STATISTICAL ANALYSIS PLAN

Study Protocol Number:	BGB-3111-214
Study Protocol Title:	A Phase 2, Open-Label Study of Zanubrutinib (BGB-3111) in Patients with Relapsed or Refractory Marginal Zone Lymphoma
Date:	10-June-2020
Version:	Final

NCT03846427

SIGNATURE PAGE

Author:

[redacted]

BeiGene Co., Ltd.

[redacted]

Approval

[redacted]

BeiGene Co., Ltd.

[redacted]

BeiGene Co., Ltd.

[redacted]

TABLE OF CONTENTS

LIST OF ABBREVIATIONS AND DEFINITIONS OF TERMS

Abbreviation	Definition
AE	adverse event
ATC	Anatomical Therapeutic Chemical
AUC	area under the curve
BTK	Bruton tyrosine kinase
CI	confidence interval
C_{max}	maximum plasma concentration
CR	complete response
CRu	complete remission unconfirmed
CT	computed tomography
CRF	case report form
CV	coefficient of variation
DOR	duration of response
ECG	electrocardiogram
ECOG	Eastern Cooperative Oncology Group
eCRF	electronic case report form
EORTC QLC-C30	European Organisation for Research and Treatment of Cancer Quality of Life Questionnaire-C30
EQ 5D-5L	5-level EQ-5D version
FDA	Food and Drug Administration
GeoMean	geometric mean
GeoCV	geometric coefficient of variation
IC_{50}	50% maximum inhibitory concentration
IRC	independent review committee
MedDRA	Medical Dictionary for Regulatory Activities
MZL	marginal zone lymphoma
MALT	mucosa-associated lymphoid tissue
NCI-CTCAE	National Cancer Institute Common Terminology Criteria for Adverse Events

NHL	non-Hodgkin lymphoma
ORR	overall response rate
OS	overall survival
PD	progressive disease
PET	positron emission tomography
PFS	progression-free survival
PK	pharmacokinetic
PP	per-protocol
PR	partial response
PRO	patient-reported outcome
PT	preferred term
Q1	quartile 1
Q3	quartile 3
R/R MZL	relapsed or refractory marginal zone lymphoma
SAP	statistical analysis plan
SD	stable disease
SOC	system organ class
TEAE	treatment-emergent adverse event
TTF	time to failure
TTR	time to response
WHO-DD	World Health Organization Drug Dictionary

1 INTRODUCTION

The purpose of this statistical analysis plan (SAP) is to describe the procedures and statistical methods that will be used to analyze and report results for the protocol BGB-3111-214 amendment 3.0 dated 03-Jun-2020 titled: *A Phase 2, Open-label Study of Zanubrutinib (BGB-3111) in Patients with Relapsed or Refractory Marginal Zone Lymphoma.* The focus of this SAP is for the planned primary and final analyses specified in the study protocol.

2 STUDY OVERVIEW

This is a multi-center, single-arm, open-label, phase 2 clinical study to evaluate the safety and efficacy of zanubrutinib in patients with relapsed or refractory marginal zone lymphoma (R/R MZL). The study is composed of an initial screening phase (up to 35 days), a single-arm treatment phase, and a follow-up phase with a safety follow-up and a long-term follow-up.

Approximately 65 patients will be enrolled. The primary efficacy analysis will be conducted when mature response rate data have been observed, estimated as no later than 6 months after the last patient receives the first dose of study drug; and the final analysis will be conducted when mature data of secondary endpoints are available. For the primary efficacy analysis, response will be evaluated by an Independent Review Committee (IRC) using the 2014 Lugano Classification for Non-Hodgkin's Lymphoma (NHL) criteria (Cheson et al 2014).

Computed tomography (CT) with contrast will be performed at screening, at 12, 24, 36, and 48 weeks, followed by every 24 weeks thereafter (72 weeks, 96 weeks, etc.) until disease progression, withdrawal of consent, death, lost to follow-up, or end of study, whichever occurs first. Magnetic resonance imaging (MRI) may be used in patients who have severe allergy to CT contrast, but if used, should be used consistently throughout the study. Positron emission tomography (PET) scan will be performed at screening, and repeated for those with PET-avid disease at 12, 24, 36, 48, and 72 weeks. For patients with PET-avid disease, an assessment of complete response (CR) or progressive disease (PD) must be confirmed by PET scan. Patients should remain on study treatment until disease progression is confirmed by investigator assessment using available radiographic and clinical findings based on the Lugano criteria (Cheson et al 2014). Bone marrow biopsy and aspirate will be repeated at time of suspected CR for those patients who had bone marrow involvement by MZL at screening. For patients with gastrointestinal involvement who had an endoscopy performed during the screening period, a follow up endoscopy is required to confirm CR.
After informed consent has been signed but prior to the administration of the study drug, only serious adverse events (AEs) should be reported. Thereafter, all AEs and serious AEs, regardless of relationship to study drug, will be reported until 30 days after the last dose of study drug. After this period, the investigator should report any serious AEs that are believed to be related to prior study drug treatment.

Screening Phase: All screening procedures must be performed within 35 days prior to the first dose of zanubrutinib, unless noted otherwise; assessments not completed within this interval must be repeated. Patients will sign the informed consent form prior to any screening evaluations.

