarized by background therapy, dose level/Japan cohort and visit as well as to the end of treatment.

Results of the 12-lead ECGs, including change from baseline, will be summarized with descriptive statistics by background therapy, dose level/Japan cohort, visit and timepoint. The overall interpretation of the 12-lead ECG and the ECG diagnosis will also be summarized by background therapy, dose level/Japan cohort and visit. Frequency and shift tables for the number of patients by interpretation of ECG as normal or abnormal will be performed by background therapy, dose level/Japan cohort at all visits and analysis time points. A listing of clinically significant abnormal 12-lead ECG findings will be provided.

Results of MUGA/echocardiogram will be summarized with descriptive statistics by background therapy, dose level/Japan cohort and visit, including change from baseline where appropriate.

The overall interpretation of the echocardiogram and the corresponding diagnosis will be summarized by background therapy, dose level/Japan cohort and visit.

Frequency and shift tables for number of patients by interpretation of echocardiogram as normal, abnormal clinically significant (CS) or abnormal not CS will be performed at all visits. A listing of clinically significant abnormal cardiological findings will be provided.

Vital sign values and change from baseline values will be summarized descriptively by background therapy, dose level/Japan cohort and visit.

The number of patients with abnormal post-dose blood pressure will be displayed by visit and post-dose hypertension grade.

6.4 Pharmacokinetics

The data from the Safety run-in part and Phase III part will be pooled for PK analysis. Individual concentration-time data of copanlisib and its metabolite 1 as well as rituximab will be listed only.

6.5 Additional Analyses Planned to Be Reported Outside the Main Report

All biomarker analyses will be considered exploratory. There will be no overall or sensitivity analysis of central pathology data. Biomarker data collected in the clinical database will be listed. Retrospective exploratory biomarker analyses including evaluation of relationship between biomarker data and clinical outcomes will be performed by or under the supervision of the sponsor's biomarker statistical expert. Results of these analyses will be reported separately, under a separate cover.

7. Document history and changes in the planned statistical analysis

7.1 Document history

Version	Date	Action	Comment
Version 0.1 draft	28 Jan 2016	Review	Safety run-in only
Version 0.2 draft	21 Jul 2016	Update and Review	Safety run-in only
Version 1.0 final	28 Aug 2016	Signed initial version	Safety run-in only
Version 1.01 draft	14 Jul 2022	Update and Review	Safety run-in only
Version 2.0 final	22 Jul 2022	Signed second version	Safety run-in only
Version 2.1 draft	13 Sep 2023	Update and review	Safety run-in only
Version 2.2 draft	15 Sep 2023	Updated with comments from review of draft version 2.1 Section 7.2 added	Safety run-in only
Version 3.0 final	22 Sep 2023	Signed third version	Safety run-in only

7.2 Changes in the planned statistical analysis

The definition of the secondary efficacy variable DCR has been amended in the SAP to remove references to "within 35 days after termination of study treatment" and "within the following 25 weeks". This change has been made to be consistent with the ORR and CRR assessment period, by including in the DCR analysis, radiological assessments performed during the Active follow-up.

Definition of DCR as defined in current CSP, Amendment 09, protocol Version 8.0, dated 21 MAR 2023

DCR is defined as the proportion of patients who have a best response rating of CR, VGPR, PR, MR, or stable disease (SD) (excluding unconfirmed early SD) that is achieved during treatment or within 35 days after termination of study treatment. An unconfirmed early SD is defined as SD on or before Study Day 76, with no additional tumor assessment of SD or better within the following 25 weeks.

SAP definition of DCR as specified in Section 6.2 of this SAP

DCR is defined as the proportion of patients who have a BOR of CR, unconfirmed CR, VGPR, PR, unconfirmed PR, MR, or confirmed stable disease (SD). Unconfirmed SD is defined as SD on or before Study Day 76, which is not confirmed by an additional tumor assessment of SD, or better, on Study Day 77 or later.

8. References

1. Cheson, BD, Fisher, RI, Barrington, SF, Cavalli, F, Schwartz, LH, et al. Recommendations for Initial Evaluation, Staging, and Response Assessment of Hodgkin and Non-Hodgkin Lymphoma: The Lugano Classification. J Clin Oncol. 2014; 32(27):3059-67.

2. Owen, RG, Kyle, RA, Stone, MJ, Rawstron, AC, Leblond, V, et al. Response assessment in Waldenstrom macroglobulinaemia: update from the VIth International Workshop. British journal of haematology. 2013; 160(2):171-6.

9. Appendix

9.1 Imputation rules for partial dates

9.1.1 Algorithm for handling partial initial diagnosis date and progression date

The following rules are used to impute partial initial diagnosis date and progression date,

A. If partial date has day and month missing, then January 01 will be assigned to the missing fields.

B. If partial date has missing day only, then the first day of the month will be assigned to the missing day.

C. If imputed progression date from A or B above is before the Randomization date, then Randomization date will be assigned to overwrite the imputed date from A or B. If the date is completely missing, no imputation will be applied for initial diagnosis date. Use day 1 for progression date.

9.1.2 Imputation rules for partial/missing Adverse Event (AE)/Concomitant medication (CM) dates

Table 9-1: Imputation rules for partial/missing dates

Partial Dates Imputation Rule	Impute partial AE/CM Start Date	Impute partial AE/CM Stop Date
#1. The day missing only	IF AESTDT year and months is same as TRTSDT year and months, then impute AESTDT= TRTSDT	IF AEENDT year and months is same as LKAD year and months, then impute AEENDT= LKAD
	ELSE IF AESTDT year and month is before TRTSDT year and months, then AESTDT= last date of the months	ELSE impute AEENDT= last date of the months
	ELSE IF AESTDT year and month is after TRTSDT year and months, then AESTDT= first date of the months	
#2. Both day and months missing	IF AESTDT year is same as TRTSDT year, then impute AESTDT=TRTSDT	IF AEENDT year is same as LKAD year, then impute AEENDT= LKAD
	ELSE IF AESTDT year is before TRTSDT year, then impute AESTDT=31DECYYYY	ELSE impute AEENDT=31DECYYYY
	ELSE IF AESTDT year is after TRTSDT year, then impute AESTDT=01JANYYYY	
#3. Completely missing	No need to impute, try to query the sites by DM	No need to impute, try to query the sites by DM
#4 Additional criteria to meet	1. AE/CM start date <= AE/CM stop date ## After the imputation of start/stop date per above rules #1~#3, further action is needed if the (imputed) AE/CM start date > the (imputed) AE/CM stop date. We will update the (imputed) AE/CM start date = the (imputed) AE/CM stop date since the imputed stop date is more reliable. Because per imputation partial stop date rules #1~#3, the (imputed) AE/CM stop date will be the maximum date based on the available data unless there's data issue. If it's a data issue, we can leave the imputation as it is and report the data issue to the DM. ## 2. The imputed dates <= LKAD 3. If TRTSDT is missing, use RANDDT as reference date	

9.1.3 : Imputation rules for partial/missing anti-cancer treatment start date

Table 9-2: Imputation rules for partial/missing anti-cancer treatment start dates

Partial Dates Imputation Rule	Impute partial anti-cancer treatment start date
#1. The day missing only	IF *CMSTDTC year is after the year of the last visit, then impute DD to be '01' ELSE IF *CMSTDTC year is the same as the year of the last visit and *CMSTDTC month is after the month of the last visit, then impute DD to be '01' ELSE IF *CMSTDTC year and month is the same as the year and month of the last visit, then impute day of the last visit
#2. Both day and months missing (MM-DD)	IF *CMSTDTC year is the same as the year of the last visit, then impute MM-DD to be the month and day of the last visit' ELSE IF *CMSTDTC year is after the year of the last visit, then impute MM-DD to be '01-01'
#3. Completely missing (YYYY-MM-DD)	IF the last visit date is known, then impute the last visit date

*CMSTDTC is the start date/time of the anti-cancer treatment start date

Title page

A Phase III, randomized, double-blind, controlled, multicenter study of intravenous PI3K inhibitor copanlisib in combination with standard immunochemotherapy versus standard immunochemotherapy in patients with relapsed indolent B-cell non-Hodgkin's lymphoma (iNHL) – CHRONOS-4

Copanlisib and standard immunochemotherapy in relapsed iNHL

Bayer study drug:	BAY 80-6946 / Copanlisib		
Study purpose:	Dose-finding, efficacy, and safety of copanlisib		
Clinical study phase:	III	**Date:**	22 SEP 2023
Study No.:	17833	**Version:**	4.0
Author:	PPD		

This Statistical Analysis Plan is produced on a word-processing system and bears no signatures.

The approval of the Statistical Analysis Plan is documented in a separate Signature Document.

Table of Contents

Table of Tables

Table of Figures

List of Abbreviations

ADRG	Analysis Data Reviewer's Guide
AE	Adverse event
ATC	Anatomical therapeutic chemical
AUC	Area under the curve
B	Bendamustine hydrochloride
BMI	Body mass index
BRM	Blind review meeting
BRR	Blind review report
BSA	Body surface area
CI	Confidence interval
CIL	Confidence interval limit
CMV	Cytomegalovirus
COVID-19	Corona virus disease 2019
CSP	Clinical study protocol
CR	Complete response
CRR	Complete tumor response rate
CSR	Clinical study report
CTCAE	Common terminology criteria adverse event
DCR	Disease control rate
DM	Demography
DMC	Data monitoring committee
DOR	Duration of response
DRS-E	Disease-related symptoms - emotional
DRS-P	Disease-related symptoms – physical
DS	Disposition
ECG	Electrocardiogram
ECOG	Eastern Cooperative Oncology Group
eCRF	Electronic case report form
EDC	Electronic Data Capture
e.g.	For example (*exempli gratia*)
eGFR	Estimated glomerular filtration rate
EOT	End-of-treatment
EX	Exposure
FAS	Full analysis set
FDA	Food and Drug Administration
FL	Follicular lymphoma
FLymSI-18	National Comprehensive Cancer Network-Functional Assessment of Cancer, Lymphoma Symptom Index-18 Questionnaire
FWB	Function and well-being
H_0	Null hypothesis
H_1	Alternative hypothesis
HA	Health authority
HbA1c	Glycated hemoglobin
HLGT	High level group term
HLT	High level term
HR	Hazard ratio
IA-SAP	Integrated Analysis-Statistical Analysis Plan
ID	Identifier
i.e.	That is (*id est*)
IgM	Immunoglobulin M
iNHL	indolent non-Hodgkin's lymphoma
ITT	Intent-to-treat
IxRS	Interactive voice/web response system
kg	Kilogram
LB	Laboratory
LKAD	Last known alive date

LPL	Lymphoplasmacytoid lymphoma
m	Meter
M-1	Metabolite 1
MedDRA	Medical Dictionary for Regulatory Activities
mg	Milligram
min	Minutes
MLG	Bayer MedDRA labeling groupings
MR	Minor response
MUGA	Multiple gated acquisition
MZL	Marginal zone lymphoma
M&S	Modeling and Simulation
NCI	National Cancer Institute
NE	Non-evaluable
NHL	Non-Hodgkin's lymphoma
NYHA	New York Heart Association
ODAC	Oncologic Drugs Advisory Committee
OEE	Overall extent of exposure
ORR	Objective response rate
OS	Overall survival
PBMQ	Product specific Bayer MedDRA Query
PD	Progressive disease
PFS	Progression-free survival
PI3K	Phosphatidylinositol 3-kinase
PK	Pharmacokinetics
PKS	Pharmacokinetics analysis set
PT	Preferred term
PR	Partial response
QoL	Quality of life
R	Rituximab
R-B	Rituximab and bendamustine
R-CHOP	Rituximab, cyclophosphamide, doxorubicin, vincristine, and prednisone
RS	Tumor response
SAE	Serious adverse event
SAF	Safety analysis set
SAP	Statistical analysis plan
SAS	Statistical Analysis System
SD	Stable disease
SLL	Small lymphocytic lymphoma
SMQ	Standardized MedDRA query
SOC	Primary system organ class
SV	Subject visits
TEAE	Treatment-emergent adverse event
TESAE	Treatment-emergent serious adverse event
TLF	Tables, listings, and figures
TSE	Treatment side-effect
TTNT	Time to next anti-lymphoma treatment
TTP	Time to progression
US, USA	United States (of America)
VGPR	Very good partial response
vs.	As opposed to (*versus*)
WHO-DD	World Health Organization - Drug Dictionary
WM	Waldenström macroglobulinemia

1. Introduction

Non-Hodgkin's lymphomas (NHLs) are a heterogeneous group of lymphoproliferative malignancies typically originating in the lymph nodes but can involve almost any organ tissue [1]. NHLs can be divided according to their clinical behavior in two main prognostic groups: indolent NHL (iNHL) and aggressive NHL. Aggressive lymphomas are characterized by an aggressive clinical course and may evolve into a lethal presentation if not immediately treated. Indolent NHLs have a relatively good prognosis with a median survival longer than 10 years, but they are incurable with current available therapeutic options, especially in advanced stages. While they are highly responsive to standard chemotherapy regimens and to radiotherapy, their natural history is characterized by a continuous pattern of relapses, which can generally be treated with success. But the time to next relapse progressively decreases each time, finally evolving into a refractory disease or in a transformation into an aggressive histologic type.

Optimal treatment of advanced stages of iNHL is controversial because of low cure rates with the current therapeutic options.

The usual front-line treatment for iNHL consists of the anti-CD20-MAb rituximab given together with an alkylating agent bendamustine (R-B), or a chemotherapy combination containing an alkylating agent [cyclophosphamide, doxorubicin, vincristine, and prednisone (CHOP)]. While the R-B combination is preferred in the US, R-CHOP remains the first option in the EU. In 2nd line, patients with R-sensitive disease (the large majority) would usually receive R-B following R-CHOP, and R-CHOP following R-B.

Rituximab is a CD20-directed cytolytic antibody indicated for the treatment of patients with NHL and widely used as a single drug or in combination regimens [2-6]. Retreatment with rituximab alone or in combination has been shown to be feasible and active [7,8].

For patients with relapsed disease, the same regimen (rituximab in combination with an alternative chemotherapy) can be used. The safety and efficacy of R-B and R-CHOP in this setting were evaluated, showing satisfactory rates of complete response (CR) and overall survival (OS) [9,10]. However, it is known that patients will invariably relapse, and further active and well-tolerated agents are needed to be used in combination with current therapy.

Copanlisib is a small molecule phosphatidylinositol 3-kinase (PI3K) inhibitor and showed excellent anti-tumor activity in pre-clinical models with up-regulated PI3Kα pathway. The PI3K/ Protein kinase B / mammalian target of rapamycin pathway is one of the prominent pathways that promote cellular survival and is constitutively activated in many types of cancers [11,12].

Copanlisib has activity as monotherapy in patients with relapsed or refractory iNHL (shown in the results of studies 12781- and 16349-Part A). Results from the Phase II pivotal study 16349 Part B were compelling enough for the Food and Drug Administration (FDA) to grant copanlisib (Aliqopa) accelerated approval in the treatment of adult patients with relapsed follicular lymphoma (FL) who have received at least two prior therapies. In addition, based on the current metabolism knowledge of each compound there is low potential for drug-drug interactions between copanlisib and the combination partners. It can therefore be expected that copanlisib would increase the activity of standard 1st and 2nd line immunochemotherapy (R-B/R-CHOP) in patients with relapsed, R-sensitive iNHL warranting re-treatment with a rituximab-based regimen. The clinical benefit should be expressed in a prolonged PFS, and symptom control in comparison to the standard immunochemotherapy. In study 17833, the

Safety run-in part for copanlisib in combination with R-B/R-CHOP has been completed. The Data Monitoring Committee (DMC) reviewed the safety data and approved that a copanlisib dose of 60 mg is safe to be used in combination with both R-B and R-CHOP in the Phase III part.

This Statistical Analysis Plan (SAP) is based on the integrated clinical study protocol (CSP, Amendment 09, protocol Version 8.0, dated 21 MAR 2023) and describes the primary analysis of the study 17833 Phase III part to be included in the clinical study report (CSR). The final analysis of the study 17833 Safety run-in part is described in a separate SAP (version 3.0, 22 SEP 2023).

2. Study objectives

The primary objective of the Phase III part is:

- To evaluate whether copanlisib in combination with standard immunochemotherapy, is superior to standard immunochemotherapy in prolonging progression-free survival (PFS) in patients with relapsed iNHL, who have received at least one, but at most three lines of treatment, including rituximab and/or rituximab biosimilars, and/or anti-CD20 monoclonal antibody-based immunochemotherapy and alkylating agents, and for whom the combination of rituximab with either bendamustine (B) or CHOP represents a valid therapeutic option.

The secondary objectives of the Phase III part are to evaluate:

- Other radiological and clinical indicators of treatment efficacy (objective response rate (ORR), duration of response (DOR), complete response rate (CRR), disease control rate (DCR), time to progression (TTP), time to next anti-lymphoma treatment (TTNT), overall survival (OS, 5-year survival rate), time to improvement and the time to deterioration in disease-related symptoms – physical (DRS-P)).
- Safety and tolerability of copanlisib in combination with R-B/R-CHOP.

The further objectives of the Phase III part are to evaluate:

- Pharmacokinetics (PK) of copanlisib
- Biomarkers of efficacy, mode-of-action-related effect, safety, and / or the pathomechanism of the disease.

Table 2–1 gives an overview of the primary and secondary efficacy objectives and the relevant variables to be analyzed.

Table 2–1: Overview of efficacy objectives and variables

Objective	Variable
Primary objective	Progression-free survival (PFS)*, described in Section 6.2.1.
Secondary objective: Response rate	Objective response rate (ORR)*, described in Section 6.2.2.
Secondary objective: Characteristics of disease-related symptoms	Time to deterioration in DRS-P of at least 3.0 points**, Time to improvement in DRS-P of at least 3.0 points**, described in Section 6.2.2
Secondary objective: Other radiological and clinical indicators	Duration of response (DOR), Disease control rate (DCR), Complete response rate (CRR), Time to progression (TTP), Time to next anti-lymphoma treatment (TTNT), Overall survival* (OS, 5-year survival rate), described in Section 6.2.2.
Other efficacy variables	National Comprehensive Cancer Network-Functional Assessment of Cancer, Lymphoma Symptom Index-18 (FLymSI-18) total and subscale scores (DRS-P, DRS-E, TSE, F/WB), described in Section 6.2.4

*Variable is part of the United States as well as Europe and rest of world confirmatory testing strategy (see Section 6.2.3)
**Variable is only part of the Europe and rest of world confirmatory testing strategy (see Section 6.2.3)

3. Study design

This is a randomized double-blind, two-arm Phase III study in patients with relapsed, rituximab-sensitive iNHL to evaluate efficacy and safety of copanlisib in combination with standard immunochemotherapy (R-B or R-CHOP) in comparison to standard immunochemotherapy (R-B or R-CHOP). The patients must have relapsed or progressed after at least one, but at most three, previous lines of therapy; previous treatments must have included rituximab and/or rituximab biosimilars, and/or anti-CD20 monoclonal antibody-based immunochemotherapy and alkylating agents given concomitantly; and patients must not have had a lack of response, or progression within 6 months of the last date of treatment with a rituximab containing regimen.

Patient randomization was to be stratified by:

- Prior to base treatment regimen[1]: R-chemotherapy to R-B, R-B to R-CHOP, and R-B to R-B[2]
- NHL histology: FL histology and other iNHL histology (Small lymphocytic lymphoma [SLL], Marginal zone lymphoma [MZL], lymphoplasmacytoid lymphoma [LPL] / Waldenström macroglobulinemia [WM])
- Duration of progression-free interval on the last rituximab, or rituximab biosimilars, or anti-CD20 monoclonal antibody containing regimen: 6-12 months and >12 months

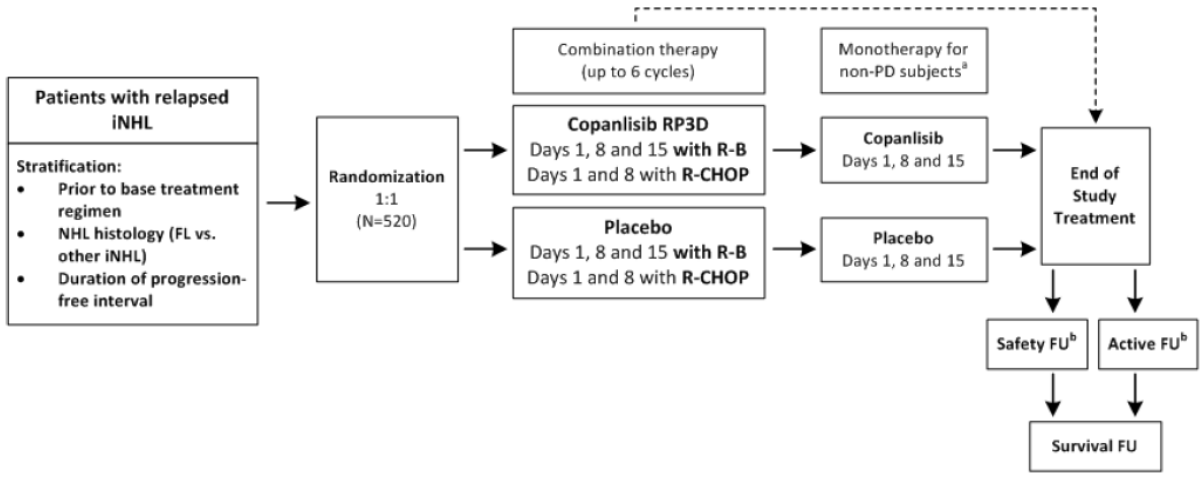
After the copanlisib immunochemotherapy combinations were deemed safe in the Safety run-in part of the trial, approximately 520 (including FL and other iNHL) patients who met the eligibility criteria were to be randomly assigned in a 1:1 ratio to one of the double-blinded treatment arms (approximately 260 patients per arm) in the Phase III part:

Arm 1: copanlisib plus R-B or R-CHOP and

Arm 2: placebo plus R-B or R-CHOP.

- Patients who never received R-B as a previous line of therapy were to be randomized to copanlisib/placebo + R-B treatment group.
- Patients who received R-B as a previous line of therapy were to be randomized to:
  - copanlisib/placebo + R-CHOP treatment group, or
  - copanlisib/placebo + R-B treatment group if ≥ 24 months progression-free interval after the last R-B treatment

Figure 3–1: Overall study design of Phase III part



FL = follicular lymphoma; FU = follow-up; iNHL = indolent non-Hodgkin's lymphoma; N = total number of patients; PD = progressive disease; R-B = rituximab and bendamustine; R-CHOP = rituximab, cyclophosphamide, doxorubicin, vincristine, and prednisone; RP3D = recommended Phase III dose

a: The maximum duration of treatment with copanlisib/placebo is 12 months (including combination therapy and

[1] Where prior treatment regimen refers to treatment received prior to entering the trial, and base treatment regimen refers to the base treatment received on the trial.

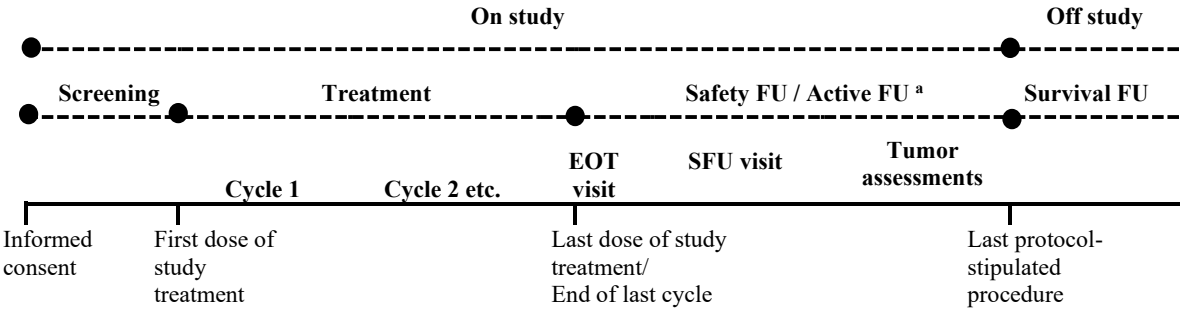
[2] If ≥ 24 months progression-free interval after the last R-B treatment.

monotherapy). Study treatment will be continued until occurrence of PD (per central independent blinded radiological review), clinical progression, unacceptable toxicity, or until another criterion is met for withdrawal from study or up to 12 months whichever comes first.

b: Safety follow-up (FU) for patients who discontinue study treatment due to PD; Active FU for patients who complete 12 months' study treatment without PD or discontinue study treatment for any other reason than PD (includes Safety FU).

An overview of the study periods is presented in Figure 3–2. The study is comprised of the following periods: Screening, Treatment, Safety follow-up, Active follow-up (if applicable) and Survival follow-up.

