Statistical Analysis Plan for Study M13-367 (Portion 2)

A Phase 1/2 Study Evaluating the Safety and Pharmacokinetics of ABT-199 in Subjects with Relapsed or Refractory Multiple Myeloma

Date: 11 January 2022

Version 1.1

Table of Contents

1.0 Introduction 5
2.0 Study Design and Objectives 5
2.1 Objectives and Hypotheses 5
2.2 Study Design Overview 6
2.3 Treatment Assignment and Blinding 7
2.4 Sample Size Determination 7
3.0 Endpoints 7
3.1 Primary Endpoint(s) 8
3.2 Secondary Endpoint(s) 8
3.2.1 Secondary Efficacy Endpoints 8
3.2.2 Secondary Patient Reported Outcomes 12
3.3 Other Efficacy Endpoint(s) 13
3.4 Safety Endpoint(s) 15
4.0 Analysis Populations 15
5.0 Subject Disposition 16
6.0 Study Treatment Duration and Compliance 16
6.1 Exposure to Study Treatment 16
6.2 Compliance 19
7.0 Demographics, Baseline Characteristics, Medical History, and Prior/Concomitant Medications 19
7.1 Demographics and Baseline Characteristics 19
7.2 Medical History 20
7.3 Prior and Concomitant Medications 20
8.0 Efficacy Analyses 21
8.1 General Considerations 21
8.2 Primary Efficacy Endpoint(s) and Analyses 22
8.2.1 Primary Efficacy Endpoint(s) 22
8.2.2 Primary Efficacy Analysis 22
8.3 Secondary Efficacy Analyses 22
8.4 Additional Efficacy Analyses 22
8.5 Efficacy Subgroup Analyses 23

9.0 Safety Analyses 23
9.1 General Considerations 23
9.2 Adverse Events 23
9.2.1 Treatment-Emergent Adverse Events 23
9.2.2 Adverse Event Overview 24
9.2.3 Treatment-Emergent Adverse Events by SOC and/or PT 25
9.2.4 SAEs (Including Deaths) and Adverse Events Leading to Study Treatment Discontinuation, Dose Reduction, or Dose Interruption 26
9.2.5 Deaths 26
9.2.6 Listing of Adverse Events 26
9.2.7 Adverse Events of Special Interest 27
9.3 Analysis of Laboratory Data 27
9.4 Analysis of Vital Signs 29
10.0 Interim Analyses 30
10.1 Data Monitoring Committee 30
11.0 Version History 30
12.0 References 31

List of Tables

Table 1. Event and Censoring Rules for PFS 9
Table 2. Event and Censoring Rules for DOR 10
Table 3. Event and Censoring Rules for TTR 10
Table 4. Event and Censoring Rules for TTP 11
Table 5. Laboratory Criteria for TLS 29
Table 6. SAP Version History Summary 30

List of Figures

Figure 1. Study Schematic 6

List of Appendices

Appendix A. Protocol Deviations 33
Appendix B. Definition of Adverse Events of Special Interest Selected Adverse Events 34
Appendix C. Potentially Clinically Important Criteria for Safety Endpoints 35
Appendix D. PRO assessments 36

1.0 Introduction

This Statistical Analysis Plan (SAP) describes the statistical analyses for Venetoclax (ABT-199) Phase 2 Portion of Study M13-367 A Phase 1/2 Study Evaluating the Safety and Pharmacokinetics of ABT-199 in Subjects with Relapsed or Refractory Multiple Myeloma. A separate SAP is available in eDocs for Phase 1 Portion of the Study M13-367.

Study M13-367 examines the efficacy and safety of venetoclax in subjects with Relapsed or Refractory Multiple Myeloma. The SAP will not be updated in case of administrative changes or amendments to the protocol unless the changes impact the analysis.

Unless noted otherwise, all analyses will be performed using SAS Version 9.4 (SAS Institute Inc., Cary, NC 27513) or later under the UNIX operating system.

Changes from the SAP version 1.0 are summarized in Section 11.0.

2.0 Study Design and Objectives

2.1 Objectives and Hypotheses

Primary Objective:

The primary objective of the Phase 2 Portion is to further evaluate the overall response (ORR) and very good partial response or better rate (VGPR+) in subjects with t(11;14)-positive multiple myeloma.

Secondary Objectives:

The secondary objectives are to monitor safety, progression free survival (PFS), duration of response (DOR), time to response (TTR), time to disease progression (TTP), and overall survival (OS) and to evaluate Patient Reported Outcomes (PRO) including Worst Pain (Brief Pain Inventory – Short Form [BPI-SF]), Physical Functioning and Global Health Status/Quality of Life (GHS/QoL) (European Organization for Research and

Treatment of Cancer Quality of Life Questionnaire Core [EORTC QLQ-C30]), and Fatigue (Patient Reported Outcomes Measurement Information System [PROMIS] Cancer Fatigue Short Form [SF]).

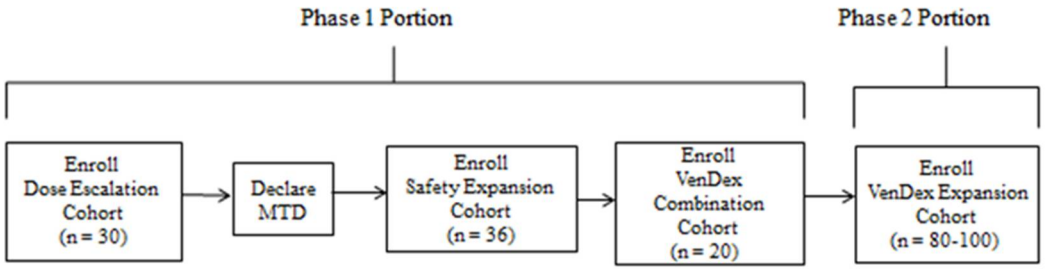
Tertiary objectives:

The Tertiary objectives are to assess other PRO endpoints (remaining subscales/items from BPI-SF, EORTC QLQ-C30, European Organization for Research and Treatment of Cancer Quality of Life Questionnaire Multiple Myeloma Module [EORTC QLQ-MY20], and EuroQol-5 Dimensions-5 Level (EQ-5D-5L).

2.2 Study Design Overview

The Phase 2 Portion of the study is an expansion of the combination cohort which further evaluate ORR and VGPR+ of venetoclax and dexamethasone in approximately 80 – 100 relapsed or refractory subjects with t(11;14)-positive MM. Disease assessment for each post-baseline IMWG assessment will be performed by the investigator. Subjects in Phase 2 Portion will receive venetoclax daily (Day 1 – 21) with dexamethasone orally Day 1, 8, and 15 per the prescribing information. Subjects will receive study treatment for up to 2 years following the first dose of the last subject enrolled on study provided they continue to tolerate venetoclax, have no evidence of disease progression, and do not meet any criteria for subject discontinuation. The schematic of the Study M13-367 is shown in Figure 1.

Figure 1. Study Schematic



2.3 Treatment Assignment and Blinding

The Phase 2 Portion of the Study M13-367 is not randomized. Eligible subjects will be enrolled into a single arm.

2.4 Sample Size Determination

The primary objective of the Phase 2 portion of the study is to test

a. H_{0a}: ORR $\leq$ 30% against H_{1a}: ORR > 30% and
b. H_{0b}: VGPR+ rate $\leq$ 12% against H_{1b}: VGPR+ rate > 12%.

These hypotheses will be tested using a family-wise one-sided type-I error rate of 0.025. The Hochberg procedure will be used to control for multiplicity. The Phase 2 Portion of the study is powered for a true ORR of 50% and a true VGPR+ rate of 30%. Under the considerations above, if at least 80 subjects are enrolled in the Phase 2 portion of the study, there will be over 95% power to reject H_{0a}, 96% power to reject H_{0b}, 92% power to reject both H_{0a} and H_{0b}, and 99% power to declare efficacy (due to high ORR, high VGPR+ rate, or high values of both).

3.0 Endpoints

Best overall response (BOR) will be derived for each patient based on timepoint-wise disease assessment performed by the investigator and ordered as: stringent complete response (sCR) > complete response (CR) > very good partial response (VGPR) > partial response (PR) > minimal response (MR) > stable disease (SD) > progressive disease (PD). BOR for a patient will be sCR if there are two consecutive disease assessments with sCR assessed, otherwise will be CR if there are two consecutive disease assessments CR (or better) assessed, otherwise will be VGPR if there are two consecutive disease assessments with VGPR (or better) assessed, otherwise will be PR if there are two consecutive disease assessments with PR (or better) assessed, otherwise will be MR if there are two consecutive disease assessments with MR (or better) assessed, otherwise will be PD if the best timepoint-wise disease assessment is PD (confirmation is needed if the initial PD is

based on lab – serum M-protein, urine M-protein, or serum FLC), otherwise will be NE if the best timepoint-wise disease assessment is unconfirmed lab-based PD, otherwise will be SD.

3.1 Primary Endpoint(s)

Overall Response Rate (ORR): defined as the proportion of subjects with BOR of sCR, CR, VGPR or PR based on IMWG criteria.

Very Good Partial Response or Better Rate (≥ VGPR Rate): defined as the proportion of subjects with BOR of sCR, CR, or VGPR based on IMWG criteria.

3.2 Secondary Endpoint(s)

MM-related death is based on investigator opinion and is collected in Death eCRF page.

3.2.1 Secondary Efficacy Endpoints

Progression-free Survival (PFS)

The PFS is defined as the number of days from the date of the first dose of study treatment to the date of first documented PD or death due to any cause, whichever occurs first.

PFS (days) = Date of first documented PD or death due to any cause –

Date of the first dose of study treatment + 1

Details on event and censoring rules for PFS are provided in Table 1.