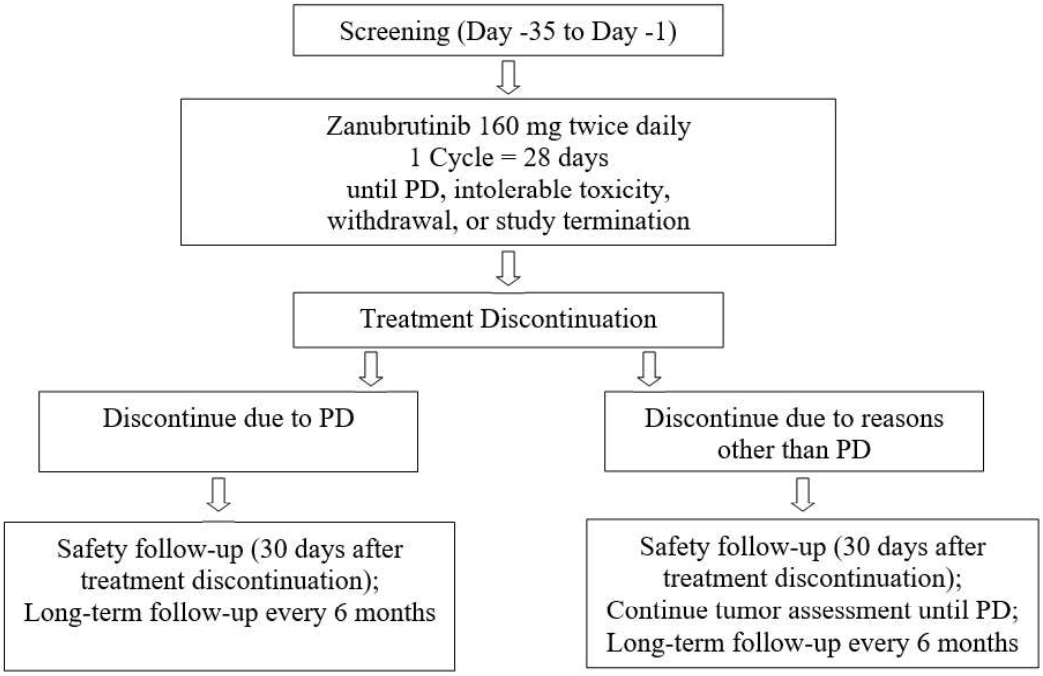
Treatment Phase: Patients will receive the first dose of zanubrutinib at Cycle 1 Day 1. All

patients will be treated with zanubrutinib at 160 mg orally twice daily (two 80-mg capsules orally twice daily) to be continued until disease progression, unacceptable toxicity, treatment consent withdrawal, or study termination. Each cycle consists of 28 days.

Safety Follow-up: All patients who permanently discontinue study drug will have a Safety Follow-up visit approximately 30 days after the last dose of study drug to collect AEs, including AEs that may have occurred or been ongoing after the patient discontinued study treatment. The investigator or his/her designee will also continue to collect information on new anti-cancer therapy given after the last dose of study drug. A laboratory assessment is only required if the patient had an ongoing laboratory abnormality at the previous visit that the investigator considered to be related to study drug.

Long-term Follow-up: Long-term Follow-up visits will be conducted every 24 weeks after the last dose of study drug until withdrawal of consent, lost to follow-up, death, or study termination. It includes monitoring survival status and subsequent therapies for MZL and may also include imaging and tumor response assessments for patients who have not yet had confirmed radiographic progression. For patients who permanently discontinue study drug treatment before radiographic progression is documented, tumor assessments (including radiographic imaging) will continue until radiographic progression is identified and confirmed by the investigator. Imaging assessments should occur every 12 weeks until what would have been Week 48 and then approximately every 24 weeks thereafter.

Figure 1 Schema for Study BGB-3111-214



3 STUDY OBJECTIVES

3.1 PRIMARY OBJECTIVES

To evaluate the efficacy of zanubrutinib in R/R MZL as measured by overall response rate (ORR) in accordance with the Lugano Classification (Cheson et al 2014) determined by IRC

3.2 SECONDARY OBJECTIVES

- To evaluate the efficacy of zanubrutinib in R/R MZL in accordance with the Lugano Classification (Cheson et al 2014) as measured by the following:
  - ORR determined by investigator
  - Progression-free survival (PFS) determined by IRC and by investigator
  - Overall survival (OS)
  - Duration of response (DOR) determined by IRC and by investigator
  - Time to response (TTR) determined by IRC and by investigator
  - Time to treatment failure (TTF)
  - Time to next line of therapy for MZL
- To evaluate patient-reported outcomes (PROs)
- To determine the safety of zanubrutinib
- To determine pharmacokinetics (PK) parameters of zanubrutinib

3.3 EXPLORATORY OBJECTIVES

- To evaluate ORR in accordance with Cheson 1999 as determined by IRC

4 STUDY ENDPOINTS

4.1 PRIMARY ENDPOINTS

The primary endpoint of the study is ORR, defined as the proportion of patients achieving a best overall response of either partial response (PR) or complete response (CR) as determined by IRC according to the 2014 Lugano Classification for Non-Hodgkin's lymphoma (NHL) criteria (Cheson et al 2014) at any time on study drug.