Figure 3–2: Study periods



EOT = End of treatment; FU = Follow up; SFU = Safety follow-up

a: SFU for patients who discontinue study treatment due to PD; Active FU for patients who complete 12 months' study treatment without PD or discontinue study treatment for any other reason than PD (includes Safety FU).

Screening period

The start of the screening period is defined by signing of the informed consent form. After a screening period of up to 28 days with a minimum of 7 days, patients who met the entry criteria were to be randomized to one of the two arms (only in the Phase III part) and start treatment. The start of the treatment period is defined by the first administration of study treatment.

Re-screening of patients who failed screening was only to be allowed once, following discussion with the sponsor's designated medical representative and after approval by the sponsor. Re-screened patients had to re-sign the informed consent form, even if it had not changed since the patient's previous screening.

Treatment period

The start of the treatment period is defined by the first administration of study treatment. Combination therapy (copanlisib/placebo with R-B or R-CHOP) was to be administered for a maximum of 6 cycles (C1-C6). Copanlisib/placebo (study drug) monotherapy was to be administered from C7 onwards. Tumor assessments with the same modality were to be performed every 12 weeks (± 7 days) from Cycle 1 Day 1.

For a detailed description of study treatment dosage and administration refer to Section 7.4 of the study protocol. An overview is presented in Table 3–1.

Table 3–1: Dosing of copanlisib/placebo during the study

Copanlisib/Placebo (C/P) dosing	Cycle length (d)	D1	D8	D15
Cycles 1-6				
C/P in combination with R-B	28	X	X	X
C/P in combination with R-CHOP	21	X	X	
Cycle 7 onwards				
C/P monotherapy	28	X	X	X

D = day

All patients were to be treated until the occurrence of progressive disease (PD) (per central independent blinded radiology review) as defined by the Lugano Classification [13] (or Owen criteria for patients with LPL/WM [14]) or clinical progression (e.g., Eastern Cooperative Oncology Group [ECOG] performance status of ≥ 3), unacceptable toxicity occurred, until another criterion was met for withdrawal from study or until completion of monotherapy, whichever came first. The maximum duration of treatment with copanlisib/placebo was to be 12 months (including combination therapy and monotherapy).

An end-of -treatment (EOT) visit was to be performed no later than 7 days after the decision was made to discontinue study treatment or when the patient completes 12 months of treatment (maximum duration of therapy).

Safety follow-up period

All patients except for those who died, withdrew consent, or were lost to follow-up were to have a Safety follow-up visit that would take place 30 days (a time window of +5 days is allowed) after the last administration of study treatment.

Active follow-up period

Only patients who completed 12 months of study treatment without PD or discontinuing study drug for reasons other than PD were to enter the Active follow-up period (which also served as a Safety follow-up), except for patients who specifically objected to entering the Active follow-up and patients who objected to all further data collection after study drug discontinuation. Patients in the Active follow-up are to have tumor assessments by central independent blinded review as outlined in the protocol (every 12 weeks [± 7 days] from Cycle 1 Day 1 during Years 1 and 2; every 24 weeks [± 7 days] during Years 3, 4, and 5; and every 24 weeks [±14 days] beyond Year 5, until either PD is centrally evaluated and documented, up to 2 years after primary completion date, end of study or a new anti-tumor treatment is administered, whichever occurs first).

Date of last visit

A patient's last visit date is defined as their last visit/contact prior to any Survival follow-up visits/contact.

Occasionally FLymSI-18 quality of life (QoL) questionnaires were completed by patients after their final radiological assessment in the Active follow-up. In such cases, if the patient ended AFU for reasons other than PD or new anti-lymphoma therapy, and the questionnaire's completion date was prior to entering the Survival follow-up, then the patient's last visit date will be the completion date of the QoL questionnaire.

Survival follow-up period

Except for patients who objected to follow-up data collection in the Survival follow-up period, all patients are to be followed off-study for overall survival at 3-monthly (± 14 days) intervals, until death or until the end of study, whichever occurs first.

End of Study

The study will end when one of the following conditions is met:

- final contact with the last remaining patient has occurred,
- 10 years after the last patient started study treatment.

Efficacy analyses will be performed when approximately 280 centrally evaluated PFS events are observed in the study. Evaluations from central independent blinded review will be used for the primary efficacy analyses of primary and secondary endpoints containing radiological tumor assessments. For patients who withdraw prematurely from the treatment period due to reasons other than PD or death, tumor assessments from the Active follow-up will be included in the primary and secondary analyses. For details of radiological tumor assessment and schedule see Section 9.4.2 of the study protocol.

4. General statistical considerations

4.1 General principles

Statistical analyses will be conducted by or under the supervision of the sponsor's Study Statistician, except for the analysis of biomarker data and pharmacokinetics/pharmacodynamics data, which will be performed by or under the direction of the sponsor's Biomarker Statistical Expert and PK experts.

The statistical evaluation will be performed by using the software package "Statistical Analysis System" (SAS), release 9.2 or higher (SAS Institute Inc., Cary, NC, USA). All variables will be summarized by descriptive statistical methods. The number of data available and missing data, mean, standard deviation, minimum, quartiles, median, and maximum will be calculated for numeric data. Frequency tables will be generated for categorical data.

A statistical evaluation of this study will be performed when approximately 280 PFS events (radiological progressive disease confirmed by independent central review, or biomedical progression in patients with LPL/WM without measurable lesion in the baseline investigator radiological assessment or death if death is before progressive disease) are observed.

Definition of efficacy and safety endpoints, analysis strategies, structure of analysis datasets and layout of analysis data displays are following Bayer standards as documented in the Bayer standard system: Oncology Therapeutic Area Standards, Global Medical Standards and will comply with the Oncology Common Statistical Language document. The given ordering reflects the priority of the different standards, where specifications of the latter standards are to be followed only if not specified in the earlier standards. Study-specific specifications may be included in addition to the project standards, if needed.

The Blind Review Report (BRR) may contain decisions which are relevant for statistical evaluation, for details see Section 4.10.

4.2 Handling of dropouts

A patient who discontinues study participation prematurely (i.e., prior to disease progression confirmed by independent central review or death) for any reason is defined as a “dropout” if the patient has already been randomized, even if no study drug has been taken. Patients who drop out will not be replaced. Refer to Section 6.4 of the study protocol for details regarding the withdrawal of patients from the study.

A patient who signed the informed consent but for any reason terminated the study before randomization is regarded a “screening failure”.

Patients who discontinue study drug due to reasons other than death, PD confirmed by independent central review or biomedical progression in patients with LPL/WM without measurable lesion in the baseline investigator radiological assessment, will enter the Active follow-up period and will not be defined as a “dropout”, except for those who specifically object to entering the Active follow-up or object to all further data collection after study drug discontinuation. The patients in the Active follow-up will have follow-up tumor assessments until disease progression is documented or new anti-tumor treatment is administered, whichever occurs first.

4.3 Handling of missing data

To achieve a well conducted clinical trial in accordance with Good Clinical Practice, every effort will be made to collect all data. However, despite best efforts, it may be inevitable that missing or incomplete data are reported. If a patient indicates they do not want to “continue”, investigators must determine whether this refers to discontinuation of study treatment, unwillingness to attend follow-up visits, unwillingness to either have telephone contact, contact with study personnel, or to allow contact with a third party (e.g., family member, doctor). In all cases, every effort should be made to continue to follow the patient for tumor assessment until patients develop radiological progression by independent central review or biomedical progression in patients with LPL/WM without investigator assessed baseline lesions evaluable by imaging. Survival information including survival status, death date and/or contact date should be determined for all patients during the Survival follow-up. All missing or partial data will be presented in the patient data listing as they are recorded on the electronic Case Report Form (eCRF). Unless specified, missing data will not be carried forward or otherwise imputed in any statistical analysis.

Missing or non-evaluable tumor observations (including a scheduled assessment that was not done and an incomplete assessment that does not result in an unambiguous tumor response or detection of new lesions) will not be imputed or used in the calculation of derived efficacy variables related to tumor assessments, e.g., response. No imputation will be performed for missing tumor assessment and response. For example, if a patient misses a scan visit and PD is documented at the next available scan visit, the actual visit date of the first documented PD will be used to calculate PD-related endpoints. For complete definitions of efficacy parameters refer to Section 6.2.

Partial dates may occur. To impute values, available data from other sources should be utilized. The last known alive date (LKAD) is derived from the main data sources and may be used in statistical analyses to impute death or tumor assessment dates which are partially or completely missing. Except for death (with a date of death on or prior to the data cut-off) data reported during cleaning, after the data cut-off date, will not be considered in the derivation of LKAD. For such cases the data cut-off date will be used as the LKAD. The last available date

across all key data panels (laboratory [LB], tumor response [RS], demography [DM], subject visits [SV], exposure [EX], disposition [DS], procedures [XP], bone marrow assessments [BM], ECG assessments [EG], microbiology [MB], pharmacokinetics [PC], tumor lesion measurements [TR] and tumor/lesion identification[TU]) will be used to determine survival status and the LKAD for each patient. If the day of the LKAD is missing, then the earliest plausible day of the month should be imputed. If both the day and month of the LKAD are missing, then the earliest plausible day of the year should be imputed.

When appropriate, the following rules will be implemented to avoid excluding patients from statistical analyses due to missing or incomplete data:

Date of death

If a death date is partially or completely missing, it will be imputed following the imputation rules below:

1. If there is an adverse event (AE) with "Fatal" outcome, the date of death will be imputed by the stop date of the AE.
2. If there is no AE with outcome "Fatal" and only the day of death is missing, it will be imputed by the first day of the month unless the LKAD is later than this date; in which case the LKAD will be used for imputation. If both day and month and/or year are missing, then the LKAD will be used for imputation.

Date of tumor scan

If a tumor assessment date is partially missing, it will be imputed following the imputation rules below:

1. If only the day is missing, it will be imputed by the first day of the month.
2. If day and month are missing, it will be imputed by the date of the last previous tumor assessment plus 28 days unless the LKAD is earlier than this date; in which case the LKAD will be used for imputation[3] .

Response rates

If a patient has no post-baseline tumor assessment available (or no post-baseline laboratory/clinical tests available for LPL/WM patients without investigator assessed radiologically measurable lesion[s]), i.e., the overall best response assessment is missing, the patient will be non-evaluable (NE), but will be included into denominator for calculation of ORR, DCR and CRR.

Patient-reported outcomes

Physical symptoms of lymphoma are assessed by the FLymSI-18 questionnaire. Missing responses will not be replaced. Missing individual items will be handled in accordance with the scoring instructions for the FLymSI-18 questionnaire (Appendix 9.2). Specifically, if there are missing items, subscale scores are prorated. This is done by multiplying the sum of the subscale by the number of items in the subscale, then dividing by the number of items answered.

[3] All tumor assessments, with complete and partial dates, will be ordered sequentially by visit number. If both day and month are missing from the final tumor assessment and the penultimate tumor assessment is the patients LKAD then the date of the penultimate assessment plus 28 days will be imputed for the final tumor assessment.

When there are missing data, prorating by subscale in this way is acceptable if at least 50% of the items were answered (e.g., a minimum of 5 of 9 items, 2 of 4 items, etc.). The total score is considered to be an acceptable indicator of patient quality of life as long as overall item response rate is greater than 80%. This is not to be confused with individual subscale item response rate, which allows a subscale score to be prorated for missing items if greater than 50% of items are answered. In addition, a total score should only be calculated if all of the component subscales have valid scores.

Time to deterioration in DRS-P of at least 3 points (see Appendix 9.3): Patients without deterioration will be censored at the date of their last tumor evaluation, if the reason for stopping treatment is not related to PD or deaths.

Time to improvement in DRS-P of at least 3 points (see Appendix 9.3): Patients will be censored at the date of their last tumor evaluation, if the reason for stopping the study (i.e., not being in treatment or Active follow-up) is not related to PD. Patients dropping out due to progression-related reasons (e.g., an AE related to PD) or experiencing a PD event (as defined for the primary analysis of PFS) or death due to any reason will be censored at the largest observation time (of events and censoring in all patients evaluated for improvement), plus 1 day.

Missing or partially missing start or end dates for AEs and concomitant medications will be imputed based on the conventions described in Appendix 9.5. The purpose of the imputation for AEs is to determine whether the event is treatment-emergent, as defined in Section 6.3.1. Similarly, the purpose of the imputation for concomitant medications is to determine if the medication was given concomitantly with study drug and, in the case of new anti-lymphoma therapies, to avoid excluding patients with partial anti-lymphoma therapy start dates from statistical analyses due to missing or incomplete data.

4.4 Handling of differing cycle lengths

Due to the differing cycle lengths of copanlisib/placebo+R-B compared to copanlisib/placebo+R-CHOP, by visit summaries will only be tabulated by background therapy if 5 or more patients are randomized to receive the background therapy. If a background therapy has less than 5 patients assigned the equivalent data will be listed in Section 10 of the CSR. In Section 10 of the CSR patients assigned to the non-tabulated background therapy will be listed first.

4.5 Handling of investigator assessments when PD is not confirmed by independent central review

The analysis of efficacy, based on investigator assessments, will exclude response assessments collected after the first investigator reported PD, regardless of whether or not the PD was confirmed by independent central review.

4.6 Handling of clinical laboratory values with qualifiers

Clinical laboratory values which have qualifiers (for example, > or <) will be included in clinical safety analyses and summaries with the qualifier removed.

4.7 Multiple testing

Except for the analysis of the primary efficacy endpoint and confirmatory statistical testing strategy, all other analyses will be descriptive. Any resulting p-values from these descriptive analyses will be nominal as there will be no multiplicity adjustments for multiple testing.

4.8 Interim analyses and data monitoring

No interim analyses are planned for this study's primary endpoint of PFS. The primary PFS analysis will be performed when approximately 280 centrally evaluated PFS events are observed in the study. Further details on the primary PFS analysis are included in Section 6.2.1 and details on the confirmatory statistical testing strategy are provided in Section 6.2.3.

The number of PFS events will be monitored in the blinded database throughout the study. When the required numbers of events are reached, the database will be cut and cleaned for analysis. Data reported during cleaning after the data cut-off will not be considered in the primary analysis except for any death reported during data cleaning with a date of death on or prior to the data cut-off.

An interim analysis of OS will also be performed at the time of the primary PFS analysis. The survival status will further be collected quarterly during the Survival follow-up period up to 10 years after the last patient started study treatment, at which point the final descriptive analysis of OS at study end will take place.

Twelve months after the primary analysis of the primary and secondary efficacy variables, and at yearly intervals thereafter until the final descriptive analysis of OS at study end, further descriptive analyses of OS (each a "follow-up OS analysis") will be performed. These analyses will be updates of the analysis of OS performed on data available at the time of analysis of the primary endpoint, PFS.

A DMC was established to ensure that patients were not exposed to undue risk during this study and regularly reviewed study data, as outlined in the DMC charter [15], to provide an independent recommendation on the advisability of continuing this study. Investigators, patients, and the sponsor remain blinded to treatment.

Until all patients discontinued study treatment, the report for the DMC included tables, listings, and figures, and was generated by an independent statistician from a Statistical Analysis Center. The format and content of these data summaries have been specified separately from this study SAP.

Decisions on trial termination, amendment, or cessation of patient recruitment based on risk/benefit assessment, were made after recommendations from the DMC were assessed by the sponsor.

4.9 Data Rules

4.9.1 Time intervals

If time intervals are to be displayed other than days in statistical evaluations, then one year is considered to have 365.25 days (average length of a year, including leap years), one month is considered to have 30.44 days (average length of a month, including leap years), one week is considered to have 7 days, and one cycle is considered to have 28 days (i.e., 4 weeks). In terms of data display, this generally also applies for patients treated with CHOP as base regimen. All time conversions from days will be displayed to 1 decimal place.

4.9.2 Baseline

Baseline is defined as the last measurements performed prior to the first study drug administration in Cycle 1. If the actual time is not available but date is available for certain measurements, the baseline value is defined as the last non-missing value collected on or before the date of the first dose of study treatment. For patients who have been randomized but not treated with any dose, randomization date will be used as the reference date for baseline value calculation.

Also consider:

- The time part for baseline flagging. Study data tabulation model/analysis data model datasets for exposure and laboratory data both capture times. If the time in either one is missing, use the dates in that case.
- Baseline can be either a scheduled or unscheduled visit.
- If the patient has a measurement on Cycle 1 Day 1 (scheduled), then this measurement will be considered as the baseline.
- If the patient has no measurements on Cycle 1 Day 1 but has a measurement at screening visit (scheduled or unscheduled), then the screening visit measurement will be considered as the baseline.
- If multiple measurements were taken at the same Cycle 1 Day 1 visit or same screening visit and all with non-missing assessment dates, then the measurement with the latest assessment date/time will be considered as baseline, whether scheduled or unscheduled.
- If both scheduled and unscheduled measurements exist for the same visit and cannot be decided by the above rules, then the measurement taken at scheduled visit will be considered as the baseline.

4.9.3 Repeated measures

If there are repeated measurements per time point (e.g., laboratory values, vital signs, etc.), the following rules will be used (unless otherwise specified):

- Before the start of the study drug administration (i.e., for screening and baseline value), the latest non-missing measurement at scheduled visits will be used. Unscheduled visits will be used if there are no measurements at scheduled visits. If the latter is the case, the last unscheduled visit will be used.
- In case of repeated measurements at any post baseline time point, the first measurement at scheduled visits will be used.
- When two records have the same patient identifier (ID), results, date, time, and visit/cycle number, except for page number or record ID, the two records will be considered duplicates and only one record will appear in the corresponding analysis dataset.

4.9.4 Overall extent of exposure

As a rule, and in accordance with the Oncology Therapeutic Area Standard, leading “0 mg” (prior to the first positive amount of drug) and trailing “0 mg” records (not followed by any positive amount of drug), will not be included in the calculation of any drug duration or amount. Similarly, trailing “drug interruptions” will not be used in statistical tables. A

footnote will be included, stating that "Interruption becoming permanent study treatment discontinuation before resumption of study treatment is not counted as an interruption".

Overall extent of exposure (OEE) for copanlisib/placebo, is defined as the time from first respective study drug intake (day_{first}) until last study drug intake (day_{last}), including 7 additional days to account for the weekly dosing regimen, and is calculated as:

$OEE = day_{last} - day_{first} + 7$

OEE for R-CHOP, is defined as the time from first respective study drug intake (day_{first}) until last study drug intake (day_{last}), including 21 additional days to account for the "every 3 weeks" (Q3W) dosing regimen and prednisone/prednisolone dosing up to Day 6, and is calculated as:

$OEE = day_{last} - day_{first} + 21 - [day_{(last)} - day_{(first)}$ of last cycle of R-CHOP]

OEE for R-B, is defined as the time from first respective study drug intake (day_{first}) until last study drug intake (day_{last}), including 28 additional days to consider the Q4W dosing regimen and B dosing up to Day 2, and is calculated as:

$OEE = day_{last} - day_{first} + 28 - [day_{(last)} - day_{(first)}$ of last cycle of R-B]

If the respective treatment ends with dose interruptions, the day of the last actual dose will be assumed to be day_{last}.

4.9.5 FLymSI-18 questionnaire

FLymSI-18 QoL questionnaires are patient-administered and should be completed by the patient as scheduled in the CSP, at the start of the study visit prior to any contact with the investigator or site personnel. If the completion date of a questionnaire is not within ± 7 days of the corresponding scheduled visit, it will be excluded from the analysis of efficacy. If a patient withdraws consent for radiological assessments but consents to completing QoL questionnaires, the questionnaires will be included in the analysis of efficacy if they were completed within ± 7 days of the scheduled visit and were not completed during the patient's Survival follow-up.

4.9.6 Stratification

In case of discrepancies between stratification factors entered in the interactive voice/web response system (IxRS) and information entered in eCRF, the information from IxRS will be used for analysis. This especially includes analyses that are performed separately for the subgroup of FL patients, which is based on one of the stratification factors.

The stratification that was assigned in IxRS will be reflected in the patient's ID number. The mapping of patient ID/randomization numbers to assigned IxRS strata is summarized in Table 4–1.

Table 4–1: Stratum per patient ID/randomization number

Randomization Start Number	Description of strata[1,2]
PPD – PPD	01: R-chemotherapy to R-B, FL, treatment-free / progression-free interval on the last rituximab, or rituximab biosimilars, or anti-CD20 monoclonal antibody containing regimen 6-12 months
PPD – PPD	02: R-chemotherapy to R-B, FL, treatment-free / progression-free interval on the last rituximab, or rituximab biosimilars, or anti-CD20 monoclonal

Randomization Start Number	Description of strata[1,2]
	antibody containing regimen >12 months
PPD PPD	03: R-chemotherapy to R-B, other iNHL, treatment-free / progression-free interval on the last rituximab, or rituximab biosimilars, or anti-CD20 monoclonal antibody containing regimen 6-12 months
PPD – PPD	04: R-chemotherapy to R-B; other iNHL, treatment-free / progression-free interval >12 months
PPD – PPD	05: R-B to R-CHOP, FL, treatment-free / progression-free interval 6-12 months
PPD – PPD	06: R-B to R-CHOP, FL, treatment-free / progression-free interval >12 months
PPD – PPD	07: R-B to R-CHOP, other iNHL, treatment-free / progression-free interval =6-12 months
PPD – PPD	08: R-B to R-CHOP, other iNHL, treatment-free / progression-free interval > 12 months
PPD – PPD	09: R-B to R-B, FL, treatment-free / progression-free interval 6-12 months
PPD – PPD	10: R-B to R-B, FL, treatment-free / progression-free interval >12 months
PPD – PPD	11: R-B to R-B, other iNHL, treatment-free / progression-free interval 6-12 months
PPD – PPD	12: R-B to R-B, other iNHL, treatment-free / progression-free interval >12 months

iNHL: indolent non-Hodgkin's Lymphoma; FL: Follicular Lymphoma; R = Rituximab; R-B = Rituximab and bendamustine; R-CHOP = Rituximab, cyclophosphamide, doxorubicin, vincristine, and prednisone

treatment-free / progression-free interval: duration of treatment-free / progression-free interval on the last rituximab, or rituximab biosimilars, or anti-CD20 monoclonal antibody containing regimen.

[1]Stratification factor "duration of treatment-free interval (6-12 months vs. >12 months)" was amended with CSP Amendment 06, Version 6.0, dated 25 SEP 2018 to "duration of progression-free interval on the last rituximab containing regimen (6-12 months vs. >12 months)

[2]Stratification category "other iNHL" includes SLL, MZL and LPL/WM.

This study is using 12 strata (defined by 3 stratification variables of which 2 are binary stratification variables) with 524 patients randomized in a 1:1 ratio.

4.9.7 Region

For demographics overview and subgroup by region efficacy analyses, patients will be grouped together, summarized, and analyzed for two separately defined groupings of geographic regions: Geographic Regions 1 and Geographic Regions 2. These groupings are defined as follows:

Geographic Regions 1:

- US

- Europe (Belgium, Bulgaria, Czech Republic, Denmark, Finland, France, Germany, Greece, Hungary, Ireland, Israel, Italy, Poland, Portugal, Romania, Slovakia, Spain, Ukraine, and United Kingdom)
- Rest of the world (All other countries/regions)

Geographic Regions 2:

- North America (Canada, Mexico, and United States)
- Asia Pacific, excluding Australia, (China, Japan, Korea, Taiwan, Hong Kong, Singapore, and Thailand),
- Rest of the world (All other countries/regions)

4.9.8 As treated

An IxRS will be used for drug distribution. Despite the obligatory use of the IxRS, patients may erroneously receive the wrong study treatment (i.e., copanlisib instead of placebo or vice versa). For the "as treated" safety analyses, using a conservative approach, patients that have received at least one dose of copanlisib will be considered for the copanlisib treatment group. Patients who have exclusively received placebo (and R-B or R-CHOP) and no copanlisib will be considered for the placebo treatment group.

4.10 Validity Review

During study conduct, important protocol deviations and validity findings affecting the assignment of patients to analysis sets will be identified through ongoing data reviews. A Blind Review Meeting (BRM) will be scheduled, just prior to unblinding and data release, in which the study team will review and finalize the list of important deviations and validity findings. Patient validity and details relevant to the statistical evaluation will be discussed and agreed upon during the BRM and documented in the BRR. Validity findings which will affect the assignment of a patient's data to analysis sets, and important deviations which may significantly affect patient safety, patient rights or may affect study data completeness, accuracy, and/or reliability, will be listed in the BRR and Protocol Deviation Document. Any changes to the statistical analysis prompted by the results of the BRM will be documented in an amendment and, if applicable, in a supplement to this SAP.

5. Analysis Sets

5.1 Assignment of analysis sets

Final decisions regarding the assignment of patients to analysis sets will be made during the BRM based on the validity findings in the final list of important deviations and other findings and documented in the BRR (see Section 4.10).

Patients who signed the informed consent but for any reason terminate the study before randomization will be considered screening failures. These patients will be listed separately.