Table 1. Event and Censoring Rules for PFS

Situation	Date of Censor or Event	Outcome
PD/death due to any cause occurred not more than 12 weeks after the last adequate disease assessment	Date of PD/death	Event
PD/death due to any cause occurred more than 12 weeks after the last adequate disease assessment	Date of the last adequate disease assessment prior to PD/death due to any cause	Censored
No post-baseline assessment	Date of the first dose of study treatment	Censored
No PD/ no death due to any cause	Date of the last adequate disease assessment	Censored
Start of new anti-MM therapy prior to PD/death	Date of the last adequate disease assessment prior to start of new anti-MM therapy	Censored

Duration of Response (DOR)

DOR is defined as the number of days from the date of the first confirmed response of PR or better until date of first documented PD or death due to MM, whichever occurs first, i.e.,:

DOR (day) = Date of first documented PD or MM-related death – Date of the first confirmed response (PR or better) + 1

Only subjects who achieved a BOR of PR or better will be included in this analysis. Details on event and censoring rules for DOR are provided in Table 2.

Table 2. Event and Censoring Rules for DOR

Situation	Date of Censor or Event	Outcome
PD/death due to MM occurred not more than 12 weeks after the last adequate disease assessment	Date of PD/death	Event
PD/death due to MM occurred more than 12 weeks after the last adequate disease assessment	Date of the last adequate disease assessment prior to PD/death due to MM	Censored
No PD/no death due to MM	Date of the last adequate assessment	Censored
Start of new anti-MM therapy prior to PD/death	Date of the last adequate disease assessment prior to start of new anti-MM therapy	Censored

Time to Response (TTR)

TTR is defined as the number of days from the date of the first dose of study treatment until the date of the first confirmed response of PR or better, i.e.:

TTR (days) = Date of the first confirmed response – Date of the first dose of study treatment + 1

Details on event and censoring rules for TTR are provided in Table 3.

Table 3. Event and Censoring Rules for TTR

Situation	Date of Censor or Event	Outcome
Confirmed response of PR or better	Date of first confirmed response of PR or better	Event
No post-baseline assessment	Date of the first dose of study treatment	Censored
No confirmed response of PR or better	Date of the last adequate disease assessment	Censored
Start of new anti-MM therapy prior to first confirmed response of PR or better	Date of the last adequate disease assessment prior to start of new anti-MM therapy	Censored

Time to Progression (TTP)

TTP is defined as the number of days from the date of the first dose of study treatment until the date of first documented PD or death due to MM, whichever occurs first, i.e.,:

TTP (days) = Date of first documented PD or MM-related death – Date of the first dose of study treatment + 1

Details on event and censoring rules for TTP are provided in Table 4.

Table 4. Event and Censoring Rules for TTP

Situation	Date of Censor or Event	Outcome
PD/death due to MM occurred not more than 12 weeks after the last adequate disease assessment	Date of PD/death	Event
No post-baseline disease assessment	Date of the first dose of study treatment	Censored
PD/death due to MM occurred more than 12 weeks after last adequate disease assessment	Date of the last adequate disease assessment prior to PD/death due to MM	Censored
No PD/no death due to MM	Date of the last adequate disease assessment	Censored
Start of new anti-MM therapy prior to PD/death	Date of the last adequate disease assessment prior to start of new anti-MM therapy	Censored

Overall Survival (OS):

For a given subject, OS is defined as the number of days from the date of the first dose of study treatment to the date of death due to any cause.

OS (days) = Date of death due to any cause – Date of the first dose of study treatment + 1

If a subject is not known to be died, OS will be censored at the date of last known alive date.

All events of death will be included, regardless the event occurred while the subject was still taking study treatment or after the subject discontinued study treatment.

3.2.2 Secondary Patient Reported Outcomes

Mean changes from baseline to each appliable post-baseline visit will be summarized for the secondary Patient Reported Outcomes as follows:

Worst Pain (Brief Pain Inventory – Short Form [BPI-SF])):

For Worst Pain, scores will be computed according to procedures outlined in BPI-SF scoring manual, available at https://www.mdanderson.org/documents/Departments-and-Divisions/Symptom-Research/BPI_UserGuide.pdf. The patient-reported outcome instruments are described in detail in Appendix D.

Physical Functioning (EORTC QLQ-C30):

For Physical Functioning scale, scores will be computed according to procedures outlined in the EORTC QLQ-C30 scoring manual, available at http://groups.eortc.be/qol/manuals. The patient-reported outcome instruments are described in detail in Appendix D. Time to deterioration in physical functioning as measured by the physical functioning domain of the EORTC QLQ-C30. Median time to deterioration will be defined as ≥ 10-point deterioration in physical functioning score from baseline.

GHS/QoL (EORTC QLQ-C30):

The GHS/QoL scale specific scores will be calculated from the 22 items in EORTC QLQ-C30 for each subject based on the QLQ-C30 instruction. Higher scores are indicative of better HRQOL. The QLQ-C30 is described in detail in Appendix D. Time to deterioration in GHS/QoL score as measured by the GHS/QoL domain of the EORTC QLQ-C30. Median time to deterioration will be defined as ≥ 10-point deterioration in GHS/QoL score from baseline.

Fatigue (PROMIS Cancer Fatigue SF):

For PROMIS Cancer Fatigue SF 7a, scores will be computed according to the procedures outlined in the PROMIS Fatigue scoring manual, available at https://www.assessmentcenter.net/Manuals.aspx. Higher scores are indicative of worse levels of fatigue. The Cancer Fatigue SF is described in detail in Appendix D.

3.3 Other Efficacy Endpoint(s)

Mean changes from baseline to each appliable post-baseline visit will be summarized for the Patient Reported Outcomes as follows

EORTC QLQ-MY20:

For EORTC QLQ-MY20, the endpoints are scores change from baseline. Scores will be summarized for four scales:

- Future perspective
- Body image
- Disease symptoms– overall domain and individual domain items listed below
  - Bone aches or pain
  - Back pain
  - Hip pain
  - Arm or shoulder pain
  - Chest pain
  - Pain increasing with activity
- Side effects of treatment

Scores for each scale will be computed according to procedures outlined in EORTC QLQ-MY20 scoring manual, available at http://www.eortc.be/qol/files/ScoringInstructions/MY20_summary.pdf and Appendix D.

EQ-5D-5L:

For EQ-5D-5L, the endpoints are scores change from baseline. Scores will be summarized for six scales:

- Mobility
- Self-care
- Usual activities
- Pain/Discomfort
- Anxiety/Depression
- EQ VAS score

Scores for each scale will be computed according to procedures outlined in EQ-5D-5L scoring manual, available at https://euroqol.org/wp-content/uploads/2016/09/EQ-5D-5L_UserGuide_2015.pdf.

Remaining Subscales of BPI-SF and EORTC QLQ-C30

Remaining subscales of BPI-SF are Pain Severity and Pain Interference. Scores for each subscale will be computed according to BPI-SF scoring manual, respectively. The endpoints are score change from baseline.

Remaining subscales of EORTC QLQ-C30 are

- Role Function
- Emotional Functioning
- Cognitive Functioning
- Social Functioning
- Fatigue
- Nausea and Vomiting
- Pain
- Dyspnea

- Insomnia
- Appetite Loss
- Constipation
- Diarrhea
- Financial Difficulties

Scores for each subscale will be computed according to EORTC QLQ-C30 scoring manual, respectively.

3.4 Safety Endpoint(s)

The following are safety endpoints.

- Adverse Events (AEs)
- Serious Adverse Events (SAEs)
- Deaths
- Electrocardiography (ECG)
- Vital signs
- Laboratory results

4.0 Analysis Populations

The following population sets will be used for the analyses.

The Intent-to-Treat (ITT) population includes all enrolled subjects who received at least 1 dose of study treatment. The ITT population will be used for all efficacy and baseline analyses.

The Safety Analysis Set consists of all subjects who received at least 1 dose of study treatment. The Safety Analysis Set will be used for all safety analyses.

5.0 Subject Disposition

The total number of subjects who were screened, enrolled, and treated will be summarized. Reasons for exclusion, including screen failure, will be summarized.

A summary of subject accountability will be provided where the number of subjects in each of the following categories will be summarized:

- Treated subjects
- Discontinued study treatment
- Discontinued venetoclax
- Discontinued dexamethasone
- Discontinued from study
- Ongoing (Treatment, Follow Up)

Number and percentage of subjects who discontinued study treatment (primary and all reasons separately for venetoclax and dexamethasone, respectively) and study will be summarized.

6.0 Study Treatment Duration and Compliance

Analyses of study treatment exposure and compliance listed below will be performed for study treatment components (venetoclax and dexamethasone) separately on the safety analysis set.

6.1 Exposure to Study Treatment

Duration of study treatment exposure

Duration of study treatment exposure will be calculated as the number of days between the start date and end date of study treatment.

Duration of study treatment exposure (days) = date of last dose of study treatment – date of first dose of study treatment + 1

For Risk Management Plan: Patient-years of exposure

Patient-years of exposure is computed as:

Patient-years of exposure (pt-yrs) = Sum of duration of study treatment for all subjects (day)/365.25

Duration of study treatment component exposure

Duration of study treatment component exposure will be calculated as the number of days between the start date and end dates for each study treatment component

Duration of study treatment component exposure (days) =
Date of last dose study treatment component – Date of the first dose of study treatment component + 1

Cumulative dose

The cumulative dose is defined as the total dose given during the study treatment exposure and will be summarized by study treatment component and treatment group. For patients who did not take any drug, the cumulative dose is 0.

Dose intensity (DI) and relative dose intensity (RDI)

The following DI or planned DI (PDI) (per protocol) calculations apply for both IV and oral dosing. In the case of IV dosing, the dose was given based on BSA calculation of DI or PDI involves BSA in the division. The RDI will be summarized separately for each study treatment component.

venetoclax and dexamethasone

$$\text{DI(mg/day)} = \frac{\text{Cumulative dose [mg]}}{\sum_{c=1}^{k} \text{Actual duation of cycle c [day]}}$$

where k = total number of cycles.

