BeiGene

STATISTICAL ANALYSIS PLAN

Study Protocol Number:	BGB-3111-1002
Study Protocol Title:	A phase I clinical study to investigate the safety, tolerability and pharmacokinetics/ pharmacodynamics of BTK inhibitor BGB-3111 in Chinese patients with B-cell lymphoma
Date:	November 1, 2017
Version:	2.0

SIGNATURE PAGE

Author:

[redacted]

Signature

Date: _____

[redacted]

Signature

Date: _____

Approval

[redacted]

Signature

Date: _____

[redacted]

Signature

Date: _____

SIGNATURE PAGE

Author:

[redacted]

[redacted]
Signature

Date: Jan 5, 2018

[redacted]

[redacted]
Signature

Date: Dec. 5, 2017

Approval

[redacted]

Signature

Date: _____

[redacted]

Signature

Date: _____

SIGNATURE PAGE

Author:

[redacted]

Signature

Date: _____

Signature

Date: _____

Approval

[redacted]

[redacted]

Date: 19 Dec 2017

Signature

Date: _____

SIGNATURE PAGE

Author:

[redacted]

Signature

Date: _____

[redacted]

Signature

Date: _____

Approval

[redacted]

Signature

Date: _____

[redacted]

Date: 1/3/18

TABLE OF CONTENTS

LIST OF ABBREVIATIONS AND DEFINITIONS OF TERMS

Abbreviation	Term
AE	Adverse event
ATC	Anatomical Therapeutic Chemical
AUC	Area under the concentration-time curve
BID	Twice a Day
BMI	Body Mass Index
BTK	Bruton tyrosine kinase
C_{max}	Maximum observed plasma concentration
CL/F	Plasma drug clearance rate
CRF	Case report form
CR	Complete response
CTCAE	Common Terminology Criteria for Adverse Events
CLL	Chronic lymphocytic leukemia
DLBCL	Diffuse Large B Cell Lymphoma
DLT	Dose-limiting toxicity
DOR	Duration of response
ECG	Electrocardiogram
ECOG	Eastern Cooperative Oncology Group
FL	Follicular Lymphoma
Hb	Hemoglobin
HCL	Hairy Cell Leukemia
LPL	Lymphoplasmacytic Lymphoma
MedDRA	Medical Dictionary for Regulatory Activities
MCL	Mantle Cell Lymphoma
MRD	Minimal Residual Disease
MZL	Marginal Zone Lymphoma
NCI	National Cancer Institute
ORR	Overall response rate
OS	Overall survival
PBMCs	Peripheral blood mononuclear cells
PD	Progressive disease, disease progression
PD	Pharmacodynamics
PK	Pharmacokinetic

PFS	Progression-free survival
PT	Preferred Term
Q1	25% quantile
Q3	75% quantile
QD	Daily
RP2D	Recommended Phase II dose
SAE	Serious adverse event
SAP	Statistical Analysis Plan
SD	Standard deviation
SLL	Small Lymphocytic Lymphoma
SMC	Safety Monitoring Committee
SOC	System Organ Class
TEAE	Treatment-emergent adverse event
$t_{1/2}$	Terminal half-life
t_{max}	Time corresponding to maximum observed plasma concentration
WHO-DD	World Health Organization Drug Dictionary
WM	Waldenström's macroglobulinemia

SUMMARY OF MAJOR CHANGES FROM PREVIOUS VERSION

Section	Version 1.0 (Previous)	Version 2.0 (Current)	Rationale
Structure			To comply to BeiGene SAP standard.
1. Introduction: Purpose 6.2 Data Analysis General Considerations	To describe the Dose Limit Toxicity (DLT) analysis of Part I only.	To describe the final analysis of both Part I and Par II.	To meet analysis needs when study proceeds to a final stage.
6.3 Subject Characteristics 6.5 Safety Analysis	To plan concise summaries serving the DLT determination purpose.	To plan full scope summaries with details, serving the final analysis purpose.	To meet analysis needs for final analysis; to comply to BeiGene SAP standard.
6.4 Efficacy Analysis	Not available.	To plan full scope summaries with details, serving the final analysis purpose.	To meet analysis needs for final analysis; to comply to BeiGene SAP standard.
6.6 PK Analysis	To plan concise summaries serving the DLT determination purpose.	To plan full scope summaries with details, serving the final analysis purpose.	To meet analysis needs for final analysis.
References, Appendices	To support content.	To support content.	To support updated content, comply to BeiGene SAP standard, and reflect latest safety interests for the product.

1 INTRODUCTION

The purpose of this statistical analysis plan (SAP) is to describe the procedures and the statistical methods that will be used to analyze and report results for BGB-3111-1002. The focus of this SAP is for the final planned analysis specified in the study protocol. This SAP is based on version 2.0 of the protocol, dated February 6, 2017.

The analysis details for Pharmacogenomics and Biomarker analyses are not described within this SAP.

2 STUDY OVERVIEW

This phase I study is to investigate the dose safety, tolerability, pharmacokinetics (PK) and Pharmacodynamics (PD) of BGB-3111 in Chinese patients with B-cell lymphoma, and to determine the recommended phase II dose (RP2D), based on the multi-dose, dose escalation phase I trials which has completed in Australia and New Zealand.