Full analysis set (FAS)

The primary population for all efficacy analyses is the FAS (considering the Intent-to-treat [ITT] principle) population, which is defined as all randomized patients in the Phase III part. The FAS will be used for the display of efficacy variables and for the display of demographics, baseline characteristics and exposure. Following the ITT principle, patients

will be analyzed as randomized, meaning, even if a patient was randomized and received no study drug or if randomized and received incorrect drug, these patients will still be analyzed for efficacy under FAS, as randomized.

Safety analysis set (SAF)

All randomized patients in the Phase III part who received at least one dose of study drug (copanlisib, placebo, R-B or R-CHOP). The SAF will be used for the analyses of safety variables. Patients will be analyzed as treated.

In case the SAF and FAS differ, depending on the degree of difference it will be considered whether displays for baseline characteristics, demographics and exposure will be repeated in the SAF.

Pharmacokinetic analysis set (PKS)

All patients in the Phase III part with at least one valid PK measurement (i.e., patients who took at least one dose of study drug and had at least one PK sample collected and measured) will be included in the evaluation of PK concentration and parameters.

6. Statistical Methodology

6.1 Population characteristics

Population characteristics will be summarized overall and by randomized treatment group (copanlisib in combination with R-B or R-CHOP versus placebo in combination with R-B or R-CHOP). Unless otherwise specified, analyses will be performed in the FAS. In case the FAS and SAF differ, depending on the degree of difference it will be considered whether displays for baseline characteristics, demographics and exposure will be repeated in the SAF. Patients will be analyzed as randomized for the FAS and as treated for the SAF.

6.1.1 Disposition

The number of countries/regions and centers enrolling patients will be presented and will include the number of centers enrolling patients by country/region and pooled country/region. The number of patients enrolled, randomized, and valid for the different analysis sets FAS, SAF and PKS will be summarized overall, by treatment groups, by country/region and by pooled countries/regions. Individual countries/regions will be pooled into regions and these regions will be pooled into geographic regions as defined in Section 4.9.7.

The number of screening failures (patients discontinuing the screening period) together with primary reasons for discontinuation will be presented overall. Re-screened patients will not be counted as screening failures unless they failed re-screening. A listing of re-screened patients, which will include all patient IDs associated with the patient, all dates of informed consent, screening visits and last visits, and reasons for failing initial screening, will be provided. If patients are screen failures, the failed inclusion/exclusion criteria will also be included. “Other” reasons for premature discontinuation from screening will be listed by treatment group and patient ID.

The number of patients assigned to treatment, in addition to the number of patients discontinuing the treatment, Safety follow-up, Active follow-up and Survival follow-up periods, together with the primary reason for discontinuation will be presented by treatment group and overall. Discontinuations due to death after the data cut-off date, reported during

data cleaning, will be included in the disposition summaries. “Other” reasons for premature discontinuations from these study epochs will be listed by treatment group and patient ID.

Time to end of study treatment and time to cause specific treatment discontinuation are calculated as time from first study drug intake to the last day of study drug intake (i.e., the last non 0 mg dose). A Kaplan-Meier plot for “Time to end of study treatment” will be provided by treatment group. Cumulative incidence function curves will be provided for “Time to cause specific treatment discontinuation” by treatment group and reason for discontinuing treatment[4].

For time to end of study treatment all patients permanently discontinuing treatment, including those that complete 12 months of treatment per protocol, will be treated as events.

For time to cause specific treatment discontinuation, patients who complete 12 months of treatment per protocol will be considered as events, while patients that permanently discontinue treatment prior to 12 months will be considered as competing risk events.

Time to treatment discontinuation due to AEs not associated with clinical disease progression will additionally be reported separately by AE severity[5] and treatment group, using cumulative incidence function curves. Only patients who discontinue treatment permanently due to AEs not associated with clinical disease progression will be included in this analysis.

Following the inverse event indicator method [16], PFS and OS events censor the true but unknown observation time of an individual and censoring in these cases are classified as events. The unobservable follow-up time of a patient that has progressed or died is interpreted as the follow-up time that potentially would have been obtained had that patient not progressed or died.

PFS follow-up time and rate, as assessed by investigator and independent central review, will be determined by treatment group using the inverse event indicator method. Patients who have progressed or died (where no progression is documented) during treatment or active follow-up, will be classified as censored and follow-up time for PFS will be calculated as time from date of randomization until date of progression or date of death. Otherwise, patients will be classified as having had events and follow-up time for PFS will be calculated as time from date of randomization until date of last evaluable tumor assessment (or last biochemical assessment for patients with LPL/WM).

OS follow-up time and rate will be determined by treatment group using the inverse event indicator method. Follow-up time for OS will be calculated as time from date of randomization until date of death or LKAD (see Section 4.3) if there is no date of death. Patients who have died will be classified as censored, and patients who have not died will be classified as events.

The frequency and percentage of patients:

- Not completing follow-up on the primary endpoint PFS:

[4] Reasons for treatment discontinuation will be grouped as follows: completed 12 months of treatment, progressed or died during treatment period (including clinical progression and AEs related to clinical progression), discontinued treatment due to AEs (not related to progression, including additional primary malignancies), discontinued treatment due to patient or physician decision, lost to follow-up (including logistical reason), or discontinued treatment due to other reasons (including protocol violations and non-compliance with study drug).
[5] AE severity will be grouped as follows: CTCAE grade 1 or 2, CTCAE grade 3 or 4 and CTCAE grade 5.

  - censored for PFS prior to primary cut-off date.
- Completing follow-up on the primary endpoint PFS until either:
  - having a PFS event
  - still ongoing in Active follow-up without a PFS event at the data cut-off date.

will be tabulated by treatment group and overall, for both blinded independent central review and investigator assessments of PFS.

A similar table will summarize the frequency and percentage of patients:

- Not completing follow-up on the secondary efficacy/safety endpoint OS:
  - censored for OS prior to primary cut-off date.
- Completing follow-up on the secondary efficacy/safety endpoint OS:
  - having an OS event (i.e., death)
  - still ongoing in either Active or Survival follow-up without an OS event at the data cut-off date.

In addition, the number of randomized patients with important protocol deviations will be presented overall and by country/region for treatment group. The frequency of important deviations, validity findings and number of patients valid for each analysis set will be presented by treatment group and overall.

A separate patient listing of protocol deviations related to the coronavirus disease 2019 (COVID-19) pandemic (e.g., delay of tumor assessments, delay, interruption, or discontinuation of the treatment due to the impact of COVID-19 pandemic) will be provided. The number of patients with protocol deviations associated with COVID-19 pandemic will be summarized overall and by protocol deviation category (important vs. other). The number of patients affected by COVID-19 pandemic related study disruptions (protocol deviations, premature discontinuations, and AEs) will also be summarized in a separate table. These protocol deviations will be entered in the Rave Electronic Data Capture (EDC) system as Manual protocol deviations and identified from the specification of "COVID".

Similarly, a separate patient listing of protocol deviations related to the Ukraine-Russia conflict (e.g., delay of tumor assessments, delay, interruption, or discontinuation of the treatment due to the impact of the conflict) will be provided. These protocol deviations will be entered in the Rave EDC system as Manual protocol deviations and identified from the specification of "Regional Crisis".

6.1.2 Demographics and other baseline characteristics

Demographic variables and other baseline characteristics will be summarized by treatment group and overall. In addition, summary statistics will be presented for continuous variables. Frequency tables will be presented for categorical variables.

These will include, but may not be limited to, demographic variables (age, sex, race, ethnicity, body mass index (BMI), systolic and diastolic blood pressure, height, weight, heart rate, temperature, New York Heart Association (NYHA) classification, estimated Glomerular Filtration Rate (eGFR), glycated hemoglobin (HbA1c), medical history of diabetes and medical history of hypertension). Age, BMI, eGFR and HbA1c will be analyzed as continuous variables and in addition categorized with the following categories:

- Age group 1 (years): <65, ≥65
- Age group 2 (years): <80, ≥80
- Age group 3 (years): <65, 65-74, 75-84, ≥85
- BMI group (kg/m²): <18.5, ≥18.5- <30, ≥30
- eGFR group (mL/min/1.73 m^2): Normal: ≥ 90 vs. Mild impairment: 60 - <90 vs. Moderate impairment: 30 - <60 vs. Severe impairment: 15 - <30)
- HbA1c group at baseline: <5.7% vs ≥5.7% – <6.5% vs ≥6.5% – ≤8.5% vs >8.5%

Medical history of diabetes and hypertension, identified by Product specific Bayer MedDRA Queries (PBMQ), will be summarized as follows:

- Diabetic History: [PBMQ] Medical history of diabetes (no/yes)
- Hypertension History: [PBMQ] Medical history of hypertension (no/yes)

In addition, the three stratification factors as reported on the randomization page in the eCRF will be summarized.

The following additional baseline characteristics will be summarized:

- Histology of tumor per investigator's assessment (FL, MZL overall and including subtype of MZL, SLL, and LPL/WM)
- Staging and size (sum of products of diameters) of tumor at study entry
- FL grade (for patients with FL) at study entry
- ECOG performance status at baseline
- Time from date of initial diagnosis to the date of randomization (months)
- Time from most recent progression to the date of randomization (weeks)
- Time from first progression to the date of randomization (months)
- Number of lesions (number of target and number of non-target lesions) at baseline
- Number of LPL/WM patients with and without measurable disease at baseline
- Serum Immunoglobulin M (IgM) level (for LPL/WM patients) at baseline
- Reason for treatment allocation
    - Prior to base treatment regimen
        - R-chemotherapy to R-B
        - R-B to R-CHOP
        - R-B to R-B
    - Duration of treatment-free/progression-free interval
        - 6-12 months
        - >12 months

In addition, demographic and other baseline characteristics will be summarized for each histology (FL, MZL, SLL, and LPL/WM) in the FAS.

The retrospective evaluation of histopathological diagnosis at baseline will be performed centrally. These data are exploratory and will be listed for the CSR appendix.

6.1.3 Medical history

Medical history findings (as defined in protocol Section 9.3.2) will be coded by Medical Dictionary for Regulatory Activities (MedDRA) codes, v25.1 or higher. Medical history will be presented for each MedDRA Primary System Organ Class (SOC) and Preferred Term (PT) by treatment group and overall. Ongoing conditions will be presented in a similar way but will also include National Cancer Institute (NCI) Common Terminology Criteria Adverse Event (CTCAE) grade at baseline. Version 4.03 (or higher) of the NCI-CTCAE dictionary will be used.

6.1.4 Prior and concomitant medication

Prior and concomitant medications will be coded by the Anatomical Therapeutic Chemical (ATC) classification system according to the World Health Organization Drug Dictionary (WHO-DD). Summaries will be provided by ATC class and subclass. The number of patients taking at least one prior medication will be reported in addition to the number of patients who have taken medication relating to each relevant ATC class and subclass. The same will be reported for concomitant medications. Note that the same medication can appear multiple times in the table as it can have several ATC codes.

A medication that has been stopped after first administration of study treatment is considered as concomitant, while those that have been stopped before first administration are considered as prior. Medications with missing start and stop date but flagged as being ongoing at end of study will be considered to have started prior to start of study treatment and ended after stop of study treatment. Medications starting more than 30 days after the last day of study treatment will be excluded from the summary of concomitant medications.

Prior and concurrent anti-cancer therapy

The mean, minimum, median, standard deviation, and maximum number of prior systemic anti-cancer therapy lines as well as number of patients with 1, 2, ≥3 lines of prior therapy will be summarized by treatment group and overall, for FAS. Time since last systemic anti-cancer therapy[6] will be summarized using descriptive statistics by treatment group and overall, for FAS. The time between the last day of the last course of systemic anti-cancer therapy and the day of confirmation of the most recent progression will also be displayed as a classification table, showing proportions of patients with ≤ 6 months vs. > 6 to <12 months vs. ≥ 12 months, by treatment group.

In addition, the following frequency tables will be provided by treatment group and overall, for the FAS:

- Prior anti-cancer therapy and therapeutic procedures
- Prior radiotherapies
- Prior systemic anti-cancer therapies (displayed by 1. ATC classes, 2. Drug names)

[6] Time between last day of last systemic anti-cancer therapy and randomization date.

- Type of prior systemic anti-cancer therapy
- Prior treatment with PI3K inhibitors (refer to the Tables, Listings and Figures (TLF) specification document for details of how prior PI3K inhibitors are identified)
- Prior diagnostic and/or therapeutic procedures for iNHL
- Concurrent diagnostic and/or therapeutic procedures
- Radiotherapies during Survival follow-up
- Systemic anti-cancer therapies during Survival follow-up (displayed by 1. ATC classes, 2. Drug names)
- Type of systemic anti-cancer therapy during Survival follow-up

6.1.5 Treatment duration and exposure

Investigator-recorded cycles for study drug exposure will be mapped to a calculated cycle. The analysis of treatment duration and exposure will use the mapped, calculated cycle. A description of how investigator defined cycles are mapped to analysis cycles can be found in the study document "Analysis Data Reviewer's Guide (ADRG)".

Descriptive statistical summaries will be provided separately for copanlisib/placebo, R-B and R-CHOP by treatment group for the following variables:

- Overall extent of exposure (OEE, as defined in section 4.9.4) for FL, MZL, SLL, LPL/WM and all SAF patients
- Overall extent of exposure, for FL, MZL, SLL, LPL/WM and all SAF patients, by categories:
    - 0-90 days
    - 91-180 days
    - 181-270 days
    - 271-365 days
    - >365 days
- Number of cycles for FL, MZL, SLL, LPL/WM and all SAF patients (statistical summaries in addition to number of patients with each number of cycles), determined by patients' last calculated cycle.
- Number of infusions during treatment period for FL, MZL, SLL, LPL/WM and all SAF patients. For patients in the R-CHOP combination arm, the number of infusions will be based on rituximab, cyclophosphamide, doxorubicin, and vincristine infusions as prednisone is taken orally.

In addition, the following analyses will be provided over all cycles and by calculated cycle for FL, MZL, SLL, LPL/WM and all SAF patients:

- Total amount of dose actually administered: sum of actual dose (mg) per calculated cycle for each patient
  - For copanlisib/placebo, prednisone/prednisolone, the actual dose per timepoint is defined as

    Prescribed dose [mg] × (Total amount of dose administered [mg OR mL] / Total amount of dose prior to administration [mg OR mL])
  - For rituximab, cyclophosphamide, doxorubicin, vincristine and B, the actual dose per timepoint is defined as

    Prescribed dose [mg] / Body surface Area (BSA) [m^2] × (Total amount of dose administered [mL] / Total amount of dose prior to administration [mL])
- Percent of planned dose received = Actual dose [mg] / Planned dose [mg] × 100%.
  - For copanlisib/placebo, a standard of 60 mg is set for all cycles, i.e., infusion days (D1, D8 and D15 of each cycle), and for planned dose (might however be modified by individual dose-reductions or interruptions). Completed cycles therefore have a planned dose of 180 mg. For incomplete cycles, the planned dose depends on the number of days (d) in that calculated cycle:
    - 60 mg if $0< d <7$
    - 120 mg if $7\leq d <14$
    - 180 mg if $14\leq d \leq 28$
  - For rituximab, a standard of 375 mg/m² body surface is set for all cycles for planned dose.
  - For cyclophosphamide, a standard of 750 mg/m² body surface is set for all cycles for planned dose.
  - For doxorubicin, a standard of 50 mg/m² body surface is set for all cycles for planned dose.
  - For vincristine, a standard of 1.4 mg/m² body surface is set for all cycles for planned dose and a maximum of 2.0 mg/m² is also previewed.
  - For B, a standard of 90 mg/m² body surface is set for all cycles for planned dose (might however be modified by individual dose-reductions or interruptions), i.e., infusion days (D1 and D2 of each cycle). For incomplete cycles, the planned dose depends on the number of days (d) in that calculated cycle:
    - 90 [mg/ m^2] ×BSA [m^2] if $0< d <2$
    - 90 [mg/ m^2] ×BSA [m^2] ×2 if $2\leq d \leq 28$
  - For prednisone/prednisolone, a standard of 100 mg is set for all cycles, i.e., oral days (D2 to D6 of each cycle), and for planned dose (might however be modified by individual dose-reductions or interruptions). Completed cycles therefore have a planned dose of 500 mg. For incomplete cycles, the planned dose depends on the number of days (d) in that calculated cycle:
    - 0 mg if $0< d <2$
    - 100 mg if $2\leq d <7$ and increase 100 mg per day

- 500 mg if 7≤ d ≤21

For patients in the SAF population with dose reduction and re-escalation (only for copanlisib/placebo), interruption (dose administered is non-zero), or delay (zero dose is reported), the number of dose reductions, interruptions, or delays per patient and their reasons will be summarized separately for copanlisib/placebo, rituximab[7], cyclophosphamide, doxorubicin, vincristine, prednisone/prednisolone and B. Interruptions with the same start date will be counted as one interruption.

The duration of dose delays in days will be summarized for copanlisib/placebo. The duration is defined as the difference between the date of delay and the date of next exposure information (infusion, interruption, or delay). Due to the treatment regimen with one week break at the end of each cycle, delays on day 15 of R-CHOP combination therapy, or on day 22 of R-B combination therapy, are capped at 7 days, in case the calculated duration is greater than 7.

Patient listings will be provided for dose modifications.

6.2 Efficacy

Descriptive evaluations of all efficacy variables will be performed. Summaries of tumor response by independent central review and investigator's assessment will be provided.

Disease progression in the context of statistical efficacy evaluation is radiological progression, as assessed by independent central review. For LPL/WM patients without investigator assessed radiologically measurable disease at baseline, progression per Owen criteria (assessed locally) will be used.

Efficacy analyses will be performed when approximately 280 centrally evaluated PFS events (PD by independent central review, progression per Owen criteria or death from any cause before PD) are observed in the study, see Section 6.6.

Evaluations from independent central blinded review (defined by Lugano Classification [13]), and Owen criteria (assessed locally) for LPL/WM patients, will be used for the efficacy analyses of primary and secondary endpoints containing radiological tumor assessments.

Efficacy analyses will be performed using the FAS, i.e., patients will be analyzed as randomized, as described in Section 5.1. For time-to-event analyses the censoring mechanism is assumed to be non-informative. Patients will be handled as right-censored in time-to-event analyses, if applicable. Censoring of efficacy variables will be handled according to Sections 6.2.1, 6.2.2, 6.2.1.4.5, 6.2.1.4.6 and 6.2.1.4.7. Missing or non-evaluable tumor observations and missing overall best response assessment will be managed as described in Section 4.3.

6.2.1 Primary efficacy variable

Definition of progression-free survival (PFS)

The primary efficacy variable is PFS, defined as the time (in days) from randomization to PD or death from any cause (if no progression is documented), whichever occurs earlier. The primary PFS analysis will be based on independent central review. Radiological progression will be confirmed in real-time by independent central review. Biochemical progression in patients with LPL/WM without measurable lesions evaluable by imaging at baseline according to investigator assessment will be assessed locally, and the confirmation of PD by

[7] Per the CSP, there will be no dose reductions for rituximab in any cycle.

independent central review is not needed[8]. However, because on progression these patients may have developed measurable disease by imaging without simultaneous increase in IgM, the imaging material will be submitted for independent central review and PD confirmation. LPL/WM patients with a radiologically measurable lesion at Screening will continue to have radiological assessments and, in addition, have laboratory tests performed on the same days. PFS for patients without PD or death at the time of analysis will be censored at the date of last evaluable tumor assessment or last biochemical assessment for patients with LPL/WM without measurable lesions evaluable by imaging.

For ease of reading, tumor assessment in the following context refers to either 1) tumor imaging assessment or 2) biochemical assessment for patients with LPL/WM without measurable lesions evaluable by imaging.

The actual date of tumor assessments or date of death will be used for this calculation. If no assessment was done at all or the assessment was incomplete at a particular time point, in general, the patient is non-evaluable (NE) for this time point. If the examination was incomplete and only a subset of measurements could be made but fulfill the criteria for progression and/or new lesions are detected, this patient would be considered evaluable with assessment of PD. If a tumor assessment is performed over more than one day (e.g., scans for chest and abdomen done on different days for a same evaluation) and the response is progression (PD), the earliest date /occurrence of the observed progression (i.e., as an event) will be used for the calculation of the PFS.

PFS for patients who have neither tumor assessments nor death after baseline will be censored at Day 1. PFS for patients without baseline tumor assessments (e.g., non-existence of a measurable lesion at baseline confirmed by independent central review) will be censored at Day 1.

If a tumor assessment date is partially missing, it will be imputed following the imputation rules detailed in Section 4.3.

Table 6–1 lists the censoring rules for PFS events.

Table 6–1: Progression-free survival (PFS) censoring rules

Situation	End Date/Day	Censored	Reason for Censoring	Detailed Reference
No baseline tumor assessment and no death, or death more than 25 (12+12+1) weeks after randomization	Reference Date/Day 1	Yes	No baseline tumor assessment	
No baseline tumor assessment and death within first 25 (12+12+1) weeks after randomization	Date of death	No	N/A	
No post-baseline tumor assessment and no death	Reference Date/Day 1	Yes	No post-baseline tumor assessment and no death	

[8] If independent and local assessments are both available for these patients, both will be used to determine efficacy endpoints. The patient's last available tumour assessment in this case will be their last assessment, regardless of whether this was assessed locally or by independent central review.

Table 6–1: Progression-free survival (PFS) censoring rules

Situation	End Date/Day	Censored	Reason for Censoring	Detailed Reference
No post-baseline tumor assessment and died after the first 25 (12+12+1) weeks after randomization	Reference Date/Day 1	Yes	Death happened after time window specified in Appendix 9.4.2	Appendix 9.4.2
Death after randomization during treatment, Active follow-up period or early enough in Survival follow-up to be within the time window for a PFS event	Date of death	No	N/A	Appendix 9.4.2
Death too late in the Survival follow-up to be within the time window for a PFS event	Last Tumor assessment date	Yes	Death happened after time window specified in Appendix 9.4.2	Appendix 9.4.2
Death occurs after two or more consecutive missed tumor assessments	Last tumor assessment date	Yes	Two or more consecutive missed tumor assessments immediately prior to death	
New anti-cancer treatment other than the study treatment prior to observing progression[1,2]	Last evaluable tumor assessment date prior to the initiation of anti-cancer treatment	Yes	New anti-cancer treatment	Appendix 9.4.3
The progression occurs at the next tumor assessment after two or more consecutive missing or non-evaluable assessments	Last tumor assessment date before the missing tumor assessments	Yes	Two or more consecutive missed tumor assessments immediately prior to progression	Appendix 9.4.1
Post-baseline tumor assessments but discontinued or withdrew early from study without documented disease progression	Last tumor assessment date	Yes	Discontinued treatment or Active follow-up. Given Active follow-up may be continued despite premature EOT, patients will be censored at last tumor assessment date (which might be in Active follow-up).	Appendix 9.4.4
Progression without two or more consecutive missed tumor assessments immediately prior to progression	Date of progression	No	N/A	
Neither PD nor death and no other criteria fulfilled	Last tumor assessment date	Yes	Patient continuing in treatment period or Active follow-up in the absence of experiencing a PFS event	

[1]Or surgery affecting target/non-target lesions prior to observing progression.
[2]If progression occurs on the same date as start of new anti-cancer treatment it will be assumed that progression was prior to the start of new anti-cancer treatment and PFS will not be censored.

6.2.1.1 Primary analysis of PFS

All randomized patients (FAS) will be included in the primary efficacy analysis.

The analysis will be performed when approximately 280 PFS events (PD by independent central review, biochemical progression in patients with LPL/WM without measurable lesions evaluable by imaging at baseline, or death if death is before PD) in the FAS are observed.

The primary efficacy analyses will evaluate whether PFS in the copanlisib in combination with R-B or R-CHOP group is longer compared with PFS in the placebo in combination with R-B or R-CHOP group in the FAS.

Descriptive summaries (frequency and percentage of patients) will be provided by treatment group for PFS event types and PFS censoring reasons.

The two treatment groups will be compared using a stratified log-rank test with one-sided alpha of 0.025 for the FAS. The primary analysis will include all three stratification factors, unless the smallest group within a stratification factor is less than 5% of the total number of randomized patients. If this occurs the stratification factor subgroup will not be included in the stratified log-rank test but suitably combined into another subgroup. This will also counteract the small possibility of exclusive randomization to placebo or copanlisib in groups of a small size. If after combining subgroups a stratification factor has no subgroups, then the stratification factor will not be included in the stratified log-rank test. A sensitivity log-rank test analysis including all groups of the three stratification factors will be performed, see Section 6.2.1.4 (Sensitivity Analysis 1). Details of the stratification factors used in randomization can be found in Section 4.9.6

The following test procedure will be applied:

According to the size of this study it is justified to assume that under $H_{0,\,PFS}$, the one-sided log-rank test [17] is a sufficiently close approximation to the normal distribution. The actual normal distributed one-sided log-rank test will be developed by dividing the test statistic with the standard error and comparing the resulting p-value against the study-wise alpha of 2.5%.