PDI is defined as follows:

1. The last cycle length is not shorter than planned cycle length

$$\text{PDI(mg/day)} = \frac{\text{Cumulative planned dose [mg]}}{\sum_{c=1}^{k} \text{Planned duration of cycle c [day]}}$$

2. The last cycles length is shorter than planned cycle length

$$\text{PDI(mg/day)} = \frac{\begin{array}{c}\text{Total planned dose for first k} - 1 \text{ cycles [mg]} \\ + \text{total planned dose for actual duration of the last cycle k}\end{array}}{\sum_{c=1}^{k-1} \text{Planned duration of cycle c} + \text{Actual duration of the last cycle [day]}}$$

k is the actual total number of cycles.

RDI is defined as

$$\text{RDI}[\%] = \frac{\text{DI [mg/day]}}{\text{PDI [mg/day]}} \times 100$$

The RDI will be summarized based on data collected on the drug administration eCRF pages. Descriptive statistics, including the mean, standard deviation, median, minimum, and maximum will be presented for duration of exposure and RDI. Number and percentage of subjects who have dose reductions or interruptions, as well as their reasons, will be provided. Number of cycles of study treatment received will be provided. In addition, summary statistics of cumulative dose and frequencies for RDI (%) will be provided by treatment group for the following categories:

- 0 to < 50%
- 50% to < 70%
- 70% to < 90%
- 90% to < 110%
- ≥ 110%

6.2 Compliance

To assess compliance, a listing of protocol deviations will be given.

7.0 Demographics, Baseline Characteristics, Medical History, and Prior/Concomitant Medications

Demographics, baseline or disease characteristics, medical history, and prior and concomitant medications will be summarized for the ITT population. Categorical variables will be summarized with the number and percentage of subjects; percentages will be calculated based on the number of non-missing observations. Continuous variables will be summarized with descriptive statistics (number of non-missing observations, mean and standard deviation, median, minimum and maximum).

7.1 Demographics and Baseline Characteristics

Continuous demographic and baseline variables include age, body weight, body height, time since diagnosis and number of lines of prior therapy.

Categorical demographic and baseline variables include

- Sex (Female, Male);
- Race (White, Black, Asian, American Indian/Alaska Native, Native Hawaiian or Other Pacific Islander, Other, Multi Race);
- Ethnicity (Hispanic or Latino, No Ethnicity);
- Age (< 40, $40 - < 60$, ≥ 60);
- Type of myeloma (IgG, non-IgG, free light chain);
- ISS status (I, II, III);
- t(11;14) status (negative, positive);
- 17P status (deleted, not deleted, undetermined);
- 13Q status (deleted, not deleted, undetermined);
- Type of myeloma (IgG, non-IgG, free light chain);
- Prior exposure to proteasome inhibitors (PI) (naive, sensitive, or refractory);

- Prior exposure to an immunomodulatory drug (IMiD) (naive, sensitive, or refractory);
- Prior exposure to an anti-CD38 monoclonal antibody (naive, sensitive, or refractory);
- Prior allogenic stem cell transplant (yes, no);
- Prior autologous stem cell transplant (yes, no);
- Prior bortezomib (yes, no);
- Bortezomib refractory status (yes, no);
- Prior lenalidomide (yes, no);
- Lenalidomide refractory status (yes, no);
- Lines of prior therapy (< 1, 1 – 3, 4 – 6, 7 – 9, > 9);
- ECOG Performance Status (grade 0, grade 1, grade 2, grade 3, grade 4).

7.2 Medical History

Medical history data will be coded using the Medical Dictionary for Regulatory Activities (MedDRA). The actual version of the MedDRA coding dictionary will be noted in the statistical tables and clinical study report. The number and percentage of subjects in each medical history category (by MedDRA system organ class and preferred term) will be summarized. The system organ class (SOC) will be presented in alphabetical order, and the preferred terms will be presented in alphabetical order within each SOC. Subjects reporting more than one condition/diagnosis will be counted only once in each row (SOC or preferred term).

7.3 Prior and Concomitant Medications

Prior and concomitant medications will be summarized by generic name. A prior medication is defined as any medication taken prior to the date of the first dose of study treatment. A concomitant medication is defined as medications (other than study treatment) taken after the first dose of any component of study treatment and within 30 days of the last dose of any component of study treatment. For summaries of concomitant medications, if an incomplete start date was collected for a medication, the

medication will be assumed to be a concomitant medication unless there is evidence that confirms that the medication was not a concomitant medication (e.g., the medication end date was prior to the first dose of study treatment).

The number and percentage of subjects taking medications will be summarized by generic drug name based on the World Health Organization (WHO) Drug Dictionary for both prior and concomitant medications. The prior MM medications taken by the treatment experienced subjects will be summarized separately from other prior medications.

8.0 Efficacy Analyses

8.1 General Considerations

All efficacy analyses will be conducted in the ITT Population.

For time-to-event endpoints, survivorship functions will be estimated by using Kaplan-Meier product-limit methodology. Estimated survival curves will be presented. If reached, median time to event and its two-sided 95% confidence interval will be presented.

Study Drug and Study Treatment

Study drug: venetoclax

Study treatment: venetoclax + dexamethasone

Study treatment components:

- venetoclax
- dexamethasone

Baseline

"Baseline" refers to the last non-missing observation before the first administration of study treatment.

Last Adequate Disease Assessment

"Last adequate Disease Assessment" refers to the last evaluable disease response per investigator assessment that is not more than 12 weeks after the previous evaluable disease response per investigator assessment or prior to the start of new anti-MM therapy.

8.2 Primary Efficacy Endpoint(s) and Analyses

8.2.1 Primary Efficacy Endpoint(s)

The primary efficacy endpoints are ORR (the proportion of subjects with BOR of sCR, CR, VGPR or PR) and ≥ VGPR Rate (the proportion of subjects with BOR of sCR, CR, or VGPR) based on IMWG criteria.[1]

8.2.2 Primary Efficacy Analysis

Point estimates of ORR and VGPR+ rates and 95% confidence intervals using exact method will be provided.

8.3 Secondary Efficacy Analyses

The distribution of PFS, OS, TTR, TTP and DOR will be estimated for all subjects using Kaplan-Meier product-limit methodology on the ITT population. If reached, median time to event and a corresponding 95% confidence interval will be estimated.

For PROs, change in score from baseline will be evaluated for the following domains: Worst Pain (BPI-SF), Physical Functioning (EORTC QLQ-C30), Fatigue (PROMIS-Fatigue), and Global Health Status/QoL (EORTC QLQ-C30).

In addition, reason for censoring will be provided for PFS and OS.

8.4 Additional Efficacy Analyses

For PROs, descriptive statistics will be used to summarize change from baseline for remaining subscales of BPI-SF, EORTC QLQ-C30, European Organization for Research

and Treatment of Cancer Quality of Life Questionnaire Multiple Myeloma Module [EORTC QLQ-MY20], and EQ-5D-5L.

8.5 Efficacy Subgroup Analyses

Subgroup analyses will be performed for the following efficacy endpoints: ORR and VGPR+ rates, PFS, OS, and DOR if there are at least 5 subjects within the subgroup level.

The following subgroups will be analyzed:

- age (< 65 years or ≥ 65 years);
- ISS stage (I or II/III);
- Lines of prior therapy (< 1, 1 – 3, 4 – 6, 7 – 9, > 9)

9.0 Safety Analyses

9.1 General Considerations

Safety data will be summarized for the Safety Analysis Set.

9.2 Adverse Events

Adverse events (AEs) will be summarized and presented using primary MedDRA System Organ Classes (SOCs) and preferred terms (PTs) according to the version of the MedDRA coding dictionary used for the study at the time of database lock. The actual version of the MedDRA coding dictionary used will be noted in the AE tables and in the clinical study report. Specific adverse events will be counted once for each subject for calculating percentages, unless stated otherwise. In addition, if the same adverse event occurs multiple times within a subject, the highest severity and level of relationship to investigational product will be reported.

9.2.1 Treatment-Emergent Adverse Events

Treatment-emergent AEs are defined as any AE with the onset that is after the first dose of study treatment and no more than 30 days after the last dose of study treatment. Events

where the onset date is the same as the study treatment start date are assumed to be treatment-emergent. All treatment-emergent AEs will be summarized overall, as well as by primary MedDRA SOC and Preferred Term. The SOCs will be presented in alphabetical order, and the PTs will be presented in alphabetical order within each SOC.

The number and percentage of subjects experiencing treatment-emergent AEs will be summarized.

9.2.2 Adverse Event Overview

An overview of AEs will be presented consisting of the number and percentage of subjects experiencing at least one event for each of the following AE categories:

- Treatment-emergent AEs;
- Treatment-emergent AEs with a "reasonable possibility" of being related to venetoclax;
- Treatment-emergent AEs with a "reasonable possibility" of being related to dexamethasone;
- Treatment-emergent AEs with NCI-CTCAE toxicity grade 3 or 4;
- Serious treatment-emergent AEs;
- Serious treatment-emergent AEs with a "reasonable possibility" of being related to venetoclax;
- Serious treatment-emergent AEs with a "reasonable possibility" of being related to dexamethasone;
- Treatment-emergent AEs leading to death;
- Treatment-emergent AEs leading to venetoclax discontinuation;
- Treatment-emergent AEs leading to dexamethasone discontinuation;
- Treatment-emergent AEs leading to venetoclax dose reduction;
- Treatment-emergent AEs leading to dexamethasone dose reduction;
- Treatment-emergent AEs leading to venetoclax interruption;
- Treatment-emergent AEs leading to dexamethasone interruption;
- All deaths

- Deaths occurring ≤ 30 days after last dose of study treatment
- Deaths occurring > 30 days after last dose of study treatment.