4.2 SECONDARY ENDPOINTS

The secondary endpoints are:

- ORR (CR + PR) in accordance with the Lugano Classification (Cheson et al 2014) determined by investigator assessment
- ORR (CR + PR) in accordance with the Lugano Classification (Cheson et al 2014) using PET assessment data for patients with FDG–avid disease determined by IRC
- PFS in accordance with the Lugano Classification (Cheson et al 2014) determined by IRC and by investigator assessment
- OS
- DOR in accordance with the Lugano Classification (Cheson et al 2014) determined by IRC and by investigator assessment
- TTR in accordance with the Lugano Classification (Cheson et al 2014) determined by IRC and by investigator assessment
- TTF
- Time to next line of therapy for MZL
- PROs measured by the EQ-5D-5L and EORTC QLQ-C30 questionnaire
- Safety parameters including AEs, Serious AEs, laboratory tests, physical exams, and vital signs
- PK parameters such as apparent clearance of the drug from plasma (CL/F) and $AUC_{0\text{-}12}$

4.3 EXPLORATORY ENDPOINTS

- ORR (CR + unconfirmed CR [CRu] +PR) in accordance with Cheson 1999 as determined by IRC

5 SAMPLE SIZE CONSIDERATIONS

Assuming a null hypothesized ORR of 30%, a sample size of 65 patients will provide 82% power for the alternative hypothesized ORR of 48%, at a 1-sided alpha level of 0.025 and using the exact binomial test. The alternative ORR is based on the observed IRC-assessed ORR for the ibrutinib study in R/R MZL (Noy et al 2017). For an observed ORR of 48% (31/65), the 95% exact binomial confidence interval is (35%, 60%).

6 STATISTICAL METHODS

6.1 ANALYSIS SETS

The Safety Analysis set includes all patients who are enrolled and receive any dose of zanubrutinib. This will be the population of primary interest for safety analyses.

The Efficacy Analysis set consists of all patients in the Safety Analysis set with confirmed diagnosis of MZL. This set will be the primary analysis set for efficacy analyses.

The PK Analysis set includes all patients who have at least one PK sample collected (have at least one post-dose PK concentration) according to the protocol and laboratory manual.

6.2 DATA ANALYSIS GENERAL CONSIDERATIONS

6.2.1 Definitions and Computations

Study drug: the study drug in this study is zanubrutinib.

Study day: Study day will be calculated in reference to the date of the first dose of study drug. For assessments conducted on or after the date of the first dose of study drug, study day will be calculated as (assessment date – date of first dose of study drug + 1). For assessments conducted before the date of the first dose of study drug, study day is calculated as (assessment date – date of first dose of study drug). There is no study day 0.

Treatment duration: The treatment duration will be calculated as (date of the last dose of study drug – date of first dose of study drug + 1).

Baseline: Unless otherwise specified, a baseline value is defined as the last non-missing value collected on/before the first dose date of study drug.

Retests and Unscheduled Visits: Unscheduled measurements will not be included in by-visit table summaries and graphs, but will contribute to best/worst case value where required (e.g. shift table). Listings will include scheduled, unscheduled and retest data.

All calculations and analyses will be conducted using SAS version 9.4 or higher.

6.2.2 Conventions

Unless otherwise specified, the following conventions will be applied to all analyses:

- 1 year = 365.25 days. Number of years is calculated as (days/365.25) rounded up to 1 significant digit.
- 1 month = 30.4375 days. Number of months is calculated as (days/30.4375) rounded up to 1 significant digit.
- P-values will be rounded to 4 decimal places. P-values that round to 0.0000 will be presented as '< 0.0001' and p-values that round to 1.000 will be presented as '> 0.9999'.
- Missing efficacy or safety data will not be imputed unless otherwise specified.
- Time-to-event or duration of event endpoints will be based on the actual date the radiograph was obtained rather than the associated visit date.

- For by-visit observed data analyses, percentages will be calculated based on the number of patients with non-missing data as the denominator, unless otherwise specified.
- For continuous endpoints, summary statistics will include n, mean, standard deviation, median, quartile 1(Q1), quartile 3(Q3), and range (minimum and maximum).

6.2.3 Handling of Missing Data

Missing data will not be imputed unless otherwise specified elsewhere in the SAP. Missing dates or partially missing dates will be imputed conservatively for prior/concomitant medications/procedures, subsequent anti-cancer therapies, adverse events and deaths. Specific rules for handling of missing or partially missing dates for prior/concomitant medications/procedures, subsequent anti-cancer therapies, adverse events, and deaths are provided in Appendix A.

By-visit summary of variables with missing data will use only non-missing data, not imputed one, unless otherwise specified. Unscheduled visits will not be included in by-visit summaries.

6.2.4 Adjustment for Covariates

Not applicable.

6.2.5 Multiplicity Adjustment

Not applicable.

6.2.6 Data Integrity

Before any pre-specified statistical analysis begins, the integrity of the eCRF data and all derived values should be reviewed to ensure that the data is accurate and complete up to a pre-specified cutoff date. Consistency checks and appropriate source data verification should be completed.

6.3 PATIENT CHARACTERISTICS

6.3.1 Patient Disposition

The summary of patient disposition will be based on all patients enrolled. The following patient disposition information will be summarized:

- Number of enrolled patients
- Number (%) of treated patients
- Number (%) of treated patients who remain on treatment
- Number (%) of treated patients who discontinue from treatment
- Reasons that patients discontinue from treatment
- Number (%) of treated patients who remain in study
- Number (%) of treated patients who discontinue from study

- Reasons that patients discontinue from study

Study follow-up time will be defined as the time from first dose date to the death date or end of study date (whichever occurs earlier) for patients discontinued from study, or the database cutoff date for ongoing patients. Treatment follow-up time is defined as the time from the first dose date to the end of treatment date for patients discontinued from the treatment, or the database cutoff date for ongoing patients. Study and treatment follow-up time will be summarized descriptively.