The SAS code to generate the required test statistic will be comparable to the following pseudo code, with variable names taking their obvious meaning:

```
PROC LIFETEST DATA=dataset;
  TIME event_time*censorny(cens_value);
  STRATA {strata variables}/ group=treatmgr test=(logrank);
  ODS OUTPUT homtests=loguni;
RUN;
```

Technically, the binary treatment variable will be designed with “1” representing copanlisib treatment, to allow the correct direction of the one-sided log-rank test (and “0” representing the placebo group).

For the purpose of the analysis, patients who before Protocol amendment 6 were stratified as “duration of treatment-free interval ≥6-12 months after completion of the last rituximab containing treatment” will be combined with those who after Protocol amendment 6 were stratified as “duration of progression-free ≥6-12 months after completion of the last rituximab,

or rituximab biosimilars, or anti-CD20 monoclonal antibody containing treatment"; and patients who before Protocol amendment 6 were stratified as "duration of treatment-free interval >12 months after completion of the last rituximab containing treatment" will be combined with those who after amendment 6 were stratified as "duration of progression-free >12 months after completion of the last rituximab, or rituximab biosimilars, or anti-CD20 monoclonal antibody containing treatment".

PFS will be tested based on hierarchy test, see Section 6.2.3.

6.2.1.2 Additional analyses of PFS

The hazard ratio (HR), (including 95% confidence interval), will be derived for the total study population from the Cox proportional hazards models that are stratified by the same factors as used for the primary efficacy analysis. SAS code comparable to the following pseudo code will be used:

```
PROC PHREG DATA=dataset;
  MODEL event_time*censorny(cens_value) = treatmgr;
  STRATA {strata variables};
RUN;
```

Kaplan-Meier estimates of median PFS (including 95% two-sided confidence intervals) and Kaplan-Meier curves will be presented for each treatment group using the following pseudo code:

```
PROC LIFETEST DATA=dataset;
  TIME event_time*censorny(cens_value);
  STRATA treatmgr;
RUN;
```

Subgroup analysis for PFS will be performed by histology (FL, MZL, SLL, and LPL/WM), see Section 6.2.5.

6.2.1.3 Exploratory analyses of PFS

Analysis of concordance and discordance between radiological progression evaluation by independent central blinded review and by investigator assessment during the blinded study phase will be performed via cross tabulations in FAS for overall and for each treatment group.

To assess the impact of the stratification variables, Kaplan-Meier curves and median PFS times by treatment groups will be estimated for each of the 3 stratification variables (if the smallest group within one stratification factor is less than 5% of the total number of randomized patients then groups of the stratification factor will be suitably combined), comparable to the following SAS pseudo-code:

```
PROC LIFETEST DATA=dataset;
BY {strata variable};
  TIME event_time*censorny(cens_value);
  STRATA treatmgr;
RUN;
```

To assess the potential impact of clinical progression, the investigator-based evaluation of PFS will also be performed including clinical progression (either direct "clinical progression"

or "AE related to clinical progression") as a progression event[9]. The first PD assigned by the investigator will be used as event.

Exploratory follow-up analyses of PFS and other efficacy endpoints may be repeated at timepoints for follow-up OS analyses (see Section 6.2.2 for details of OS follow-up timepoints).

6.2.1.4 Sensitivity analyses of PFS

The following sensitivity analyses for the primary PFS analysis will be performed using similar methodology as the primary analysis by independent central review unless otherwise specified:

6.2.1.4.1 Sensitivity Analysis 1 (only if all groups within the 3 stratification factors are not used in the primary PFS analysis)

If the smallest group within one stratification factor is less than 5% of the total number of randomized patients a sensitivity analysis for PFS will be conducted using stratified log-rank test with all three stratification factors at randomization. The three stratification factors are: prior to base treatment regimen (R-chemotherapy to R-B, R-B to R-CHOP, and R-B to R-B), NHL histology (FL vs. other iNHL) and duration of progression-free interval on the last rituximab, or rituximab biosimilars, or anti-CD20 monoclonal antibody containing regimen (6-12 months vs. >12 months).

6.2.1.4.2 Sensitivity Analysis 2

PFS will be evaluated with the unstratified log-rank test and Cox model.

6.2.1.4.3 Sensitivity Analysis 3

PFS will be evaluated by log-rank test and Cox model and stratified by the following regions:

- North America (Canada, Mexico, and United States),
- Asia Pacific, excluding Australia (China, Taiwan, Hong Kong, Japan, Korea, Singapore, and Thailand)
- Rest of the world (All other countries/regions).

9 If the clinical progression is at end of treatment, the last date of treatment, or if there is an AE related to clinical progression, the earlier date of either is used. If the clinical progression is at end of active follow-up, the last visit date (as provided on the eCRF), or if there is an AE related to clinical progression, the earlier date of either is used. If both radiological and clinical progression are documented, the earlier date of radiological progression or the derived date of clinical progression is used. In case of death without prior progression, the date of death is used, and the patient is considered to have an event.

6.2.1.4.4 Sensitivity Analysis 4

In case of many discrepancies between stratification factors entered in the IxRS and information entered in eCRF, PFS will be evaluated with the stratified log-rank test and Cox model based on stratification information entered in eCRF. In the case that this is not evaluable the prior to base treatment regimen categories R-B to R-CHOP and R-B to R-B, will be combined in the analysis. In the case that a patient has an eCRF entry for both "progression-free survival >12 months" and for "treatment-free survival >12 months", progression-free survival >12 months will be used. Patients with the important protocol deviation rituximab resistance (defined as lack of response, or progression within 6 months of the last course of treatment with a rituximab containing regimen, including rituximab maintenance) will be excluded from the sensitivity analysis as their stratification factor, duration of progression-free interval on the last rituximab, or rituximab biosimilars, or anti-CD20 monoclonal, is entered incorrectly in both IxRS and the eCRF.

The censoring mechanism of patients without PFS events at the time of analysis is assumed to be non-informative for the primary efficacy analysis. Sensitivity analyses will be performed, assessing the impact of a potential informative censoring of such patients.

6.2.1.4.5 Sensitivity Analysis 5

PFS will be evaluated considering all patients that are censored, and not at risk, to have events in both arms (Kaplan-Meier curves and Cox model for HR). This excludes patients ongoing in Active follow up at the time of analysis, as these patients will still be "administratively censored" for this sensitivity analysis.

6.2.1.4.6 Sensitivity Analysis 6

PFS will be evaluated considering any change as a progression event, e.g., treating initiation of a new anti-cancer agent[10] as an event at the date of start of new anti-cancer agent, disease progression as an event at the date of progression ignoring scheduled missing assessments, premature treatment discontinuation as an event at the date of discontinuation [18]. Table 6–2 lists the censoring order and rules for this sensitivity analysis. If a patient satisfies multiple criteria in Table 6–2, then the earliest criterion will be considered.

[10] Or surgery affecting target/non-target lesions prior to observing progression

Table 6–2: PFS censoring rules for sensitivity analysis 6 (any change or premature discontinuation from treatment considered as a progression event)

Situation and censoring order	End Date/Day	Censored	Reason for Censoring	Detailed Reference
All patients regardless of whether they have a baseline or post-baseline tumor assessment				
New anti- cancer treatment prior to observing progression[1,2]	Date of start of new anti-cancer treatment[3]	No	N/A	
Discontinued or withdrew early from treatment without documented disease progression	Date of treatment discontinuation (i.e., date of last non 0mg dose)	No	N/A	
Patients with neither baseline nor post-baseline tumor assessment				
Death within 25 (12+12+1) weeks after randomization	Date of death	No	N/A	
Death more than 25 (12+12+1) weeks after randomization	Reference Date/Day 1	Yes	Death happened after time window specified in Appendix 9.4.2	Appendix 9.4.2
Patients without a baseline tumor assessment				
No baseline tumor assessment and no death	Reference Date/Day 1	Yes	No baseline tumor assessment and no death	
No baseline tumor assessment and death within 25 (12+12+1) weeks after randomization	Date of death	No	N/A	
No baseline tumor assessment and death during treatment period or Active follow-up	Date of death	No	N/A	
Death early enough in Survival follow-up to be within the time window for a PFS event.	Date of death	No	N/A	
Death too late in the Survival follow-up to be within the time window for a PFS event	Last Tumor assessment	Yes	Death happened after time window specified in Appendix 9.4.2	Appendix 9.4.2
Patients without a post-baseline tumor assessment				
No post-baseline tumor assessment and no death	Reference Date/Day 1	Yes	No post-baseline tumor assessment and no death	
No post-baseline tumor assessment and death during treatment period or Active follow-up	Date of death	No	N/A	
No post-baseline tumor assessment and death during Survival follow-up, within 25 (12+12+1) weeks after randomization	Date of death	No	N/A	
No post-baseline tumor assessment and death during Survival follow-up, but more than 25 (12+12+1) weeks after randomization	Reference Date/Day 1	Yes	Death happened after time window specified in Appendix 9.4.2	Appendix 9.4.2

Table 6–2: PFS censoring rules for sensitivity analysis 6 (any change or premature discontinuation from treatment considered as a progression event)

Situation and censoring order	End Date/Day	Censored	Reason for Censoring	Detailed Reference
Patients with both baseline and post-baseline tumor assessment				
Death or progression occurred during treatment, Active follow-up or early enough in Survival follow-up to be within the time window for a PFS event, regardless of missed tumor assessments	Earliest of: • Death date or • Date of progression	No	N/A	
Death too late in the Survival follow-up to be within the time window for a PFS event	Last Tumor assessment	Yes	Death happened after time window specified in Appendix 9.4.2	Appendix 9.4.2
Neither PD nor death and no other criteria fulfilled	Last tumor assessment	Yes	Patient continuing in treatment period or Active follow-up in the absence of experiencing a PFS event	

[1] Or surgery affecting target/non-target lesions prior to observing progression.
[2]If progression occurs on the same date as start of new anti-cancer treatment it will be assumed that progression was prior to the start of new anti-cancer treatment and PFS will not be censored.
[3]Or date of surgery.

6.2.1.4.7 Sensitivity Analysis 7

PFS will be evaluated more conservatively than in Sensitivity analysis 6 by additionally treating premature discontinuations from the Safety follow-up or Active-follow-up as PFS events on the date of discontinuation [18]. Table Table 6–3 lists the rules for this sensitivity analysis.

Table 6–3: PFS censoring rules for sensitivity analysis 7 (any change or premature discontinuation from treatment or Active follow-up considered as a progression event)

Situation and censoring order	End Date/Day	Censored	Reason for Censoring	Detailed Reference
All patients regardless of whether they have a baseline or post-baseline tumor assessment				
New anti- cancer treatment prior to observing progression[1,2]	Date of start of new anti-cancer treatment[3]	No	N/A	
Discontinued or withdrew early from treatment, Safety follow-up or Active follow-up without documented disease progression	Date of discontinuation (i.e., date of last non 0mg dose)	No	N/A	
Patients with neither baseline nor post-baseline tumor assessment				
Death within 25 (12+12+1) weeks after randomization	Date of death	No	N/A	
Death more than 25 (12+12+1) weeks after randomization	Reference Date/Day 1	Yes	Death happened after time window specified in Appendix 9.4.2	Appendix 9.4.2

Table 6–3: PFS censoring rules for sensitivity analysis 7 (any change or premature discontinuation from treatment or Active follow-up considered as a progression event)

Situation and censoring order	End Date/Day	Censored	Reason for Censoring	Detailed Reference
Patients without a baseline tumor assessment				
No baseline tumor assessment and no death	Reference Date/Day 1	Yes	No baseline tumor assessment and no death	
No baseline tumor assessment and death within 25 (12+12+1) weeks after randomization	Date of death	No	N/A	
No baseline tumor assessment and death during treatment period or Active follow-up	Date of death	No	N/A	
Death early enough in Survival follow-up to be within the time window for a PFS event.	Date of death	No	N/A	
Death too late in the Survival follow-up to be within the time window for a PFS event	Last Tumor assessment	Yes	Death happened after time window specified in Appendix 9.4.2	Appendix 9.4.2
Patients without a post-baseline tumor assessment				
No post-baseline tumor assessment and no death	Reference Date/Day 1	Yes	No post-baseline tumor assessment and no death	
No post-baseline tumor assessment and death during treatment period or Active follow-up	Date of death	No	N/A	
No post-baseline tumor assessment and death during Survival follow-up, within 25 (12+12+1) weeks after randomization	Date of death	No	N/A	
No post-baseline tumor assessment and death during Survival follow-up, but more than 25 (12+12+1) weeks after randomization	Reference Date/Day 1	Yes	Death happened after time window specified in Appendix 9.4.2	Appendix 9.4.2
Patients with both baseline and post-baseline tumor assessment				
Death or progression occurred during treatment, Active follow-up or early enough in Survival follow-up to be within the time window for a PFS event, regardless of missed tumor assessments	Earliest of: • Death date or • Date of progression	No	N/A	
Death too late in the Survival follow-up to be within the time window for a PFS event	Last Tumor assessment	Yes	Death happened after time window specified in Appendix 9.4.2	Appendix 9.4.2
Neither PD nor death and no other criteria fulfilled	Last tumor assessment	Yes	Patient continuing in treatment period or Active follow-up in the absence of experiencing a PFS event	

[1] Or surgery affecting target/non-target lesions prior to observing progression.
[2] If progression occurs on the same date as start of new anti-cancer treatment it will be assumed that progression was prior to the start of new anti-cancer treatment and PFS will not be censored.
[3]Or date of surgery.

6.2.1.4.8 Sensitivity Analysis 8

A listing of dose delays ≥ 60 days due to COVID-19 or the Ukraine-Russia conflict will be presented. Additional sensitivity analyses of PFS related to COVID-19 and/or the Ukraine-Russia conflict will be performed if treatment delays due to either are observed and PFS evaluations are thought to be at risk of being affected. If further analyses are deemed necessary, these will be detailed in a supplement to the SAP.

6.2.2 Secondary efficacy variables

Definition of secondary efficacy variables

ORR

Assessed in all patients, ORR is defined as the proportion of patients who have a best response rating while on treatment or in Active follow-up (i.e., until time of analysis of PFS) of CR or partial response (PR) according to the Lugano Classification and for patients with LPL/WM a response rating of CR, very good partial response (VGPR), PR, or minor response (MR) according to the Owen Criteria. As initiation of new anti-tumor therapy ends the Active follow-up, any tumor assessments taking place after the initiation of new anti-cancer therapy will not be included in the calculation of ORR.

DOR

Assessed in responders only, DOR is defined as the time (in days) from first observed tumor response (CR, VGPR, PR, or MR) until PD or death from any cause, whichever is earlier. Patients without PD or death at the time of analysis will be censored at the date of their last evaluable tumor assessment if available or at day 1 of response, otherwise. DOR will only be defined for patients with at least one CR, VGPR, PR, or MR. The conventions for defining progressive disease after 2 missing assessments are the same as for PFS (see Table 6–1).

DCR

DCR is defined as the proportion of patients who have a BOR of CR, unconfirmed CR, VGPR, PR, unconfirmed PR, MR, or confirmed stable disease (SD). Unconfirmed SD is defined as SD on or before Study Day 76, which is not confirmed by an additional tumor assessment of SD, or better, on Study Day 77 or later.

CRR

Assessed in all patients, CRR is defined as the proportion of patients who have a best response rating of CR while on treatment or in Active follow-up (i.e., until the time of analysis of PFS) according to Lugano classification and for patients with LPL/WM, a response rating of CR according to the Owen criteria.

TTP

TTP is defined as the time (in days) from randomization to PD or death related to PD, whichever is earlier. Death related to PD is any death except for:

a) Death due to an AE unrelated to progression
b) Death with a specification of "other" as reason (which excludes PD)

TTP for patients without PD at the time of analysis or with death not related to progression will be censored at the date of their last tumor evaluation. TTP for patients who have neither tumor assessments nor death related to PD after baseline will be censored at Day 1. TTP for patients without baseline tumor assessments, and no death related to PD within the first 25 weeks after randomization, will be censored at Day 1. The conventions for calculation of TTP are the same as for PFS (Section 6.2.1) with the exception that death not related to PD will not be considered as an event.

The actual dates of tumor assessments will be used for this calculation.

TTNT

TTNT is defined as the time from date of randomization to start of new anti-lymphoma therapy, where date of randomization is day 1. Start of new anti-lymphoma therapy is a component of the Survival follow-up assessments, and it is assumed that, unless documented, patients did not receive concomitant anti-lymphoma therapy during treatment/Active follow-up (as this would trigger the EOT/end of Active follow-up period), or in Safety follow-up (if a patient who stopped treatment did not enter the Active follow-up). New anti-lymphoma therapy is any new systemic anti-cancer treatment, or radiotherapy for lymphoma with a consolidation intent. The start date of the first new anti-lymphoma therapy (systemic or radiotherapy) will be used as the start date of new anti-lymphoma therapy.

Table 6–4 lists the censoring rules for TTNT.

Table 6–4 TTNT censoring rules

End Date/Day	Censored	Situation	Reason for Censoring
Start date of new anti-cancer treatment	No	New anti-cancer treatment received by patient	N/A
Last visit date	Yes	Patient randomized but did not take study medication, did not enter active follow-up and no Survival follow-up assessment of new anti-cancer treatment	No survival assessment of new anti-cancer treatment
		Alive at EOT, did not enter Active follow-up and no Survival follow-up assessment of new anti-cancer treatment	No survival assessment of new anti-cancer treatment.
		Discontinued Active follow-up for reasons other than death, did not enter Survival follow-up	No survival assessment of new anti-cancer treatment.
		Still ongoing in Survival follow-up, never entered Active follow-up and has no assessments of new anti-cancer treatment recorded	No survival assessment of new anti-cancer treatment.
		Still ongoing in Survival follow-up, entered Active follow-up	No survival assessment of new anti-cancer treatment

Table 6–4 TTNT censoring rules			
End Date/Day	**Censored**	**Situation**	**Reason for Censoring**
		and no assessment of new anti-cancer treatment recorded	
		Discontinued from Survival follow-up for reasons other than death, never entered Active follow-up and no assessment of new anti-cancer treatment recorded	No survival assessment of new anti-cancer treatment.
		Discontinued from Survival follow-up for reasons other than death, entered Active follow-up and no assessment of new anti-cancer treatment recorded	No survival assessment of new anti-cancer treatment
		Died during Survival follow-up, never entered Active follow-up and no assessment of new anti-cancer treatment recorded	No survival assessment of new anti-cancer treatment.
		Died during Survival follow-up, entered Active follow-up and no assessment of new anti-cancer treatment recorded	No survival assessment of new anti-cancer treatment
Data cut-off date	Yes	Still ongoing in Active follow-up	Still on study so assumption is they have not started a new anti-cancer treatment.
Date of death	Yes	Died during treatment, Safety follow-up or Active follow-up	Died before endpoint reached.
Date of last survival assessment of new anti- cancer treatment	Yes	Still ongoing in Survival follow-up with survival assessments of new anti-cancer treatment but no new therapy reported	Endpoint not reached yet.
		Discontinued from Survival follow-up for reasons other than death, was assessed for new anti- cancer treatment but no new therapy reported	Discontinued before endpoint reached.
		Died during Survival follow-up, was assessed for new anti-cancer treatment but no new therapy reported and date of death after the date of the last survival assessment of new anti- cancer treatment.	Died before endpoint reached.

OS

OS is defined as the time (in days) from randomization until death from any cause. OS of patients alive at the time of analysis will be censored at their LKAD. Death after the data cut-off, reported during cleaning, will be considered for establishing LKAD at data cut-off.

See Section 4.3 for details regarding derivation of LKAD and imputation methods for partial/missing death dates.

Time to deterioration in DRS-P of at least 3 points

Time to deterioration in DRS-P of at least 3 points, as measured by the FLymSI-18 questionnaire, will be evaluated for all patients. It is defined as the time (in days) from randomization to DRS-P decline, PD (as defined for the primary analysis of PFS), or death due to any reason, whichever is earlier. DRS-P deterioration censoring rules are listed in Table 6–5.

Table 6–5: DRS-P deterioration censoring rules

Situation	End Date/Day	Censored	Reason for Censoring
No baseline DRS-P assessment	Reference Date/Day 1	Yes	No baseline or post-baseline DRS-P assessment
No post-baseline DRS-P assessment and no PD or death within time window	Reference Date/Day 1	Yes	No post-baseline DRS-P assessment and no PD/death within time window
Patient has DRS-P deterioration and has PD	Min (DRS-P deterioration date, PD)	No	N/A
Patient has DRS-P deterioration and no PD	DRS-P deterioration date	No	N/A
Patient has no DRS-P deterioration but has PD before the last DRS-P assessment	PD date	No	Has event, as PD occurred (even if no deterioration)
Patient has no DRS-P deterioration but has PD or death within the time window	PD date or death date	No	
Patient has no DRS-P deterioration, no PD, or no death within time window	Last DRS-P assessment date	Yes	Patient no longer “on study” and event for DRS-P deterioration not reached.
New anti-cancer treatment other than the study treatment prior to observing DRS-P deterioration (and no PD before switch to new anti-cancer treatment)	Last evaluable DRS-P assessment date prior to the initiation of anti-cancer treatment	Yes	New anti-cancer treatment
Discontinued or withdraw early from the study without documented DRS-P deterioration or PD or Death within time window	Last DRS-P assessment date	Yes	Discontinued
Neither DRS-P deterioration, nor PD/death and no other criteria fulfilled	Last DRS-P assessment date	Yes	Patient regularly ongoing in trial without DRS-P event occurred

PD as defined for the primary analysis of PFS.
Similar time window for PFS will be used for the derivation.

Time to improvement in DRS-P of at least 3 points

Time to improvement in DRS-P of at least 3 points, as measured by the FLymSI-18 questionnaire, will be evaluated for all patients. It is defined as the time (in months) from randomization to DRS-P improvement of at least 3 points. The patient will generally be censored at the last non-missing post-baseline DRS-P score, if no DRS-P improvement, PD

(as defined for the primary analysis of PFS), or death occurs. If a PD or death occurs before a potential DRS-P improvement (or no DRS-P improvement is documented at all), the patient will be censored at the largest observation time (of events and censoring in all patients evaluated for improvement), plus 1 day. Regular censoring rules, apart from the special rule for PD/Death will be comparable to the censoring for deterioration of DRS-P. The DRS-P improvement censoring rules are listed in Table 6–6.

Table 6–6: DRS-P improvement censoring rules

Situation	End Date/Day	Censored	Reason for Censoring
Patient has DRS-P improvement before PD or death	DRS-P improvement date	No	N/A
Patient has PD or death (including drop out due to PD) before DRS-P improvement	Largest observation time (of events and censoring in all patients evaluated for improvement), plus 1 day	Yes	PD / death occurs before improvement and thus improvement becomes impossible for these patients
Patient has PD or death (including drop out due to PD) after DRS-P improvement	DRS-P improvement date	No	Same rule as Line 1, as PD / death after DRS-P improvement has no impact
No baseline or post-baseline DRS-P assessment	Reference Date/Day 1	Yes	No baseline or post-baseline DRS-P assessment
Patient has no DRS-P improvement, no PD, and no death	Last DRS-P assessment	Yes	Event for DRS-P improvement not reached
New anti-cancer treatment other than the study treatment prior to observing DRS-P improvement (and no PD before last DRS-P assessment before switch to other than study treatment)	Largest observation time (of events and censoring in all patients evaluated for improvement), plus 1 day	Yes	New anti-cancer treatment
Discontinued or withdraw early from the study without documented DRS-P improvement or PD/Death	Last DRS-P assessment date	Yes	Discontinued
Neither DRS-P improvement, nor PD/death and no other criteria fulfilled	Last DRS-P assessment	Yes	Patient regularly ongoing in trial without DRS-P event occurred

PD (as defined for the primary analysis of PFS)

FLymSI-18 questionnaire

The FLymSI-18 questionnaire contains 18 items, each of which utilizes a Likert scale with 5 possible responses ranging from 0 'Not at all' to 4 'Very much'. Nine items reflect the DRS-P, and the responses to the items are summed to calculate a DRS-P subscale score. Four items represent disease-related emotional symptoms (DRS-E), and the responses to those items may be used to calculate a DRS-E subscale score. Three items represent treatment side effects, and the responses to these items may be summed to calculate a treatment side-effect (TSE) subscale score. Finally, two items represent function and well-being (FWB), and responses to those items may be summed to calculate a FWB subscale score. The questionnaire allows for calculation of total score and four subscales: DRS-P, DRS-E, TSE and FWB. A summary of the scoring guideline is included in Appendix 9.2.