The study consists of two parts, the first part being the safety evaluation of two dose schedules, and the second part being the dose expansion. At the time of this SAP's composition, the first part was completed and 160 mg BID was chosen as the RP2D. Based on the results obtained, dose expansion studies were performed afterwards in patients with follicular lymphoma (FL) and marginal zone lymphoma (MZL).

Part I: Safety evaluation

According to the results of preclinical toxicological studies and the results of the phase I clinical study conducted in Australia and New Zealand, two regimens of BGB-3111 320 mg daily (160 mg BID, administered in the morning and at night, or 320 mg QD) and "3+3" design was adopted for the assessment. Three eligible subjects were enrolled for each dose regimen: first, if 1/3 (1 case) patients experienced dose-limiting toxicity (DLT), another 3 subjects were to be enrolled for continued observation. According to the preliminary results of the clinical study in Australia and New Zealand, two regimens of BGB-3111 320 mg dose group are both safe and tolerable. If none of the first 3 enrolled patients experienced DLT or only one patient among the 6 patients (including the supplemented 3 patients) experienced DLT, then it was presumed that the dose of BGB-3111 (160 mg BID and/or 320 mg QD) is safe and tolerable. Next each dose group would recruit new patients until about 10 patients per group (including the patients enrolled previously) to further assess the safety, tolerability, PK and preliminary PD of BGB-3111.

If more than two (inclusive) from 3 subjects or 6 subjects (including the supplemented 3 patients) experienced DLT within the DLT observation window (Day 1-28), it would be

concluded that the dose of BGB-3111 (160 mg BID and/or 320 mg QD) had exceeded the maximum tolerated dose (MTD) and dose would be decreased to 80 mg BID or 160 mg QD. Three eligible patients would be enrolled into each new dose group for re-assessment. If 80 mg BID or 160 mg QD is safe, the sponsor and investigator would analyze and discuss whether to move forward with an intermediate dose group (120 mg BID or 240 mg QD) according to the available safety data and PK data. If neither 160 mg QD and/or 80 mg BID was tolerable, the trial would be terminated.

During the trial, the sponsor, leading investigator and investigators would establish a safety monitoring committee (SMC) for ongoing safety assessment. SMC would decide the dose level and regimen for next group or if enroll more eligible patients, based on the efficacy data from the former dose level.

The toxicity observation period of BGB-3111 was 28 days and treatment observation period is from day 29 to three years. A subject will be considered end of study if he/she shows disease progression, intolerance to toxicity, death, withdrawal from the study, lost to follow-up, study termination by the sponsor, or having completed all 3 years of treatment, whichever occurs first. If a subject is judged to still have clinical benefit by the investigator at the end of 3-year treatment, he/she may be transferred to the extension study to continue BGB-3111 treatment until disease progression, intolerance, death, or withdrawal from study, under the approval by the sponsor's medical monitor.

The first study analysis was conducted after all the subjects have completed DLT assessment, and this analysis focused on the safety and tolerability of BGB-3111. PK and PD were also included in this analysis. According to the safety data of the Part I and the results of the Australian Phase I clinical trial, the dose of 160 mg BID has been selected as the RP2D. For subjects in part I, if the initial dose was 320 mg QD, after the completion of DLT assessment and with prior approval from investigator and sponsor's medical monitor, the administration regimen could be changed to 160 mg BID.

Part II: Dose expansion

To further evaluate the preliminary anti-tumor effects of BGB-3111 in Chinese subjects with FL or MZL, approximately 20 subjects with relapsed or refractory FL or MZL will be enrolled. The RP2D determined in Part I, i.e. 160 mg BID, will be used in the Part II.

To evaluate the preliminary efficacy without harming subject's benefit, the anti-tumor activity assessment will be conducted every 12 weeks (conducted approximate every 24 weeks 1 year later) after first dose. Subjects who do not experience disease progression may continue treatment until intolerable to toxicity, death, progression of the disease, withdrawal of informed

consent, or discontinuing treatment by the investigator when the risk is greater than the benefit. If a subject is judged to still have clinical benefit by the investigator at the end of 3-year treatment, he/she may be transferred to the extension study to continue BGB-3111 treatment until disease progression, intolerance, death, or withdrawal from study.

3 STUDY OBJECTIVES

3.1 PART I: SAFETY EVALUATION

3.1.1 Primary Objective

- To evaluate the safety and tolerability of BGB-3111 in Chinese patients with B-cell lymphoma and determine RP2D.

3.1.2 Secondary Objective

- To characterize the PK of single- and multiple- dose of BGB-3111 orally in Chinese patients with B-cell lymphoma.
- To evaluate the inhibitory effect of BGB-3111 on BTK in peripheral blood mononuclear cells (PBMCs).

3.1.3 Exploratory Objective

- [redacted]

3.2 PART II: DOSE EXPANSION

3.2.1 Primary Objective

- To evaluate the preliminary anti-tumor activity of BGB-3111 in subjects with FL or MZL.