9.2.3 Treatment-Emergent Adverse Events by SOC and/or PT

Adverse Events by SOC and PT

The following summaries of AEs by SOC and PT will be generated:

- Treatment-emergent AEs;
- Treatment-emergent AEs with a "reasonable possibility" of being related to venetoclax;
- Treatment-emergent AEs with a "reasonable possibility" of being related to dexamethasone;
- Treatment-emergent AEs with NCI-CTCAE toxicity grade 3 or 4.

Adverse Events by SOC/PT and by NCI-CTCAE Toxicity Grade

AEs will also be summarized by SOC/PT and by maximum NCI-CTCAE toxicity grade. If a subject has an adverse event with an unknown NCI-CTCAE toxicity grade, then the subject will be counted in the category of "unknown," even if the subject has another occurrence of the same adverse event with a grade present. The only exception is if the subject has another occurrence of the same AE with the highest grade level (Grade 5). In this case, the subject will be counted under the "Grade 5" category.

Specific adverse events will be counted once for each subject for calculating percentages, unless stated otherwise. In addition, if the same adverse event occurs multiple times within a subject, the highest severity and level of relationship to investigational product will be reported.

Adverse Events by PT

The following summaries of treatment-emergent AEs tabulated according to PT and sorted by decreasing frequency will be generated:

- Treatment-emergent AEs;
- Treatment-emergent AEs with a "reasonable possibility" of being related to venetoclax;
- Treatment-emergent AEs with NCI-CTCAE toxicity grade 3 or 4;
- Treatment-emergent AEs of Grade 3 or higher with a "reasonable possibility" of being related to venetoclax.

9.2.4 SAEs (Including Deaths) and Adverse Events Leading to Study Treatment Discontinuation, Dose Reduction, or Dose Interruption

The number and percentage of subjects experiencing SAEs (including deaths), SAE with a "reasonable possibility" of being related to venetoclax, SAE with "reasonable possibility" of being related to dexamethasone and AEs leading to venetoclax discontinuation, dose reduction, or dose interruption will be summarized by SOC and PT.

9.2.5 Deaths

The number of subject deaths will be summarized 1) for deaths occurring while the subject was still receiving study treatment in this study, 2) for deaths occurring off treatment within 30 days after the last dose of study treatment and 3) for all deaths in this study regardless of the number of days after the last dose of study treatment.

9.2.6 Listing of Adverse Events

The following additional summaries of adverse events will be prepared:

- All serious AEs;
- Treatment-emergent SAEs;

- Treatment-emergent AEs with NCI-CTCAE toxicity grade 3 or 4;
- Treatment-emergent AEs leading to death;
- Subject deaths occurring ≤ 30 days after last dose of study treatment;
- Subject deaths occurring > 30 days after last dose of study treatment.

9.2.7 Adverse Events of Special Interest

Adverse events of special interest will be summarized by PT and will be based on standardized or company MedDRA queries (SMQs or CMQs). Detailed information about the search criteria are provided in Appendix B.

Listings of AEs of special interest meeting the search criteria will be provided.

9.3 Analysis of Laboratory Data

Data collected from central and local laboratories, including additional laboratory testing due to an SAE, will be used in all analyses, except for Baseline where SAE-related laboratory assessments on or before the first dose of study treatment will be excluded. The hematology and chemistry laboratory tests defined in the protocol Section 5.3.1.1 Table 13 will be summarized.

The baseline value for clinical laboratory tests will be the last non-missing measurement on or before the day of the first dose of study treatment. Values on Day 1 must also be before the time of the first dose of study treatment if time is available. The same baseline value will be used for analyses of the baseline and post-baseline period.

Shift Tables

Changes in laboratory parameters will be tabulated using shift tables. For hematology and chemistry laboratory parameters with NCI CTCAE 4.03 toxicity grade exist, baseline and post-baseline values will be assigned a toxicity grade of 0, 1, 2, 3, or 4. For laboratory values either within the normal range or outside the normal range but in the direction opposite the test will be classified as grade 0 values.

The baseline value for clinical laboratory tests will be the last non-missing measurement on or before the day of the first dose of study treatment. The maximum post-baseline NCI toxicity grade value is the value with highest NCI toxicity grade collected after the first dose of study treatment and within 30 days following the last dose of study treatment. If multiple values are collected on the same day, the maximum grade value will be used as the value for that day.

For each variable, cross tables will be generated for the number of subjects with baseline values of Grade 0, Grade 1, Grade 2, Grade 3, Grade 4, or missing grade, versus maximum post-baseline observations of Grade 0, Grade 1, Grade 2, Grade 3, Grade 4, missing grade. All subjects in the safety analysis set will be included in the cross tabulation regardless whether baseline or post-baseline measurements are collected.

Assessment of Drug-Induced Liver Injury

The number and percentage of subjects in each treatment group who have at least one observed post-baseline value meeting the following criteria will be tabulated

- Alanine Amino Transferase (ALT) > 3 × Upper Limit Normal (ULN)
- AST > 3 × ULN
- AST or ALT > 3 × ULN
- Total bilirubin value > 2 × ULN
- ALT or AST > 3 × ULN AND Total bilirubin > 2 × ULN within 72 hours.

A listing of all observed ALT, AST, and total bilirubin values will be generated for subjects with an observed value meeting any of these criteria.

Laboratory Assessments for Tumor Lysis Syndrome

To determine if a subject's laboratory values qualify for Tumor Lysis Syndrome (TLS), the Howard criteria[2] will be assessed. The Howard definition for laboratory TLS requires

≥ 2 of the metabolic abnormalities specified in Table 5 post-baseline and within 24 hours of each other.

Table 5. Laboratory Criteria for TLS

Element	Value
Uric Acid	> 476 µmol/L
Potassium	> 6.0 mmol/L
Inorganic Phosphorus	> 1.5 mmol/L
Corrected Calcium	< 1.75 mmol/L

The following summaries of laboratory criteria will be provided:

- Number and percentage of subjects meeting the definition of laboratory TLS (at least two values meeting the criteria in Table 5, occurring within 24 hours of each other).
- Listing of all values for these four analyses for each subject meeting the definition of laboratory TLS at least once during treatment.

9.4 Analysis of Vital Signs

Vital sign measurements of systolic and diastolic blood pressure, pulse rate, and body temperature will be summarized.

The baseline value for vital signs will be the last measurement on or before the day of the first dose of study treatment. The same baseline value will be used analyses of the baseline and post-baseline period.

Vital sign variables will be evaluated based on potentially clinically important (PCI) criteria (Appendix C). For each vital sign PCI criterion, the number and percentage of subjects who have an on-treatment vital sign value meeting the criteria will be summarized. A post-baseline value must be more extreme than the baseline value to be considered a PCS finding. Listings will be provided to summarize subject-level vital sign data for subjects meeting PCI criteria.

10.0 Interim Analyses

There will be no interim analysis. Final analysis will occur after all enrolled subjects have completed the study. Data will be locked after performing appropriate data cleaning.

10.1 Data Monitoring Committee

This study does not require a separate data monitoring committee (DMC) to review either safety or efficacy data. A Safety Review Committee (SRC) will periodically review safety data across all studies with venetoclax in MM that do not have a study-specific independent monitoring committee. This review committee will be responsible for periodic, regular reviews to assess the safety of the interventions during the trial. A separate charter will be prepared outside of the protocol outlining the SRC member responsibilities, frequency of data reviews, and relevant data to be assessed.

11.0 Version History

Table 6. SAP Version History Summary

Version	Date	Summary
1.0	06 April 2020	Original version
1.1	11 January 2022	Changes from SAP version 1.0 • Clarify the description of event and censoring rules of PFS in Table 1 • Clarify the description of event and censoring rules of DOR in Table 2 • Specify the analysis TTR to be conducted in the ITT population. Add the description of event and censoring rules of TTR in Table 3 • Renumber Table 3 to Table 4. Clarify the description of event and censoring rules of PFS in Table 4 • Define last adequate disease assessment in Section 8.1 • Correct minor typos

12.0 References

1. Kumar, S, Paiva B, Anderson KC, et al. International Myeloma Working Group consensus criteria for response and minimal residual disease assessment in multiple myeloma. Lancet Oncol. 2016;17(8):e328-46.
2. Howard SC, Jones DP, Pui CH. The tumor lysis syndrome. N Engl J Med. 2011;364(19):1844-54.
3. Cleeland CS. Pain assessment in cancer. In Effect of Cancer on Quality of Life. Osaba D, editor. Boca Raton: CRC Press; 1991:21.
4. Atkinson TM, Mendoza TR, Sit L, et al. The brief pain inventory and its "pain at its worst in the past 24 hours" item: clinical trial endpoint considerations. Pain Med. 2010;11(3):337-46.
5. Aaronson NK, Ahmedzai S, Bergman B, et al. The European Organization for Research and Treatment of Cancer QLQ-C30: a quality-of-life instrument for use in international clinical trials in oncology. J Natl Cancer Inst. 1993;85(5):365-76.
6. Cocks K, Cohen D, Wisloff F, et al. An international field study of the reliability and validity of a disease-specific questionnaire module (the QLQ-MY20) in assessing the quality of life of patients with multiple myeloma. Eur J Cancer. 2007;43(11):1670-8.
7. Cella D, Yount S, Rothrock N, et al. The patient-reported outcomes measurement information system (PROMIS): progress of an NIH roadmap cooperative group during its first two years. Med Care. 2007;45 (5 Suppl 1):S3-S11.
8. Garcia SF, Cella D, Clauser SB, et al. Standardizing patient-reported outcomes assessment in cancer clinical trials: a patient-reported outcomes measurement information system initiative. J Clin Oncol. 2007;25(32):5106-12.
9. Cella D, Riley W, Stone A, et al. The patient-reported outcomes measurement information system (PROMIS) developed and tested its first wave of adult self-reported health outcome item banks: 2005-2008. J Clin Epidemiol. 2010;63(11):1179-94.