Analyses of secondary efficacy variables

All secondary efficacy endpoints will be analyzed in the FAS at the time of the analysis of the primary efficacy variable. Image based secondary efficacy endpoints will be analyzed based on blinded independent central review data, except for patients with LPL/WM without measurable lesions evaluable by imaging at baseline investigator assessment, whose local assessments will be used in place of centrally reviewed data. Image based secondary efficacy endpoints analyses will also be analyzed and presented based on investigator assessments. PET-CT combined investigator assessment will be derived following the principles detailed in Appendix 9.1.

The ORR will be analyzed using the Cochran-Mantel-Haenszel test stratified by the same stratification factors used in the PFS analysis. The null hypothesis is defined as:

$H_{0,\,ORR}$: $ORR_{Copanlisib+R\text{-}B/R\text{-}CHOP} \leq ORR_{Placebo+R\text{-}B/R\text{-}CHOP}$

The alternative hypothesis is:

$H_{1,\,ORR}$: $ORR_{Copanlisib+R\text{-}B/R\text{-}CHOP} > ORR_{Placebo+R\text{-}B/R\text{-}CHOP}$

In addition, the point estimate as well as 95% two-sided confidence intervals for the Mantel-Haenszel weighted treatment difference [19] will be calculated.

The time to deterioration and time to improvement in DRS-P subscale of FLymSI-18 will be analyzed analogously to PFS, using similar hypotheses, decision rules, and significance level, but will only be included in the confirmatory testing for “Europe and the rest of the world” (Section 6.2.3).

Analysis of OS

OS will be analyzed in terms of safety and efficacy. Further details with regards to the OS safety analysis are included in Section 6.3.5, while OS efficacy analyses are detailed in this section.

For efficacy, the analysis of OS will be performed initially as an interim analysis at the time of primary completion of the primary endpoint, PFS. The final formal statistical analysis of OS will take place 5 years after the last patient started study treatment, when it is predicted that approximately 176 OS events will have occurred. In both instances, OS will be analyzed using stratified log-rank tests analogue to the analysis for the primary analysis of the primary endpoint. Further details regarding the division of study-wise alpha are included in Section 6.2.3. The final descriptive analysis of OS will take place at the end of the study, up to 10 years after the last patient started study treatment.

The survival status of each patient will be collected quarterly during the Survival follow-up period, up to 10 years after the last patient started study treatment. At this point the final descriptive analysis of OS at study end will take place.

Approximately 12 months after analysis of the primary and secondary efficacy variables, and at yearly intervals thereafter, further descriptive analyses of OS (each a “follow-up OS analysis”) will be performed. These analyses will be updates of the analysis of OS performed on data available at the time of analysis of the primary endpoint, PFS.

Analysis of other secondary efficacy variables

TTP, DOR, TTNT as well as CRR will not be included into the confirmatory testing strategy but analyzed supportively only.

TTP, DOR and TTNT will be analyzed using stratified log-rank tests analogue to the analysis for the primary analysis of the primary endpoint. CRR will be analyzed using the Cochran-Mantel-Haenszel test and estimated difference analogous to the analysis of ORR. The test will be adjusted for the same stratification factors as used for the primary analysis of PFS.

Additional analyses of secondary efficacy variables

ORR

Exploratory analysis will be performed to summarize the concordance and discordance between best overall response results by independent central review and investigator's assessment.

Subgroup analysis for ORR will be performed by histology (FL, MZL, SLL, and LPL/WM), see Section 6.2.5.

DRS-P

As a sensitivity analysis, an analysis of "worsening of DRS-P alone", considering patients as "censored" at the date of their last DRS-P measurement, ignoring PD or death (that is, not considering PD or death as events), will be performed. The sensitivity analysis censoring rules, for the worsening of DRS-P alone, are listed in Table 6–7.

Table 6–7: Worsening of DRS-P alone sensitivity analysis censoring rules

Situation	End Date/Day	Censored	Reason for Censoring
No baseline DRS-P assessment	Reference Date/Day 1	Yes	No baseline or post-baseline DRS-P assessment
Patient has DRS-P deterioration	DRS-P deterioration date	No	N/A
Patient has no DRS-P deterioration	Last DRS-P assessment	Yes	
New anti-cancer treatment other than the study treatment prior to observing DRS-P deterioration	Last evaluable DRS-P assessment prior to the initiation of anti-cancer treatment	Yes	New anti-cancer treatment
Discontinued or withdraw early from the study without documented DRS-P deterioration	Last DRS-P assessment date	Yes	Discontinued
No DRS-P deterioration and no other criteria fulfilled	Last DRS-P assessment	Yes	Patient regularly ongoing in trial without DRS-P event having occurred

6.2.3 Confirmatory statistical test strategy

Separate statistical testing strategies will be conducted for (1) the United States and (2) Europe and the rest of the world, as outlined below.

United States

For the United States, to control the study-wise alpha, a fixed-sequence multiple testing strategy will be implemented. The sequence of the testing of primary and secondary endpoints will be as follows and as shown in Figure 6–1:

The primary endpoint PFS will be tested in the FAS population (tested at one-sided alpha=0.025).

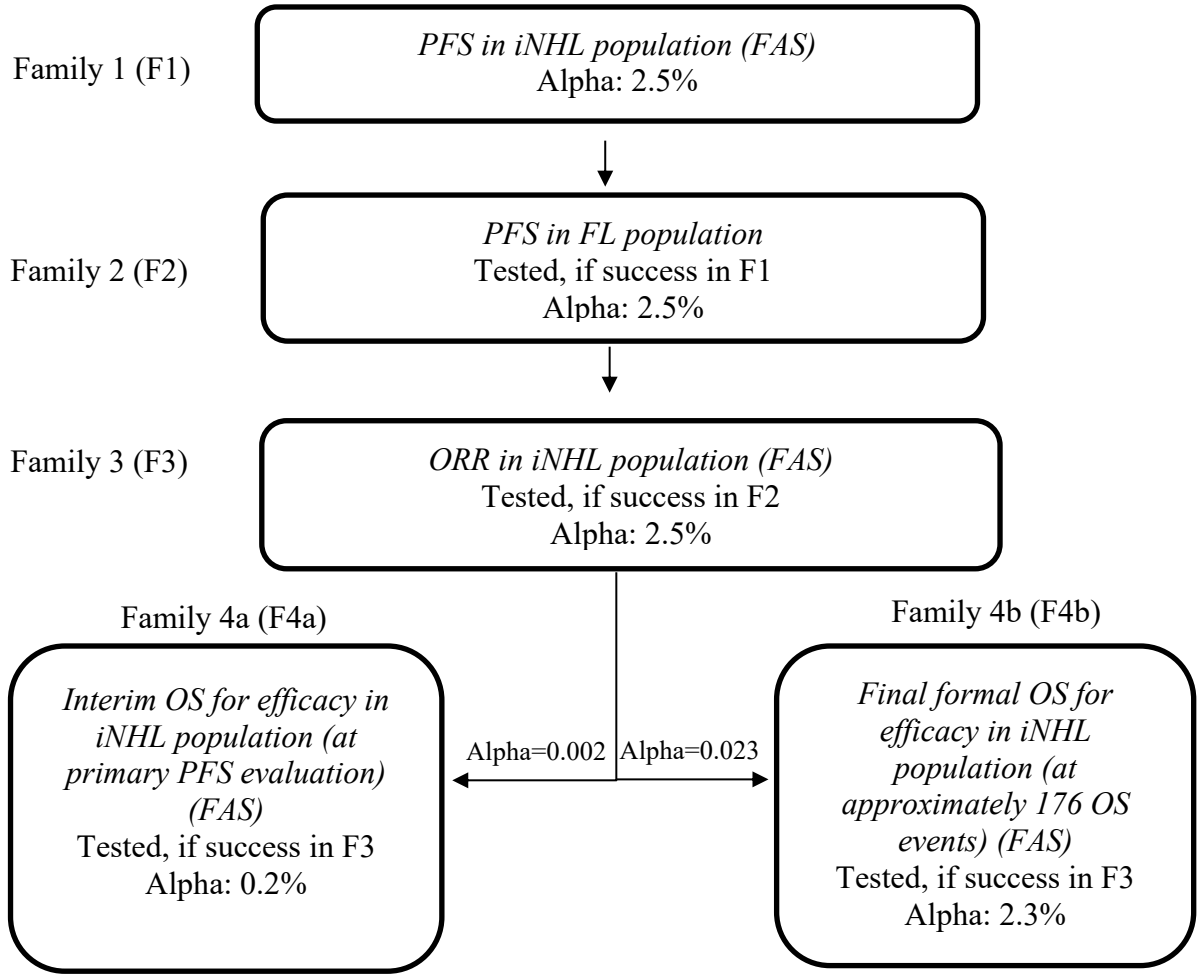
If the PFS test in the FAS population is successful,

a) a second test on PFS in the FL histology population will be performed (tested at one-sided alpha=0.025); if this is successful,

b) a third test on the ORR in the FAS population will be performed (tested at one-sided alpha=0.025); if this is successful,

c) a fourth test on the OS in the FAS population will be performed. An interim OS analysis (tested at one-sided alpha=0.002) and a final formal OS analysis (tested at one-sided alpha=0.023) will be performed. This would conclude the confirmatory test procedure.

Figure 6–1 Confirmatory test strategy based on four test families for United States



Europe and the rest of the world

For Europe and the rest of the world, to control the study-wise alpha, a fixed-sequence multiple testing strategy will be implemented. In this strategy, the sequence of the testing of primary and secondary endpoints is as follows and as shown in Figure 6–2:

The primary endpoint PFS will be tested in the FAS population (tested at one-sided alpha=0.025).

If the PFS test in the FAS population is successful,

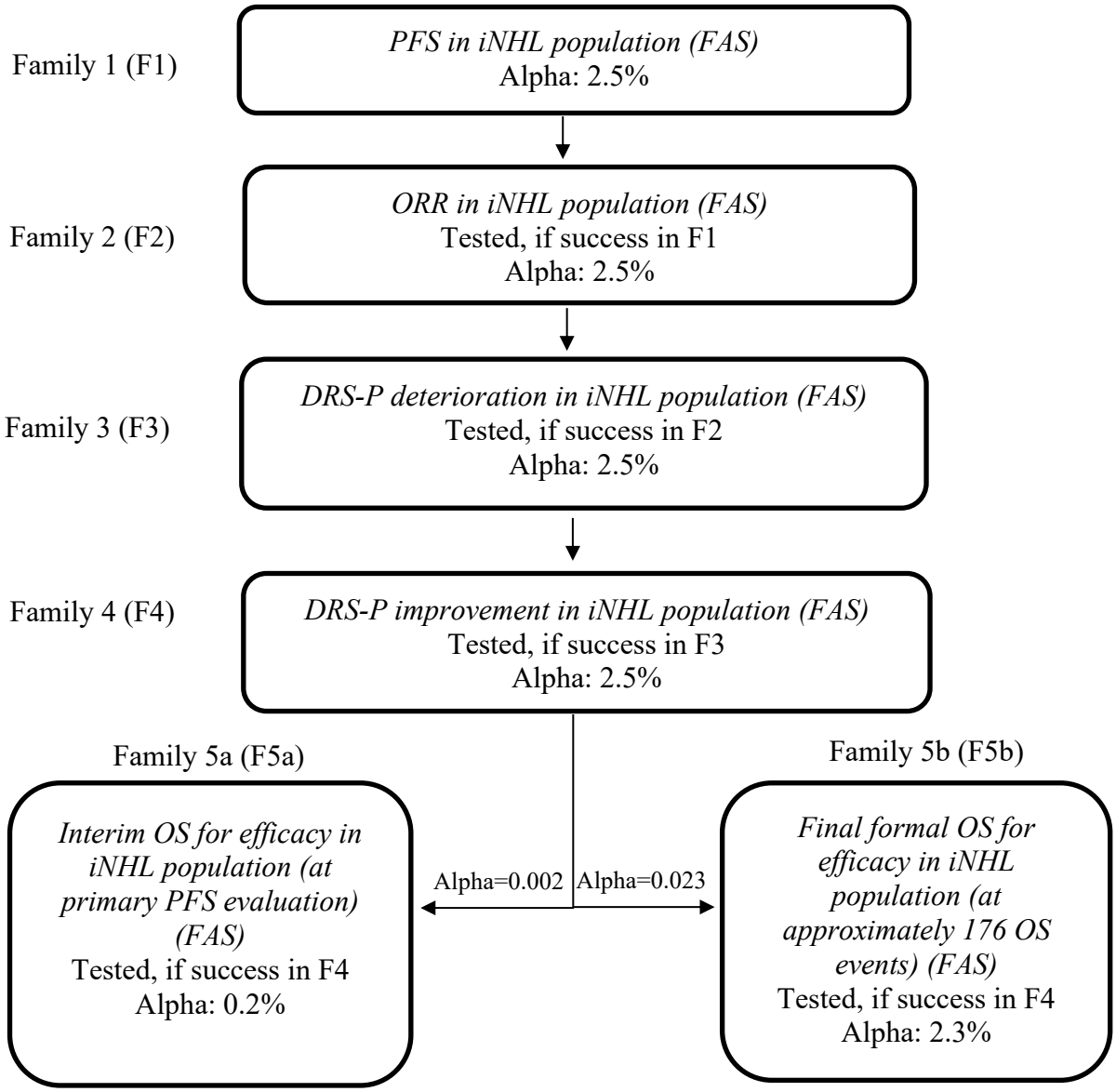
a) a second test on the ORR in the FAS population will be performed (tested at one-sided alpha=0.025); if this is successful,

b) a third test on the DRS-P deterioration in the FAS population will be performed (tested at one-sided alpha=0.025); if this is successful,

c) a fourth test on the DSR-P improvement in the FAS population will be performed (tested at one-sided alpha=0.025); if this is successful,

d) a fifth test on the OS in the FAS population will be performed. An interim OS analysis (tested at one-sided alpha=0.002) and a final formal OS analysis (tested at one-sided alpha=0.023) will be performed. This would conclude the confirmatory test procedure.

Figure 6–2: Confirmatory Test Strategy Based on Five Test Families for "Europe and the rest of the world"



Conditional power with respect to PFS and ORR in the full study population, and PFS in the FL subpopulation, is determined based on published studies that evaluate 2nd line immuno-chemotherapy combinations in iNHL populations: Rummel, 2005 [10], Rummel, 2016 [24] and Rueda, 2018 [25], which are also described in Section 9.6.

Conditional statistical power assessments for the test in ORR, under different scenarios, is provided in Table 6–8.

Table 6–8 Assessment of Statistical Power for ORR in the Overall FAS

Assumed ORR		Assumed risk difference	Resulting conditional power
Placebo	**Copanlisib**		
0.80	0.90	0.1	87.7%
0.85	0.90	0.05	36.0%
0.85	0.95	0.1	96.7%
0.90	0.95	0.05	53.5%

ORR = objective response rate.
Note: Calculated using binomial enumeration, for Fisher's exact test, using PASS, 2022.

Table 6–8 shows that the conditional statistical power for the test in ORR is expected to range between 36% and 97%.

As iNHL is an indolent condition, there may not be a meaningful number of OS events at the time of the primary analysis of PFS. Therefore, the study-wise alpha, assigned for the analysis of OS, will be divided between an interim and a final formal OS analysis, as detailed in Section 4.8.

The formal interim analysis of OS for efficacy is planned at the time of primary analysis for PFS (F4a in the United States and F5a in "Europe and the rest of the world"), if all other key secondary endpoints are significant, using a small portion of the study-wise alpha (0.002). At the time point of primary analysis, it is assumed that approximately 151 OS events have occurred. In that case, the considered scenarios would detect a significant OS effect with statistical (conditional) power ranging from 6.6% (if true OS HR=0.8) to 60.3% (if true OS HR=0.6).

A final formal analysis of OS for efficacy (F4b in the United States and F5b in "Europe and the rest of the world") is planned to be performed after all patients have been followed up for at least 5 years, when approximately 176 OS events are expected to have occurred and will use the remaining portion of the study-wise alpha (0.023). At this analysis the considered scenarios would detect a significant OS effect with statistical conditional power ranging from 30.3% (if true OS HR=0.8) to 91.8% (if true OS HR=0.6).

If the formal interim analysis of OS is significant, this final formal analysis of OS for efficacy will be for descriptive purposes only.

Beyond the final formal analysis of OS, there might be additional follow-up analyses not formally applying alpha control.

Power results for all scenarios and formal analyses are presented in Table 6–9

Table 6–9 Scenarios for formal OS evaluations for efficacy

True underlying HR for OS	Power to detect significant treatment effect in OS vs. HR=1	
	With 151 events and alpha of 0.002	With 176 events and alpha of 0.023
0.60	60.3%	91.8%
0.65	40.9%	80.6%
0.70	24.6%	64.5%
0.75	13.3%	46.5%
0.80	6.6%	30.3%

Given current information, the analysis considering a Hazard Ratio for OS of 0.75 is considered the most appropriate. To power a hypothetical new study with 80% at the final formal analysis of OS with HR=0.75 as true hazard ratio, 380 OS events would be needed.

Additional assessments of OS data and scenarios for OS (answering an HA request)

In Appendix 9.6, an assessment of OS events needed to detect a difference in OS compared with the placebo group, for plausible HR assumptions, over a range of power levels from 60% to 90% is described.

In Appendix 9.7, details of the formal OS event analyses for varying primary PFS time points and OS hazard ratios are described.

In Appendix 9.8, the likelihood to rule out a harmful effect in OS with additional OS follow-up beyond the final formal (at approximately 176 OS events) analysis is simulated.

6.2.4 Other efficacy variables

Due to differing cycle lengths, by visit summaries will be reported for those assigned to R-B treatment (copanlisib+R-B and placebo+R-B) only. Equivalent data for those assigned to R-CHOP treatment will be listed.

FLymSI-18 total and subscale scores (DRS-P, DRS-E, TSE, FWB) are other efficacy variables.

1. Area under the curve (AUC) across all data of FLymSI-18 DRS-P subscale score

An additional analysis of the physical symptoms of lymphoma (as measured using the FLymSI-18 DRS-P subscale) will be performed to assess differences between treatment arms in the FAS based on AUC. AUC will be derived using the linear trapezoidal rule:

$$AUC_{0-t_n} \approx \frac{1}{2}\sum_{i=1}^{n-1}(t_{i+1} - t_i)(C_i + C_{i+1})$$

where C_i denotes the non-missing DRS-P score at timepoint t_i. Missing DRS-P scores will not be replaced. If a patient has only the baseline value, the AUC will not be calculated.

Summary statistics will be provided for AUC and for the AUC normalized over the observation time.

2. FLymSI-18 total and subscale scores (DRS-P, DRS-E, TSE, FWB), and time to meaningful worsening of physical symptoms of lymphoma based on the DRS-P subscale.

Total FLymSI-18 and subscales will be evaluated descriptively. The number and percentage of patients with a meaningful worsening of physical symptoms, defined as a decrease in DRS-P score ≥3 points, will be presented by visit and overall. The time to meaningful worsening of physical symptoms will be described.

3. ECOG performance status

ECOG performance status will be summarized by visit using descriptive statistics for the original score, as well as for the change from baseline score by treatment group in the FAS.

6.2.5 Subgroup analyses

6.2.5.1 Histology specific subgroup analysis

For selected subgroups of patients in the FAS, as defined by cancer histology, separate analyses of treatment effect for the primary and various secondary endpoints will be performed. Diagnosis of cancer histology from investigator's assessment will be used in statistical analyses.

The respective endpoints and histology can be seen from Table 6–10, which describes the planned output format.

In addition, for the time to event endpoints, histology-specific Kaplan Meier evaluations comparing treatment groups will be provided (i.e., KM plot and descriptive statistics including log-rank test).

Table 6–10 Overview of Histology Specific Subgroup Analyses

	Variable						
Histology Subgroup	**PFS**	**ORR**	**CRR**	**DCR**	**Time to deterioration* DRS-P**	**Time to improvement* DRS-P**	**OS**
FL	c)	b)	b)	b)	a)	a)	a)
MZL	a)	b)	b)	b)	a)	a)	a)
Extranodal MZL	a)	b)	b)	b)	a)	a)	a)
Nodal MZL	a)	b)	b)	b)	a)	a)	a)
Splenic MZL	a)	b)	b)	b)	a)	a)	a)
SLL	a)	b)	b)	b)	a)	a)	a)
LPL/WM	a)	b)	b)	b)	a)	a)	a)

* of at least 3 points
FL: follicular lymphoma; MZL: marginal-zone lymphoma; SLL: small lymphocytic lymphoma; LPL: lymphoplasmacytic lymphoma; WM: Waldenström macroglobulinemia
a) Evaluated by use of a non-stratified Cox-PH model.
Results displayed in table: Number of events per treatment group, HR, 95% CI.
b) Evaluated by use of a chi-square test. Results displayed in table: Number patients per treatment group, estimated treatment difference and 95% CI.
c) Evaluated by use of stratified Cox-PH model (stratified by prior to base treatment regimen and duration of progression free interval) and also a non-stratified Cox-PH model. Results displayed in table: Number of events per treatment group, HR, 95% CI.

The same subgroup analyses will be conducted for the secondary endpoints DOR, TTP and TTNT by histology subgroups FL and MZL, separately.

6.2.5.2 Other subgroup analyses

Subgroup analyses will include forest plots of response/HRs as well as treatment-interaction analyses. The forest plots of response/HRs will be provided for the primary efficacy endpoint (PFS) as well as the secondary efficacy endpoints ORR, time to deterioration in DRS-P and time to improvement in DRS-P. The treatment-interaction analysis will be conducted for the primary efficacy endpoint (PFS), and the secondary efficacy endpoints time to deterioration in DRS-P and time to improvement in DRS-P. Kaplan-Meier estimates of median times to PFS (including 95% CI) and Kaplan-Meier curves will be provided for each subgroup level for each treatment group. Subgroup levels without a sufficient number of events observed will be excluded from corresponding subgroup analysis. All subgroup analyses will be done on non-stratified Cox model and log-rank test. For efficacy endpoint ORR, chi-square test will be used.

For the primary efficacy endpoint (PFS), subgroup levels to be analyzed are:

- iNHL histology (FL vs. other iNHL histology)
- iNHL histology (FL vs. MZL vs. SLL vs. LPL/WM) treatment interaction test only will be performed; for subpopulation analysis see Section 6.2.5.1
- Prior to base treatment regimen (R-chemotherapy to R-B vs. R-B to R-CHOP vs. R-B to R-B). Prior to base treatment regimen categories R-B to R-CHOP and R-B to R-B may be combined due to small patient numbers.
- Duration of progression free interval on the last rituximab, or rituximab biosimilars, or anti-CD20 monoclonal antibody containing regimen (6-12 months vs. >12 months)
- Prior lines of systemic anti-cancer therapy (1 vs. 2 vs. ≥3)
- Geographic regions 1 (US, Europe, Rest of world)
- Geographic regions 2 (North America, Asia Pacific, Rest of the world)
- Age group (years) (<65 vs. ≥65)
- ECOG performance status (0 vs. 1 vs. 2)
- Sex (Male vs. Female)
- Race
- Ethnicity (Hispanic or Latino vs. non-Hispanic and non-Latino)

For secondary efficacy endpoints ORR, DCR, CRR, time to deterioration of DRS-P, and time to improvement in DRS-P, the same subgroup levels will be analyzed excluding sex.

Subgroup analyses may be performed for OS at the time of the final formal OS analysis (approximately 5 years after the last patient started study treatment).

6.3 Safety

6.3.1 Adverse events (AEs)

AEs will be coded by MedDRA codes and NCI-CTCAE version 4.03 (or higher) dictionary. The final versions used will be stored in the clinical database and provided in a footnote to the respective tables. Severity of AEs and hematological/biochemical toxicities based on laboratory measurements will be graded using NCI-CTCAE dictionary. AEs will be classified as related or not related to study drug, as determined by the investigator.

A treatment-emergent AE (TEAE) is defined as any event occurring or worsening after start of study drug administration until 30 days after the last study drug intake.

Descriptive summary tables (frequency and percentage of patients, not of events) will be presented by treatment group and overall, by MedDRA SOC/PT and by worst NCI CTCAE grade for the following:

- Overview of TEAEs
- TEAEs
- TEAEs of infusion related reaction
- Drug-related TEAEs (copanlisib/placebo or R-B/R-CHOP)
- Treatment-emergent serious AEs (TESAEs)
- Drug-related TESAEs
- TEAEs of special interest (as detailed in Section 6.3.3).

Exposure adjusted TEAEs will be summarized by treatment group and by MedDRA SOC/PT. The exposure adjusted incidence rate per 100 patient years is defined as 100 times the number of patients with at least one TEAE divided by the sum of exposure times in years, where exposure time is the time to first occurrence if a TEAE occurs, or the treatment duration plus time at risk after treatment ends if no TEAE occurs. The time at risk after treatment is the minimum of 30 days or the days of death relative to the last treatment.