3.2.2 Secondary Objective

- To further assess the safety and tolerability of BGB-3111 in patients with FL or MZL.
- To further characterize the PK of single- and multiple- dose of BGB-3111 orally in Chinese patients with B-cell lymphoma.

3.2.3 Exploratory Objective

- [redacted]

4 STUDY ENDPOINTS

4.1 PART I: SAFETY EVALUATION

4.1.1 Primary Endpoints:

- Safety and tolerability of BGB-3111: The occurrence of adverse events (AEs) and serious adverse events (SAEs) of each subject will be monitored according to NCI-CTCAE 4.03

grading criteria during the whole study and the safety of BGB-3111 regimen will be evaluated.

4.1.2 Secondary Endpoints:

- PK parameters of BGB-3111 after a single dose and multiple doses
- The inhibitory effect of BGB-3111 on BTK activity in PBMC will be assessed by measuring the proportion of BTK occupied in PBMC.

4.1.3 Exploratory Endpoints:

- [redacted]

4.2 PART II: DOSE EXPANSION

4.2.1 Primary Endpoints:

- ORR, CRR, PRR, duration of response (DOR) and PFS of BGB-3111 in subjects with FL or MZL.

4.2.2 Secondary Endpoints:

- Safety and tolerability of BGB-3111 in FL and MZL: The occurrence of AEs and SAEs of each subject will be monitored according to NCI-CTCAE 4.03.
- PK parameters of BGB-3111 after a single dose and multiple doses

4.2.3 Exploratory Endpoints:

- [redacted]

5 SAMPLE SIZE CONSIDERATIONS

Part I (Safety evaluation): The final sample size of the trial will depend on the assessed dose levels and toxicity occurrence of BGB-3111. It is expected that about 20 subjects will be enrolled in order to assess the safety and tolerability of BGB-3111 and determine RP2D.

Part II (Dose expansion): Approximately 20 subjects with relapsed or refractory FL and MZL will be enrolled to evaluate the anti-tumor activities of BGB-3111 at the RP2D level (160 mg BID). The sample size calculation is based primarily on the accuracy level of the ORR estimate. When the sample size is 20 subjects, if the ORR is 35%, the 95% exact confidence interval would be (15.4%, 59.2%), thus the anti-tumor activity of BGB-3111 in refractory FL or MZL will be clearly demonstrated.

6 STATISTICAL METHODS

6.1 ANALYSIS POPULATIONS

DLT Evaluable Population: Includes all subjects who received at least 75% of the planned doses of treatment during the DLT observation period (Day 28). This population will be used in the DLT summary.

Safety Population: Includes all subjects who received at least one dose of BGB-3111. The safety population will be used for all summaries (except DLT and PK).

PK Population: Includes all subjects who received at least one dose of BGB-3111 and at least one post dose PK concentration. Subjects who had major protocol deviations that would impact the absorption, distribution, metabolism and excretion of BGB-311 will be excluded.

Per-protocol (PP) Population includes subjects who received at least one dose of BGB-3111 and had no major protocol deviations. Criteria for exclusion from the PP will be determined and documented before the database lock. This may be used as a supplementary to the analysis of the primary (efficacy) endpoints of the dose expansion part if more than 15% of the safety population have major protocol deviations.

Summary of analysis populations will provide the number and percentage of subjects in each analysis population, out of all the subjects in the Safety Population.

6.2 DATA ANALYSIS GENERAL CONSIDERATIONS

Part I of the study is completed and this SAP describes the final analysis to be performed after study completion ("primary analysis" in the protocol). Study will be completed if all subjects experience disease progression, intolerance to toxicity, death, withdrawn from study, termination by sponsor or having completed all 3-year treatment. The final analysis will include the safety, tolerability, PK, PD, and efficacy of both regimens of BGB-3111. The results of these analyses will be included in the CSR.

Except for the DLT summary which is to be performed for Part I patients only, all other summaries will be performed for Part I patients only, Part II patients only, and all patients pooled. In Part I summaries, patients will be grouped by the two dose level/schedules (i.e. 160 mg BID and 320 mg QD) individually, then possibly pooled. A subject will be included in the dose level/schedule group per the first dose he/she received. They will also be grouped by disease type. In Part II summaries, all patients received the RP2D of 160 mg BID, therefore they will be grouped by disease type FL, MZL and then pooled. In the pooled summaries of both Part I and Part II patients, patients will be grouped by disease type only.

All centers will be pooled together in all analyses.

Statistical programming and analyses will be performed using SAS® (SAS Institute, Inc., Cary, NC, USA), version 9.3.

6.2.1 Definitions and Computations

Study treatment (study drug) for this study is BGB-3111.

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

In the situation where the event date is partial or missing, the date will appear partial or missing in the listings; Study day and any corresponding durations will be presented based on the imputations specified in Appendix A.

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

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

Definitions of efficacy variables are provided in Section 6.4 and definitions of safety variables are provided in Section 6.5.

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.
- Age will be calculated as the integer part of (date of informed consent – date of birth + 1)/365.25
- Time to and duration of image-based event endpoints (such as PFS and DFS) will be based on the actual date the radiograph was obtained rather than the associated visit date.
- Missing efficacy or safety data will not be imputed unless otherwise specified.
- For by-visit observed data analyses, percentages will be calculated based on the number of patients with nonmissing data as the denominator, unless otherwise specified.