6.3.2 Protocol Deviations

Criteria for important protocol deviations will be established and patients with important protocol deviations will be identified and documented before the database lock. Important protocol deviations will be summarized for all patients in the Safety Analysis set. They will also be listed.

6.3.3 Demographic and Other Baseline Characteristics

Demographic and baseline characteristics will be summarized descriptively for patients in the Safety Analysis set as follows:

- Age (years) and age (years) categorized as <65, ≥65 and <75, and ≥75 years
- Gender
- Race
- Ethnicity
- Country
- Baseline Eastern Cooperative Oncology Group（ECOG) performance status
- Baseline vital signs (height, weight, temperature, systolic blood pressure, diastolic blood pressure, heart rate)

A listing of demographic will be provided.

6.3.4 Disease History and Characteristics

The number (percentage) of patients reporting a history of disease and characteristic will be summarized for the Safety Analysis set. Disease characteristics include:

- Time since first diagnosis of MZL to study entry (months)
- Disease subtype (extranodal MZL of mucosa-associated lymphoid tissue (MALT), nodal MZL, or splenic MZL)
- Disease status (relapsed or refractory)
- Sites of disease (gastric, non-gastric/non-cutaneous, and cutaneous) for MALT subtype
- Evidence of FDG avid disease by IRC

- Lymphoma involvement in the bone marrow
- Number (%) of patients with prior radiotherapy
- Time from the end of last radiotherapy to study entry (months)
- Number (%) of patients with prior stem cell transplant

A listing of disease history will be provided.

6.3.5 Prior Anti-Cancer Drug Therapies

The following information related to prior anti-cancer therapy will be summarized for the Safety Analysis set:

- Number (%) of patients with prior systematic therapy for cancer
- Number of lines of prior therapies and number of lines of prior therapies categorized as 1, 2, 3, etc.
- Number of prior regimens and number of prior regimens categorized as 1, 2, 3, etc.
- Number of prior regimens with CD20-directed therapy and number of prior regimens with CD20-directed therapy categorized as 1, 2, 3, etc.
- Time from the end of last therapy to study entry (months)
- Time from end of last disease progression to study entry (months)
- Reason last therapy ended
- Duration of last therapy (months)
- Best overall response to last therapy
- Prior treatment by regimen

A listing of prior anti-cancer drug therapies will be provided.

6.3.6 Prior and Concomitant Medications

Prior and concomitant medications will be coded using the World Health Organization Drug Dictionary (WHO DD) drug codes of the version currently in effect at Beigene at the time of database lock, and will be further classified to the appropriate Anatomical Therapeutic Chemical (ATC) code.

The number (percentage) of patients reporting prior and concomitant medications will be summarized by ATC medication class and WHO DD preferred term for the Safety Analysis set. Prior medications will be defined as medications that started before the first dose of study drug. Concomitant medications will be defined as medications that (1) started before the first dose of study drug and were continuing at the time of the first dose of study drug, or (2) started on or after the date of the first dose of study drug up to 30 days after the patient's last dose or initiation of a new anti-cancer therapy. For the purpose of determining if a medication should be noted as a concomitant medication, the imputation rules stated in Appendix A will be used.

6.3.7 Medical History

Medical history will be mapped to a preferred term (PT) and a system organ class (SOC) using the Medical Dictionary for Regulatory Activities (MedDRA) of the version currently in effect at BeiGene at the time of database lock. The number (percentage) of patients reporting a history of any medical condition, as recorded on the CRF, will be summarized by SOC and PT for the Safety Analysis set.

6.4 EFFICACY ANALYSIS

The set of primary interest for efficacy analyses is the Efficacy Analysis set.

6.4.1 Primary Efficacy Endpoint

ORR by IRC

The primary efficacy endpoint is ORR according to the Lugano Classification (Cheson et al 2014) as assessed by IRC. ORR is defined as the proportion of patients achieving a best overall response of CR or PR. The point estimate and corresponding two-sided Clopper-Pearson 95% CI for ORR will be presented.

The best overall response is defined as the best response recorded from the start of zanubrutinib throughout the study. Patients with no post-baseline response assessment (due to any reason) will be considered non-responders for the purposes of this analysis. The proportion for each response category (CR, PR, stable disease [SD], and PD) will be presented. The primary efficacy analysis will be conducted when mature response rate data have been observed, estimated as no later than 12 months after the last patient received the first dose of study drug.

6.4.2 Secondary Efficacy Endpoints

ORR by Investigator

ORR by investigator is defined as the as the proportion of patients achieving a best overall response of CR or PR at any time on study drug assessed by investigator according to the Lugano Classification (Cheson et al 2014) and it will be analyzed using the same statistical methods as ORR by IRC.

ORR using PET assessment in FDG-avid patients by IRC

ORR accordance with the Lugano Classification (Cheson et al 2014) using PET assessment data for patients with FDG–avid disease determined by IRC will be analyzed the same way as ORR by IRC.

PFS by IRC

PFS by IRC is defined as the time (months) from study treatment start to PD, as determined by IRC according to the Lugano Classification (Cheson et al 2014), or death of any cause, whichever occurs first:

PFS = (the earlier of PD or death date – the date of first study dose +1)/30.4375

PFS will be right-censored based on rules provided in Table 1. The PFS censoring rule will follow FDA Guidance for Industry Clinical Trial Endpoints for the Approval of Cancer Drugs and Biologics (FDA 2018).