Incidences of patients' TEAEs, drug-related (copanlisib/placebo or R-B/R-CHOP, respectively) and/or TESAEs by MedDRA SOC/PT and worst NCI-CTCAE grade will be repeated by the following criteria:

- age group (Years): (<65 vs. ≥65)
- sex (Male vs. Female)
- history of hypertension: [PBMQ = Standardized MedDRA Query (SMQ)_90001275 in MedDRA v25.1 or higher] Medical history of hypertension (yes vs. no) and
- history of diabetes: [PBMQ = SMQ_90001276 in MedDRA v25.1 or higher] Medical history of diabetes (yes vs. no).

Furthermore, all TEAEs and drug-related TEAEs with incidence of at least 10% will be summarized by MedDRA SOC/PT and worst NCI-CTCAE grade. In addition, all TEAEs and drug-related TEAEs with incidence of at least 5% will be summarized by MedDRA for worst NCI-CTCAE grade and for NCI-CTCAE grade 3-4. All TEAEs will be summarized by MedDRA SOC/PT for NCI-CTCAE grade 5.

The number and percentage of patients who discontinued study treatment due to a TEAE/TESAE (discontinued all study treatments) or required a dose reduction (copanlisib/placebo/R-B/R-CHOP) due to a TEAE or interruption caused by a TEAE will be summarized separately for copanlisib/placebo and R-B/R-CHOP by MedDRA SOC/PT and worst NCI-CTCAE grade. The incidences of these TEAEs will be presented also separately by drug relatedness (copanlisib/placebo or R-B/R-CHOP) and by MedDRA SOC/PT and worst NCI-CTCAE grade.

In addition, the following AE summaries by cancer histology (according to investigator pathology) will be provided:

- Overview of TEAEs
- TEAEs
- Drug-related TEAEs (copanlisib/placebo or R-B/R-CHOP)
- TESAEs
- Drug-related TESAEs

Deaths (including all grade 5 TEAEs) will be summarized by treatment group and overall. As "AE" is not captured as a primary cause of death in the Survival follow-up, for the overall summary of deaths, deaths with "AE" or "other" as the primary cause will be pooled together under the primary cause "other".

AEs occurring prior to starting treatment and AEs occurring at least 30 days after the last dose of study treatment (i.e., post-treatment AEs) will also be summarized by treatment group and overall.

In addition, patient listings will be provided by treatment group for the following AEs:

- Patients who died during study treatment up to 30 days after last dose of study treatment: patient ID, sex, age, race, histology, sponsor AE identifier, start and stop dates of study drug, combination therapy (R-B/R-CHOP), date of death, days within first and last dose date, cause of death, MedDRA SOC and PT and relationship to study drug (related to copanlisib/placebo, related to R-B/R-CHOP, related to both, and no).
- Patients who died later than 30 days after last dose of study treatment (including patients with a death date after the data cut-off date, reported during data cleaning): patient ID, sex, age, race, histology, sponsor AE identifier, start and stop dates of study drug, combination therapy (R-B/R-CHOP), date of death, days within first and last dose date, cause of death, MedDRA SOC and PT, and relationship to study drug (related to copanlisib/placebo, related to R-B/R-CHOP, related to both, and no).
- Treatment-emergent serious adverse events (TESAEs): patient ID, sex, age, histology, race, sponsor AE identifier, reported term, MedDRA SOC and PT, CTCAE toxicity grade, reason for seriousness, combination therapy (R-B/R-CHOP), days within first

and last dose date, start and stop dates of AE, relationship to study drug (related to copanlisib/placebo, related to R-B/R-CHOP, related to both, and no), protocol required procedure related (yes/no), outcome, actions taken with each study drug.

- Study drug-related SAEs will be provided separately for copanlisib/placebo and R-B/R-CHOP, respectively, with similar information.
- TEAEs of special interest: patient ID, sex, age, histology, race, sponsor AE identifier, reported term, MedDRA SOC and PT, CTCAE toxicity grade, days within first and last dose date, combination therapy (R-B/R-CHOP), start and stop dates of AE, relationship to study drug (related to copanlisib/placebo, related to R-B/R-CHOP, related to both, and no), protocol required procedure related (yes/no), serious (yes/no), reason for seriousness, outcome, actions taken with each study drug.
- TEAEs leading to permanent discontinuation of study treatment, a dose reduction (copanlisib/placebo/R-B/R-CHOP) or interruption/delay will be listed separately for copanlisib/placebo and R-B/R-CHOP, respectively: patient ID, sex, age, histology, race, sponsor AE identifier, reported term, MedDRA SOC and PT, CTCAE toxicity grade, days within first and last dose date, start and stop dates of AE, relationship to study drug (related to copanlisib/placebo, related to R-B/R-CHOP, related to both, and no), protocol required procedure related (yes/no), serious (yes/no), reason for seriousness, outcome, actions taken with each study drug.

6.3.2 Adverse events by time-interval

TEAEs, drug-related TEAEs, and TESAEs will be presented for each study drug according to time-of-onset (new onset or worsening).

The following time intervals will be presented for copanlisib and placebo:

- Day 1 – Day 90: to include all copanlisib/placebo-treated patients
- Day 91 – Day 180: to include all copanlisib/placebo-treated patients, treated for at least 3 months
- Day 181 – Day 270: to include all copanlisib/placebo-treated patients, treated for at least 6 months
- Day 271 – Day 360: to include all copanlisib/placebo-treated patients, treated for at least 9 months
- Day 361 or greater: to include all copanlisib/placebo-treated patients, treated for at least 12 months.

The following time intervals will be presented for R-B and R-CHOP:

- Day 1 – Day 90: to include all R-B/R-CHOP-treated patients
- Day 91 – Day 180: to include all R-B/R-CHOP-treated patients, treated for at least 3 months
- Day 181 or greater: to include all R-B/R-CHOP-treated patients, treated for at least 6 months.

6.3.3 Adverse events of special interest

Non-infectious pneumonitis/interstitial lung disease

Non-infectious pneumonitis is defined as AE of special safety interest (protocol Section 6.4.1). The SMQ 'Interstitial lung disease' with narrow search (i.e., category 2A) will be used to select these AEs of special interest.

Patient listings will be provided by treatment group for TEAEs of pneumonitis leading to copanlisib/placebo treatment discontinuation: patient ID, sex, age, histology, race, sponsor AE identifier, MedDRA SOC and PT, reported term, CTCAE toxicity grade, days within first and last dose date, reason for seriousness, combination therapy (R-B/R-CHOP), start and stop dates of AE, relationship to study drug (related to copanlisib/placebo, related to R-B/R-CHOP, related to both, and no), protocol required procedure related (yes/no), outcome, actions taken with each study drug.

A descriptive summary of treatment-emergent and copanlisib/placebo-related treatment-emergent MedDRA PTs within the SMQ “Interstitial lung disease” (frequency and percentage of patients, not of events) will be presented by treatment group and cancer histology (total iNHL and histological subpopulations according to investigator pathology).

For treatment emergent events under the SMQ “Interstitial lung disease”, the incidence proportions of worst CTCAE grades, seriousness, outcomes, and actions taken will be presented for the total iNHL and histological subpopulations (according to investigator pathology). This table will also be provided for the subgroups of patients with and without history of respiratory disorders (as defined in Section 6.3.4.1).

For copanlisib/placebo-related treatment emergent events under the SMQ “Interstitial lung disease”, the incidence proportions of worst CTCAE grade, seriousness, outcome, and action taken will be presented for the total iNHL and histological subpopulations (according to investigator pathology). This table will also be provided for the subgroups of patients with and without history of respiratory disorders (as defined in Section 6.3.4.1).

Tabular and graphical displays of time-of-onset (treatment emergent new onset or worsening of non-infectious pneumonitis) using the time intervals:

- Day 1 – Day 90: to include all copanlisib/placebo-treated patients
- Day 91 – Day 180: to include all copanlisib/placebo-treated patients treated for at least 3 months
- Day 181 – Day 270: to include all copanlisib/placebo-treated patients treated for at least 6 months
- Day 271 – Day 360: to include all copanlisib/placebo-treated patients treated for at least 9 months
- Day 361 or greater: to include all copanlisib/placebo-treated patients treated for at least 12 months

will be presented by treatment group and cancer histology (total iNHL and histological subpopulations according to investigator pathology).

Tabular and graphical displays of time-of-onset of copanlisib/placebo related non-infectious pneumonitis TEAEs (treatment emergent new onset or worsening of non-infectious

pneumonitis consider related to copanlisib/placebo), applying the same time intervals above, will be presented by treatment group for total iNHL.

Summary of patients with treatment-emergent non-infectious pneumonitis/Interstitial lung disease requiring corticosteroids, antibiotics, or both will be displayed descriptively. Refer to the TLF specification document for details of how corticosteroids and antibiotics are identified.

6.3.4 Adverse event groupings

Table 6–11 resents definitions for AE groupings. The overall strategy for the MedDRA searches is aimed at achieving high specificity. When considered to be appropriate, narrow versions of SMQs will be used. Additional Bayer-specific MedDRA groupings will be used for searches when SMQs would not adequately cover the targeted search. The Bayer-specific MedDRA groupings include MedDRA Labeling Groupings (MLG). These groupings are created and centrally maintained by a dedicated coding group, according to requests by end-users from the medical and statistical functions. They are therefore potentially subject to change, e.g., through MedDRA up-versioning.

A table of AE groupings, their definition (MLG/SMQ/MedDRA high level group term (HLGT)/ MedDRA high level term (HLT)) and the observed PTs under this grouping will be provided.

Table 6–11: Adverse Event Groupings

Adverse Event Grouping	MedDRA Search criteria
Arterial hypertension	MLG: Hypertension
Hyperglycemia	MLG: Hyperglycemia
Lymphopenia	MLG: Lymphopenia plus PT: CD4 lymphocytes decreased.
Myocardial ischemia	MLG: Myocardial ischemia and angina pectoris
Neutropenia (incl. Febrile neutropenia)	MLG: Neutropenia plus PT: Febrile neutropenia
Opportunistic Infections	SMQ: Opportunistic infections (narrow, i.e., category 2A)
Respiratory tract infections	HLGT: Respiratory tract infections plus HLT Lower respiratory tract and lung infections, plus HLT: Upper respiratory tract infections
Thrombocytopenia	MLG: Thrombocytopenia
Colitis (excl. infective)	HLT: Colitis (excl infective)
Diarrhoea (excl. infective)	HLT: Diarrhoea (excl infective)
Cytomegalovirus (CMV) infections	HLT: Cytomegaloviral infections plus PT: Cytomegalovirus test positive

Descriptive summary tables (frequency and percentage of patients, not of events) of MedDRA PTs (within each AE grouping) will be presented by treatment group and cancer histology (total iNHL and histological subpopulations) for TEAEs and copanlisib/placebo related TEAEs.

An additional table will be provided for the AE grouping “Opportunistic infections” which will present the incidence of TEAEs within the grouping by HLGT and MedDRA PT, by treatment group and cancer histology (total iNHL and histological subpopulations).

For each of the AE groupings in Table 6–11, descriptive summaries (frequency and percentage of patients, not of events) for treatment-emergent events within the grouping will

be presented by treatment group and cancer histology (total iNHL and histological subpopulations) and will include:

- Incidence proportions of worst CTCAE grades, of seriousness, of outcomes, and of actions taken, 95% Clopper Pearson (exact) confidence interval (CI) for the overall incidence proportion and for the category of serious AEs.
- Incidence proportions of time-of-onset (new onset or worsening), including tabular and graphical displays; with time intervals as follows:
    - Day 1 – Day 90: to include all copanlisib/placebo-treated patients
    - Day 91 – Day 180: to include all copanlisib/placebo-treated patients treated for at least 3 months
    - Day 181 – Day 270: to include all copanlisib/placebo-treated patients treated for at least 6 months
    - Day 271 – Day 360: to include all copanlisib/placebo-treated patients treated for at least 9 months
    - Day 361 or greater: to include all copanlisib/placebo-treated patients treated for at least 12 months

For each of the AE groupings in Table 6–11, descriptive summaries (incidence and percentage of patients, not of events) for copanlisib/placebo-related treatment-emergent events within the grouping will be presented by treatment group and cancer histology (total iNHL and histological subpopulations) and will include:

- Incidence proportions of worst CTCAE grade, of seriousness, of outcomes, and of actions taken, 95% Clopper Pearson (exact) CI for the overall incidence proportion and for the category of serious AEs.
- Incidence proportions of worst CTCAE grade, of seriousness, of outcomes, and of actions taken for the following AE groupings and subgroups:
    - respiratory tract infections by history of respiratory disorders [MedDRA System Organ Class] (no/yes)
    - arterial hypertension by history of hypertension [PBMQ=SMQ]
    - hyperglycemia by history of diabetes,

  as defined above and in Section 6.3.4.1.

6.3.4.1 Subgroups

The following subgroups will be presented, for the AE groupings described in Section 6.3.4:

- sex (Male and Female)
- age group (<65 years and ≥65 years)
- BMI group (<18.5 kg/m², 18.5 - <30 kg/m² and ≥30 kg/m²)
- Duration of progression-free interval on the last rituximab, or rituximab biosimilars, or anti-CD20 monoclonal antibody containing regimen (6-12 months and >12 months)
- Prior to base treatment regimen (R-chemotherapy to R-B vs. R-B to R-CHOP vs. R-B to R-B). Prior to base treatment regimen categories R-B to R-CHOP and R-B to R-B may be combined due to small patient numbers.

6.3.5 Safety OS analysis

In the briefing document for the 21 APR 2022 Oncologic Drugs Advisory Committee (ODAC) meeting [23], the FDA indicates the need for future randomized studies of PI3K inhibitors in hematologic malignancies to assess "the impact [of study drug] on [overall] survival as both a safety and efficacy metric" in order to have a "high likelihood that when a study is completed it can indicate an acceptable benefit-risk profile".

While this CHRONOS-4 study had been planned prior to the 21 APR 2022 ODAC meeting, an OS safety analysis is planned to be performed at the time of primary completion for the primary endpoint of PFS.

The analysis planned is to be considered descriptive/exploratory in nature and is to be used as a pre-specified measure to identify whether caution is warranted regarding the risk of a substantial detrimental OS treatment effect.

This analysis will also be independent of OS as assessed as an efficacy measure and thus is considered to not affect the study alpha.

The analysis performed will analyze OS until 6 months after the end of study treatment. Thus, only OS events until 6 months (i.e., 183 days) after the end of the respective study treatment of patients will be considered. Patients with no OS event at this time will be censored at either their LKAD (if they are lost to follow-up before 183 days after EOT) or at 183 days after the end of study treatment.

Analogous to the efficacy analysis of OS, the stratified cox regression model will be considered. In contrast to the efficacy analysis of OS, the safety analysis of OS will be performed on the SAF.

The CHRONOS-4 study is not powered to detect an OS statistical difference and the restriction to consider only OS events until 6 months after the end of study treatment further restricts statistical power. Thus, 80% CIs will be provided for the HR and the model will be considered as descriptive/exploratory in nature.

In addition to the results of the Cox regression, Kaplan-Meier plots and statistics will be provided for this analysis.

6.3.6 COVID-19 pandemic-relevant adverse events

The COVID-19 pandemic-relevant AEs will be identified using the SMQ "Covid-19" with a broad search. A patient listing of these AEs will be provided. The serious AEs and those leading to treatment discontinuation and/or deaths will be flagged in the listing.

6.3.7 Pregnancies

The results of pregnancy tests will be listed. Any pregnancy and pregnancy outcome will be documented.

6.3.8 Clinical laboratory parameters

Due to differing cycle lengths, by visit summaries will be reported for those assigned to R-B treatment (copanlisib+R-B and placebo+R-B) only. Equivalent data for those assigned to R-CHOP treatment will be listed.

Results of the clinical laboratory evaluations (coagulation, hematology, chemistry and urinalysis) will be summarized by visit, treatment group and overall, with worst CTCAE grade post-baseline. For tables displaying treatment-emergent laboratory abnormalities,

patient specimens collected between the start of treatment and 30 days after EOT (i.e., the date of last study drug) will be included.

Summary statistics on the values and changes from baseline will also be presented for each quantitative clinical laboratory variable by treatment group and overall, at each post-baseline visit. Unscheduled laboratory assessments will not be included in summary analyses over time (by time point), unless they are the last measurement prior to first study drug intake, but will be included in other analyses not summarized by time point, such as shift tables, incidence of abnormalities, etc. For these variables, the number and percentage of patients with transitions from worst grade at baseline to worst grade at post-baseline visits relative to the respective laboratory's reference ranges will be presented by study visit.

Changes in worst grade of laboratory toxicities at last pre-treatment value compared to worst grade post-baseline value and changes in worst grade of laboratory toxicities under treatment will also be provided. For the change in worst grade of laboratory toxicities under treatment, observations until the last on treatment visit will be considered. In case there is more than one observation for the same timepoint, the worst grade will be taken. Unscheduled laboratory assessments will be included in the patient listings of clinical laboratory parameters.

In addition to the above summary, specific glucose evaluations will be described in the next section.

6.3.9 Further safety parameters

Due to differing cycle lengths, by visit summaries will be reported for those assigned to R-B treatment (copanlisib+R-B and placebo+R-B) only. Equivalent data for those assigned to R-CHOP treatment will be listed.

6.3.9.1 Hyperglycemia

Hyperglycemia AEs using specific MedDRA PT grouping "MLG Hyperglycemia" will be summarized by treatment group and overall, by worst NCI CTCAE grade for the following:

- TEAEs
- Drug-related TEAEs
- TESAEs
- Drug-related TESAEs

Glucose measurements on Day 1 of Cycle 1 will be displayed by CTCAE grade of pre copanlisib/placebo infusion and post-dose, 1 hour and 2 hours after the end of copanlisib/placebo infusion, at the end of rituximab infusion and at the end of B infusion[11].

For all other Day 1, Day 8 and Day 15 infusions, glucose measurements will be displayed by CTCAE grade of pre-copanlisib/placebo infusion and post-dose 1 hour after the end of copanlisib/placebo infusion, at the end of rituximab infusion (Day 1), and at the end of B infusion if the patient has received corticosteroid premedication prior to the B infusion[11].

Glucose measurements on Day 2 of Cycles 1 to 6 will be displayed by CTCAE grade prior to B infusion and at the end of the B infusion (within 30 min) if a patient has received corticosteroid premedication prior to B infusion[11]; and for the R-CHOP group prior to

[11] If patient is in the R-B group and receives corticosteroid premedication prior to the bendamustine infusion

rituximab infusion and at the end of rituximab infusion (within 30 min) if the patient has received corticosteroid premedication prior to the rituximab infusion.

Changes from respective pre-dose (defined as the last pre-dose glucose measurement in each visit) in glucose will be summarized using descriptive statistics at each visit and analysis time point by pre-dose fasting status and overall. Unspecified glucose values will not be summarized.

The following plots will be provided overall and for fasting[12] and non-fasting status:

- mean pre-dose glucose values by visits and treatment group
- mean glucose values 1 hour post end of copanlisib/placebo infusion by visits and treatment group
- mean change, 1 hour post end of copanlisib/placebo infusion, from respective pre-dose glucose value, by visits and treatment group

Patients will be categorized according to their HbA1c baseline values as:

- <5.7%
- ≥5.7% – <6.5%
- ≥6.5% – ≤8.5%
- >8.5%.

Within each group, the maximum HbA1c values by cycle and EOT visit[13], as well as 3 months after last dose (if available) will be summarized by descriptive statistics. A plot of mean change from baseline HbA1c value, by visits and treatment group, will be provided for each HbA1c baseline category. Shift tables from baseline to the EOT visit (or last test if this is within 4 weeks preceding the EOT visit), and 3 months after last dose (where available) for each category of HbA1c will be presented.

For patients with hyperglycemia AEs, the number of patients using antihyperglycemic treatment with a drug start date on or after their first TEAE will be summarized overall and separately for the WHO-DD ATC classes for: a) insulin and analogues and b) blood glucose-lowering medications, excluding insulins. In addition, these summaries will be presented separately for the subgroups of patients with versus without a history of diabetes (as defined in Section 6.3.1).

6.3.9.2 Electrocardiogram (ECG)

Results of ECGs will be summarized with descriptive statistics by visit and timepoint, including change from baseline when appropriate.

12-lead ECG evaluation (including QT interval corrected for heart rate – Bazett and QT interval corrected for heart rate – Fridericia) will be performed at screening and at each visit.

[12] Fasting status was collected prior to the infusion so “Pre-dose” refers to fasting status.
[13] HbA1c test should be performed at EOT if not performed within 4 weeks preceding the EOT visit. If no value is given at EOT visit, but their last value is within the allowed 4 weeks prior to the EOT visit, their last value will be used here.

Summary of ECG absolute QTc interval prolongation or QTc increase from baseline will be presented by treatment groups. The criteria are:

QTc prolongation: QTc interval ≤ 450 msec, QTc interval 451-480 msec, QTc interval 481-500 msec and QTc interval > 500 msec.

Increase from baseline in QTc interval increases from baseline < 20 msec, increases from baseline ≥ 20 - < 30 msec, increases from baseline ≥ 30 - < 60 msec and increases from baseline ≥ 60 msec.

The overall interpretation of the 12-lead ECG and the ECG diagnosis will be summarized by visit.

Frequency and shift tables for the number of patients by interpretation of ECG as normal or abnormal will be performed at all visits and analysis time points. A listing of clinically significant abnormal 12-lead ECG findings will be provided.

6.3.9.3 Vital signs

Blood pressure on treatment days will be measured at pre-dose, 30 min (after the start of infusion), right after the end of infusion, and 1 hour and 2 hours after the end of infusion, on infusion days. In addition, on Day 1, blood pressure will be measured 30 minutes after the start of rituximab infusion and at the end of rituximab infusion (window of ± 10 min is allowed[14]) from Cycle 1-6 in the R-B treatment groups.

Results of vital signs will be summarized with descriptive statistics by visit, including change from baseline when appropriate. The number of patients with abnormal post-dose blood pressure will be displayed by visit and post-dose hypertension grade.

Changes from respective pre-dose in blood pressure will be summarized using descriptive statistics at each visit and analysis time point by diastolic and systolic blood pressures (mmHg) and overall.

The following descriptive analysis box plots will be provided for diastolic blood pressure and systolic blood pressure:

- mean pre-dose values by cycle and treatment group
- mean values 30 minutes after start of copanlisib/placebo infusion by cycle and treatment group
- mean values at end of copanlisib/placebo infusion by cycle and treatment group
- mean values 1 hour post end of copanlisib/placebo infusion by cycle and treatment group
- mean values 2 hours post end of copanlisib/placebo infusion by cycle and treatment group
- mean change, 30 minutes after start of copanlisib/placebo infusion, from respective pre-dose value, by cycle and treatment group
- mean change, at end of copanlisib/placebo infusion, from respective pre-dose value, by cycle and treatment group

[14] A time window of ±10 min is allowed for all post-dose blood pressure measurements.

- mean change, 1 hour post end of copanlisib/placebo infusion, from respective pre-dose value, by cycle and treatment group
- mean change, 2 hours post end of copanlisib/placebo infusion, from respective pre-dose value, by cycle and treatment group

6.3.9.4 Echocardiogram

Results of multiple gated acquisition (MUGA)/echocardiogram will be summarized with descriptive statistics by visit, including change from baseline where appropriate.

The overall interpretation of the echocardiogram and the corresponding diagnosis will be summarized by visit.

Frequency and shift tables for number of patients by interpretation of echocardiogram as normal, abnormal CS or abnormal not CS will be performed at all visits. A listing of clinically significant abnormal cardiological findings will be provided.

6.3.9.5 Hypertension

Hypertension AEs will be presented using MLG grouping of MLG Hypertension for the following:

- TEAEs
- Drug-related TEAEs
- TESAEs
- Drug-related TESAEs

For patients with hypertension AEs, the number of patients using antihypertensive treatment with a drug start date on or after their first TEAE of hypertension will be summarized descriptively. In addition, the number of patients requiring new antihypertensive treatment will be presented separately for the subgroups of patients with versus without a history of hypertension (as defined as Section 6.3.1). The number of patients with abnormal blood pressure (according to CTCAE grading) will be displayed by visit and worst post-infusion hypertension grade.

6.3.9.6 Integrated analysis

Details of the pooling strategy will be described in an Integrated Analysis - Statistical Analysis Plan (IA-SAP). Results of these analyses will be presented in a separate report.

6.4 Pharmacokinetics/pharmacodynamics

The data from the Safety run-in part and Phase III part will be pooled for PK analysis.

Individual concentration-time data of copanlisib and its metabolite M-1 as well as rituximab will be listed only. Further analyses will be described in a separate Modeling & Simulation (M&S) Plan and performed by a PK expert (pharmacometrics) and will include population pharmacokinetics, and exposure-response analysis.

The pharmacokinetic serum exposure of rituximab will be compared to predictions from a published population PK model of rituximab [20]. The results will be reported together with the population PK and exposure-response of copanlisib in the M&S Report.