- For continuous endpoints, summary statistics will include n, mean, standard deviation (SD), median, Q1, Q3 and range (minimum and maximum). The minimum and maximum will be reported to the same number of decimal places as the raw data recorded in the database. The mean, median, Q1 and Q3 will be reported to one more decimal place than the raw data recorded in the database. The SD will be reported to two more decimal places than the raw data recorded in the database. In general, the maximum number of decimal places reported shall be four for any summary statistic.
- For discrete endpoints, summary statistics will include frequencies and percentages. Percentages will be presented to one decimal place. Percentages will not be presented for zero counts.
- Unless specified otherwise, the visit window in the table 2 of protocol will be used in analysis. All assessment out of the visit window will be considered as unscheduled assessment and not included in by visit analysis. For by visit analysis, if multiple assessments happened in the same visit window, the closest assessment to the plan date/time (if collected) will be used in analysis. If more than one records have same distance to the plan date/time or multiple assessments on the planned date/time, the last non-missing one will be used.

6.2.3 Handling of Missing Data

Missing data will not be imputed unless otherwise specified. Missing dates or partially missing dates will be imputed conservatively for adverse events and prior/concomitant medications/procedures as provided in Appendix A.

When summarizing categorical variables, subjects with missing data are generally included in the denominator to calculate percentages unless otherwise specified. When needed, the category of "Missing" is created and the number of patients with missing data is presented.

When summarizing continuous variables, subjects with missing data are not included in calculations. No imputations are made.

By-visit endpoints will be analyzed using observed data, unless otherwise specified. For observed data analyses, missing data will not be imputed and only the observed records will be included.

If the start day of a subsequent anti-cancer therapy is missing, it will be assumed to be the first day of the month.

If only the day of death date is missing, the death will be assumed to be on the first day of the month if the last known alive data is earlier. If the last known alive date is later than the first day of the month, then the death date will be assumed to be the last known alive date plus 1 day.

No imputation of AE grades will be performed. TEAEs with missing CTCAE grade will only be summarized in the all-grades column.

If the assessment of the relationship of an AE to study treatments is missing, then the AE is assumed to be related to the study treatments in the safety analysis, but no imputation should be done at the data level.

6.2.4 Adjustment for Covariates

No adjustments for covariates are planned. Baseline factors may be used in the model as covariates as supportive analyses for endpoints.

6.2.5 Multiplicity Adjustment

Not applicable.

6.2.6 Data Integrity

Before pre-specified statistical analysis begins, the integrity of the data should be reviewed to assure fit-for-purpose. The data set for analysis should be an accurate and complete representation of the subjects' relevant outcomes from the clinical database. All data should be complete and reviewed up to a pre-specified cutoff date. Consistency checks and appropriate source data verification should be complete.

6.3 SUBJECT CHARACTERISTICS

6.3.1 Subject Disposition

The number (percentage) of subjects screened, treated, discontinued from study drug and discontinued from the DLT phase and the entire study will be summarized. The primary reason for end of treatment (study drug discontinuation) and end of study will be summarized by categories. If data cut-off date is before study completion, the number (percentage) of subjects ongoing in the treatment period and in the study at cut-off will also be summarized.

6.3.2 Protocol Deviations

Major protocol deviation criteria will be established and subjects with major protocol deviations will be identified and documented before the database lock. Major protocol deviations will be summarized by category. They will also be listed by each category.

6.3.3 Demographic and Other Baseline Characteristics

Demographic and other baseline characteristics will be summarized using descriptive statistics.

Continuous variables include age, height, weight, body mass index (BMI), vital signs; categorical variables include sex, age group (<65, ≥65), Eastern Cooperative Oncology Group (ECOG) performance score, serum virology and echocardiogram.

The number (percentage) of subjects reporting a history of disease and characteristic, as recorded on the CRF, will be summarized. Disease characteristics may include B-Cell Lymphoid Malignancies type, stage of disease at initial diagnosis and at study entry, genotype status for WM (MYD88, CXCR4), CLL (del(17p), del(11q), del(13q), p53 deletion, IgVH mutation), presence or absence of disease-related constitutional symptoms, lymphadenopathy, hepatomegaly, splenomegaly, time (years) from first diagnosis of B-Cell Lymphoid Malignancies to first dose date, time (years) since most recent of progression relapse to first dose date, baseline bone marrow cellularity and baseline bone marrow tumor cell infiltration. A listing of disease history will be provided.

6.3.4 Prior Anti-Cancer Drug Therapies and Surgeries

The number (percentage) of subjects with prior anti-B-cell lymphoma drug therapies, prior anti-B-cell lymphoma radiotherapy and prior anti-B-cell lymphoma surgeries will be summarized, as well as the number of these prior therapies. The therapies and surgeries with the same sequence/regimen number are counted as one prior therapy/surgery.

6.3.5 Prior and Concomitant Medications

Prior and concomitant medications will be coded using the World Health Organization Drug Dictionary (WHO DD) drug codes version September 2016, and will be further classified to the appropriate Anatomical Therapeutic Chemical (ATC) code.