Table 1. Censoring Rules for Analysis of Progression-Free Survival

No.	Situation	Date of Progression or Censoring	Outcome
1	No baseline and/or post-baseline disease assessments	Date of the first dose	Censored
2	Death or PD between planned disease assessments	Date of death or first disease assessment showing PD, whichever occurs first	Progressed
3	Alive without documented disease progression at the time of data cut-off or withdrawal from study (including lost-to-follow-up without disease progression)	Date of last disease assessment	Censored
4	New anticancer treatment started before documented disease progression or death	Date of last disease assessment prior to date of new anticancer treatment	Censored
5	Death before first disease assessment	Date of death	Progressed
6	Death or progression after more than one missed scheduled disease assessment	Date of last disease assessment without documented disease progression before missed tumor assessments	Censored

Note that the frequency of disease assessments is about every 12 weeks in the first 48 weeks and every 24 weeks thereafter. Therefore, missing more than one disease assessment will be interpreted as gaps longer than 24 weeks in the first 96 weeks or gaps longer than 48 weeks thereafter for censoring purposes of PFS.

Kaplan Meier methodology will be used to estimate the median and other quantiles of PFS. Kaplan Meier curves will be constructed to provide a visual description of the PFS change with time. Two-sided 95% CIs of median and other quartiles, if estimable, will be constructed with a generalized Brookmeyer and Crowley method (Brookmeyer 1982; Klein 1997) with log-log transformation. PFS rates at selected landmark time points (e.g. 6-month) will be provided with corresponding 95% CIs calculated based on the Greenwood's formula (Kalbfleisch and Prentice 1980) with log-log transformation.

The duration of the follow-up for PFS will be determined by reverse Kaplan-Meier method (Schemper 1996).

PFS by Investigator

PFS by investigator will be analyzed using the same statistical methods as PFS by IRC.

OS

OS is defined as the time (months) from study treatment start to death due to any cause. Patients who remain alive before data cutoff or discontinuation of the study (discontinued study due to reasons other than "Death") will be censored at the time of data cutoff or the last date the patient is known to be alive. OS will be analyzed using the similar method as described for PFS.

DOR by IRC

DOR by IRC for responders (CR or PR) is defined as the time (months) from the date of the earliest qualifying response (PR or better) to the date of PD, as determined by IRC according to the Lugano Classification (Cheson et al 2014), or death for any cause, whichever occurs earlier. The censoring rules for DOR will follow the PFS censoring rules. Kaplan-Meier methods will be used to estimate median and other quartiles and 95% CIs for DOR. Only responders will be included in this analysis.

DOR by Investigator

DOR by investigator will be analyzed using the same statistical methods as DOR by IRC.

TTR by IRC

TTR for responders (CR or PR) by IRC according to the Lugano Classification (Cheson et al 2014) is defined as the time (months) from study treatment start to the date of the earliest qualifying response (PR or better) as determined by IRC. TTR by IRC will be summarized by sample statistics such as mean, median, and standard deviation for responders only.

TTR by Investigator

TTR by investigator will be analyzed using the same statistical methods as TTR by IRC.

TTF

TTF is defined as the time (months) from study treatment start to the date of discontinuation of study drug due to any reason. TTF will be censored at the data cutoff for the patients who don't discontinue study treatment. The distribution of TTF will be summarized by the Kaplan-Meier method similarly as described for the PFS analysis.

Time to Next Line of Therapy for MZL

Time to next line of therapy is defined as the time (months) from study treatment start to the start of the first subsequent therapy for MZL. Time to next line of therapy for MZL will be censored as of date of study discontinuation or most recent contact date for patients who receive no subsequent therapy for MZL. Time to next line of therapy for MZL will be summarized by the Kaplan-Meier method similarly as described for the PFS analysis.

Table 2. Censoring Rules for Analysis of Time to Next Line of Therapy for MZL

Situation	Date of Next Line of Therapy for MZL or Censoring	Outcome
Study discontinuation without any next line of therapy for MZL	Date of study discontinuation	Censored
Administration of any next line of therapy for MZL (including radiotherapy)	Date of first administration of next line of therapy for MZL	Event
No new administration of any next line of therapy for MZL	Last day known alive	Censored

Patient-Reported Outcomes (PROs)

The PROs questionnaires will be summarized on Cycle 1 Day 1, the end of Cycle 3, then every 12 weeks for 12 months, followed by every 6 months thereafter, and will continue to be assessed until disease progression, death, or withdrawal of consent, regardless of study treatment discontinuation. The scoring of European Organization for Research and Treatment of Cancer Quality of Life Questionnaire-C30 (EORTC QLQ-C30) and 5-level EQ-5D version (EQ-5D-5L) will follow their corresponding manuals (Fayers et al. 2001; EuroQol Group 1990; EuroQol Group 2015; Herdman et al 2011).

- EORTC QLQ-C30 (version 3.0)

The EORTC QLQ-C30 is a questionnaire developed to assess the quality of life of cancer patients which includes 30 separate questions (items) resulting in 1 Global Health Status/QoL scale, 5 functional scales (Physical Functioning, Role Functioning, Emotional Functioning, Cognitive Functioning, and Social Functioning), 3 symptom scales (Fatigue, Nausea and Vomiting, and Pain), and 6 single items (Dyspnea, Insomnia, Appetite Loss, Constipation, Diarrhea, and Financial Difficulties).