6.5 Biomarker evaluation

All biomarker analyses will be considered exploratory. There will not be an overall or sensitivity analysis of central pathology data. Biomarker data collected in the clinical database will be listed. Retrospective exploratory biomarker analyses including evaluation of relationship between biomarker data and clinical outcomes will be performed by or under the supervision of the sponsor's biomarker statistical expert. Results of these analyses will be reported separately, under a separate cover.

6.6 Determination of sample size

EAST 6.4 was used to calculate sample size. The sample size calculation is based on the primary efficacy endpoint, PFS, in the FAS and was performed exclusively for the Phase III part.

The sample size calculation assumes that the median PFS in the control arm (base regimen either R-B or R-CHOP) is 24 months. With regard to R-B, a Phase 2 study in relapsed indolent B-cell (82% of patients) and mantle cell (18% of patients) lymphoma was considered for median PFS [21]. This study had, for the full population, a median PFS of 22.9 months, but the median PFS was not reported per indication. As the study population was a mixture of B-cell and mantle cell lymphoma, a more conservative median PFS of 24 months was assumed for R-B treatment. With regard to R-CHOP, a study in 16 relapsed FL patients was considered for the median PFS assumption due to the limited data available for R-CHOP in iNHL patients at the time of this assessment. The median PFS in the 13 FL patients not receiving autologous stem cell transplant was extrapolated based on the disease survival curve of patients who achieved a CR. Based on this extrapolation, a median PFS of around 20 months was assumed for R-CHOP in FL patients [22]. Building on these studies and considering that majority of the patients in this study are on R-B, a combined median PFS of 24 months was considered as a reasonable assumption for the control arm.

The HR was developed by considering copanlisib monotherapy data and publicly available information as of JAN 2016 (study inception), in studies in which treatments had been evaluated in similar indications. Sample size planning for CHRONOS-4 targeted an increase of 50% in median PFS time in the copanlisib + R-B/R-CHOP combination arms. This equates to a targeted median PFS time of 36 months and results in a targeted HR of 0.667.

Assuming a median PFS of 24 months under the control treatment, a 1-sided alpha of 0.025, a power of 90%, and a randomization ratio of 1:1 between the experimental and control arms, 256 events (progression based on blinded independent central review or death if death occurs before progression) are required to detect a 50% increase in PFS. The monthly dropout (loss to follow-up and unevaluable for tumor assessment) rate is assumed to be 0.48%, which corresponds to a total number of 52 (10%) unevaluable/dropout patients over the duration of the study through to 256 PFS events.

A total of approximately 520 patients were to be accrued and randomized in the two treatment groups combined. The actual total number of patients enrolled and randomized was 524.

Due to health authority (HA) interest in a powered result for PFS in the FL population that was received after all randomized patients had completed study treatment, primary completion has been extended to when approximately 280 PFS events are reached in the iNHL population. At this point it is expected that there will be sufficient PFS events to achieve an

adequately powered comparison between the experimental and control arms, with an expected number of approximately 192 PFS events in the FL population.

The expected study duration to reach approximately 280 PFS events is at least 52 months, and the expected accrual duration was 33 months with a 15-month linear ramp-up (maximum accrual rate being approximately 20 patients per month).

7. Document history and changes in the planned statistical analysis

7.1 Document history

- SAP final version 1.0, dated 24 FEB 2021
- SAP final version 2.0, dated 08 AUG 2022
  - Section 1, deletion of unnecessary information regarding R-B and R-CHOP completion times.
  - Section 3, amended to past tense and to include definitions of prior to base treatment regimen, start of the treatment period and patient's last visit date. Reordered study periods section and added clarification that re-screened patients must re-sign the informed consent form. Detail on tumor assessments from Active follow-up being included in the primary and secondary analyses has also been moved from Section 4.2. The text "the estimated median OS time will not change in further follow-ups, i.e., all patients censored by the cut-off date have a censored OS time greater than the estimated median OS time in both arms" has been removed as an of "End of Study" condition.
  - Section 4.1, addition of reference to Oncology Common Statistical Language document. Clinical Copanlisib Project Standards removed as they do not exist.
  - Section 4.2, clarification added that those entering Active follow-up will not be defined as a "dropout".
  - Section 4.4 and throughout, wording "approximately 256 centrally evaluated PFS events' updated to 'at least 256 centrally evaluated PFS events".
  - Section 4.5.3, amended to include decision on how to handle duplicate records in analysis datasets.
  - Section 4.5.4, amended to include option of prednisolone.
  - Section 4.5.5, rules added for including QoL questionnaires in analysis of efficacy
  - Section 4.5.6, stratification factors expanded in Table 4-1 for clarity and consistency with Section 3.
  - Section 4.5.7, amended to include the term "regions" in addition to rest of world countries.
  - Section 5.1, reordered final two paragraphs to include these under "Assignment of analysis sets' and "Safety analysis set (SAF)" headings.
  - Section 6.1, moved specification of treatment groups from Section 6.1.2.

- Section 6.1.1, clarified that the number of patients assigned to treatment will be presented. Additional text added regarding how the number of countries/regions and centers enrolling patients will be presented. Tables added for duration of follow-up and summarizing number of patients completing endpoint follow-up, having an event, and being censored. Amended section to include definitions for time to end of study treatment and time to cause specific treatment discontinuation, and cumulative incidence function curves of time to cause specific treatment discontinuations and time to treatment discontinuation due to AEs. Additionally, further details have been provided for the COVID-19 pandemic related protocol deviations listing and a listing of protocol deviations related to the Ukraine-Russia conflict added.
- Section 6.1.2, amended age group ordering, numbered each age group, and removed duplication of "Histology group by investigator at baseline". Hepatic function at baseline added to demography. NYHA classification, HbA1c group, eGFR group, medical history of diabetes and medical history of hypertension moved from baseline cancer characteristics to demography.
- Section 6.1.3, on-going medical history conditions added.
- Section 6.1.4, amended to include mean number of prior systemic anti-cancer therapy lines, the definition of time since last systemic anti-cancer therapy and further detail on what will be reported for prior and concomitant medications. Last category of prior lines has also been amended from '≥4 lines of prior therapy' to "≥3 lines of prior therapy". Prior and concurrent therapies bullet points have been reordered for clarity, with "Radiotherapies during long term follow-up" changed to 'Radiotherapies during Survival follow-up' as palliative radiotherapies only collected during Survival follow-up. The following text was added for clarity "Medications starting more than 30 days after the last day of study treatment will be excluded from the summary of concomitant medications".
- Section 6.1.5, amended to state that number of patients with each number of cycles will also be summarized, summaries by histology added and OEE categories have been amended from "271-360 days" and "361-390 days" to "271-365 days" and ">365 days" respectively. By cycle summaries changed to by calculated cycle (see ADRG).
- Section 6.2.1, amended to include footnote detailing how independent and local assessments for patients with LPL/WM will be used, and text split and assigned to subsections 6.2.1.1, 6.2.1.2, 6.2.1.3 and 6.2.1.4.
- Table 6-1, Table 6-2 and Table 6-3, censoring rules updated.
- Section 6.2.1.1, amended to include descriptive summary of PFS event types and PFS censoring reasons and to clarify how stratification factors will be included in the analysis of PFS.
- Section 6.2.1.4, Sensitivity Analysis 6 amended to specify that only patients not at risk will be considered as events, Table 6-2 has been amended to add more scenarios and Sensitivity Analysis 8 has been added.
- Section 6.2.2, TTNT updated, and table of censoring rules added.

- Section 6.2.2, time to improvement in DRS-P definition changed from days to months and OS follow-up analysis timing changed from 24 months after analysis of the primary and secondary efficacy variables, to 12 months after.
- Section 6.2.2 and 6.2.3, amended to include OS in confirmatory hierarchical testing.
- Section 6.2.4, "time to onset of physical symptoms on the DRS-P subscale" changed to "time to onset of meaningful worsening of physical symptoms on the DRS-P subscale" as this is a more precise definition. The content of the variable has not changed.
- Section 6.2.5.2, prior lines of therapy categories updated to "1 vs. 2 vs. ≥3" from "1 vs. 2 vs. 3 vs. ≥4".
- Section 6.3.1, summaries of non-infectious pneumonitis/interstitial lung disease removed as pneumonitis/interstitial lung disease are now in Section 6.3.3, AEs of special interest. Clarification added that all grade 5 TEAEs will be included when summarizing deaths, and AEs occurring prior to starting treatment will also be summarized. Table of deaths updated as deaths with "AE" or "other" as the primary cause will be pooled together under the primary cause "other". Additional detail on what will be reported in each patient listing has also been added. History of hypertension and history of diabetes mellitus PBMQs updated from MedDRA 21.1 to MedDRA v25.0. Drug-related listings of death added.
- Section 6.3.2, amended to include footnote specifying study procedure related AEs and SAEs occurring during Active follow-up will also be included in each time-interval. Typo of 'Day 181 or greater' has also been corrected to 'Day 181 – Day 270'. Interval summaries and for R-B and R-CHOP related TEAEs and SAEs added.
- Section 6.3.3, new section on adverse events of special interest added.
- Section 6.3.4, new section on AE groupings added.
- Section 6.3.5, Safety OS analysis added.
- Section 6.3.6, amended details on how COVID-19 relevant AEs will be identified and reported.
- Section 6.3.7, amended to add pregnancy outcomes will also be reported.
- Section 6.3.8, clarified that EOT is defined as the date of last study drug.
- Section 6.3.9.1, added footnote and details within text to clarify which HbA1c test values should be used for the EOT time point. Additional summary plots added.
- Section 6.3.9.3, summary of the number of patients with abnormal post-dose blood pressure by visit and post-dose hypertension grade added.
- Section 6.3.9.2, moved echocardiogram related summaries to Section 6.3.9.4.
- Section 6.3.9.2, 6.3.9.3 and 6.3.9.4, amended to align text with what will be presented.
- Section 6.3.9.3, additional summary plots added.
- Section 6.3.9.6, "Integrated analysis" added.
- Section 6.6, amended to include additional detail on sample size assumptions for the median time on control arm and the targeted HR.
- Section 8, amended to include 2 references added to Section 6.6,1 reference moved from Section 6.4 and 1 reference added to Section 6.3.5.

- Section 9.3.2 The following text was added “A death event is a PFS event if the date of death is ≤ 25 (12+12+1) weeks after randomization.”.

- SAP final version 3.0, dated 04 APR 2023
  - Section 1, version of integrated clinical study protocol updated to Amendment 09, protocol Version 8.0, dated 21 MAR 2023.
  - Section 2, Table 2-1 updated to reflect two testing hierarchies, one for United States, and one for Europe and rest of world.
  - Section 3, details of schedule of tumor assessment in active follow-up, end of study definition and the number of PFS events required for primary completion analysis, updated based on integrated clinical study protocol (CSP, Amendment 09, protocol Version 8.0, dated 21 MAR 2023), to reflect tumor assessments beyond year 5, the potential for patients to be followed off-study for up to 10 years (previously 5 years) and that approximately 280 centrally evaluated PFS events (previously 256) are required for the primary completion analysis.
  - Section 4.1, updated based on integrated clinical study protocol (CSP, Amendment 09, protocol Version 8.0, dated 21 MAR 2023), to reflect statistical evaluation of this study will be performed when approximately 280 centrally evaluated PFS events are observed (previously 256).
  - Section 4.3, updated to list all key panels used to determine survival status and LKAD. Text added to the explain the method and purpose of imputing missing or partially missing start or end dates for adverse events and concomitant medications.
  - Section 4.4, updated, based on integrated clinical study protocol (CSP, Amendment 09, protocol Version 8.0, dated 21 MAR 2023), to reflect the potential for patients to be followed off-study for up to 10 years (previously 5 years) and the amended number of PFS events required for the primary completion analysis. Text relating to DMC updated to past tense.
  - Section 4.5.6 text detailing relationship between strata and patient ID updated for clarity, after FDA review comment.
  - Section 4.5.7, further detail added to describe groupings of geographic regions and how these will be analyzed.
  - Section 6.1, text regarding differing FAS and SAF updated to align with Section 5.1.
  - Section 6.1.1, text describing number of patient summaries updated for clarity.

    Text added to avoid double counting patients, who passed re-screening, in number of screening failures and number of randomized patients.

    Additional listings of re-screened patients and “Other” reasons for premature discontinuations from epochs added.

    Summary of “time to treatment discontinuation due to AEs” amended to “time to treatment discontinuation due to AEs not associated with clinical disease progression”.

Duration of study participation to last visit and to LKAD replaced with summary of PFS follow-up and OS follow-up by treatment group using the inverse event indicator method.

Tables added summarizing number of patients with protocol deviations associated with COVID-19 pandemic, overall and by protocol deviation category number, and number of patients affected by COVID-19 pandemic related study disruptions.

- Section 6.1.2, "other" added to "baseline characteristics", and definition of moderate impairment changed from "40- <60" to "30- <60".

  Description of diabetes and hypertension medical history identification updated for clarity.

  Number of LPL/WM patients with and without measurable disease at baseline added to summary of other baseline characteristics.

  Analysis population defined (FAS) for by histology (FL, MZL, SLL, and LPL/WM) summaries of demographic and other baseline characteristics.

- Section 6.1.4, text amended to clarify the definition of prior medication.

  Frequency table added for prior anti-cancer therapy and therapeutic procedures.

- Section 6.1.5, added text to describe patients in the R-CHOP arm will have number of infusions calculated based on rituximab, cyclophosphamide, doxorubicin, and vincristine.

- Section 6.2, updated, based on integrated clinical study protocol (CSP, Amendment 09, protocol Version 8.0, dated 21 MAR 2023), to amend number of PFS events required for the primary completion analysis.

- Section 6.2.1, removal of "progression-free survival" from heading, removal of text describing imputation method when tumor assessment date is partially missing and replaced with a reference to Section 4.3, as this is repeated text.

  First censoring scenario updated to include "or death more than 25 (12+12+1) weeks after randomization".

  Censoring scenario updated from "discontinued or withdrew early from treatment" to "discontinued or withdrew early from study".

  Explanatory footnote added to Table 6-1 for the scenario where progression occurs on the same date as start of new anti-cancer treatment.

- Section 6.2.1.1, updated, based on integrated clinical study protocol (CSP, Amendment 09, protocol Version 8.0, dated 21 MAR 2023), to amend number of PFS events required for the primary completion analysis.

- Section 6.2.1.3, text added to detail follow-up analyses of PFS and other efficacy endpoints at timepoints for follow-up OS.

- Section 6.2.1.4.7, 6.2.1.4.8, Explanatory footnote added to Table 6-2 and Table 6-3 for the scenario where progression occurs on the same date as start of new anti-cancer treatment.

- Section 6.2.2, added a clarification that new anti-cancer therapy ends the Active follow-up period.

Added text “and no death related to PD within the first 25 weeks after randomization” under TTP heading.

Added the word “study” to “treatment” under TTNT heading.

Reason for censoring added to scenario in Table 6-5, to distinguish row from “Neither DRS-P, nor PD/death and no other criteria fulfilled”.

Removed repetition of “censoring rules” from time to improvement section.

Removed subheading “Analysis of ORR, time to deterioration, time to improvement in DRS-P” and wording updated to reflect image based secondary efficacy endpoints.

Added clarification that local assessments will be used in place of centrally reviewed data for LPL/WM patients without measurable disease at baseline.

Added reference to Appendix 9.1, detailing how PET-CT investigator assessments will be combined.

Updates to analysis of OS section, to include a final descriptive analysis of OS at end of study.

Added censoring rules for DRS-P sensitivity analysis to Table 6-7.

- Section 6.2.2 and 6.2.3, amended to include two testing hierarchies: one for US and one for Europe and rest of world. US hierarchy updated to include PFS in FL population and to remove time to deterioration and time to improvement in DRS-P.
- Section 6.2.3, added reference to Appendices 9.6, 9.7 and 9.8, which describe further assessments of OS data and scenarios for OS (answering an HA request).
- Section 6.2.4, 6.3.8 and 6.3.9, clarified that by visit summaries will be reported for those assigned to R-B only.
- Section 6.2.5.2, updated in include a stratified Cox-PH model for the FL histology subgroup.
- Section 6.3.1, all references to MedDRA v25.0 updated to v25.1.

  The number and percentage of patients who discontinued study treatment due to a TESAE (discontinued all study treatments) was added.
- Section 6.3.2, updated text to add treatment emergent AEs, as well as treatment emergent drug-related AEs and TESAEs.

  Removal of footnote 12, as all treatment emergent AEs will be summarized.
- Section 6.3.4, addition of CMV infections to AE Groupings.
- Section 6.3.4.1, definition of “Prior to base treatment regimen” updated to reflect same definition as in Section 6.2.5.
- Section 6.3.9.1, text relating to blood pressure removed and added to Section 6.3.9.3.
- Section 6.6, sample size text updated to reflect changes to primary completion definition in integrated CSP amendment 09, version 8.0, dated 21 MAR 2023.

Text added to specify 192 events in the FL population are expected at the time of primary completion.

- Appendix 9.1, principles for combination of PET and CT/MRI responses, added.
- Appendices 9.4.1 and 9.4.2, details of schedule of tumor assessment in active follow-up updated based on integrated clinical study protocol (CSP, Amendment 09, protocol Version 8.0, dated 21 MAR 2023), to reflect tumor assessments beyond year 5.
- Appendix 9.5.1, algorithm added for handling partial initial diagnosis date and progression date.
- Appendix 9.5.2, imputation rules added for partial/missing Adverse Event/Concomitant medication dates.
- Appendix 9.5.3, imputation rules added for partial/missing anti-cancer treatment start dates.
- Appendix 9.6, added to provide reassurance, regarding OS safety, following an HA request.
- Appendix 9.7, added to provide reassurance, regarding OS safety, following an HA request.
- Appendix 9.8, added to provide reassurance, regarding OS safety, following an HA request.

- SAP draft version 4.0, dated 22 SEP 2023
  - MedDRA v25.1 updated to “v25.1 or higher” as v25.1 has been superseded.
  - Specified throughout that measurable lesions assessed by investigator are required for LPL/WM patients.
  - Section 4.3, additional datasets added to determine survival status and the LKAD.
  - Section 4.4, added in case a background therapy has less than 5 randomized patients assigned.
  - Section 4.5, added in case further investigator assessments are taken following investigator reported PD.
  - Section 4.6, added to clarify how clinical laboratory values with qualifiers will be included in analyses and summaries.
  - Section 4.7, added to clarify that p-values are nominal, except for the primary efficacy endpoint and confirmatory statistical testing strategy.
  - Section 4.9.4, R-B and R-CHOP overall extent of exposure updated to handle study drug combinations with differing start and end dates in the final cycle.
  - Section 6.1.5, amended to clarify interruptions are distinguished from delays by non-zero dose administered, and delays from interruptions by zero dose administered. Text added to explain that interruptions with the same start date will be counted as one interruption. Duration of interruption has been removed from the summaries, only duration of delay will be summarized.

The number of days used to calculate the planned dose for incomplete cycles has been amended to take in to account the '-1 to +2 days' time window for study drug administration.

- Section 6.2.1.3, footnote (9) amended to use a corrected derivation for the date of clinical progression.
- Section 6.2.1.4, Sensitivity Analysis 5 moved to end of section and re-numbered as Sensitivity Analysis 8. Dose delays due to the Ukraine-Russia conflict and related analyses added to re-numbered Sensitivity Analysis 8. Previous Sensitivity Analyses 6, 7 and 8 updated and re-numbered accordingly.

  Table 6–2 and Table 6–3 updated to correct Appendix references.
- Section 6.2.2, definition of DCR amended to remove references to "within 35 days after termination of study treatment" and "within the following 25 weeks". Also refer to Section 7.2 ("Changes in the planned statistical analysis").

  Definition of TTNT amended from "stop of study treatment to start of new anti-lymphoma therapy" to "date of randomization to start of new anti-lymphoma therapy" and censoring rules updated. This change was made because the definition of TTNT in Section 10.3.2.3 of CSP, Amendment 09, protocol Version 8.0, dated 21 MAR 2023 is incorrect. Also refer to Section 7.2 ("Changes in the planned statistical analysis"). Table 6–4 also updated to remove "LKAD" when patient died during treatment, Safety follow-up or Active follow-up.
- Section 6.2.3, additional information on conditional power added.
- Section 6.3.3, tabular and graphical displays added for time-of-onset of copanlisib/placebo related non-infectious pneumonitis TEAEs.
- Section 9.5.3, Table 9–2, updated and footnote added.
- Section 9.7, updated to use a benchmark number of OS events under 80% power to calculate the information fractions instead of a target number of events.
- Section 9.8, Table 9–5, probability of point estimate / upper bound exceeding 1 added for final formal OS analysis, projected number of events as a percentage is recalculated relative to benchmark number of events (380) instead of expected number of events at final formal OS analysis (176), results from additional simulations with a true HR=0.75 added.

7.2 Changes in the planned statistical analysis

- The secondary efficacy variable TTNT is incorrectly defined in Section 10.3.2.3 of CSP, Amendment 09, protocol Version 8.0, dated 21 MAR 2023.

Incorrect definition of TTNT as defined in current CSP, Amendment 09, protocol Version 8.0, dated 21 MAR 2023

TTNT is defined as the time from stop of study medication to start of new anti-lymphoma therapy.

Corrected definition of TTNT as specified in Section 6.2.2 of this SAP

TTNT is defined as the time from date of randomization to start of new anti-lymphoma therapy, where date of randomization is day 1.

The intention is to correct the TTNT definition in the next CSP amendment. However, if primary completion occurs prior to the next CSP amendment the correct definition will be specified in the SAP and used in the analysis, with the divergence from the CSP definition also documented in the SAP and CSR as "changes in the planned statistical analysis".

- The definition of the secondary efficacy variable DCR has been amended in the SAP to remove references to "within 35 days after termination of study treatment" and "within the following 25 weeks". This change has been made to be consistent with the ORR and CRR assessment period, by including in the DCR analysis, radiological assessments performed during the Active follow-up.

  Definition of DCR as defined in current CSP, Amendment 09, protocol Version 8.0, dated 21 MAR 2023

  DCR is defined as the proportion of patients who have a best response rating of CR, VGPR, PR, MR, or stable disease (SD) (excluding unconfirmed early SD) that is achieved during treatment or within 35 days after termination of study treatment. An unconfirmed early SD is defined as SD on or before Study Day 76, with no additional tumor assessment of SD or better within the following 25 weeks.

  SAP definition of DCR as specified in Section 6.2.2 of this SAP

  DCR is defined as the proportion of patients who have a BOR of CR, unconfirmed CR, VGPR, PR, unconfirmed PR, MR, or confirmed stable disease (SD). Unconfirmed SD is defined as SD on or before Study Day 76, which is not confirmed by an additional tumor assessment of SD, or better, on Study Day 77 or later.

8. References

1. Shankland KR, Armitage JO, Hancock BW. Non-Hodgkin lymphoma. Lancet. 2012; 380(9844):848-57.
2. Hiddemann W, Kneba M, Dreyling M, Schmitz N, Lengfelder E, Schmits R, et al. Frontline therapy with rituximab added to the combination of cyclophosphamide, doxorubicin, vincristine, and prednisone (CHOP) significantly improves the outcome for patients with advanced-stage follicular lymphoma compared with therapy with CHOP alone: results of a prospective randomized study of the German Low-Grade Lymphoma Study Group. Blood. 2005;106(12):3725-32.
3. Marcus R, Imrie K, Solal-Celigny P, Catalano JV, Dmoszynska A, Raposo JC, et al. Phase III study of R-CVP compared with cyclophosphamide, vincristine, and prednisone alone in patients with previously untreated advanced follicular lymphoma. J Clin Oncol. 2008;26(28):4579-86.
4. Rummel M, Kaiser U, Balser C. Bendamustine plus rituximab versus fludarabine plus rituximab in patients with relapsed follicular, indolent and mantle cell lymphomas - final results of the randomized phase III study NHL 2-2003 on behalf of the StiL (Study Group Indolent Lymphomas, Germany) Blood. 2010(116).
5. Hainsworth JD, Litchy S, Burris HA, 3rd, Scullin DC, Jr., Corso SW, Yardley DA, et al. Rituximab as first-line and maintenance therapy for patients with indolent non-Hodgkin's lymphoma. J Clin Oncol. 2002;20(20):4261-7.
6. McLaughlin P, Grillo-Lopez AJ, Link BK, Levy R, Czuczman MS, Williams ME, et al. Rituximab chimeric anti-CD20 monoclonal antibody therapy for relapsed indolent lymphoma: half of patients respond to a four-dose treatment program. J Clin Oncol. 1998;16(8):2825-33.
7. Johnston A, Bouafia-Sauvy F, Broussais-Guillaumot F, Michallet AS, Traulle C, Salles G, et al. Retreatment with rituximab in 178 patients with relapsed and refractory B-cell lymphomas: a single institution case control study. Leuk Lymphoma. 2010;51(3):399-405.
8. Hainsworth JD, Litchy S, Shaffer DW, Lackey VL, Grimaldi M, Greco FA. Maximizing therapeutic benefit of rituximab: maintenance therapy versus re-treatment at progression in patients with indolent non-Hodgkin's lymphoma--a randomized phase II trial of the Minnie Pearl Cancer Research Network. J Clin Oncol. 2005;23(6):1088-95.
9. Domingo-Domènech E, González-Barca E, Estany C, Sureda A, Besalduch J, Fernández de Sevilla A. Combined treatment with anti-CD20 (rituximab) and CHOP in relapsed advanced-stage follicular lymphomas. Haematologica. 2002;87(11):1229-30.
10. Rummel MJ, Al-Batran SE, Kim SZ, Welslau M, Hecker R, Kofahl-Krause D, Josten KM, Dürk H, Rost A, Neise M, von Grünhagen U, Chow KU, Hansmann ML, Hoelzer D, Mitrou PS. Bendamustine plus rituximab is effective and has a favorable toxicity profile in the treatment of mantle cell and low-grade non-Hodgkin's lymphoma. J Clin Oncol. 2005 May 20;23(15):3383-9.

