Parexel International

Secura Bio, Inc.

VS-0145-229

A Phase 2, Randomized, Open-label, 2-Arm Study Comparing 2 Intermittent Dosing Schedules of Duvelisib in Subjects with Indolent Non-Hodgkin Lymphoma (iNHL)

Statistical Analysis Plan

Version: 3.0

Parexel Project Number: 243661

SPONSOR SIGNATURE PAGE

Approved by: [redacted] _____
Date
(Interim)
Secura Bio, Inc.

PAREXEL SIGNATURE PAGE

Signature(s) below confirm that the Statistical Analysis Plan was developed in accordance with SOP-GDO-WW-019 and that it is approved for release.

This document has been approved and signed electronically on the final page by the following:

	Signatory
Author	[redacted]

TABLE OF CONTENTS

REVISION HISTORY

Version No.	Effective Date	Summary of Change(s)
0.1	29 May 2019	New document
0.2	12 August 2019	Incorporate internal reviewer's comments
1.0	22 August 2019	Incorporate final comments, finalize for signatures
1.1	26 August 2021	Updates per protocol v4.0, document sponsor change from Verastem to Secura Bio., add summaries for COVID-19 related protocol deviations
2.0	31August 2021	Incorporate all changes, finalize for signatures.
2.1	19May 2023	Updates per dry run comments: Add Kaplan Meier and Forest plots for efficacy analyses Add AESI for COVID-19 Change TEAE summaries currently by SOC, HLGT, and PT to by SOC and PT
3.0	09 August 2023	Incorporate final comments, finalize for signatures.

LIST OF ABBREVIATIONS

Abbreviation / Acronym	Definition / Expansion
AE	Adverse event
ALT	Alanine aminotransferase
AST	Aspartate aminotransferase
AT	All-treated
ATC	Anatomical Therapeutic Chemical
BID	Twice daily
BLQ	Below the limit of quantification
BMI	Body mass index
BP	Blood pressure
CI	Confidence interval
CLL	Chronic lymphocytic leukemia
CR	Complete response
CS	Clinically significant
CTCAE	Common Terminology Criteria for Adverse Events
ctDNA	circulating tumor DNA
DNA	Deoxyribonucleic acid
DOR	Duration of response
ECG	Electrocardiograms
ECHO	Echocardiogram
ECOG	Eastern Cooperative Oncology Group
eCRF	Electronic case report form
EQ-5D-3L	EuroQoL 5-dimensions 3-level
FL	Follicular lymphoma
HLGT	High-level group term
ICF	Informed consent form
iNHL	Indolent Non-Hodgkin's Lymphoma
IWG	International Working Group
KM	Kaplan-Meier

Abbreviation / Acronym	Definition / Expansion
LLQ	Lower limit of quantitation
LNRR	Lymph node response rate
MedDRA	Medical Dictionary for Regulatory Activities
mITT	Modified intent-to-treat
MUGA	Multigated acquisition
MZL	Marginal zone lymphoma
NCI	National Cancer Institute
NCS	Not clinically significant
ORR	Overall response rate
OS	Overall survival
PD	Progressive disease
PE	Physical examination
PFS	Progression-free survival
PK	Pharmacokinetics
PP	Per-protocol
PR	Partial response
PT	Preferred Term
QoL	Quality of Life
RBC	Red blood cell
RNA	ribonucleic acid
R/R	Relapsed or refractory
SAE	Serious adverse event
SAP	Statistical analysis plan
SD	Standard deviation
SI	International System of Units
SLL	Small lymphocytic lymphoma
SoA	Schedule of assessments
SOC	System organ class
TEAE	Treatment-emergent adverse event
TTF	Time to treatment failure

Abbreviation / Acronym	Definition / Expansion
TTFR	Time to the first response
VAS	Visual Analog Scale
WHODrug	World Health Organization - Drug Dictionary

1 INTRODUCTION

Dosing with duvelisib 25 mg twice daily (BID) has been shown to be efficacious in multiple Phase 1-3 studies including subjects with indolent non-Hodgkin lymphoma (iNHL). The efficacy, safety and tolerability of this dose supported the U.S. approval of duvelisib for the treatment of adult subjects with relapsed or refractory (R/R) chronic lymphocytic leukemia (CLL)/small lymphocytic lymphoma (SLL) after at least 2 prior therapies and accelerated approval in adult subjects with R/R follicular lymphoma (FL) after at least 2 prior systemic therapies. The current study will evaluate whether efficacy can be achieved and maintained, with acceptable or improved safety and tolerability, by inclusion of pre-specified 2-week drug holidays in 2 different 25 mg BID schedules in subjects with R/R iNHL.

This statistical analysis plan (SAP) provides a detailed description of the statistical methods and analyses to be carried out in support for study VS-0145-229.

The following will be collected but not entered into the clinical study database:

- Pharmacokinetic data (PRA Health Sciences)
- Pharmacodynamic data (Primity Bio)
- Exploratory biomarkers (Covance)

The analyses described in this SAP are based upon the following study documents:

- Study Protocol, Version 1.0 (April 4, 2019)
- Study Protocol, Version 4.0 (May 19, 2021)
- electronic Case Report Form (eCRF), Version 1.0 (August, 21, 2019)
- electronic Case Report Form (eCRF), Version 9.0 (23 Sep 2022)

2 STUDY OBJECTIVES

2.1 Primary Objective(s)

Evaluate the efficacy of duvelisib administered with prescribed drug holidays in subjects with iNHL.

2.2 Secondary Objective(s)

- Evaluate supportive efficacy measures of duvelisib administered with prescribed drug holidays in subjects with iNHL.
- Evaluate the safety and tolerability of duvelisib administered with prescribed drug holidays in subjects with iNHL.
- Characterize the pharmacokinetics (PK) of duvelisib metabolite(s) and duvelisib when duvelisib is administered with prescribed drug holidays in subjects with iNHL.

2.3 Exploratory Objective(s)

- Evaluate Quality of Life (QoL) in subjects treated with duvelisib with prescribed drug holidays.
- Evaluate potential biomarkers of clinical efficacy and/or safety of duvelisib administered with prescribed drug holidays in subjects with iNHL.

3 INVESTIGATIONAL PLAN

3.1 Overall Study Design and Plan

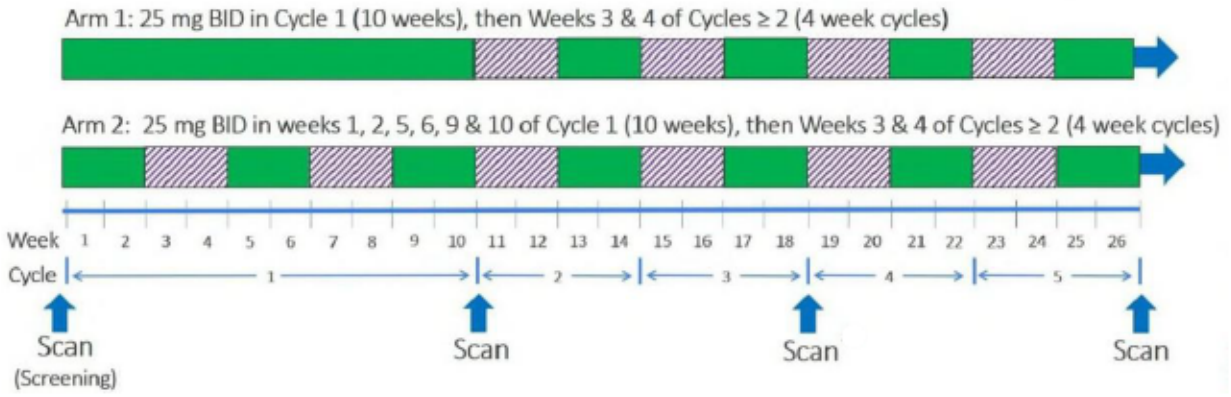
This is a Phase 2, randomized, open-label, multicenter, international 2-arm study designed to evaluate the efficacy and safety of alternative dosing schedules of duvelisib treatment in subjects with R/R iNHL who have received at least 1 prior systemic therapy.