The number (percentage) of subjects reporting prior and concomitant medications will be summarized by ATC medication class and WHO DD preferred term in the safety population. Prior medications are defined as medications that started before the first dose date. Concomitant medications will be defined as medications that (1) started before the first dose of study treatment and were continuing at the time of the first dose of study treatment, or (2) started on or after the date of the first dose of study treatment up to 30 days after the subject's last dose or initiation of a new anti-cancer therapy. A listing of prior and concomitant medications will be provided.

6.3.6 Medical History

Medical History will be coded using MedDRA (version 19.1). The number (percentage) of subjects reporting a history of any medical condition, as recorded on the CRF, will be summarized by System Organ Class (SOC) and Preferred Term (PT) in the safety population. A listing of medical history will be provided.

6.4 EFFICACY ANALYSIS

This section describes the efficacy analysis of both the exploratory endpoints of Part I and the primary endpoints of part II to be performed at study completion.

Response and progression status will be determined by the investigator using the evaluation criteria (IWCLL-CLL, Lugano criteria for NHL, IWWM-WM) detailed in protocol Appendix 3. A patient's best overall response is the best confirmed response recorded prior to the date of progressive disease, data cut, or start of new anti-cancer treatment, whichever is the earliest.

Sensitivity analyses will be performed for all the efficacy endpoints in the PP population, if more than 15% of the safety population have major protocol deviations.

Primary efficacy endpoints of part II will be summarized in FL and MZL patients respectively.

Overall Response Rate (ORR), Complete Response Rate (CRR), Partial Response Rate (PRR)

'Overall Response' in ORR is defined as a subject's best overall response being complete response (CR) or partial response (PR) for the NHL subjects; CR, CRi, nPR, PR, or partial response with lymphocytosis (PR-L) for the CLL subjects; and CR, very good partial response (VGPR), PR, or minor response (MR) for the WM subjects. ORR is defined as the proportion of subjects who achieve the 'overall response'. CRR is defined as the proportion of subjects who achieve the CR as the best overall response. PRR is defined as the proportion of subjects who achieve the PR or higher (or PR and VGPR for WM subjects) as the best overall response.

The number and proportion of subjects who achieve each best overall response in the evaluation criteria and who achieve 'overall response' as defined above will be summarized. Patients without postbaseline tumor assessment will be considered as non-responders. The 2-sided 95% exact binomial confidence interval of ORR, CRR, PRR and other proportions will be determined using the Clopper-Pearson method.

Progression Free Survival (PFS)

PFS is defined as the time (in months) from the date of first study treatment to disease progression or death (due to any cause), whichever occurs first. For purposes of calculating PFS, the start date of progressive disease is the date at which progression was first observed.

The duration of PFS will be right-censored for patients who met 1 of the following conditions: 1) no baseline disease assessments; 2) starting a new anti-cancer therapy before documentation of disease progression or death; 3) death or disease progression immediately after more than 1 consecutively missed disease assessment visit; and 4) alive without documentation of disease progression before the data cutoff date. For such patients, the primary analysis of PFS will be

right-censored according to the convention described in Table 1. These conventions are based on 2012 国家食品药品监督管理局《抗肿瘤药物临床试验终点技术指导原则》.

Table 1: Date of Progression or Censoring for Progression-free Survival

Situation	Date of Progression or Censoring	Outcome
No baseline disease assessments	Date of first dose	Censored
New anti-cancer treatment started before documentation of PD or death	Date of last disease assessment prior to start of a new anti-cancer treatment	Censored
Death or PD immediately after more than 1 consecutively missed disease assessment visit*	Date of last disease assessment visit without documentation of PD that is before the first missed visit	Censored
Alive and without PD documentation	Date of last disease assessment	Censored
Death or PD between planned disease assessments	Date of death or first disease assessment showing PD, whichever occurs first	Progressed
Death before first disease assessment	Date of death	Progressed

* the frequency of disease assessments is every 12 weeks in the first 48 weeks and every 24 weeks thereafter.

The distribution of PFS, including median PFS and PFS rate at selected timepoints such as 12 and 24 weeks, will be summarized descriptively using the Kaplan-Meier method. The 95% confidence interval for median and other quartiles of PFS will be generated by using Brookmeyer method (Brookmeyer and Crowley 1982), whereas the 95% confidence interval for PFS rate at selected timepoints will be generated by using Greenwood formula (Greenwood 1926). Duration of follow-up for PFS will be estimated by reverse Kaplan-Meier method (Schemper and Smith 1996). Kaplan-Meier curves will be constructed to provide a visual description of the PFS change with time.

Duration of Response (DOR)

Duration of response for responders (those who achieve overall response PR or above) is defined as the time interval (in number of days) between the date of the earliest qualifying response and the date of PD or death for any cause (whichever occurs earlier). Duration of response analysis will only include responders. Censoring rule for DOR will follow PFS censoring rule.

The distribution of DOR will be summarized descriptively using the Kaplan-Meier method. Median follow-up for DOR will be estimated according to the Kaplan-Meier estimate of potential follow-up also termed "reverse Kaplan-Meier" (Schemper and Smith 1996).