The scores at each assessment timepoint and changes from baseline in Global Health Status/QoL scale, 5 functional scales, and 9 symptom scales/items will be summarized descriptively (e.g., mean, standard deviation, median, Q1, Q3, minimum, maximum).

- EQ-5D-5L

The EQ-5D-5L comprises a descriptive system for use as a measure of health outcome with following five dimensions: mobility, self-care, usual activities, pain/discomfort, and anxiety/depression, and an EQ Visual Analogue scale (EQ VAS). Each dimension has five levels: no problems, slight problems, moderate problems, severe problems, and extreme problems. The EQ VAS records the respondent's self-rated health on a 0 to 100 scale, with 100 labelled 'the best health you can imagine' and 0 'the worst health you can imagine'.

The number (%) of each level of all five dimensions at each assessment timepoint will be summarized. The VAS score at each assessment timepoint and changes from baseline will be summarized descriptively (e.g., mean, standard deviation, median, Q1, Q3, minimum, maximum).

6.4.3 Subgroup Analyses

Primary and selected secondary endpoints will be summarized descriptively in the specified subgroups, as appropriate (i.e., when there is sufficient number of patients in the subgroup, otherwise relevant subgroups may be combined): sex, age group (<65 vs. ≥65), ECOG performance status (0 vs. ≥1), prior line of therapy for MZL (<3 vs. ≥3), and MZL subtypes. Within-group summary statistics (e.g., ORR and medians for PFS, DOR with their 95% CIs) will be presented in forest plots. The subgroup variables and the cutoff values are subject to change if warranted to better represent the data.

6.4.4 Exploratory Efficacy Endpoints

ORR (CR + CRu + PR) in accordance with Cheson 1999 as determined by IRC will be analyzed using the same methods as described for ORR by IRC.

6.5 PHARMACOKINETIC ANALYSES

The analysis set for patients with PK samples will contain all patients who had at least one PK sample collected according to the protocol and laboratory manual.

Actual collection times will be used for the calculation of PK parameters in the final analysis and reporting. For intensive PK profile, PK parameters such as C_{max}, $AUC_{0-\infty}$, AUC_{0-t} and half-life and other appropriate PK parameters after single and multiple doses will be derived using the standard non-compartmental method and will be computed in WinNonlin® Enterprise v.5.2 or higher. Estimates for these parameters will be summarized (e.g., n, mean, standard deviation, coefficient of variation [CV%], median, min, max, geometric mean [GeoMean], and geometric coefficient of variation [GeoCV%]).

For intensive PK data, individual and mean plasma zanubrutinib concentration-time data will be summarized and displayed in both tabular and graphical form.

The following table lists PK parameters planned to be estimated. Additional PK parameters may be calculated if deemed appropriate.

$AUC_{0-\infty}$	Area under the plasma concentration-time curve from zero to infinity
AUC_{0-t}	Area under the plasma concentration-time curve from zero to t
AUC_{last}	Area under the plasma concentration-time curve from zero to the last measurable concentration
C_{max}	Maximum plasma concentration
t_{max}	Time to maximum plasma concentration
$t_{1/2}$	Half life
CL/F	Apparent plasma clearance
Vd/F	Apparent volume of distribution

6.6 SAFETY ANALYSES

All safety analyses will be performed based on the Safety Analysis set.

The incidence of treatment-emergent adverse events (TEAEs) and serious TEAEs will be summarized. Laboratory test results, vital signs, and their changes from baseline will be summarized using descriptive statistics (e.g., n, mean, standard deviation, median, Q1, Q3, minimum, maximum for continuous variables; n [%] for categorical variables). Abnormal values will be flagged.

6.6.1 Extent of Exposure

Extent of exposure to the study drug will be summarized descriptively with respect to the following:

- Number of treatment cycles received: defined as the total number and percentage of treatment cycles in which at least one dose of the study drug is administered
- Duration of exposure (months): defined as the duration (months) from the date of the first dose to the last dose of the study drug
- Cumulative dose received per patient (g): defined as the cumulative dose (g) of the study drug during the treatment period of the study
- Actual dose intensity (mg/day): defined as the total dose of the study drug (mg) received by a patient divided by the duration of exposure (days)
- Relative dose intensity (%): defined as the ratio of the actual dose intensity (mg/day) and the planned dose intensity (mg/day), where the planned dose is (160*2) mg/day

The number (percentage) of patients with dose modification (including reduction and interruption) will be summarized with the respective reasons. The cycles in which the first dose reduction/interruption occurs will be summarized. Frequency of dose reduction/interruption will be summarized descriptively. Duration of dose interruption (days) will be summarized descriptively. Patient data listings will be provided for all dosing records, and for the above calculated summary statistics.

6.6.2 Adverse Events

AEs will be graded by the investigators using NCI CTCAE v4.03 (NCI 2010). Verbatim description of AEs (as recorded by the investigator on the eCRF) will be classified into standardized PT and SOC using MedDRA.

TEAE is defined as an AE that has an onset date or a worsening in severity from baseline (pre-treatment) on or after the first dose of study drug and up to 30 days following study drug discontinuation, or the start of new anticancer therapy, whichever comes first. Only those AEs that are treatment-emergent will be included in summary tables. All AEs, treatment-emergent or otherwise, will be presented in patient data listings.