Subjects will be stratified by the number of prior therapies (1 or > 1), bulky disease status (longest diameter of baseline lesion < 5 cm or ≥ 5 cm) and time since last recurrence (≥ 24 months or < 24 months).

Subjects will be randomized to Arm 1 or Arm 2 in a 1:1 fashion.

- Arm 1: duvelisib 25 mg BID for one 10-week cycle followed by 25 mg BID on Weeks 3 and 4 of each subsequent 4-week cycles until disease progression, unacceptable toxicity, or withdrawal.
- Arm 2: duvelisib 25 mg BID on Weeks 1, 2, 5, 6, 9 and 10 of one 10-week cycle, then on Weeks 3 and 4 of each subsequent 4-week cycles until disease progression, unacceptable toxicity, or withdrawal (Figure 1).

Figure 1: Treatment Scheme



Solid sections: duvelisib 25 mg BID; hashed sections: drug holiday. Subjects will be randomized 1:1 to Arms 1 and 2. Subjects will be assessed for disease response on Day 1 of Cycles 2, 4, 6, 9, 13, 17 and every fourth cycle thereafter until disease progression, unacceptable toxicity or withdrawal.

Abbreviations: BID: twice daily.

The study has a 2-stage design. For each arm: in the first stage, 15 subjects will be enrolled. Responses will be assessed by the Investigator according to the 2007 revised IWG criteria (Cheson 2007) (primary endpoint) and the 2014 Lugano criteria (Cheson 2014). The evaluation of the first stage will take place after the enrolled subjects have been followed for a minimum of 3 cycles. If there are fewer than 6 responders (partial response (PR) or complete response (CR)) among the first 15 subjects per arm, consideration may be given to stopping that arm. A treatment arm may also not proceed to the second stage if the totality of the data is compelling (e.g., accumulating enough data in Stage 1 to conclude the overall response rate (ORR) exceeds 30%). Otherwise, approximately 36 additional subjects will be enrolled for a total of 51 per arm. Enrollment will be continuous between Stage 1 and Stage 2. If fewer than 22 (43%) responders (PR or CR) are observed in 51 subjects, consideration may be given to concluding that the treatment schedule does not warrant further study in this population. Each arm will be evaluated separately.

It is anticipated that study enrollment will occur over approximately 2 years. The last subject visit will occur 2 years after the last subject is randomized. If the study ends prior to all subjects discontinuing treatment, any subject continuing to receive benefit will be provided the opportunity to continue to receive duvelisib. Long-term continued treatment may be provided via an expanded access program consistent with local regulations, such as the Secura Bio NPP. Study procedures and their timing are summarized in the schedule of assessments (SoA) (Protocol Section 1.3)

3.2 Endpoints

Primary Endpoint

Overall Response Rate (ORR) according to the 2007 revised International Working Group (IWG) Criteria (Cheson 2007)

Secondary Endpoints

- ORR according to 2014 Lugano criteria (Cheson 2014)

Note: with the exception of time to treatment failure, all additional secondary response endpoints listed below as well as progression-free survival (PFS), will be assessed separately using 2007 revised IWG criteria (Cheson 2007; referred to as "revised IWG criteria" for subsequent references in this document) and 2014 Lugano criteria (Cheson 2014; referred to as "Lugano criteria" for subsequent references in this document).

- ORR, at 6, 12, 18, and 24 months after first dose of study intervention
- PFS
- Time to treatment failure (TTF)
- Duration of response (DOR)
- Overall survival (OS)
- Lymph node response rate (LNRR)
- Time to the first response (TTFR)
- Adverse events (AEs), serious AEs (SAEs), vital signs, physical examinations, and clinical laboratory values
- PK parameters for duvelisib and metabolite(s)

Exploratory Endpoints

- Eastern Cooperative Oncology Group (ECOG) Performance Score
- EuroQoL 5 dimensions 3-level (EQ-5D-3L) questionnaire responses
- Blood assessments of immune cell populations, chemokines, cytokines, proteins and/or circulating tumor (ctDNA)
- Assessments of protein, DNA and or RNA in fecal microbiota
- Tumor biopsy evaluation of biomarkers such as gene and copy variation, RNA expression, protein expression, and/or immune cell content

3.2.1 Efficacy Variables

Assessment of disease response and progression status in all subjects will be evaluated by the Investigator using 2007 revised IWG criteria (Cheson 2007) and 2014 Lugano criteria (Cheson 2014). Response and progression will be evaluated using the following procedures, on Day 1 of Cycles 2, 4, 6, 9, 13, 17 and every fourth cycle thereafter, as shown in the SoA (Protocol Section 1.3).

- Focused physical examination (PE) (including assessment of liver and spleen and review of disease-related constitutional symptoms)
- Imaging scans of chest, abdomen, and pelvis
- Bone marrow biopsy/aspirate (as necessary to confirm CR)

These evaluations will be done until progressive disease (PD) is documented, other anticancer therapy is initiated, or death occurs. For subjects who discontinue study intervention for reasons other than radiologic disease progression, response assessments will be performed during protocol-specified long-term follow-up.

3.2.2 Safety Variables

- Physical examinations, including assessment of liver/spleen size, clinical assessment of tumor masses (if evaluable by PE), review of disease-related constitutional symptoms (B symptoms: symptoms of fever [i.e., temperature > 38°C/100.4°F] without evidence of infection, weight loss, and drenching night sweats without evidence of infection), neurologic examination, height
- Vital signs (semi-supine systolic and diastolic blood pressure [BP] and pulse, temperature, respiratory rate)
- 12-lead electrocardiograms (ECGs): PR interval, QRS interval, RR interval, QT interval and Fridericia's correction [QTcF])
- Echocardiogram (ECHO)/multigated acquisition (MUGA) scans (may be performed as clinically indicated per Investigator's judgement)
- Clinical laboratory tests (hematology, clinical chemistry, urinalysis and immunoglobulins parameters)
- Adverse event (AE) assessments
- Concomitant medication assessments

3.3 Exploratory Variables

- ECOG Performance Status
- EQ-5D-3L questionnaire (Subject Self-reported Health-related QoL)
- Blood assessments of immune cell populations, chemokines, cytokines, proteins and/or circulating tumor (ctDNA)
- Assessments of protein, DNA, and or RNA in fecal microbiota
- Tumor biopsy evaluation of biomarkers such as gene and copy number variation, RNA expression, protein expression, and/or immune cell content

4 STATISTICAL METHODS

4.1 Data Quality Assurance

All tables, figures and data listings to be included in the report will be independently checked for consistency, integrity and in accordance with standard Parexel procedures.