Minimal Residual Disease (MRD)

MRD is measured in patients with CR only. MRD clearance rate is defined as the proportion of patients with negative MRD status among all CR patients, and will be estimated with 95% exact binomial confidence interval. The summary will be provided separately for the assessment of MRD in the bone marrow and in the peripheral blood.

6.5 SAFETY ANALYSES

The incidence of treatment-emergent adverse events (TEAEs) will be summarized. Laboratory test results, vital signs, physical exams and ECG measurements 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.

6.5.1 Extent of Exposure

The following measures of overall extent of study drug exposure will be summarized.

- Duration of BGB-3111 treatment (weeks), defined as (date of last dose – date of first dose + 1) divided by 7
- Number (%) of patients dosed with BGB-3111 by time intervals, where a patient will be considered to have been dosed in an interval if the patient receives at least one dose of study drug in that interval
- Cumulative dose of BGB-3111 (mg) administered
- Average daily dose (dose intensity) of BGB-3111 (mg/day) administered, defined as the cumulative dose (mg) divided by the duration of BGB-3111 treatment (days) and calculated based on methods which account for treatment delays on the calculated dose intensity results (Hryniuk 1990, Longo 1990)
- Relative dose intensity of BGB-3111 (%, actual vs planned), defined as the ratio of the cumulative administered dose to the prescribed starting daily dose times the duration of BGB-3111 treatment (days)
- Number (%) of patients with BGB-3111 permanently discontinued
- Reasons for permanent discontinuation of BGB-3111
- Number (%) of patients with dose interruption for BGB-3111 and number of dose interruptions
- Number (%) of patients with dose reductions for BGB-3111 and number of dose

reductions

6.5.2 Adverse Events

AEs will be graded by the investigators using CTCAE v4.03. The AE verbatim descriptions (investigator terms from the CRF) will be classified into standardized medical terminology using the Medical Dictionary for Regulatory Activities (MedDRA). Adverse events will be coded to the MedDRA (Version 19.1) lower level term closest to the verbatim term. The linked MedDRA preferred term (PT) and primary system organ class (SOC) are also captured in the database.

6.5.2.1 DLT Analysis

Dose-limiting toxicity (DLT) in the toxicity observation period is defined in protocol Section 4.2 and were evaluated by the Safety Monitoring Committee (SMC) and used to determine the RP2D. No DLT was observed in Part I of this study and no summary or listing will be generated.

6.5.2.2 Other Adverse Events Analysis

A treatment emergent adverse event (TEAE) is defined as an AE that had an onset date on or after the date of the first dose of study treatment through 30 days after the last dose (permanent discontinuation of study treatment) or initiation of new anti-cancer therapy, whichever is earlier. Only those AEs that were treatment emergent will be included in summary tables. All AEs, treatment emergent or otherwise, will be presented in subject data listings.

An overview table, including the incidence of and the number of subjects with TEAEs, treatment-emergent serious adverse events (SAEs), treatment-related TEAEs, TEAEs with grade 3 or above, treatment-related TEAEs with grade 3 or above, treatment-related SAEs, TEAEs that led to death, treatment-related TEAEs that led to death, and TEAEs that led to treatment discontinuation, or dose interruption will be provided. Treatment-related AEs include those events considered by the investigator to be definitely related, possibly or probably related to study treatment or with missing assessment of the causal relationship.

The incidence of TEAEs will be reported as the number (percentage) of subjects with TEAEs by SOC and PT (in decreasing frequency of the pooled group). A subject will be counted only once by the highest severity grade according to CTCAE v.4.03 within an SOC and PT, even if the subject experienced more than 1 TEAE within a specific SOC and PT.

The following TEAE categories will be summarized by SOC and PT:

- all TEAEs
- treatment-emergent SAEs
- treatment-related TEAEs

- TEAEs with grade 3 or above
- treatment-related TEAEs with grade 3 or above
- treatment-related SAEs
- TEAEs that led to death
- TEAEs that led to treatment discontinuation, or dose interruption

The following TEAE categories will be summarized by PT in descending order of incidences:

- all TEAEs
- treatment-related TEAEs
- TEAEs with grade 3 or above

The following TEAE categories will be summarized by SOC, PT, and maximum severity:

- all TEAEs
- treatment-related TEAEs.

Incidence of TEAEs of interest, TEAEs of interest with grade 3 or above, treatment-related TEAEs of interest will be summarized by category and PT. TEAEs of interest will also be summarized by category, PT, and maximum severity. TEAEs of interest are defined in Appendix B.

Subject data listings of all AEs, SAEs, treatment-related AEs, grade 3 or above AEs, AEs that led to death and AEs that led to treatment discontinuation will be provided.

All deaths and causes of death will be summarized, including those occurred during the study treatment period and those reported during the survival follow-up period after treatment completion/discontinuation.

6.5.2.3 Exposure-Adjusted Incidence Rates (EAIR)

Given the long BGB-3111 treatment duration, an exposure adjusted analysis is also planned to analyze AEIs (TEAE of interest). The analysis restricts on the occurrence of the first event per patient and ignores the existence of later (multiple) events as these cannot be assumed to occur independent of previous events.