Treatment-related AEs include those events considered by the investigator to be related to study drug or with missing assessment of the causal relationship.

An overview of patients with at least one TEAE will be presented with the incidence of:

- Patients with at least 1 TEAE
- Patients with at least 1 TEAE with grade 3 or higher
- Patients with at least 1 Serious TEAE
- Patients with at least 1 TEAE leading to death
- Patients with at least 1 TEAE leading to treatment discontinuation
- Patients with at least 1 TEAE leading to dose modification, including dose interruption and dose reduction
- Patients with at least 1 treatment-related TEAE
- Patients with at least 1 treatment-related TEAE with grade 3 or higher
- Patients with at least 1 treatment-related Serious TEAE
- Patients with at least 1 treatment-related TEAE leading to death
- Patients with at least 1 treatment-related TEAE leading to treatment discontinuation
- Patients with at least 1 treatment-related TEAE leading to dose modification, including dose interruption and dose reduction
- Patients with TEAE of special interest

The incidence of following TEAEs and Serious TEAEs will be reported by SOC and PT (and by worst grade if specified), sorted by decreasing frequency of SOC and PT (a patient with multiple occurrences of the same event within a SOC and PT will be counted only once by the worst grade according to CTCAE v4.03):

- TEAE (any grade) by SOC/PT and by SOC/PT/worst grade
- TEAE with grade 3 or higher by SOT/PT and by SOC/PT/worst grade
- Serious TEAE by SOC/PT and by SOC/PT/worst grade
- TEAE leading to death by SOC/PT
- TEAE leading to treatment discontinuation by SOC/PT/worst grade
- TEAE leading to treatment modification by SOC/PT/worst grade
- Treatment-related TEAE by SOC/PT/worst grade
- Treatment-related TEAE with grade 3 or higher by SOC/PT/worst grade
- Treatment-related Serious TEAE by SOC/PT/worst grade
- Treatment-related TEAE leading to death by SOC/PT
- Treatment-related TEAE leading to treatment discontinuation by SOC/PT/worst grade
- Treatment-related TEAE leading to treatment modification by SOC/PT/worst grade

The incidence of following TEAEs will be reported by PT:

- TEAE (any grade)
- TEAE with grade 3 or higher
- Serious TEAE
- TEAE leading to treatment discontinuation
- TEAE leading to dose modification
- TEAE leading to death
- Treatment-related TEAE

Incidence of and time to TEAEs of special interest such as severe bleed (defined as ≥ grade 3 bleed of any site or central nervous system bleed of any grade), and atrial fibrillation (both new onset and exacerbation of existing atrial fibrillation) will be summarized. Exposure adjusted incidence rate of TEAEs of special interest will be summarized as well. Incidence of TEAE related to liver, arrhythmia and diarrhea will be summarized.

A summary of the number of deaths and the cause of death, classified by deaths within 30 days of last dose of study drug and deaths more than 30 days after the last dose, will be provided.

Listings of all AEs, treatment-related AE, grade 3 or above AEs, Serious AEs, AEs leading to treatment discontinuation, AEs leading to treatment modification, and AEs leading to death will be provided.

6.6.3 Laboratory Values

CBC and serum chemistry values will be evaluated for each laboratory parameter. Abnormal laboratory values will be flagged and identified as those outside (above or below) the normal range. Descriptive summary statistics (n, mean, standard deviation, median, minimum, maximum for continuous variables; n [%] for categorical variables) for selected laboratory parameters and their changes from baseline will be provided by visit.

Laboratory parameters that are graded in NCI-CTCAE (v.4.03) will be summarized by CTCAE grade. Shift tables will be used to assess the change of each laboratory parameter from its toxicity grade at baseline to the worst post-baseline toxicity grade including the unscheduled results. Parameters with CTCAE grading in both high and low directions (e.g., calcium, glucose, magnesium, phosphorus, potassium, sodium) will be summarized separately.

A summary of patients with a change of 2 or more toxicity grade compared to baseline and patients with postbaseline toxicity grade of 3 or more will be provided for laboratory parameters of interest. Listings of selected laboratory parameters will be provided.

6.6.4 Vital Signs

Descriptive statistics for the actual values of vital signs parameters, including resting diastolic and systolic BP, heart rate, temperature, and weight, and changes from baseline will be presented by visit. Baseline vital signs will be listed by patient.

6.6.5 Electrocardiogram (ECG)

A 12-lead electrocardiogram (ECG) will be performed locally in triplicate at screening for all patients and as clinically indicated at other timepoints. Number (percentage) of patients with postbaseline ECG parameters and change from baseline (if post-baseline ECG parameters are measured) categorized will be presented.

6.6.6 Eastern Cooperative Oncology Group Performance Status (ECOG)

ECOG performance status will be summarized as the number (percentage) of patients with each ECOG grade at the screening visit, each visit during study treatment, and at the Safety follow-up visit. Shift tables assessing the ECOG performance status at baseline versus worst performance status on study will be presented.

6.7 OTHER ANALYSES

Additional exploratory analyses may be conducted as appropriate. Any exploratory analyses that are performed will be appropriately titled/labeled as exploratory and will be clearly distinguished from planned analyses when results are reported in the Clinical Study Report.

7 INTERIM ANALYSIS

No formal interim analyses are planned for this study. Data from the study will be reviewed by the study monitoring committee on a periodic basis. Summaries and analyses of subsets of the study data may be performed on a periodic basis for submission to professional meetings and for internal decision-making.