4.2 General Presentation Considerations

Baseline is defined as the last available pre-treatment assessment (scheduled or unscheduled). 'Treatment Day' will be calculated relative to the date of randomization, i.e., Treatment Day = Assessment Date - First Dose Date + 1.

The following visit windows will be defined for the purposes of determination of ORR per Lugano and revised IWG criteria at 6, 12, 18 and 24 months:

Timepoint	Visit Window
Month 6	155 – 211 (183 ± 28 days) from first dose
Month 12	351 - 407 days (379 ± 28 days) from first dose
Month 18	463 – 519 days (491± 28 days) from first dose
Month 24	687 – 743 days (715 ± 28 days) from first dose

For other by-visit summaries, data will be categorized based on the scheduled visit (i.e. cycle) specified on the eCRF on which the data were collected.

Continuous data will be summarized in terms of the mean, standard deviation (SD), median, minimum, maximum and number of observations, unless otherwise stated. Continuous data that are expected to be skewed will be presented in terms of the maximum, upper quartile, median, lower quartile, minimum and number of observations. 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, lower quartile and upper quartile 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.

Categorical data will be summarized in terms of the number of subjects providing data at the relevant time point (n), frequency counts and percentages. Any planned collapsing of categories will be detailed in the SAP text and the data displays.

Percentages will be presented to one decimal place. Percentages will not be presented for zero counts. Percentages will be calculated using n as the denominator. If sample sizes are small, the data displays will show the percentages, but any textual report will describe frequencies only.

Changes from baseline in categorical data will be summarized using shift tables where appropriate.

P-values greater than or equal to 0.001, in general, will be presented to three decimal places. P-values less than 0.001 will be presented as "<0.001".

Confidence intervals (CIs) will be presented to one more decimal place than the raw data.

Time-to-event endpoints (PFS, DOR, OS, TTF) will be estimated using the Kaplan-Meier (KM) method. Descriptive statistics for analyses of such endpoints will include number of events, number of subjects censored, estimated 25% quartile, estimated median, estimated 75% quartile, and 95% CI for median estimate. Detailed descriptions of the time-to-event analyses to be performed for secondary endpoints are described in Section 4.10.3.

4.3 General Considerations

4.3.1 Missing Data

In general, other than for partial dates, missing safety data will not be imputed and will be treated as missing. However, in all data presentations (except listings), safety assessment values below the limit of quantitation (BLQ) will be imputed as 'lower limit of quantitation (LLQ)', where LLQ will be replaced with the value for the lower limit of quantitation. In listings, BLQ values will be reported as '<LLQ'.

4.3.1.1 Imputation of Partial Dates

For missing medication and AE start dates, the following will be applied:

- If year is missing (or completely missing), no imputation will be performed.
- If year is present but months and day are missing, or year and day are present but month is missing, the missing day/month will be imputed as 01 January.
- If year and month are present but day is missing, the missing day will be imputed as first day of the month.

For missing medications and AE end dates, the following will be applied:

- If year is missing (or completely missing), no imputation will be performed.
- If year is present but month and day are missing, or year and day are present but month is missing, the missing day/month will be imputed as 31 December.
- If year and month are present but day is missing, the missing day will be imputed as last day of the month.
- If a subject dies in the same month/year of partial stop date, the date of death will be used.

In addition, for AEs, if for a partial start date, the AE start date could (when also considering the AE end date) potentially be on the first dose of study intervention date, the AE start date will be imputed with the first dose of study intervention date to assume a "worst case" scenario; e.g. AE from UNK-Jun-2017 to 23-Jul-2017 with the first dose of study intervention date 21-Jun-2017, then the AE start date will be imputed to 21-Jun-2017.

4.3.1.2 Handling Missing Dates/Months/Years for Determination of TEAEs

Adverse events (AEs) with incomplete onset dates will be handled as follows for the purpose of determining treatment emergence according to the definition provided in Section 4.11.2.

- If the start/end date of an AE is partially missing, the date will be compared as accurately as possible with the first dose of study drug and the date of last dose of study drug plus 30 days. The AE will be assumed to be treatment emergent if it cannot be definitively shown that the AE did not start or worsen during the treatment-emergent period.
- If the start date is completely missing, an AE will be considered treatment-emergent unless the stop date is before first dose of study drug.

The original partial or missing date will be displayed in the AE listings.

4.3.1.3 Handling Missing Dates/Months/Years for Concomitant Medications

Prior or concomitant medications with incomplete start dates will be handled as follows for the purpose of determining whether a non-study medication is a concomitant medication

- If the start/stop date of a medication is partially missing, the date will be compared as accurately as possible with the first dose date of study drug and the last dose date of study drug plus 30 days. The medication will be assumed to be concomitant if it cannot be definitively shown that the stop date is before first dose date of study drug, or the start date is more than 30 days after the last dose date of study drug.
- If the start/stop dates are completely missing, a medication will be considered concomitant.

4.3.1.4 Handling Missing Dates/Months/Years for Disease History

For the purpose of calculating the duration from initial diagnosis, partial/missing dates for diagnosis will be imputed as follows:

- If both day and month are missing and the year is prior to the year of screening, the imputed date will be 01 July
- If both day and month are missing and year is the same as the year of screening, the imputed date will be the middle point between 01 January of the year and the screening date. If the middle point falls between two dates, the first of the two dates will be used.
- If day is missing and the month and year are prior to the month and year of screening, the imputed date will be the 15th day of the month.
- If day is missing and the month and year are the same as the month and year of screening, the imputed date will be the middle point between the first day of the month and the screening date. If the middle point falls between two dates, the first of the two dates will be used.
- No imputation will be performed if the year is missing.

4.3.2 Standard Dictionaries for Coding

The following dictionaries will be used to code medical verbatim terms for the clinical data below.

Types of Data	Coding Tools and Latest Version
Adverse events	Medical Dictionary for Regulatory Activities (MedDRA) Version 24.0
Prior and concomitant medications	World Health Organization Drug Dictionary (WHODrug) March 2021
Prior systemic therapies	WHODrug March 2021
Subsequent anticancer therapies	WHODrug March 2021
Medical history	MedDRA Version 24.0

4.4 Software

All report outputs will be produced using SAS® version 9.3 or a later version in a secure and validated environment.

4.5 Analysis Sets

Modified Intent-to-treat (mITT) Analysis Set

The modified intent-to-treat (mITT) analysis set will include all subjects who receive at least one dose of study drug. The mITT analysis set will serve as the primary analysis set for all efficacy endpoints and baseline characteristics.

All-treated (AT) Analysis Set

The all-treated (AT) analysis set will include all subjects who received at least one dose of study drug. The AT analysis set will be the primary analysis set for all safety endpoints.