10. EQ-5D-5L and EQ-5D-VA. Available from: http://www.euroqol.org.

Appendix A. Protocol Deviations

The number and percentage of subjects who reported at least one of the following protocol deviation categories will be provided.

- Subject entered into the study even though s/he did not satisfy entry criteria.
- Subject developed withdrawal criteria during the study and was not withdrawn.
- Subject received wrong treatment or incorrect dose of study.
- Subject took prohibited concomitant medication.

Appendix B. Definition of Adverse Events of Special Interest Selected Adverse Events

Adverse Events of Special Interest (AESI) and Selected Adverse Events will be identified using the following search criteria.

Risk	Search Criteria
Tumor Lysis Syndrome (5 searches) -AESI	1) SMQ 20000219 – "Tumour lysis syndrome" (Narrow-scope) 2) PT terms of "Hyperkalaemia" and "Blood potassium increased" 3) PT terms of "Hyperphosphataemia" and "Blood phosphorus increased" 4) PT terms of "Hypocalaemia" and "Blood calcium decreased" 5) PT terms of "Hyperuricaemia" and "Blood uric acid increased"
Neutropenia – expanded search	Neutropenia CMQ 80000171
Neutropenia	PT terms – "Neutropenia" and "Neutrophil count decreased"
Serious Infection	SAEs in the SOC of "Infections and Infestations"
Opportunistic infection	BCL-2 CMQ
Second Primary Malignancy	SMQ 20000194 – "Malignant tumours" (Narrow) and SMQ 20000217 - "Myelodysplastic syndromes" (Narrow)
Lymphopenia	PT terms – "Lymphopenia" and "Lymphocyte count decreased"
Anemia	PT terms – "Anaemia" and "Haemoglobin decreased"
Thrombocytopenia	PT terms – "Thrombocytopenia" and "Platelet count decreased"

Appendix C. Potentially Clinically Important Criteria for Safety Endpoints

The criteria for Potentially Clinically Important (PCI) vital sign findings are described in Table C.

Table C. Criteria for Potentially Clinically Important Vital Sign Values

Vital Signs Variables	Criterion	Definition of Potentially Clinically Significant
Systolic Blood Pressure	High	Value ≥ 160 mmHg
Diastolic Blood Pressure	High	Value ≥ 100 mmHg
Pulse	Low	Value < 50 bpm
	High	Value ≥ 120 bpm
Temperature	Low	Value < 36°C
	High	Value ≥ 38.5°C

Appendix D. PRO assessments

PRO assessments include: BPI-SF, EORTC QLQ-C30, EORTC QLC MY20, PROMIS Cancer Fatigue SF, and EQ-5D-5L.

BPI-SF

The BPI-SF is a pain-specific measure developed to assess patient-reported severity (or intensity) of pain (4 items) and the impact of pain on daily functioning (7 items) in patients with cancer pain.[3] The four pain severity items assess pain at its "worst," "least," "average," and "now" (current pain). For these items, patients are asked to rate their pain on an 11-point numeric rating scale with anchors of 0 (no pain) and 10 (pain as bad as you can imagine). The BPI "worst" pain severity item has been shown to be reliable and valid for use as a single item.[4] The BPI-SF also includes questions to measure the interference of pain in the patient's daily life, including general activity, mood, ability to walk, normal work (both outside the home and housework), relationships with other people, sleep, and enjoyment of life. For these items, patients are asked to describe the extent to which pain has interfered on an 11-point numeric rating scale with anchors of 0 (does not interfere) to 10 (completely interferes).

Table D-1. BPI-SF Scales

Scales	Items
Pain Severity*	BPI2-Pain Right Now
	BPI2-Pain at its Least in Last 24 Hours
	BPI2-Pain on the Average
	BPI2-Pain at its Worst in Last 24 Hours
Pain Interference**	BPI2-Pain Interfered Enjoyment of Life
	BPI2-Pain Interfered General Activity
	BPI2-Pain Interfered Walking Ability
	BPI2-Pain Interfered with Mood
	BPI2-Pain Interfered with Normal Work
	BPI2-Pain Interfered with Relations
	BPI2-Pain Interfered with Sleep

* Calculated as sum of scores of all 4 items and then dividing by 4. This gives the Pain Severity score out of 10.

** Calculated as sum of scores of all 7 items and then dividing by 7. This gives the Pain Interference score out of 10.

EORTC QLQ-C30

HRQoL, functioning, and symptoms will be assessed with the EORTC-QLQ-C30 version 3.[5] The QLQ-C30 is a 30-item subject self-report questionnaire composed of both multi-item and single scales, including five functional scales (physical, role, emotional, social, and cognitive), three symptom scales (fatigue, nausea and vomiting, and pain), a global health status/quality of life scale, and six single items (dyspnea, insomnia, appetite loss, constipation, diarrhea, and financial difficulties). The subjects rate items on a four point scale, with 1 as "not at all" and 4 as "very much." The QLQ-C30 was developed and validated for use in a cancer patient population, and its reliability and validity is highly consistent across different language cultural groups. The items and scales of the QLQ-C30 are as in Table D-2:

Table D-2. EORTC QLQ-C30 Scales

	Scale	Number of Items	Item Range	Item Numbers
Global Health Status/Quality of Life				
Global Health Status/Quality of Life	QL2	2	6	29, 30
Functional Scales				
Physical Functioning*	PF2	5	3	1 – 5
Role Function*	RF2	2	3	6 – 7
Emotional Functioning*	EF	4	3	21 – 24
Cognitive Functioning*	CF	2	3	20, 25
Social Functioning*	SF	2	3	26 – 27
Symptom Scales				
Fatigue	FA	3	3	10, 12, 18
Nausea and Vomiting	NV	2	3	14, 15
Pain	PA	2	3	9, 19
Dyspnoea	DY	1	3	8
Insomnia	SL	1	3	11
Appetite Loss	AP	1	3	13
Constipation	CO	1	3	10
Diarrhoea	DI	1	3	17
Financial Difficulties	FI	1	3	28

* Function Scales.

Note: Item range is the difference between the possible maximum and the minimum response to individual items.

Scoring algorithms for scales are as follows:

For all Scales, the Raw Score (RS) is the mean of the component items (I):

If items I_1, I_2..., I_n are included in a scale, the procedures are as follows:

$$\text{Raw Score (RS)} = (I_1 + I_2 + ... + I_n)/(\text{number of non-missing items}),$$

Then the Functional Scales:

$$\text{Score} = (1 - (RS - 1)/\text{range}) \times 100$$

where range is provided in Table 13.

For example, to calculate Emotional Functioning (EF),

$$RS_{EF} = (I_{21} + I_{22} + I_{23} + I_{24})/4$$

$$Score_{EF} = (1 - (R_{SEF} - 1)/3) \times 100$$

If a subject completed more than 50% of the items in a scale, then the raw score of that subject will contribute to the summary statistics of that scale. If a subject completed less than 50% of the items in a scale, then the raw score of that subject will be dropped from the calculation of the summary statistics of that scale.

EORTC QLQ-MY20

The EORTC QLQ-MY20 was developed as an additional module for the QLQ-C30 and is composed of 20 items specific to multiple myeloma.[6] The QLQ-MY20 includes scales for disease symptoms, side effects of treatment, future perspective, and body image. Values for each scale range from 0 to 100. The subjects rate items on a four-point scale, with 1 as "not at all" and 4 as "very much." The QLQ-MY20 is a reliable and valid instrument for measuring quality of life in myeloma patients.

Table D-3. QLQ-MY20 scales

	Scale	Number of Items	Item Range	QLQ-MY20 Item Numbers
Functional Scales/items				
Future perspective	MYFP	3	3	18 - 20
Body image	MYBI	1	3	17
Symptom Scales				
Disease symptoms	MYDS	6	3	1 - 6
Side effects of treatment	MYSE	10	3	7 - 16

PROMIS Cancer Fatigue SF

The PROMIS® is a system of highly reliable, precise measures of patient-reported health status for physical, mental, and social well-being.[7] PROMIS instruments measure concepts such as pain, fatigue, physical function, depression, anxiety and social function. Fatigue will be assessed using the PROMIS Cancer Fatigue SF that has been developed for use in oncology populations.[8,9] PROMIS Cancer Fatigue SF is a seven item questionnaire that assesses the impact and experience of fatigue over the past 7 days. All questions employ the following five response options: 1 = Not at all, 2 = A little bit, 3 = Somewhat, 4 = Quite a bit, and 5 = Very much. The score is calculated as below:

Score = Sum of score from all questions × 7/Total number of answered questions

If 4 or more questions were not being answered, then the score is considered as missing. The details of scoring method can be found in the scoring manual https://www.assessmentcenter.net/documents/PROMIS%20Fatigue%20Scoring%20Manual.pdf.

EQ-5D-5L

The EQ-5D-5L is a generic preference instrument that has been validated in numerous populations.[10] The EQ-5D-5L has five dimensions: mobility, self-care, usual activities, pain/discomfort and anxiety/depression. These dimensions are measured on a five level scale: no problems, slight problems, moderate problems, severe problems, and extreme problems. The EQ-5D-5L also contains a visual analog scale (VAS) to assess the subject's overall health. Detailed EQ-5D-5L scoring method can be found at https://euroqol.org/wp-content/uploads/2016/09/EQ-5D-5L_UserGuide_2015.pdf.