The incidence rate for a patient is derived from the duration of treatment exposure of that patient. A patient's duration of exposure is given either 1) by the time when the event has occurred (non-censored data), or 2) by the total duration of treatment in case the patient does not show the adverse event of interest (censored data). Depending on whether a patient has an adverse event or not, the duration of exposure enters the denominator in its non-censored or censored form, respectively.

The average EAIR per AEI considers the first event per patient per AEI only, and the corresponding exposure time in the denominator:

$$EAIR_{AEI} = \frac{\sum_{i=1}^{n} TEAE_{AEI,i}}{\sum_{i=1}^{n} t_{AEI,i}}$$

Whereby $TEAE_{AEI,i}$ represents the first TEAE among all AEI TEAEs of patient i and t_i as time when the TEAE occurs (non-censored data) or total duration of treatment if no event occurs (censored data).

6.5.3 Laboratory Values

Laboratory safety tests will be evaluated for selected hematology, chemistry, coagulation, immunology and urinalysis parameters described in Table 2.

Descriptive summary statistics for laboratory parameters and their changes from baseline will be summarized by visit. The analysis visit will be derived based on the rules defined in section 6.2.2. Laboratory results will be standardized and summarized using Standard International (SI) units, as appropriate.

Laboratory parameters (ALP, ALT, AST, total bilirubin, albumin, creatinine, calcium, glucose, uric acid, sodium, phosphorus, potassium, magnesium, hemoglobin, platelet counts, WBC count, neutrophil, lymphocyte, APTT and INR) that are graded in NCI CTCAE (v.4.03) will be summarized by shifts from baseline CTCAE grades to maximum post-baseline grades. In the summary of laboratory parameters by CTCAE grade, parameters with CTCAE grading in both high and low directions will be summarized separately. Number (percentage) of subjects with abnormal postbaseline laboratory values (those outside the normal range) and with grade 3 or above toxicity will be summarized by visit and across all post-baseline visits.

Subject data listings of selected hematology and serum chemistry parameters, immunology, urinalysis and coagulation will be provided.

Table 2 Laboratory Tests

Clinical biochemistry	**Hematology**	**Coagulation**	**Urinalysis**	**Immunoglobulin analysis and immunofixation**
Alkaline phosphatase (ALP)	Hemoglobin (Hgb)	Prothrombin Time (PT)	pH value	IgA
Alanine aminotransferase (ALT)	Reticulocyte count (optional test)	Activated Partial Thromboplastin Time (APTT)	Specific gravity	IgG

Aspartate aminotransferase (AST)	Platelet counts	International Normalized Ratio (INR)	Glucose	IgM
Albumin	WBC count with differential		Protein	Immunofixation[2]
Total and direct bilirubin	Neutrophil count		Ketones	
CO_2 (Carbon dioxide)	Eosinophil count		Hematuria	
Blood urea nitrogen (BUN)	Lymphocyte count		24-hour protein[1] Random urine protein	
Calcium[3]	Bands (optional test)			
Chlorine				
Creatinine				
Glucose				
Lactate dehydrogenase (LDH)				
Phosphate (optional test)				
Total protein				
Potassium				
Sodium				
Magnesium				
Phosphorus				
Uric Acid				
Creatinine clearance				

1. If urine protein is ≥2+ in urinalysis, a 24-hour urine protein will be detected with calculation of random urine protein to creatinine ratio.
2. All subjects with WM will receive the first immunofixation. If monoclonal protein exists, all immunoglobulin will be detected.
3. For hypocalcemia and hypercalcemia, serum calcium will be corrected using the formula: Corrected calcium = Serum calcium + 0.8 * (4 – serum albumin) where serum calcium is recorded in mg/dL and serum albumin is recorded in g/dL.

6.5.4 Vital Signs

Descriptive statistics for vital sign parameters (systolic and diastolic blood pressure [BP], pulse rate, temperature, weight and respiratory rate) and changes from baseline will be presented by visit. The analysis visit will be derived based on the rules defined in section 6.2.2. If the vital

signs assessment date is the same as the PK sample taken date, the pre-dose vital signs will be used in summary.

Blood pressure will also be summarized by shifts from baseline CTCAE grades to maximum post-baseline grades.

All vital signs including the post-dose assessment at the PK sample date, repeat assessments and unscheduled assessments will be listed by subjects and visits.

6.5.5 Physical Examination

Physical examination results will be listed without summary.

6.5.6 Electrocardiograms (ECG)

Descriptive statistics for ECG HR and intervals (PR, RR, QRS, QT, QTc, QTc-Bazzett [QTcB], and QTc-Fridericia [QTcF]) and changes from baseline will be presented by visit. The analysis visit will be derived based on the rules defined in section 6.2.2. If multiple readings are recorded, the average of the readings at each visit will be used for the summary. If there are both pre-dose and post-dose readings on the same date, pre-dose readings will be used in calculating change from baseline. Number and percentage of patients with at least one post baseline QTcF of > 450 msec, > 480 msec, or > 500 msec, and with the largest increase from baseline in the ≤ 30 msec, > 30 and ≤ 60 msec, and > 60 msec will be summarized. The interpretation of the ECG will be listed without summary. Any clinically significant abnormalities that are changes from baseline will be described in the CSR.