Per-protocol (PP) Analysis Set

The per-protocol (PP) analysis set includes all subjects in the mITT analysis set who do not violate the protocol in a way that would significantly affect the study outcome. Subjects who meet any of the following criteria may be excluded from the PP analysis set:

- Do not have histologically confirmed diagnosis of iNHL: FL Grades 1 to 3a, marginal zone lymphoma (MZL; splenic, nodal, or extranodal), or SLL
- Have not received at least 1 prior systemic regimen for iNHL
- Must have at least one bi-dimensionally measurable lesion ≥ 1.5 cm (which has not been previously irradiated) according to 2007 revised IWG response criteria (Cheson 2007)

- Clinical or histological evidence of transformation to a more aggressive subtype of lymphoma or Grade 3b FL or Richter's transformation or CLL
- Never received study drug
- Received concomitant anticancer therapy prior to observation of ORR

Subjects with major protocol deviations not listed above and with COVID-19 related protocol deviations will be considered for removal from the PP analysis set on a case by case basis. The PP analysis set will be a secondary analysis set for the analysis of ORR. Final determination of subjects to be included in the PP analysis set will be determined prior to database lock and unblinding of treatment assignment.

4.6 Study Subjects

4.6.1 Disposition of Subjects

A clear accounting of the disposition of all subjects who enroll in the study will be provided, from screening to study completion.

The disposition of subjects will include the number and percentage of subjects for the following categories: the number randomized, the number and percentage randomized but not dosed, and the number and percentage dosed. These categories will be summarized for each treatment arm and for the two treatment arms combined (total). The percentages will be based on all randomized subjects. An end-of-treatment disposition (still on treatment vs. discontinued from treatment) will be provided for each treatment arm and for the two treatment arms combined (total) based on the AT analysis set. The primary reason for discontinuation will be included in the table. A listing of these subjects will also be provided, broken down by center and treatment arm.

An end-of-study disposition (still on study vs. discontinued from study) will also be provided for each treatment arm and for the two treatment arms combined (total) based on the mITT analysis set. The primary reason for study discontinuation will be included in the table. A listing of these subjects will also be provided, broken down by center and treatment arm.

A summary of strata as captured in IRT will be presented by treatment arm and for the two treatment arms combined (total) based on the mITT analysis set. A summary table and a listing of discrepancies in data for stratification factors between those captured in the IRT and those reported on the eCRF will be presented. Both the summary table and listing will be displayed by treatment arm.

All subjects in the mITT population will also be summarized by site and by country.

Upon database release, protocol deviation and analysis set outputs will be produced and will be sent to Secura Bio for review. An analysis set classification meeting will be arranged to discuss the outputs and to decide which subjects and/or subject data will be excluded from certain

analyses. Decisions made regarding the exclusion of subjects and/or subject data from analyses will be made prior to unblinding and will be documented and approved by Verastem.

The following analysis set summaries will be prepared:

- The number and percentage of subjects entering and completing each phase of the study by treatment group and overall for each analysis set (Analysis set: mITT).
- A by-subject listing of analysis set details presented by treatment group that will include: center, subject identifier, inclusion/exclusion flag for each set and reason for exclusion from each set. All subjects screened should appear on this listing.

4.6.2 Protocol Deviations

Major protocol deviations are defined as those deviations from the protocol likely to have an impact on the perceived efficacy and/or safety of study treatments. The impact of major protocol deviations on the efficacy and/or safety results will be investigated by assessing the robustness of the study results and conclusions to the choice of analysis set, both including and excluding data potentially affected by major protocol deviations, if applicable.

Major protocol deviations and any action to be taken regarding the exclusion of subjects or affected data from specific analyses are defined in the Protocol Deviation Specification.

The following summaries will be presented:

- The number and percentage of subjects with a major protocol deviation by treatment group and overall and by type of deviation (Analysis set: mITT).
- The number and percentage of subjects with any COVID-19 related protocol deviation (either major or minor) by treatment group and overall and by type of deviation (Analysis set: mITT).
- A by-subject listing of major protocol deviations and COVID-19 related protocol deviations will be provided.

4.7 Demographic and Other Baseline Characteristics

Demographic variables and other baseline characteristics will be summarized for each treatment arm based on all randomized subjects (mITT analysis set) and for the two treatment arms combined (total). The variables to be summarized will include (but will not be limited to):

- Age (continuous variable)
- Age group (< 65 vs. ≥ 65)
- Gender
- Race
- Ethnicity
- Body mass index (BMI), calculated as weight (kg)/(height [m])^2

- Stratification factors used in the randomization
  - Number of prior therapies (1 or >1)
  - Bulky disease status (longest diameter of baseline lesion < 5 cm or ≥ 5 cm)
  - Time since last recurrence (≥ 24 months or < 24 months)
- ECOG performance status
- Number of prior therapies (< 3 or ≥ 3)
- Months from initial diagnosis to randomization (calculated as date of randomization – date of initial diagnosis + 1)/(365.5/12)
- Months from last recurrence (calculated as date of randomization – date last recurrence + 1)/(365.5/12)
- iNHL subtype: FL, SLL or MZL; for MZL subjects type of MZL will also be categorized: splenic, nodal, or extranodal
- Specific prior therapies received (i.e., prior surgery, radiotherapy, chemotherapy, and other systemic anticancer therapies, including timeframe in which they were received, where applicable)

Age will be calculated as the number of complete years between a subject's birth year and the date of informed consent.

4.8 Prior and Concomitant Medication

The World Health Organization - Drug Dictionary (WHODrug) will be used for concomitant and prior medications coding. Medications will be considered as prior if they stopped before the date of first dose of study drug. Medications will be considered concomitant if they were taken at any time between the first dose of study drug and 30 days after the last dose of study drug, inclusive. If the start date or end date of a medication is completely or partially missing refer to Section 4.3.1.3 for the algorithm to determine whether a medication is concomitant.

Prior medications and concomitant medications will be summarized separately. Both summaries will be based on the AT analysis set. Medications will be summarized by Anatomical Therapeutic Chemical (ATC) classification level 2, and preferred drug name for each treatment arm. The summary will be sorted alphabetically by ATC level 2 and then by decreasing frequency by preferred drug name. A subject taking the same drug multiple times will be counted only once.

A listing will be provided for all non-study medications taken on the study. An identifier will be provided to show if a medication is prior or concomitant. Medications that started more than 30 days after the last dose of study drug will be identified as post-treatment. Additional groupings of medications will be defined to identify medications administered to treat AEs specific to the duvelisib risk profile.

4.9 Medical History

Medical history will be summarized and based on the AT analysis set. Medical history will be summarized by system organ class (SOC) and preferred term (PT) using the number and percentage of subjects who had at least one occurrence of a SOC or PT. The summary will be sorted alphabetically by SOC and by decreasing frequency of PT within an SOC.

4.10 Efficacy Evaluation

4.10.1 Analysis and Data Conventions

This study will test the null hypothesis that the ORR for each arm is $\leq 30\%$ against the alternative that ORR is $\geq 55\%$. Each treatment arm will be evaluated separately; no formal comparisons between treatment arms are planned. The ORR (proportion of subjects achieving a CR or PR) will be estimated, along with the 2-sided 95% exact CI.

4.10.1.1 Multi-center Studies

Statistical analyses will not control for sites; however, subgroup analyses for region (North America or Rest of World) will be performed for the primary endpoint, as well as the secondary endpoint of ORR per Lugano criteria.