8 CHANGES IN THE PLANNED ANALYSIS

Not applicable.

9 REFERENCES

1. Brookmeyer, B et al. (1982), A confidence interval for the median survival time, Biometrics 38(1), p29-41.
2. Cheson BD, Horning SJ, Coiffier B, et al. Report of an International Workshop to Standardize Response Criteria for Non-Hodgkin's Lymphomas. J Clin Oncol 1999; 17(4): 1244.
3. Cheson B, Fisher R, Barrington S, 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.
4. Cheson BD, Ansell S, Schwartz L, et al. Refinement of the Lugano Classification lymphoma response criteria in the era of immunomodulatory therapy. Blood. 2016;128(21):2489-2496.
5. Common Toxicity Criteria Version 4.03. (2010) Cancer Therapy Evaluation Program. 14 June 2010.
6. EuroQol Group. EuroQol—a new facility for the measurement of health-related quality of life. Health Policy. 1990;16(3):199-208.
7. EuroQol Group. EQ-5D-5L User Guide: basic information on how to use the EQ-5D-5L instrument. 2015. https://euroqol.org/wp-content/uploads/2016/09/EQ-5D-5L_UserGuide_2015.pdf.
8. Fayers P., Aaronson, NK, Bjordal K., Groenvold M., et al. 2001. EORTC QLQ-C30 Scoring Manual (3rd edition). (3rd ed.) Brussels: European Organization for Research and Treatment of Cancer.
9. Food and Drug Administration Center for Drug Evaluation Research CDER and Center for Biologics Evaluation and Research (2018). FDA Guidance for Industry Clinical Trial Endpoints for the Approval of Cancer Drugs and Biologics.
10. Herdman M, Gudex C, Lloyd A, et al. Development and preliminary testing of the new five-level version of EQ-5D (EQ-5D-5L). Quality of Life Research. 2011;20(10):1727-36.
11. Kalbfleisch, J. D. and Prentice, R. L. (1980), The Statistical Analysis of Failure Time Data, New York: John Wiley & Sons.
12. Klein, J. P. and Moeschberger, M. L. (1997), Survival Analysis: Techniques for Censored and Truncated Data, New York: Springer-Verlag.
13. Noy A, de Vos S, Thieblemont C, et al. Targeting Bruton tyrosine kinase with ibrutinib in relapsed/refractory marginal zone lymphoma. Blood 2017;129(16):2224-32.
14. Schemper M., and Smith T.L., (1996), A note on quantifying follow-up in studies of failure time, Controlled Clinical Trials, 17: 343-346.

10 APPENDIX

Appendix A: Missing Data Imputation Rule

In general, missing or partial dates will not be imputed as data level. The following rules will apply for the specific analysis and summary purposes mentioned below only.

A.1 Prior/Concomitant Medications/Therapies/Procedures

When the start date or end date of a medication/therapy/procedure is partially missing, the date will be imputed to determine whether the medication/therapy/procedure is prior or concomitant. The following rules will be applied to impute partial dates for medications/therapies/procedures:

If start date of a medication/therapy/procedure is partially missing, impute as follows:

- If both month and day are missing, then set to January 01
- If only day is missing, then set to the first of the month

If end date of a medication/therapy/procedure is partially missing, impute as follows:

- If both month and day are missing, then set to December 31
- If only day is missing, then set to last day of the month

If start date or end date of a medication/therapy/procedure is completely missing, do not impute.

A.2 Adverse Events

The imputation rule for the safety analyses will be used to address the issues with partial dates. When the start date or end date of an adverse event is partially missing, the date will be imputed to determine whether the adverse event is treatment-emergent. When in doubt, the adverse event will be considered treatment emergent by default. The following rules will be applied to impute partial dates for adverse events:

If start date of an adverse event is partially missing, impute as follows:

- If both month and day are missing and year = year of treatment start date, then set to treatment start date
- If both month and day are missing and year ≠ year of treatment start date, then set to January 01
- If day is missing and month and year = month and year of treatment start date, then set to treatment start date
- If day is missing and month and year ≠ month and year of treatment start date, then set to first day of the month
- If start date is completely missing, set to treatment start date as long as adverse event end

date is not before treatment start date

If end date of an AE is partially missing, impute as follows:

- If both month and day are missing, then set to December 31
- If only day is missing, then set to last day of the month
- If end date is completely missing, do not impute.

If the imputed AE end date > last known alive date or end of study date, then set to the last known alive date or end of study date, whichever occurs first.

A.3 Deaths

In case only the day of a death date is missing, the death will be assumed to be on the 1st date of the month if the last date of patient known to be alive is earlier that the 1st date of the month, otherwise the death date will be assumed to be on 1 day after the last date of patient known to be alive.

A.4 Subsequent Anti-cancer Therapies

If the start date of a subsequent anti-cancer therapy is missing, the start date will be assumed to be on the 1st date of the month.

Appendix B: Specification on the Use of Lab Data

For determination of baseline value and for the summary of lab parameter at each scheduled visit and the change from baseline, following rule will be applied:

if a patient has either non-missing local or central lab parameter value only, that value will be treated as baseline value; if a patient has both non-missing local and central lab parameter values, the central lab parameter value will be used over the local one.

For the shift table of the summary of the worst post-baseline grade from baseline, it will be based on all data (both local and central) at all post-baseline visits (including unscheduled visits).

If the lab parameter has a missing unit at a visit, it will be not used in the related summary analysis.