1.0 Title Page

Statistical Analysis Plan — Phase 1 Subjects

Study M13-367

A Phase 1/2 Study Evaluating the Safety and Pharmacokinetics of ABT-199 in Subjects with Relapsed or Refractory Multiple Myeloma

Date: 20 Sep 2018

Version 1.0

2.0 Table of Contents

1.0 Title Page 1
2.0 Table of Contents 2
3.0 Introduction 5
4.0 Study Background 5
4.1 Objectives 5
4.1.1 Primary Objectives 5
4.1.2 Secondary Objectives 5
4.1.3 Exploratory Objectives 6
4.2 Study Design 6
4.3 Endpoints 7
4.3.1 Safety Endpoints 7
4.3.2 Efficacy Endpoints 7
4.3.2.1 Disease Progression Definition and Confirmation 7
4.3.2.2 Response Rates 7
4.3.2.2.1 Overall Response Rate (ORR) 8
4.3.2.2.2 Very Good Partial Response or Better Rate (≥ VGPR Rate) 8
4.3.2.3 Time to Event Endpoints 8
4.3.2.3.1 Time to Progression (TTP) 8
4.3.2.3.2 Time to Response (TTR) 9
4.3.2.3.3 Duration of Response (DOR) 9
4.3.2.4 Exploratory Efficacy Endpoint: Minimal Residual Disease (MRD) Negativity Rate 9
5.0 Analysis Population and Subgroup Reporting 10
5.1 Analysis Population 10
5.2 Subgroup Reporting 10
6.0 Analysis Conventions 12
6.1 Definition of Study Drug 12
6.2 Definition of Baseline 12
6.3 Definition of Final Observation 12
6.3.1 Definition of Date of Last Contact 13

6.4 Definition of Study Rx Days (Days Relative to the First Dose of Study Treatment) ... 13
6.5 Definition of Analysis Windows ... 14
7.0 Subject Disposition ... 14
8.0 Demographics, Baseline Characteristics, Medical History, and Previous/Concomitant Medications ... 15
8.1 Demographic and Baseline Characteristics ... 15
8.1.1 Categorical Demographic Characteristics ... 15
8.1.2 Continuous Demographic Characteristics ... 16
8.2 Study Drug Exposure ... 16
8.3 Medical History ... 17
8.4 Prior/Concomitant Medications and Prior Multiple Myeloma Therapies ... 17
8.4.1 Concomitant Medications ... 17
8.4.2 Prior medication ... 18
8.5 Compliance ... 18
9.0 Safety Analysis ... 18
9.1 General Considerations ... 18
9.2 Analysis of Adverse Events ... 19
9.2.1 Treatment-Emergent Adverse Events (TEAEs) ... 19
9.2.2 Serious Adverse Events ... 21
9.2.3 Deaths ... 22
9.2.4 Non-Treatment-Emergent Adverse Events ... 22
9.3 Analysis of Laboratory and Vital Signs ... 22
9.3.1 Longitudinal Analyses of Laboratory and Vital Signs Data ... 22
9.3.2 Analyses of Laboratory Data Using NCI CTCAE ... 23
9.3.3 Laboratory Search Strategy for TLS ... 24
9.3.4 Assessment of Drug-Induced Liver Injury ... 24
9.4 Analysis of Vital Signs ... 24
9.4.1 Potentially Clinically Significant Values ... 24
9.4.2 Mean Change from Baseline ... 24
9.5 Analysis of ECG Parameters ... 25
10.0 Efficacy Analysis ... 25

10.1 Response Rates 25
10.2 Time to Event Endpoints 26
11.0 Pharmacokinetic Analysis 26
11.1 Tabulations and Summary Statistics 26
11.2 Model and Test 27
11.3 Missing Values and Model Violations 28
12.0 References 28
13.0 Appendix 29
13.1 IMWG 2011[1-3] Response Criteria 29

List of Tables

Table 1. Event and Censoring Rules for TTP 8
Table 2. Event and Censoring Rules for DOR 9
Table 3. Search Criteria for Selected Adverse Events 21
Table 4. Time Windows for Analysis of Laboratory and Vital Signs Parameters – Phase 1 Monotherapy Subjects 32
Table 5. Time Windows for Analysis of Laboratory and Vital Signs Parameters – Phase 1 VenDex Combination Subjects 33
Table 6. Clinical Laboratory Tests – Phase 1 Portion 34
Table 7. Definitions of Laboratory Tumor Lysis Syndrome (TLS) 34
Table 8. Criteria for Potentially Clinically Significant Vital Signs Parameters 35

List of Figures

Figure 1. Overall Study Design 6

3.0 Introduction

This statistical analysis plan (SAP) provides details on the planning, analysis, and reporting of demographics, efficacy, and safety in venetoclax (ABT-199) Study M13-367.

Unless noted otherwise, all analyses will be performed using SAS version 9.4 or later (SAS Institute Inc., Cary, NC 27513) under the Unix operating system.

This SAP only pertains to the Phase 1 portion of Study M13-367, which includes the dose escalation, safety expansion, and venetoclax + dexamethasone (VenDex) combination cohorts. The analyses for the Phase 2 portion of Study M13-367 will be documented in a separate SAP.

4.0 Study Background

4.1 Objectives

4.1.1 Primary Objectives

The primary objectives are to assess the safety profile, characterize pharmacokinetics (PK), determine the dosing schedule, the maximum tolerated dose (MTD), and the recommended Phase 2 dose (RPTD) of venetoclax monotherapy when administered in subjects with relapsed or refractory multiple myeloma (MM). This study will also assess the safety profile and PK of VenDex in subjects with t(11;14)-positive MM.

4.1.2 Secondary Objectives

The secondary objectives are to evaluate the preliminary efficacy data regarding the effect of venetoclax monotherapy or combined with dexamethasone in regard to overall response rate (ORR), time to response (TTR), time to disease progression (TTP), and duration of response (DOR).

4.1.3 Exploratory Objectives

The exploratory objectives are to evaluate pharmacodynamic and predictive biomarkers for associations with PK, safety, and efficacy. Additionally, minimal residual disease (MRD) will be assessed in the bone marrow by next generation sequencing (NGS).

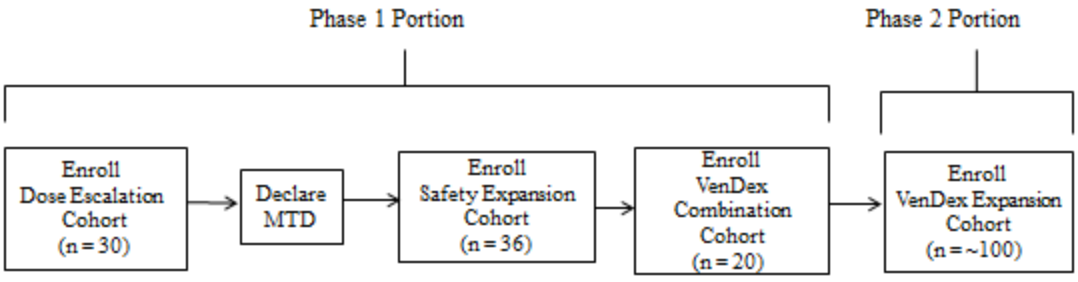
4.2 Study Design

This is a Phase 1/2, open-label, multicenter dose escalation and expansion study.

Phase 1 includes the following cohorts utilizing a 3 + 3 design:

- Dose escalation (venetoclax monotherapy, 30 subjects)
  - 300 mg (6 subjects)
  - 600 mg (9 subjects)
  - 900 mg (6 subjects)
  - 1200 mg (9 subjects)
- Safety expansion (venetoclax monotherapy, 36 subjects at designated cohort dose of 1200 mg)
- VenDex combination (800 mg venetoclax + 40 or 20 mg dexamethasone, 20 subjects, positive for t(11;14))

Figure 1. Overall Study Design



See protocol for information on dosing schedules.

Disease assessments will be based on the International Uniform Response Criteria for Multiple Myeloma (IMWG, 2011), per investigator assessment. See Appendix 13.1 for complete details.

4.3 Endpoints

4.3.1 Safety Endpoints

Safety endpoints include:

- Adverse events (AE)
- Serious adverse events (SAE)
- Deaths
- Electrocardiography (ECG, MUGA)
- Vital sign parameters
- Laboratory assessments

4.3.2 Efficacy Endpoints

4.3.2.1 Disease Progression Definition and Confirmation

Disease progression (PD) is defined as satisfying at least one of the following:

- PD due to imaging (new lesion, increase in lesion size, etc.)
- Clinical PD
- Per IMWG PD (non-imaging) that is confirmed by a consecutive assessment

4.3.2.2 Response Rates

We will report the best overall response a subject experienced before discontinuation due to any reason (PD, death, etc.). All responses must be confirmed by a subsequent visit, with the exception of SD, which is confirmed by the baseline assessment. Responses occurring after the start of a new line of anti-MM therapy will not be considered.

4.3.2.2.1 Overall Response Rate (ORR)

ORR is defined as the proportion of subjects who experienced a response of PR or better. In particular, this includes subjects with a best overall response of PR, VGPR, CR, or sCR.

4.3.2.2.2 Very Good Partial Response or Better Rate (≥ VGPR Rate)

≥ VGPR rate is defined as the proportion of subjects who experienced a response of VGPR or better. In particular, this includes subjects with a response of VGPR, CR, or sCR.

4.3.2.3 Time to Event Endpoints

4.3.2.3.1 Time to Progression (TTP)

TTP is defined as the number of days from the date of first dose until the date of PD or death due to MM, whichever occurs first, i.e.,:

TTP (days) = PD or MM-related death date – date of first dose + 1

Table 1. Event and Censoring Rules for TTP

Situation	Date of Censor or Event	Outcome
PD/death due to MM	PD/death date	Event
No adequate baseline and/or post-baseline disease assessment	Date of first dose	Censored
PD/death due to MM more than 12 weeks after last adequate disease assessment	Last adequate disease assessment date	Censored
No PD/no death due to MM	Last adequate disease assessment date	Censored
Start of new anti-MM therapy prior to PD/death	Last adequate disease assessment date prior to start of new anti-MM therapy	Censored

4.3.2.3.2 Time to Response (TTR)

TTR is defined as the number of days from the date of first dose until the date of first response of PR or better, i.e.,:

TTR (days) = date of first response – date of first dose + 1

Only subjects who achieved a response (PR or better) will be included in this analysis.