11. Ihle NT, Powis G. Take your PIK: phosphatidylinositol 3-kinase inhibitors race through the clinic and toward cancer therapy. Molecular cancer therapeutics. 2009;8(1):1-9.
12. Yuan TL, Cantley LC. PI3K pathway alterations in cancer: variations on a theme. Oncogene. 2008;27(41):5497-510.
13. Cheson, BD, Fisher, RI, Barrington, SF, Cavalli, F, Schwartz, LH, et al. Recommendations for Initial Evaluation, Staging, and Response Assessment of Hodgkin and Non-Hodgkin Lymphoma: The Lugano Classification. J Clin Oncol. 2014; 32(27):3059-67.
14. Owen, RG, Kyle, RA, Stone, MJ, Rawstron, AC, Leblond, V, et al. Response assessment in Waldenstrom macroglobulinaemia: update from the VIth International Workshop. British journal of haematology. 2013; 160(2):171-6.
15. Data monitoring committee charter. version 1.0, 24 March 2017
16. Schemper M and Smith TL. A note on quantifying follow-up in studies of failure time. Controlled Clinical Trials. 1996; 17(4):343-346.
17. Collett D. Modelling Survival Data in Medical Research: Chapman and Hall; 2003.
18. FDA. Clinical Trial Endpoints for the Approval of Non-Small Cell Lung Cancer Drugs and Biologics Guidance for Industry, April 2015. Available from: https://www.fda.gov/downloads/drugs/guidances/ucm259421.pdf (accessed date: 18 MAY 2020)
19. Koch GG, Carr GJ, Amara IA, Stokes ME, Uryniak TJ. 'Categorical Data Analysis' in Statistical methodology in the pharmaceutical sciences /edited by Berry DA, Marcel Dekker; 1990.
20. Rozman S et al. Population pharmacokinetics of rituximab in patients with diffuse large B-cell lymphoma and association with clinical outcome. Br J Clin Pharmacol 2017; 83: 1782-1790. DOI:10.1111/bcp.13271
21. Robinson KS, Williams ME, van der Jagt RH, Cohn P, Herst JA, Tulpule A, et al. Phase II multicenter study of bendimustine plus rituximab in patients with relapsed indolent B-cell and mantle cell non-Hodgkin's lymphoma. J Clin Oncol. 2008 Sep 20;26(27):4473-9.
22. Domingo-Domenech E, Gonzalez-Barca E, Estany C, Sureda A. Besalduch J, Fernandez de Sevilla A. Combined treatment with anti-CD20 (rituximab) and CHOP in relapsed advanced-stage follicular lymphomas, Haematologica. 2002 Nov;87(110);1229-30.
23. FDA. Phosphatidylinositol 3-kinase (PI3K) inhibitors in hematologic malignancies. Briefing document, Oncologic Drugs Advisory Committee Meeting, 21 APR 2022.
24. Rummel M, Kaiser U, Balser C, Stauch M, Brugger W, Welslau M, et al. Bendamustine plus rituximab versus fludarabine plus rituximab for patients with relapsed indolent and mantle-cell lymphomas: a multicentre, randomised, open-label, non-inferiority phase 3 trial. Lancet Oncology. 2016 January; 17(1):57-66.

25. Rueda A, Calvo V, Casanova M, Rodriguez-Abreu D, Aguiar D, et al. Efficacy of the combination of rituximab-bendamustine as a second-line treatment in patients with follicular lymphoma who progress after immunochemotherapy: a phase II trial of the Spanish Lymphoma Oncology Group. Leuk Lymphoma. 2019 Jun; 60(6):1576-1579. Epub 2018 Dec 5.

9. Appendix

9.1 Principles for combination of PET and CT/MRI responses

- FDG-PET response overrides, in most cases, CT/MRI response

  e.g. If CT response = PD and PET response = CMR → Overall response is CR

- If an FDG-PET assessment is not within a ±7 day time window of a CT/MRI assessment, then the FDG-PET assessment will be considered as a standalone assessment.
- After a combined response of FDG-PET and CT/MRI is derived for a particular FU time point it is highly recommended that subsequently any change of overall response that is derived from CT or MRI alone is confirmed with PET.
- Any following time points assessed with CT or MRI alone are affected by a prior FDG-PET result, with no time limit for carrying forward an FDG-PET result, i.e.
  1. Prior PET response will determine further FU overall responses, as long as CT response does not get worse compared to previous CT responses

     e.g. FU 3: PET (NMR), CT (SD) → overall response SD;

     FU 4: No PET, CT (PR) → overall response SD;

     FU 5: PET (CMR), CT (PR) → overall response CR
  2. A worse CT response will determine FU overall response

     e.g. FU 3: PET (CMR), CT (PR) → overall response CR;

     FU 4: No PET, CT (SD) → overall response SD;

     FU 5: PET (PMD), CT (SD) → overall response PD

After a combined response of FDG-PET and CT/MRI is derived for a particular FU time point PET should be performed to confirm CR and/or disease progression.

9.2 Scoring instruction for FLymSI-18

Record answers of each item response (0 = 'Not at all' to 4 = 'Very much'). In order to have a common order for the scores with 0 reflecting worst outcome and 4 reflecting best outcome, reverse the coding for the following items by subtracting the response from '4':

GP1: I have lack of energy

GP4: I have pain

C2: I am losing weight

Leu1: I am bothered by lumps or swelling in certain parts of my body

MBT6: I get tired easily

BP1: I have bone pain

HI8: I have trouble concentrating

GE6: I worry that my condition will get worse

BRM9: I have emotional ups and downs

Leu4: Because of my illness, I have difficulty planning for the future

Leu5: I feel uncertain about my future health

GP2: I have nausea

N3: I worry about getting infections

GP5: I am bothered by side effects of treatment

FLymSI-18 Total score (range 0-72)

Sum the (reversed, if applicable) responses of all items. Multiply by 18 (number of items in the FLymSI-18) and divide by the number of items answered.

FLymSI-18 DRS-P score (range 0-36)

Sum the (reversed, if applicable) responses of the items GP1, GP4, C2, Leu1, BMT6, BP1, HI8, GF5, C6. Multiply by 9 (number of items in the FLymSI-18 DRS-P score) and divide by the number of items answered.

FLymSI-18 DRS-E score (range 0-16)

Sum the reversed responses of the items GE6, BRM9, Leu4, Leu5. Multiply by 4 (number of items in the FLymSI-18 DRS-E score) and divide by the number of items answered.

FLymSI-18 TSE score (range 0-12)

Sum the reversed responses of the items GP2, N3, GP5. Multiply by 3 (number of items in the FLymSI-18 TSE score) and divide by the number of items answered.

FLymSI-18 FWB score (range 0-8)

Sum the responses of the items GF3, GF7. Multiply by 2 (number of items in the FLymSI-18 FWB score) and divide by the number of items answered.

9.3 Patient-reported outcomes

Please refer to Patient-Reported Outcomes Dossier in Support of the National Comprehensive Cancer Network—Functional Assessment of Cancer Therapy Lymphoma Symptom Index-18 (NFLymSI-18) in Indolent B-Cell Non-Hodgkin's Lymphoma (iNHL) Patients by (Department of Medical Social Sciences, Feinberg School of Medicine, Northwestern University).

9.4 PFS rules

9.4.1 Progression and PD after two or more consecutive missed tumor assessments

If the progression occurs at the next tumor assessment after two or more consecutive missing or non-evaluable assessments, PFS will be censored at the date of the last evaluable scan before the consecutive missing assessments. Considering tumor assessment are scheduled at

every 12 weeks (± 7 days) from Cycle 1 Day 1 during Year 1 and 2; every 24 weeks (± 7 days) during Year 3, 4 and 5; and every 24 weeks (±14 days) beyond Year 5, the following conditions will apply to account for the tumor assessment schedule change starting at year 2 and year 3.

For patients with progression observed before death with an evaluable post-baseline evaluation before two missing assessments:

- if date of progression is ≤ 117 weeks (104 weeks+12 weeks + 1 week) after randomization and > last evaluable assessment + 2*12 + 1 weeks, PFS will be censored at the date of the last evaluable scan BEFORE the two missing assessments.
- if date of progression is > 117 weeks (104 weeks +12 weeks + 1 week) and ≤ 141 weeks (116 weeks +24 weeks + 1 week) after randomization and > last evaluable tumor assessment + 12+24 + 1 weeks, PFS will be censored at the date of the last evaluable scan BEFORE the two missing assessments.
- if date of progression is > 141 weeks after randomization and > last evaluable tumor assessment + 2*24 + 1 weeks during Year 3, 4 and 5; or > last evaluable tumor assessment + 2*24 + 2 weeks beyond Year 5, PFS will be censored at the date of the last evaluable scan BEFORE the two missing assessments.
- otherwise, the PFS time will not be censored and be calculated as the date of progression minus the date of randomization.

9.4.2 Death

Considering tumor assessments are scheduled at every 12 weeks (± 7 days) from Cycle 1 Day 1 during Year 1 and 2; every 24 weeks (± 7 days) during Year 3, 4 and 5; and every 24 weeks (±14 days) beyond Year 5, the following conditions will apply:

A death event is a PFS event if the date of death is ≤ 25 (12+12+1) weeks after randomization.

For patients who die after the last evaluable tumor assessment, in the absence of progression, the death event is a PFS event in the following conditions:

- if date of death is ≤ 117 weeks (104 weeks+12 weeks+1 week) after randomization, and within the 24+1 week of the last evaluable tumor assessment,
- if date of death is > 117 and ≤ 141 weeks (116 weeks+24 weeks+1 week) after randomization, and within the 36+1 weeks of the last evaluable tumor assessment,
- if date of death is > 141 weeks after randomization, and within the 48 +1 weeks of the last evaluable tumor assessment during Year 3, 4 and 5, or within the 48 +2 weeks of the last evaluable tumor assessment beyond Year 5.

9.4.3 Patients with new anti-cancer therapy

For patients who change anti-cancer treatment[15] to something other than the study treatment prior to observing progression, PFS will be censored at the date of the last evaluable tumor assessment prior to the initiation of anti-cancer treatment. Independent review committee (IRC) assessments will be used in the event of IRC and investigator disagreement on the last response assessment prior to a switch to a new anti-cancer treatment. For LPL/WM patients, if

[15] Or have surgery affecting target/non-target lesions prior to observing progression

PD was assessed by an investigator based on laboratory parameters alone, no independent confirmation of PD by independent blinded review is necessary.

9.4.4 Patients who discontinue or withdraw

For patients who discontinue or withdraw early from the study without documented disease progression, PFS will be censored at the date of the last evaluable tumor assessment unless the patient dies, and death is considered as a PFS event mentioned above. Note that this applies to patients without progression based on blinded independent central radiological review but who discontinued because according to the investigator assessment he/she had progressed (if such a patient exists).

9.5 Imputation rules for partial dates

9.5.1 Algorithm for handling partial initial diagnosis date and progression date

The following rules are used to impute partial initial diagnosis date and progression date:

A. If partial date has day and month missing, then January 01 will be assigned to the missing fields.

B. If partial date has missing day only, then the first day of the month will be assigned to the missing day.

C. If imputed progression date from A or B above is before the Randomization date, then Randomization date will be assigned to overwrite the imputed date from A or B. If the date is completely missing, no imputation will be applied for initial diagnosis date. Use day 1 for progression date.

9.5.2 Imputation rules for partial/missing Adverse Event (AE)/Concomitant medication (CM) dates

Table 9–1: Imputation rules for partial/missing dates

Partial Dates Imputation Rule	Impute partial AE/CM Start Date	Impute partial AE/CM Stop Date
#1. The day missing only	IF AESTDT year and months is same as TRTSDT year and months, then impute AESTDT= TRTSDT	IF AEENDT year and months is same as LKAD year and months, then impute AEENDT= LKAD
	ELSE IF AESTDT year and month is before TRTSDT year and months, then AESTDT= last date of the months	ELSE impute AEENDT= last date of the months
	ELSE IF AESTDT year and month is after TRTSDT year and months, then AESTDT= first date of the months	
#2. Both day and months missing	IF AESTDT year is same as TRTSDT year, then impute AESTDT=TRTSDT	IF AEENDT year is same as LKAD year, then impute AEENDT= LKAD
	ELSE IF AESTDT year is before TRTSDT year, then impute AESTDT=31DECYYYY	ELSE impute AEENDT=31DECYYYY
	ELSE IF AESTDT year is after TRTSDT year, then impute AESTDT=01JANYYYY	
#3. Completely missing	No need to impute, try to query the sites by DM	No need to impute, try to query the sites by DM
#4 Additional criteria to meet	1. AE/CM start date <= AE/CM stop date ## After the imputation of start/stop date per above rules #1~#3, further action is needed if the (imputed) AE/CM start date > the (imputed) AE/CM stop date. We will update the (imputed) AE/CM start date = the (imputed) AE/CM stop date since the imputed stop date is more reliable. Because per imputation partial stop date rules #1~#3, the (imputed) AE/CM stop date will be the maximum date based on the available data unless there's data issue. If it's a data issue, we can leave the imputation as it is and report the data issue to the DM. ## 2. The imputed dates <= LKAD 3. If TRTSDT is missing, use RANDDT as reference date	

9.5.3 Imputation rules for partial/missing anti-cancer treatment start date

Table 9–2 Imputation rules for partial/missing anti-cancer treatment start dates

Partial Dates Imputation Rule	Impute partial anti-cancer treatment start date
#1. The day missing only	IF *CMSTDTC year is after the year of the last visit, then impute DD to be '01' ELSE IF *CMSTDTC year is the same as the year of the last visit and *CMSTDTC month is after the month of the last visit, then impute DD to be '01' ELSE IF *CMSTDTC year and month is the same as the year and month of the last visit, then impute day of the last visit
#2. Both day and months missing (MM-DD)	IF *CMSTDTC year is the same as the year of the last visit, then impute MM-DD to be the month and day of the last visit' ELSE IF *CMSTDTC year is after the year of the last visit, then impute MM-DD to be '01-01'
#3. Completely missing (YYYY-MM-DD)	IF last visit date is known, then impute the last visit date

*CMSTDTC is the start date/time of the anti-cancer treatment.

9.6 Hypothetical OS events needed to detect a difference in OS (answering an HA request)

Modelling of plausible range of median OS estimates for control arm

Three publications studying patients with Rituximab-Bendamustine (R-B) in second line have been identified that are describing results for OS:

a) Rummel, 2005 [10] (N=63): With a schedule of R-B for four cycles, one dose of R 7 days before first cycle and 28 days after last cycle

b) Rummel, 2016 [24] (N=116, in R-B arm): R-B for 6 cycles (R Day 1, RB Day 1,2), followed by R maintenance (every 12 weeks, for 2 years), introduced per amendment

c) Rueda, 2018 [25] (N=40): R-B for 6 cycles (R Day 1, RB Day 1,2), followed by R maintenance (every 12 weeks, for 2 years)

The studies express some variability regarding composition of proportion of Rituximab sensitive patients, NHL histologies included (FL only for c), iNHL including mantle cell lymphoma for Rummel, 2005 and Rummel, 2016)

Beside variations mentioned above, the main difference between the studies is regarding the use of R maintenance therapy:

For studies b) and c), the results can be considered confounded because of using Rituximab maintenance, as compared to how R-B is applied in the C-4 study.

As median OS was generally not reached for most studies and because Rituximab-maintenance has been applied, the OS curves were:

a) modelled using accelerated failure time models and

b) restricted data (i.e., first 12/24/36 months in separate models) was used in studies that applied maintenance

This approach indicated that it is plausible to assume that median OS is within a range of 75-85 months for RB therapy in 2-line NHL, if no maintenance therapy is applied (data on file).

Modelling of median OS estimates for blinded C-4 study data (as of 27/03/2023)

Using blinded C-4 study data, a model was fitted to ongoing OS data, which resulted in a plausible estimate of 115 months (or 9.6 years) for the median of OS in the total (blinded) population. It is considered that 100 months is a conservative and plausible estimate to use in the crude calculation of median OS for the C-R treatment arm.

Crude modelling of potential HR for OS of C-4

To generate the plausible range of median OS estimates for the C-R treatment arm (in a crude fashion), the above results are then considered to be from an exponential distribution with constant hazard, where the (crude) median for copanlisib is calculated as:

Crude median(copanlisib) = 2*median(pooled C-4) – median(control arm).

This crude median is then used to calculate a range of crude potential HRs that might be seen in C-4.

The resulting crude medians and hazard ratios are displayed in Table 9–3, together with the corresponding number of events required, for a range of power values from 60-90% and a two-sided test with alpha of 5%. The expected study duration to obtain the respective number of OS events, was determined using a simulation approach.

While acceptable HR/power combinations seem possible, the timing until sufficient events are accumulated in many cases would stretch beyond feasible calendar dates (i.e., beyond the year 2029).

Table 9–3: Crude median assumptions and resulting HR displayed with events and study duration needed to achieve specific statistical power values					
		Crude Median OS [months]		Events# needed	Date at which number of events is reached##
Power	**Crude HR**	**Placebo**	**Copanlisib**		
60%	0.600	75	125	76	N/A
	0.667	80	120	120	24.04.2023
	0.702	82.5	117.5	157	18.09.2024
	0.739	85	115	215	20.03.2027
70%	0.600	75	125	95	N/A
	0.667	80	120	151	25.06.2024
	0.702	82.5	117.5	198	08.06.2026
	0.739	85	115	270	>2029
80%	0.600	75	125	121	08.05.2023
	0.667	80	120	192	02.03.2026
	0.702	82.5	117.5	251	05.01.2029
	0.739	85	115	344	>2029
90%	0.600	75	125	162	30.11.2024
	0.667	80	120	257	05.05.2029
	0.702	82.5	117.5	336	>2029
	0.739	85	115	460[1]	>2029

N/A: Number of events already reached. [1]Number of events not reached before remaining patients dropped out in ~4% of cases.
Schoenfeld formula applied.
Median of simulations reported

9.7 Details of OS event analyses for varying primary PFS time points and OS hazard ratios (answering an HA request)

The information fraction compared to a benchmark number of 380 events[16], expected number of OS events and conditional power for various assumed HRs, at interim and final formal analysis, are shown in Table 9–4 below.

Note that no alpha spending function is used, but instead a weighted Bonferroni approach (0.002 at initial interim evaluation, 0.023 at final formal analysis) is applied.

The number of OS events is dependent on the timing of the primary evaluation of PFS. The results presented in Table 9–4 are based on the earliest and best guess estimates of the primary completion date, using the current blinded study data.

[16] Following the strategy proposed by the FDA in an information request, 380 events are needed for a successful final OS analysis, if true HR=0.75 and power is to be 80%

Table 9–4: Overview of interim and final evaluation conditional power

If primary PFS event target reached in June 2024:

Planned OS time point	Expected OS events at time point	Benchmark number of OS events[3]	Information fraction compared to bench-mark[3]	Bonferroni split of alpha (for 1-sided test)	Conditional power, if true HR= 0.6	0.65	0.70	0.75	0.80
Interim[1]	151	380	39.7%	0.002	60.3%	40.9%	24.6%	13.3%	6.6%
Final[2]	176	380	46.3%	0.023	91.8%	80.6%	64.5%	46.5%	30.3%

If primary PFS event target reached in Feb 2024:

Planned OS time point	Expected OS events at time point	Benchmark number of OS events[3]	Information fraction compared to bench-mark[3]	Bonferroni split of alpha (for 1-sided test)	Conditional power, if true HR= 0.6	0.65	0.70	0.75	0.80
Interim[1]	141	380	37.1%	0.002	56.1%	37.4%	22,3%	12.1%	6.0%
Final[2]	176	380	46.3%	0.023	91.8%	80.6%	64.5%	46.5%	30.3%

[1]Interim OS analysis at PFS primary completion.
[2]Final formal OS analysis (at approximately 176 OS events).
[3]Using benchmark of 380 events, following the strategy proposed by the FDA in an information request.

9.8 Likelihood to rule out a harmful effect in OS (answering an HA request)

The likelihood that a sufficient number of OS events will be observed, to rule out any potential harmful effect in OS, given 3 and 5 years of additional OS follow-up beyond the planned, final, formal OS analysis (at approximately 176 OS events), was assessed using the following simulation approach:

a) The original (blinded) study data was used as source data for the simulations.

b) A random treatment group assignment was simulated.

c) For patients at risk, future OS events and censoring times were predicted (using a model based on (a) above, where predictions were independent of the simulated treatment group).

d) For each HR assumption, shown in Table 9–5, individual survival times were scaled according to the patient's random treatment assignment[17] .

e) The simulated data was then administratively censored at the respective dates, 3 and 5 years after the predicted date for the final formal OS analysis, (i.e., after 3 years of additional FUP and after 5 years of additional FUP).

f) A Cox PH model was fitted to the simulated data, and a point estimate for HR as well as an upper 95% confidence interval limit (CIL) was saved.

g) This was repeated 10000 times, and the proportion of saved point estimates > 1, as well as the proportion of saved upper 95% CILs > 1, were determined.

The above steps were repeated for all HR assumptions and additional FUP combinations and reported in Table 9–5.

Table 9–5: Resulting probability of HR point estimate / upper bound exceeding 1

Time points of evaluation	Final formal OS analysis	3 years additional FUP after final formal OS analysis	5 years additional FUP after final formal OS analysis
Expected date	25-Jun-2025	25-Jun-2028	25-Jun-2030
Projected number of events (in % compared to benchmark number of 380 events[1])	176 (46.3%)	240 (63.2%)	276 (72.6%)
True HR for OS	**Probability of Point Estimate / Upper Bound Exceeding 1**		
HR=0.600	<0.1% / 7.6%	<0.1% / 3.3%	0.0% / 2.1%
HR=0.667	0.5% / 26.0%	0.1% / 14.7%	0.1% / 11.3%
HR=0.702	0.9% / 36.8%	0.3% / 25.8%	0.3% / 20.8%
HR=0.739	2.3% / 49.6%	1.3% / 40.1%	0.8% / 34.1%
HR=0.75	3.0% / 54.7%	1.5% / 43.5%	1.2% / 38.4%
HR=0.8	7.3% / 70.2%	4.9% / 62.2%	4.3% / 58.8%
HR=0.9	24.9% / 90.8%	22.5% / 88.2%	20.4% / 86.8%
HR=1.0	49.3% / 97.4%	48.5% / 97.5%	49.7% / 97.3%

[1]Following the strategy proposed by the FDA in an information request, a benchmark of 380 events is required for a successful final OS analysis, if true HR=0.75 and power is to be 80%.

At the final formal OS analysis, a harmful effect (considered as HR for OS > 1.0) can be ruled out:

a) For the point estimate with a high probability of 97% (=100%-3.0%) or higher for a true HR of 0.75 or lower, while

b) For the upper bound of the confidence interval, high probability of 92.4% (=100%-7.6%) can only be achieved for a true HR of 0.60 or lower.

[17] Assuming a constant hazard, predicted OS event times will be calculated using,

- Simulated survival time × 2 × HR/(1+HR), for patients in the simulated placebo group
- Simulated survival time × 2 × 1/(1+HR), for patients in the simulated copanlisib group

These simulations indicate that with an additional follow-up of 5 years, a harmful effect (considered as HR for OS > 1.0) can be ruled out

c) For the point estimate with a high probability of 98.8% (=100%-1.2%) or higher for a true HR of 0.75 or lower, while

d) For the upper bound of the confidence interval, high probability of 97.9% (=100%-2.1%) can only be achieved for a true HR of 0.60 or lower.

This result is considered indicative of this trial

a) having been designed for a primary endpoint of PFS, and

b) studying a patient population with anticipated long survival time leading to a long time until sufficient OS events can be collected.

The result is also considered to reinforce the result of Table 9–3, where many scenarios would result in evaluations in 2030 or beyond.