4.10.1.2 Adjustments for Covariates

Adjustments for covariates will not be performed for this study

4.10.1.3 Handling of Dropouts or Missing Data

In general, values for missing data will not be imputed unless methods for handling missing data are specified. For analyses of ORR, subjects who do not exhibit a response will be classified as non-responders. For analyses of time to event endpoints, including DOR, PFS, OS, and TTF, subjects for whom the event is not observed will be censored as described in Section 4.10.3.

4.10.1.4 Multiple Comparisons/Multiplicity

Each treatment arm will be analyzed separately with no adjustment for multiplicity.

4.10.1.5 Interim Analyses

No formal interim analysis is planned. Study data for each arm will be evaluated at the end of Stage 1 to determine whether a study arm should be terminated or continued to Stage 2. Please see additional details of the 2-stage design in Section 4.13 below.

4.10.1.6 Examination of Subgroups

The uniformity of the treatment effect for ORR (using both revised IWG and Lugano criteria) will be examined for the following subgroups:

1. Stratification factors
    - Number of prior therapies (1 vs. > 1)
    - Bulky disease status (longest diameter of baseline lesion < 5 cm vs. ≥ 5 cm)
    - Time since last recurrence (≥ 24 months vs. < 24 months)
2. Baseline ECOG score (0 vs. ≥ 1)
3. Region of world (North America vs. Rest of World)
4. Additional characterization of number of prior systemic regimens (1, 2, or ≥ 3)
5. iNHL subtype (FL, MZL, SLL)
6. Prior treatment with rituximab (yes vs. no)
7. Prior treatment with bendamustine (yes vs. no)
8. Age (< 60 years, ≥ 60 years)
9. Sex (male, female)
10. Ethnic origin (Caucasian, non-Caucasian)

Forest plots will be used to plot response rate by study Arm for each subgroup.

4.10.2 Primary Efficacy Variable – ORR

The ORR (proportion of subjects achieving a CR or PR using the number of subjects in the mITT analysis set as the denominator) will be estimated, along with the 2-sided 95% exact CI using the Clopper-Pearson exact method for binomial proportions. Responses for the primary ORR analysis will be assessed using the 2007 revised IWG criteria (Cheson 2007). Missing data will be imputed by assuming that any subjects not exhibiting a response (CR or PR) are non-responders. Responses occurring after the start of subsequent anticancer therapy will not be included in the numerator.

The primary analysis of ORR will be performed 12 months after the last subject is randomized, using the time windows specified in Section 4.2 as applicable.

A by-subject listing of the primary efficacy data will be provided.

4.10.3 Secondary Efficacy Variables

The secondary efficacy analyses will be conducted on the mITT analysis set.

- ORR according to 2014 Lugano criteria (Cheson 2014) will be analyzed in the same fashion as the primary endpoint

Note: with the exception of TTF, all additional secondary response endpoints (described below) as well as PFS will be assessed separately using 2007 revised IWG criteria (Cheson 2007) and 2014 Lugano criteria (Cheson 2014)

- ORR, at 6, 12, 18, and 24 months after first dose of study intervention, defined as the proportion of subjects achieving CR or PR at each of these time points.

- PFS will be assessed using KM methods from time of first dose of study intervention to PD or death. Median PFS will be estimated, along with the 95% CI. The 6-month and 12-month PFS estimates will also be presented. See Table 1 for event and censoring definitions. Kaplan-Meier survival curves by arm will be produced.

- TTF will be calculated as the time from first dose of study intervention until discontinuation for any reason and will be summarized using KM methods. Median TTF will be estimated, along with the 95% CI. The 6-month and 12-month TTF estimates will also be presented. Subjects who are still on treatment at time of data cutoff will be censored at their last dose date. Kaplan-Meier survival curves by arm will be produced.

- DOR will be calculated for those subjects with a CR or PR from the time of first response to PD using KM methods. Median DOR will be estimated, along with the 95% CI. The 6-month and 12-month DOR estimates will also be presented. See Table 1 for event and censoring definitions. Kaplan-Meier survival curves by arm will be produced.

- OS will be assessed using KM methods from time of first dose of study intervention to death. Median OS will be estimated, along with the 95% CI. The 6-month and 12-month OS estimates will also be presented. Subjects without documented death will be censored at their last known alive date. Kaplan-Meier survival curves by arm will be produced.

- LNRR will be calculated as the proportion of subjects achieving $\geq$ 50% decrease in the sum of the product of the diameters of target lymph nodes. A 2-sided 95% exact CI for LNRR will also be estimated. A waterfall plot of minimum percent change in the sum of the product of the diameters by arm will be produced.

- TTFR will be calculated for those subjects with a CR or PR from the time of first dose of study intervention to time of first CR or PR. TTFR will be summarized descriptively (mean, SD, median, range). Kaplan-Meier survival curves by arm will be produced.

A by-subject listing of DOR, PFS and OS data will be provided.

Table 1: Primary PFS Censoring/Event Rule

Situation	Date of Event or Censoring	Outcome
No adequate baseline disease status assessment	Date of randomization + 1 day	Censored

No adequate post-baseline disease status assessment unless death occurs prior to first scheduled post-baseline assessment	Date of randomization + 1 day	Censored
No documented progression or death	Date of last adequate disease status assessment	Censored
Documented progression with ≤1 missing scheduled disease status assessment before progression	Date of the earliest assessment that results in a finding of unequivocal progression	Event
Death before progression being documented with ≤1 missing scheduled disease status assessment before death	Date of death	Event
Documented progression or death following a gap between adequate disease status assessments equivalent to 2 missed visits	Date of last adequate disease status assessment before the gap	Censored
New anticancer treatment or procedure started before documented progression	Date of last adequate disease status assessment prior to start of new anticancer therapy	Censored

Note: Disease status assessment includes radiologic documentation (CT, PET/CT, MRI scans) (chest, abdomen and pelvis, others as clinically indicated), bone marrow biopsy/aspirate (may not be required of all patients at all scheduled disease status assessments), and focused physical examination (disease related symptoms, liver and spleen assessment). An adequate baseline disease assessment is any assessment with measurable target lesions from scans. An adequate post-baseline disease status assessment is any disease status assessment for which a disease status of CR, PR, SD or PD is arrived per revised International Working Group (IWG) response criteria for malignant lymphoma[2].

4.11 Safety Evaluation

All safety summaries and analyses will be based upon the AT Analysis Set as defined in Section 4.5.

4.11.1 Extent of Exposure

Extent of exposure will be summarized for each treatment arm based on the AT analysis set. Extent of exposure will be summarized for the following variables:

- Duration (months): (date of last dose – date of first dose + 1) divided by (365.25/12).
- Number of cycles started (continuous and categorical)
- Relative dose intensity (overall and by cycle), defined as 100% × (total dose received)/(planned cumulative dose for the duration of treatment). The planned cumulative dose for the treatment duration is defined as 25 mg BID × total number of days on which duvelisib is scheduled to be administered during the interval between first dose date and last dose date. The treatment scheme is provided in Section 3.1.
- Number and percentage of subjects with a dose reduction
- Time from first dose to first dose reduction

- Total days spent at dose reduction: last date of dose reduction – first date of dose reduction + 1
- Number and percentage of subjects with an unplanned dose hold (excluding the planned 2-week holidays for each arm)
- Number and percentage of subjects with an unplanned dose hold < 7 days, ≥ 7 days
- Time from first dose to first unplanned dose reduction
- Number and percentage of subjects for whom study drug was discontinued

A by-subject listing will be presented for exposure to study drug

4.11.2 Adverse Events

Adverse events will be coded using Medical Dictionary for Regulatory Activities (MedDRA) Version 22 or higher. The severity grade of the AE will be assessed by the Investigator according to the National Cancer Institute Common Terminology Criteria for Adverse Events (NCI-CTCAE, Version 5.0). If an AE is not included in the NCI-CTCAE Version 5.0, the severity grade of the AE will be assessed by the Investigator according to the study protocol.