4.3.2.3.3 Duration of Response (DOR)

DOR is defined as the number of days from the date of first response of PR or better until date of PD or death due to MM, whichever occurs first, i.e.,:

DOR (day) = PD or MM-related death date – date of first response (PR or better) + 1

Only subjects who achieved a response of PR or better will be included in this analysis.

Table 2. Event and Censoring Rules for DOR

Situation	Date of Censor or Event	Outcome
PD/death due to MM	PD/death date	Event
PD/death due to MM more than 12 weeks after last adequate disease assessment	Last adequate disease assessment date	Censored
No PD/no death due to MM	Last adequate disease assessment date	Censored
Start of new anti-MM therapy prior to PD/death	Last adequate disease assessment date prior to start of new anti-MM therapy	Censored

4.3.2.4 Exploratory Efficacy Endpoint: Minimal Residual Disease (MRD) Negativity Rate

MRD negativity rate is defined as the proportion of subjects who have negative MRD by bone marrow aspirate at any time point after receiving their first dose of study drug. MRD negativity will be defined at the 10^{-4}, 10^{-5}, and 10^{-6} thresholds (less than one residual myeloma cell per 10^4, 10^5, and 10^6 total nucleated cells, respectively) as

measured by centralized testing of bone marrow aspirate by next generation sequencing (NGS). Subjects who were dosed but have missing or indeterminate MRD status per given threshold will be considered as positive for the calculation of MRD negativity status.

5.0 Analysis Population and Subgroup Reporting

5.1 Analysis Population

Unless otherwise specified, all analyses will be performed using all subjects who received at least one dose of study drug.

One notable exception is that any reporting of dexamethasone related information (i.e., AEs related to dexamethasone, subject dexamethasone discontinuation, etc.) will be inherently relevant to only the VenDex treated subjects.

5.2 Subgroup Reporting

- Response rates (including ORR and ≥ VGPR rate) will be summarized for the following groups, where sufficient data is available:
  - All monotherapy subjects
  - 300 mg venetoclax monotherapy cohort
  - 600 mg venetoclax monotherapy cohort
  - 900 mg venetoclax monotherapy cohort
  - 1200 mg venetoclax monotherapy and safety expansion cohorts
  - Venetoclax monotherapy subjects with t(11;14)
  - Venetoclax monotherapy subjects that are non-t(11;14)
  - VenDex combination subjects
  - All subjects
  - t(11;14) subjects, by dosing cohort:
    a. 300 mg venetoclax monotherapy t(11;14) subjects
    b. 600 mg venetoclax monotherapy t(11;14) subjects

c. 900 mg venetoclax monotherapy t(11;14) subjects
d. 1200 mg venetoclax monotherapy and safety expansion t(11;14) subjects
e. All t(11;14) subjects

- t(11;14) subjects, by therapy and BCL-2 status:
  a. BCL2 high subjects with t(11;14) on monotherapy
  b. BCL2 low subjects with t(11;14) on monotherapy
  c. BCL2:BCL2L1 ratio high subjects with t(11;14) on monotherapy
  d. BCL2:BCL2L1 ratio low subjects with t(11;14) on monotherapy
  e. BCL2 high subjects in the VenDex combination cohort
  f. BCL2 low subjects in the VenDex combination cohort
  g. BCL2:BCL2L1 ratio high subjects in VenDex combination cohort
  h. BCL2:BCL2L1 ratio low subjects in VenDex combination cohort

- Time to event endpoints (TTP, TTR, and DOR) will be summarized for the following groups:
  - All monotherapy subjects
  - Venetoclax monotherapy subjects with t(11;14)
  - Venetoclax monotherapy subjects that are non-t(11;14)
  - VenDex combination subjects
  - All subjects
- MRD negativity rate will be reported for the following groups:
  1. All subjects
  2. Subjects who achieved a response of CR or better.
- All other reporting (demographics, safety, listings, etc.,) will be broken down into the following groups:
  - 300 mg venetoclax monotherapy
  - 600 mg venetoclax monotherapy
  - 900 mg venetoclax monotherapy

- 1200 mg venetoclax monotherapy
- Safety expansion subjects (venetoclax monotherapy)
- Dose escalation subjects (venetoclax 1200 mg monotherapy)
- All subjects on venetoclax monotherapy
- VenDex combination subjects
- All subjects

6.0 Analysis Conventions

6.1 Definition of Study Drug

Subjects who took at least one dose of study drug will be included all analyses.

For the monotherapy cohort subjects, study drug is defined as venetoclax whereas for the combination VEN + DEX subjects, study drug is defined as either venetoclax or dexamethasone.

Study drug discontinuation is defined as the last dose of study drug.

6.2 Definition of Baseline

For all analyses where baseline values are used, baseline refers to the last non-missing observation collected on or prior to the date of first dose of study drug. If there was a measurement collected on the day of the first dose of study drug, if there was no time collected, it is assumed this measurement was collected before the first dose of study drug.

6.3 Definition of Final Observation

The final observation will be defined as the last non-missing observation collected after baseline and less than or equal to 30 days after the last dose of study drug.

6.3.1 Definition of Date of Last Contact

The date of last contact will be derived from the following list of data panels:

- Vital signs
- Physical exam
- Lab variables, including SAE lab reports
- ECOG performance status
- Study drug administration
- IMWG assessments
- Transfusions
- Electrocardiogram
- Adverse event
- Concomitant Medication
- Laboratory sample collection (e.g., PK, biomarker)
- Clinical disease progression assessments

The data panels with the last recorded date will be considered the date of last contact.

6.4 Definition of Study Rx Days (Days Relative to the First Dose of Study Treatment)

Rx Days will be calculated for each time point relative to the first dose of study drug. They will be defined as the number of days between the day of the first dose of study drug and the specific time point. Different distinctions of Rx days are displayed below.

- Rx days will be negative values when the time point of interest is prior to the first study drug dose day.
- Rx days will be positive values when the time point of interest is on or after the first study drug dose day.
- The day of the first dose of study drug is defined as Rx Day 1, while the day prior to the first study drug dose is defined as Rx Day –1 (there is no Rx Day 0).

6.5 Definition of Analysis Windows

Analysis time windows are given in Table 4 and Table 5 and will be constructed using the following algorithm:

- Determine the nominal Study Rx Day for each scheduled visit
- Determine the window around the specific nominal Study Rx Day as in Table 4 and Table 5
- If more than one assessment is included in a time window, the assessment closest to the nominal day will be used. If there are two observations of equal distance to the nominal day, the latest one will be used
- If more than one post-baseline measurement exists for a subject on a particular day, an arithmetic average is calculated and used as the subject's measurement for that day. For laboratory shift tables, the CTC grade will be assigned based on all observed laboratory values and no averages will be taken
- One cycle is defined as 21 days

7.0 Subject Disposition

A listing will be provided for the subjects who:

- Screen failed
- Enrolled, but were not treated

A summary will be provided for the number and percentage of subjects who:

- Discontinued venetoclax for any by and primary reason
- Discontinued dexamethasone for any by and primary reason
- Discontinued the study for any and by primary reason (for subjects in the monotherapy portion of the trial, subjects who discontinue venetoclax also discontinue the study)
- Are active/ongoing

8.0 Demographics, Baseline Characteristics, Medical History, and Previous/Concomitant Medications

8.1 Demographic and Baseline Characteristics

8.1.1 Categorical Demographic Characteristics

Categorical demographic characteristics will be summarized by the frequency and percentage (calculated using non-missing values) of subjects. The following categorical demographic characteristics will be reported:

- Sex (Female, Male)
- Race (White, Black, Asian, American Indian/Alaska Native, Native Hawaiian or Other Pacific Islander, Other, Multi Race)
- Ethnicity (Hispanic or Latino, No Ethnicity)
- Age (< 40, 40 – < 60, ≥ 60)
- Type of myeloma (IgG, non-IgG, free light chain)
- ISS status (I, II, III)
- t(11;14) status (negative, positive)
- t(4;14) status (negative, positive)
- t(14;16) status (negative, positive)
- 17P status (deleted, not deleted, undetermined)
- 13Q status (deleted, not deleted, undetermined)
- Chromosomal Abnormality (CA) risk by FISH (High/Standard)
    - High risk subjects include those that are t(4;14) positive OR t(14;16) positive OR 17p13 deleted
    - Standard risk subjects include those that are t(4;14) negative AND t(14;16) negative AND 17p13 not deleted
    - Subjects with missing or indeterminate t(4;14), t(14;16), or 17p13 status are unable to be categorized by CA risk, and are considered missing
- Prior allogenic stem cell transplant (yes, no)
- Prior autologous stem cell transplant (yes, no)

- Prior bortezomib (yes, no)
- Bortezomib refractory status (yes, no)
- Prior lenalidomide (yes, no)
- Lenalidomide refractory status (yes, no)
- Lines of prior therapy (< 1, 1 – 3, 4 – 6, 7 – 9, > 9)
- ECOG Performance Status (GRADE 0, GRADE 1, GRADE 2, GRADE 3, GRADE 4)

8.1.2 Continuous Demographic Characteristics

Continuous demographic characteristics will be summarized by sample size, mean, standard deviation, median, minimum, and maximum values, which will be calculated using only the non-missing observations. The following continuous demographic characteristics will be reported:

- Age
- Weight
- Weight – female
- Weight – male
- Height

8.2 Study Drug Exposure

Duration of exposure to study drug will be defined for each subject as (last study drug dose date – first study drug dose date) + 1 and will be summarized by sample size, mean, standard deviation, median, minimum and maximum. The number and percentage of subjects exposed to study drug will be summarized for the following durations: 0 to 1 cycle, > 1 cycle to 2 cycles, > 2 cycles to 3 cycles, > 3 cycles to 4 cycles, > 4 cycles to 5 cycles, > 5 cycles to 6 cycles, and > 6 cycles.