6.5.7 Eastern Cooperative Oncology Group (ECOG)

ECOG scores will be summarized by visit. A shift table from baseline to worst post-baseline in ECOG performance score will be summarized.

6.6 Pharmacokinetic Analyses

The analysis population for subjects with PK samples will contain all subjects who have at least one PK sample collected according to the protocol and laboratory manual. Nominal time may be used for the interim PK analysis and actual collection times will be used in the final analysis and reporting. For intensive PK profile, pharmacokinetic parameters will be derived using standard non-compartmental methods. In addition, PK data from this study, along with data from other trials will be included in the population PK analysis and will be reported separately.

For intensive PK profile on Day 1 and Day 2 of Week 1 and Day 1 of Week 2, plasma BGB-3111 concentration-time data will be summarized and displayed in both tabular and graphical

form. Individual and mean plasma concentration versus time data will be tabulated and plotted by dose level. The PK parameter will be estimated based on noncompartmental methods and will be computed in WinNonlin® Enterprise v.5.2 or higher.

The PK parameter estimates for a single dose profile include:

$AUC_{0-\infty}$	Area under the plasma concentration-time curve from zero to infinity
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

The PK parameter estimates following multiple doses include:

$AUC_{0-\tau}$	Area under the plasma concentration-time curve at each dosing interval
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
C_{min}	Minima drug concentration (trough) during any dosing interval at steady state
RAUC	AUC accumulation (AUC of Day 1 Week 2/AUC of Day 1 Week 1)
RC_{max}	Accumulation of maximum plasma concentration (C_{max} of Day 1 Week 2/C_{max} of Day 1 Week 1)

Additional PK parameters may be calculated if deemed appropriate. Estimates for these parameters will be tabulated and summarized by the dose level, schedule and collection day (i.e., n, Mean, Standard deviation, CV%, Median, Min, and Max, GeoMean and GeoCV%).

Since only 2 dose schedules (160 mg BID and 320 mg QD, equivalent daily dose) were studied, the dose proportionality analysis will not be performed.

7 INTERIM ANALYSIS

No formal interim analyses are planned for this study. Data from Part I of the study were reviewed by the SMC at the end of the DLT observation period. Summaries and analyses of subsets of the study data will be performed on a periodic basis for submission to professional meetings and for internal decision-making.

8 CHANGES IN THE PLANNED ANALYSIS

None.

9 REFERENCES

Brookmeyer B and Crowley J. A confidence interval for the median survival time. Biometrics, 38: 29-41, 1982

Common Terminology Criteria for Adverse Events, Version 4.03. Cancer Therapy Evaluation Program.14 June 2010.

Hryniuk W and Goodyear M. The calculation of received dose intensity. Journal of Clinical Oncology 8:1935–1937, 1990

Longo D, Duffey P, DeVita V, Wesley M, Hubbard S, and Young R. The calculation of actual or received dose intensity: A comparison of published methods. Journal of Clinical Oncology 9:2042–2051, 1991

Schemper M, Smith TL. A note on quantifying follow-up in studies of failure time. Controlled Clinical Trials 17:343–346, 1996

10 APPENDIX

Appendix A: Missing Data Imputation Rule

In general, missing or partial dates will not be imputed at data level. The following rules will be applied to the specific analysis and summary described below.

A.1 Prior/Concomitant Medications/Procedures

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

If start date of a medication 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 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 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 adverse event 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.

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 subject 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 subject known to be alive.

Appendix B: Events of Interest (EOI)

	EOI Name	Search Criteria
1	Hemorrhage	Haemorrhage terms (excl laboratory terms) SMQN
2	Major hemorrhage - defined as CTCAE v4.03 Grade ≥ 3 bleeding of any site, or central nervous system (CNS) bleeding of any grade.	Haemorrhage terms (excl laboratory terms) SMQN: Select PTs of any SOC other than the Nervous system disorders SOC that are Grade ≥ 3 and PTs of the Nervous system disorders SOC that are of any grade.
3	Atrial fibrillation	Atrial fibrillation PT, Atrial flutter PT
4	Hypertension	Hypertension SMQN
5	Second primary malignancies	Malignant Tumours SMQN
6	Diarrhea	Diarrhoea PT, Diarrhoea infectious PT, Clostridium difficile colitis PT, Diarrhoea haemorrhagic PT
7	Tumor lysis syndrome	Tumour lysis syndrome SMQN
8	Infection	Infections and Infestations SOC
9	Neutropenia	Neutropenia PT, Neutrophil count decreased PT, Febrile neutropenia PT, Agranulocytosis PT, Neutropenic infection PT, Neutropenic sepsis PT
10	Thrombocytopenia	Thrombocytopenia PT, Platelet count decreased PT
11	Anemia	Anaemia PT, Haemoglobin decreased PT

SMQN: Standard MedDRA Query Narrow; PT: Preferred Term; HLGT: High Level Group Term; SOC: System Organ Class