A treatment-emergent adverse event (TEAE) is defined as any AE that emerges or worsens in the period from the first dose of duvelisib to 30 days after the last dose of duvelisib. The onset date of an AE will be compared to the first dose date and the last dose date plus 30 days (inclusive) to determine whether the AE is treatment-emergent or not. Refer to Section 4.3.1.2 for the algorithm to determine whether an AE is treatment emergent if the onset date of an AE is completely or partially missing.

TEAEs will be summarized for each treatment arm by MedDRA SOC and PT, or PT only, or SOC only. For summary tables by SOC and PT, SOC will be sorted alphabetically, and PT will be sorted by decreasing frequency in Arm 1 within each SOC. For summary tables by PT only, PT will be sorted by decreasing frequency. For selected summaries (indicated in the list below), TEAEs will be summarized by SOC, high-level group term (HLGT) and PT, sorted by most frequently occurring SOC.

If multiple TEAEs of the same PT are reported for a subject, only the maximum grade observed for this PT will be included in summary of TEAEs by grade; the subject will be counted only once in the number of subjects for this PT and only once for the number of subjects for the SOC to which the PT is classified.

An overview TEAE summary table will be provided which will include the number of subjects with AEs in selected categories. In addition, TEAEs will be summarized for the following categories, and will be tabulated by SOC and PT, unless otherwise specified.

- TEAEs (all grades, all causalities)
- Treatment-related TEAEs (all grades)
- TEAEs by maximum grade (all causalities)

- Treatment-related TEAEs by maximum grade
- Grade 3 or higher TEAEs (all causalities)
- Grade 3 or higher treatment-related TEAEs
- Treatment-emergent SAEs (all causalities)
- TEAEs resulting in discontinuation of study drug (all causalities)
- TEAEs resulting in dose hold (all causalities)
- TEAEs resulting in dose reduction (all causalities)
- TEAEs resulting in dose hold or reduction (all causalities)
- TEAEs resulting in any dose modification (dose hold, dose reduction and/or discontinuation)
- TEAEs occurring during the first 10 weeks of treatment
- Treatment-related TEAEs resulting in discontinuation of study drug
- TEAEs by resolution status (resolved/resolved with sequelae vs. not resolved)
- Grade 3 and or higher TEAEs by resolution status (resolved/resolved with sequelae vs. not resolved)
- Treatment-emergent SAEs by resolution status (resolved/resolved with sequelae vs. not resolved)
- TEAEs resulting in death (all causalities)
- Treatment-related TEAEs resulting in death
- TEAEs reported in ≥ 20% of all treated subjects by PT (all grades, all causalities)
- Grade 3 or higher treatment-emergent AEs reported in ≥ 20% of all treated subjects by PT (all causalities)
- Treatment-emergent SAEs reported in ≥ 2% of all treated subjects by PT (all causalities)
- TEAEs reported in ≥ 5% of subjects and occurred at ≥2% higher incidence in subjects on Arm 1 (all grades, all causalities)
- Grade 3 or higher TEAEs reported in ≥ 5% of subjects on either arm (all causalities)

A by-subject listing of the following AE categories will be presented.

- All AEs (TEAEs will be flagged)
- All SAEs (TEAEs will be flagged)
- All SAE reported with the term disease progression
- TEAEs resulting in dose hold
- TEAEs resulting in dose reduction
- TEAEs resulting in discontinuation of study drug
- TEAEs resulting in death (see Section 4.11.3 for full characterization of deaths)
- TEAEs resolved or resolved with sequelae
- TEAEs not resolved

4.11.2.1 Other Adverse Events of Interest

Selected MedDRA PTs will be combined to create grouped PTs or categories to further characterize risks associated with duvelisib. The following risks will be summarized as Adverse Event of Special Interest:

- Diarrhea/colitis
- Infections
  - All infections, excluding Pneumonia and COVID-19 infections
  - Pneumonia
  - COVID-19 infections
- Cutaneous reactions
- Non-infectious pneumonitis
- Neutropenia
- Hepatotoxicity (as defined by transaminase elevations)

The specific PTs that will be combined for each risk are defined in Appendix 1.
For each risk, PTs will be summarized at the grouped PT and individual PT (and grouped SOC if applicable), at the subject level:

- TEAEs (all grades, all causalities)
- Treatment-related TEAEs (all grades)
- Grade 1 or 2 only TEAEs (all causalities)
- Grade 3 or higher TEAEs (all causalities)
- Treatment-emergent SAEs (all causalities)
- TEAEs resulting in dose holds
- TEAEs resulting in dose reductions
- TEAEs resulting in dose holds or reductions
- TEAEs resulting in discontinuation of study drug (all causalities)
- Time to initial onset (defined as AE start date – first dose date + 1)

In addition, the following summaries will be created for each grouped PT within each grouped SOC, at the event level:

- Summary of dose modifications
  - Number of events resulting in dose holds
  - Number of events resulting in dose reductions
  - Number of events resulting in dose holds or reductions
  - Number of events resulting in discontinuation of study drug
- Summary of outcome
  - Number of events resolved
  - Number of events resolved with sequelae
  - Number of events resolved or resolved with sequelae
  - Number of events not resolved
- Duration of AE

- Concomitant medications used to treat AEs (a specific list of concomitant medications associated with each risk will be provided)

For each subject and each AE, the worst severity recorded will be attributed and used in the by-severity summaries. Similarly, the worst causality (most related to treatment) will be attributed and used in the by-causality summaries. If causality is missing, the worst case will be assumed (related to treatment).

A by-subject listing of all AEs (including non-TEAEs) will be provided. This listing will be presented by treatment group and will include: center, subject identifier, age, sex, race, AE (SOC, PT, and verbatim term), date of onset, date of resolution, duration, severity, seriousness, action taken, outcome and causality.

4.11.3 Deaths, Serious Adverse Events, and Other Significant Adverse Events

A complete tabulation of all deaths occurring on study will be provided, which will include:

- All reported fatal AEs
  - Treatment-emergent fatal AEs by SOC and PT
    - Deaths within 30 days of duvelisib last dose by SOC and PT
    - Deaths > 30 days from duvelisib last dose by SOC and PT
  - Treatment-related treatment-emergent fatal AEs by SOC and PT
    - Deaths within 30 days of duvelisib last dose by SOC and PT
    - Deaths > 30 days from duvelisib last dose by SOC and PT
  - Fatal AEs reported that start > 30 days post duvelisib last dose by SOC and PT

The following listings will also be presented:

- A by-subject listing of all deaths that occurred during the study will be presented and will distinguish between deaths on treatment (through 30 days after last dose of duvelisib), fatal AEs outside the 30 day window, and deaths during survival follow-up
- A by-subject listing of all SAEs
- A by-subject listing of all AEs leading to discontinuation of study treatment
- A by-subject listing of all other significant AEs

Listings will follow the format described for AEs in Section 4.11.2, if appropriate.

4.11.4 Clinical Laboratory Evaluation

Laboratory tests will be reported separately for hematology and blood chemistry.
For the purposes of presentation in both tables and listings, the following laboratory test results will be converted to the International System of Units (SI) before presentation: sodium, potassium, chloride, bicarbonate (or CO_2), albumin, total protein, creatinine, uric acid, calcium, phosphorus, magnesium, glucose, total and direct bilirubin, alkaline phosphatase, red blood cell (RBC) count, hemoglobin, hematocrit, platelets, white blood cell count with 5-part differential

performed manually or by flow cytometry (lymphocytes, neutrophils, monocytes, basophils, and eosinophils), amylase, lipase, alanine aminotransferase (ALT), and aspartate aminotransferase (AST). Laboratory tests will be graded according to CTCAE Version 5.0 or higher.
If a laboratory test value is reported using a non-numeric qualifier (e.g., less than [<] a certain value or greater than [>] a certain value), the given numeric value will be used in the summary statistics, ignoring the non-numeric qualifier.
For laboratory tests graded by CTCAE, the following summary tables will be created by treatment arm:

- Summary of worst post-baseline value on treatment, regardless of baseline value according to the following categories:
  - Any grade
  - Grades 1 - 2
  - Grades 3 - 4
  - Grade 3
  - Grade 4
- Summary of worst post-baseline value on treatment that represents a worsening in grade from baseline according to the following categories:
  - Any grade
  - Grades 1 - 2
  - Grades 3 - 4
  - Grade 3
  - Grade 4

Laboratory tests with bi-directional grades (e.g., hyperglycemia and hypoglycemia) will be presented separately for each direction within each summary table.
For laboratory tests not graded by CTCAE, the following summary tables will be created by treatment arm:

- Summary of worst post-baseline value on treatment, regardless of baseline value according to the following categories:
  - Low (less than lower limit of normal)
  - Normal (greater than or equal to lower limit of normal and less than or equal to upper limit of normal)
  - High (greater than upper limit of normal)
- Summary of worst post-baseline value on treatment that represents a worsening from baseline according to the following categories:
  - Low (less than lower limit of normal)
  - Normal (greater than or equal to lower limit of normal and less than or equal to upper limit of normal)
  - High (greater than upper limit of normal)

A subject will be defined as having a treatment-emergent laboratory abnormality if any of the following conditions are satisfied for a specific laboratory parameter:

- Laboratory result within the normal range at baseline and either a result below the lower limit of the normal range or above the upper limit of the normal range at any post-baseline time point on or before 30 days after the last dose of duvelisib.
- Laboratory result below the lower limit of the normal range at baseline and a laboratory result above the upper limit of the normal range at any post-baseline time point on or before 30 days after the last dose of duvelisib.
- Laboratory result above the upper limit of the normal range at baseline and a laboratory result below the lower limit of the normal range at any post-baseline time point on or before 30 days after the last dose of duvelisib.

A by-subject listing of all laboratory data will be provided by treatment group, with abnormal values flagged, and include center, subject identifier, age, sex, race, weight and visit. Laboratory reference ranges will also be listed.

Laboratory values (hematology, biochemistry and urinalysis) will be listed by subject and study time point including changes from baseline (with the exception of urinalysis). Baseline will be defined as described in Section 4.2.

All values outside the clinical reference ranges will be flagged in the data listings. The abnormal values will be flagged with 'L' for values below the lower limit of the clinical reference range and 'H' for values above the upper limit of the clinical reference range and included in the listings. The Investigator will assess whether the values outside the clinical reference range are clinically significant and these will be reported as abnormal not clinically significant (NCS) or abnormal clinically significant (CS). Clinically significant laboratory values will be recorded by the Investigator as AEs.

Descriptive statistics (for non-categorical data including hematology and biochemistry) will be presented by cohort for both individual values (N, mean, SD, median, minimum, and maximum) and changes from baseline.

Boxplots over time (at each scheduled visit) will be provided for each laboratory parameter.

4.11.5 Vital Signs, Physical Findings and Other Observations Related to Safety.

4.11.5.1 Vital Signs

The actual values of vital sign parameters, including temperature, heart rate, weight and systolic and diastolic BP, will be presented in a by-subject listing. Vital sign parameter reference ranges will also be listed.
Blood pressure will be summarized as follows by treatment arm:

- Descriptive statistics, for systolic and diastolic BP separately
  - Baseline value
  - Maximum on-treatment value
  - Change from baseline to maximum on-treatment value

- Categorical summary of baseline value, maximum on-treatment value for systolic and diastolic BP separately, with categories defined as follows:
    - Diastolic BP
        - < 80 mm Hg
        - ≥ 80 – 89 mm Hg
        - ≥ 90 mm Hg
    - Systolic BP
        - < 120 mm Hg
        - ≥ 120 – 129 mm Hg
        - ≥ 130 – 139 mm Hg
        - ≥ 140 mm Hg
- Categorical summary of baseline value, maximum on-treatment value for systolic and diastolic BP combined using the categories define above (i.e., diastolic <80 mm Hg and systolic <120 mm Hg, diastolic <80 mm Hg and systolic 120 – 129 mm Hg)

Heart rate will be summarized as follows by treatment arm:

- Descriptive statistics for baseline value, minimum on-treatment value, maximum on treatment value, change from baseline to maximum on-treatment value
- Categorical summary of baseline value, minimum on-treatment value, maximum on-treatment value, with categories defined as follows:
    - < 60 bpm
    - ≥ 60 – 100 bpm
    - > 100 bpm

Boxplots over time (at each scheduled visit) will be provided for each BP and heart rate measured.

4.11.5.2 ECG, ECHO/MUGA, and Physical Examinations

Per protocol, ECGs, ECHO/MUGA results, and full physical examinations findings are to be collected at baseline only. By-subject listing will be provided for each of these parameters.

4.11.6 Safety Monitoring

No data monitoring committee will be set up for monitoring the safety of study treatment.

4.12 Other Analyses

4.12.1 Pharmacokinetics

The PK data collected will be analyzed by standard population PK methods, using appropriate software. The intent of this analysis is to characterize the parameters of PK disposition. Additional exploration of the relationship between duvelisib exposure and safety or efficacy assessments may be completed. If there is only a limited amount of plasma concentration data

from this study, the data may be pooled with the results of other studies to perform the population PK analysis. A separate PK analysis plan may be developed.

4.12.2 Quality of Life (QoL)

4.12.2.1 ECOG Performance Status

Summaries of ECOG performance status will be presented by visit. Shift tables for change in ECOG performance score from baseline by worst post-baseline score will be produced.
A by-subject listing of ECOG performance status data will be provided.

4.12.2.2 EQ-5D-3L Scores

The EQ-5D-3L contains a descriptive system with one response for each of the 5 dimensions (mobility, self-care, usual activities, pain/discomfort, and anxiety/depression). The first response is coded as 1 (indicating no problems), the second response is coded as a 2 (indicating some problems), and the third response is coded as a 3 (indicating extreme problems). Ambiguous responses (e.g., more than one response in a dimension, if applicable) are treated as missing values. The EQ-5D also contains a Visual Analog Scale (VAS) for health state ranging from 0 to 100, where 0 represents worst imaginable health state and 100 represents the best imaginable health state. Ambiguous values (the line crosses the VAS twice) should be treated as missing values.
For each dimension of the descriptive system, the number and percentage of subjects with no problems, some problems, and extreme problems will be reported at each visit.
Summary statistics for the VAS of health state and the changes from baseline will be reported for each visit. By-subject listing of these data will be provided.

Also refer to the EQ-5D-3L user guide [4] for additional considerations.

4.12.3 Other Exploratory Analyses

The other exploratory endpoints, including biomarker endpoints, will be summarized descriptively. A separate analysis plan may be developed for formal analyses.

4.13 Determination of Sample Size

The sample size for each arm was calculated using a 2-stage approach.

For each arm: in Stage 1, 15 subjects will be randomized. The evaluation of the first stage will take place after the 15 enrolled subjects have been followed for a minimum of 3 cycles (18 weeks).

If there are fewer than 6 responders (PR or CR) among the first 15 subjects, consideration may be given to terminate the treatment arm. The decision to proceed to Stage 2 will be based on the totality of data observed, including types of responses, safety data, and discontinuation rates. The treatment arm may also not proceed to Stage 2 if the totality of the data is viewed as compelling;

such scenarios could include (but are not limited to) accumulating enough data in Stage 1 to conclude the ORR exceeds 30%. If the treatment arm proceeds to Stage 2, up to 36 additional subjects will be enrolled for a total of 51. If fewer than 22 (43%) responders (PR or CR) are observed in 51 subjects, consideration may be given to conclude that the treatment schedule does not warrant further study in this population. The final evaluation of each schedule will be based on the totality of data observed, including types of responses, safety data, and discontinuation rates. This design yields a type I error rate of 0.025 (1-sided) and power of 90% under the assumptions described above.

Subjects will be stratified by number of prior therapies (1 or > 1), bulky disease status (longest diameter of baseline lesion < 5 cm or ≥ 5 cm) and time since last recurrence (≥ 24 months or < 24 months).

Note: "Enrolled" means subjects', or their legally acceptable representative's, agreement to participate in a clinical study following completion of the informed consent process. Potential subjects who are pre-screened for the purpose of determining eligibility for the study but do not sign an informed consent form (ICF) are not considered enrolled.

4.14 Changes in the Conduct of the Study or Planned Analysis

Additional analyses not specified in this SAP may be conducted, if necessary. In such cases, a separate analysis plan will be created to document such analyses. The SAP will be updated as necessary if the protocol is amended.

5 REFERENCES

[1] Simon, R. Optimal two-stage designs for phase II clinical trials. Controlled Clinical Trials 1989;10(1);1-10.[1] Cheson BD, Pfistner B, Juweid ME, et al. Revised response criteria for malignant lymphoma. *Journal of clinical oncology : official journal of the American Society of Clinical Oncology.* 2007;25(5):579-586.

[2] Cheson BD, Fisher RI, Barrington SF, et al. Recommendations for initial evaluation, staging, and response assessment of Hodgkin and non-Hodgkin lymphoma: the Lugano classification. *Journal of clinical oncology : official journal of the American Society of Clinical Oncology.* 2014;32(27):3059-3068.

[3] SAS® Version 9.3 of the SAS System for Personal Computers. Copyright © 2011. SAS Institute Inc. SAS and all other SAS Institute Inc. product or service names are registered trademarks or trademarks of SAS Institute Inc., Cary, NC, USA.

[4] EQ-5D-3L User Guide (Version 6. 2018) https://euroqol.org/wp-content/uploads/2018/12/EQ-5D-3L-User-Guide_version-6.0.pdf. Accessed June 3, 2019.

6 APPENDIX

Appendix 1: List of Preferred Terms for Adverse Events of Special Interest

Grouped Term	Preferred Terms
All infections (excluding Pneumonia and COVID-19 infections)	All preferred terms under SOC = 'Infections Infections and infestations' excluding any preferred term containing the word 'pneumonia' and excluding any preferred terms included in the COVID-19 infections Grouped Term.
Diarrhea	Diarrhoea, Diarrhoea hemorrhagic, Defecation urgency
Colitis	Colitis, Colitis erosive, Enterocolitis, Enterocolitis hemorrhagic, Colitis microscopic, Necrotising colitis, Colitis ulcerative
Pneumonia	Pneumonia, any PT containing the word/string "pneumonia" except for "Pneumonia aspiration", Atypical pnuemonia, Atypical mycobacterial pnuemonia, Bronchopneumonia, Bronchopulmonary aspergillosis
Pneumonitis	Pneumonitis, Interstitial lung disease, lung infiltration, acute interstitial pneumonitis
Transaminase elevation	Alanine aminotransferase increased, Aspartate aminotransferase increased, Transaminases increased, Hypertransaminasaemia, Hepatic enzyme increased, Acute hepatic failure, Drug-induced liver injury, Hepatic failure, Hepatocellular injury, Hepatotoxicity
Cutaneous reactions	Dermatitis, Dermatitis allergic, Dermatitis bullous, Dermatitis exfoliative, Dermatitis psoriasiform, Drug eruption, Drug reaction with eosinophilia and systemic symptoms, Erythema, Erythema multiforme, Generalised erythema, Exfoliative rash, Rash, Rash generalised, Rash erythematous, Rash follicular, Rash macular, Rash maculo-papular, Rash maculovesicular, Rash papular, Rash papulosquamous, Rash pruritic, Rash pustular, Rash vesicular, Toxic skin eruption, Palmar erythema, Palmoplantar keratoderma, Palmar-plantar erythrodysesthesia syndrome, Perivascular dermatitis, Skin reaction, skin toxicity, Stevens-Johnson syndrome, Toxic epidermal necrolysis
Neutropenia	Neutropenia, Neutrophil count decreased
COVID-19 infections	Asymptomatic COVID-19, Congenital COVID-19, Coronavirus infection, Coronavirus test positive, , COVID-19 pneumonia, COVID-19 treatment, Multisystem inflammatory syndrome in children, Post-acute COVID-19 syndrome, SARS-CoV-2 RNA decreased, SARS-CoV-2 RNA fluctuation, SARS-CoV-2 RNA increased, SARS-CoV-2 sepsis, SARS-CoV-2 test false negative, SARS-CoV-2 test positive, SARS-CoV-2 viraemia, Suspected COVID-19, Vaccine derived SARS-CoV-2 infection

PAREXEL International Electronic Signature Page

This page is the manifestation of the electronic signature(s) used in compliance with PAREXEL International's electronic signature policies and procedures and in compliance with applicable regulations.

UserName: [redacted]
Title: [redacted]
Date: Thursday, 31 August 2023, 02:46 PM GMT Standard Time
Meaning: Document contents approved.