8.3 Medical History

Medical history and ongoing conditions, including MM-related conditions and symptoms will be summarized and listed. Separate summaries will be presented for ongoing and historical medical conditions. The summaries will be presented by primary system organ class and preferred term. Medical history/current medical conditions are coded using the Medical Dictionary for Regulatory Activities (MedDRA) terminology.

The body systems will be presented in alphabetical order and the conditions/diagnoses will be presented in alphabetical order within each body system. Subjects reporting more than one condition/diagnosis within a body system will be counted only once for that body system.

8.4 Prior/Concomitant Medications and Prior Multiple Myeloma Therapies

8.4.1 Concomitant Medications

- Definition: Concomitant medications are defined as medications (other than study treatment) taken after the first dose of study drug and within 30 days of the last dose of study drug that (1) was started prior to the first dose of study drug and continued to be taken after the first dose of study drug or (2) was started after the first dose of study drug, but was not started 30 days after the last dose of study drug. Concomitant medications will be reported by generic name assigned by the World Health Organization (WHO) dictionary.
- Analysis: The number and percentage of subjects who take at least one concomitant medication will be summarized. For each specific concomitant medication, the number and percentage of subjects who take at least one dose of this medication will also be summarized. For summaries of concomitant medications, if an incomplete start date was collected for a medication, the medication will be assumed to be a concomitant medication unless there is evidence that confirms that the medication was not a concomitant medication (i.e., if the medication end date was prior to the first dose of study drug).

8.4.2 Prior medication

- Definition: A prior medication (also identified as prior therapy) will be defined as any medication taken with an end date prior to the first dose of study drug.
- Analysis: The number and percentage of subjects who take at least one prior medication will be summarized. For each specific prior medication, the number and percentage of subjects who take at least one dose of this medication will also be summarized. The frequency and percentage of subjects who have prior multiple myeloma therapies will be summarized by regimen names. In addition, the best response to any therapies prior to first dose of study drug will be summarized.

8.5 Compliance

To assess compliance, a listing of protocol deviations will be given.

9.0 Safety Analysis

9.1 General Considerations

For basophils, eosinophils, lymphocytes, monocytes, neutrophils and reticulocytes, reporting in tables will be done only for absolute values. If sites have provided only a % and no absolute value, conversion from % to absolute will be performed according to the following formulas:

- Basophils [absolute value] = basophils [%] • leukocytes [absolute value], same for eosinophils, lymphocytes, monocytes, and neutrophils
- Reticulocytes [absolute value] = reticulocytes • erythrocytes [absolute value]

Reference intervals for absolute values will be used (i.e., reference intervals will not be converted).

For urea, reporting in tables will be done only as blood urea nitrogen (BUN). In case sites have provided only a urea value and no BUN value, conversion from Urea to BUN will be performed according to the following formula:

- BUN [mmol/L] = urea [mmol/l]/2.14

9.2 Analysis of Adverse Events

Adverse event intensity and laboratory evaluation changes will be assessed using National Cancer Institute (NCI) Common Terminology Criteria for Adverse Events (CTCAE) Version 4.0.

9.2.1 Treatment-Emergent Adverse Events (TEAEs)

Analyses of adverse events will include only "treatment-emergent" events, i.e., those that have an onset on or after the day of the first dose of study drug or increases severity after the first dose of study drug. Analyses will not include those that have an onset greater than 30 days after last dose of study drug. Events where the onset date was the same as the study drug start date will be assumed to be treatment-emergent, unless the study drug start time and the AE start time are collected and the AE start time is prior to the study drug start time. If an incomplete onset date is collected for an AE, the AE will be assumed to be treatment-emergent unless there is evidence that confirms that the AE was not treatment-emergent (e.g., the AE end date is prior to the date of the first dose of study drug).

Treatment-emergent adverse events will be summarized by preferred term (PT) within a system organ class (SOC) according to the MedDRA dictionary. Subjects reporting more than one AE within a SOC will be counted only once for that SOC. Subjects reporting more than one AE for a PT will be counted only once for that PT. For summaries of TEAEs by the National Cancer Institute Common Terminology Criteria for Adverse Events (NCI CTCAE) grade, at each level of summation (overall, SOC, and PT) each subject is counted only once at the maximum grade level. The possible grade levels are 1 – 5, and unknown. If a subject had an AE with unknown grade, then the subject will be

counted in the grade category of "unknown," even if the subject had another occurrence of the same event with a grade present. The only exception is if the subject had another occurrence of the same AE with a grade of 3, 4 or 5. In this case, the subject will be counted under the grade 3, 4 or 5 category based on the maximum grade level reported.

The number and percentage of subjects experiencing at least one TEAE will be further summarized by assigned venetoclax dose and overall as follows:

- Any TEAE
- Any TEAE occurring in at least 5% of subjects in any venetoclax dose level
- Any TEAE with NCI CTCAE toxicity grade 3, 4 or 5
- Any TEAE with NCI CTCAE toxicity grade 3 or 4
- Any TEAE broken down by maximum NCI CTCAE toxicity grade (severity)
- Any TEAE with reasonable possibility of related to venetoclax by the investigator
- Any TEAE with reasonable possibility of related to dexamethasone by the investigator
- Any TEAE with NCI CTCAE toxicity grade 3, 4, or 5 and reasonable possibility of related to venetoclax by the investigator
- Any TEAE with NCI CTCAE toxicity grade 3, 4, or 5 and reasonable possibility of related to dexamethasone by the investigator
- Any TEAE leading to discontinuation of venetoclax
- Any TEAE leading to discontinuation of dexamethasone
- Any TEAE leading to discontinuation of venetoclax not related to progression
- Any TEAE leading to discontinuation of dexamethasone not related to progression
- Any TEAE leading to Venetoclax dose interruption
- Any TEAE leading to dexamethasone dose interruption
- Any TEAE leading to Venetoclax dose reduction
- Any TEAE leading to dexamethasone dose reduction
- Any TEAE leading to Dose limiting toxicities

Except for the discontinuation, interruption, reduction, and DLT AE summaries, the above tables will also be summarized by SOC and PT in descending order.

Table 3. Search Criteria for Selected Adverse Events

Risk	Search Criteria
Tumor Lysis Syndrome	SMQ – "Tumour lysis syndrome" (Narrow-scope)
Neutropenia	PT terms – "Neutropenia," "Neutrophil count decreased," "Febrile neutropenia," "Agranulocytosis," "Neutropenic infection," and "Neutropenic sepsis"
Serious Infection	SAEs in the SOC of "Infections and Infestations"
Second Primary Malignancy	SMQ – "Malignant tumours" (Narrow) and "Myelodysplastic syndromes" (Narrow)
Lymphopenia	PT terms – "Lymphopenia" and "Lymphocyte count decreased"
Anemia	PT terms – "Anaemia" and "Haemoglobin decreased"
Thrombocytopenia	PT terms – "Thrombocytopenia" and "Platelet count decreased"
Drug Induced Liver Injury (DILI)	SMQ – "Drug related hepatic disorders – comprehensive search"
Peripheral Neuropathy	Peripheral Neuropathy SMQ – "narrow"
Sepsis	PTs including the word 'sepsis' + PT 'septic shock'
Serious Opportunistic infections	SAEs in the Opportunistic infection CMQ (BCL-2 inhibitor product specific)

An AE overview and a summary by PT will be provided for each selected AE search in the table above by assigned venetoclax dose and overall. In addition, a listing of events in each AE search will be provided.

9.2.2 Serious Adverse Events

Serious adverse events will be summarized by PT within an SOC according to the MedDRA dictionary. The number and percentage of subjects experiencing at least one TEAE will be further summarized by assigned Venetoclax dose and overall as follows:

- Any treatment-emergent serious adverse event (SAE).
- Any SAE with reasonable possibility of related to venetoclax by the investigator.

- Any SAE with reasonable possibility of related to dexamethasone by the investigator.

9.2.3 Deaths

The number and percentage of subject deaths will be summarized (1) for all deaths in this study regardless of the number of days after the last dose of study drug, and (2) for deaths occurring within 30 days of the last dose of study drug. Listings of subject deaths will be provided (1) for deaths occurring within 30 days of the last dose of study drug, and (2) for deaths occurring more than 30 days after the last dose of last study drug.

9.2.4 Non-Treatment-Emergent Adverse Events

The serious and non-serious adverse events occurring after the study-specific informed consent is signed but prior to the first dose of the investigational product considered to be causally related to study required procedures will be summarized for all dosed Phase 1 subjects.

9.3 Analysis of Laboratory and Vital Signs

9.3.1 Longitudinal Analyses of Laboratory and Vital Signs Data

Changes from baseline will be analyzed for each scheduled post-baseline visit and time point as well as for the final visit for blood chemistry and hematology parameters specified in Table 6. If more than one measurement exists for a subject on a particular day and time point, then an arithmetic average will be calculated. This average will be considered to be that subject's measurement for that day and time point. Post-baseline measurements more than 30 days after last study drug will not be included. Subjects that do not have a baseline measurement or do not have any post-baseline measurements will not be included.

If there are lab duplicates identified (same subject, visit, date, and time of collection) between the central (Covance) and the local lab (EDC-RAVE), we will utilize the central lab over the local lab (EDC-RAVE data) as opposed to taking the average.

Mean changes from baseline at each scheduled post-baseline visit and time point as well as at Final Visit will be summarized with the baseline mean, visit mean, change from baseline mean, standard deviation, and median.

Laboratory tests to be summarized are included in Table 6.

9.3.2 Analyses of Laboratory Data Using NCI CTCAE

The maximum NCI CTCAE toxicity grade value will be the value with the highest NCI CTCAE toxicity grade collected after the first dose of study drug and within 30 days following last dose of study drug. In cases where multiple values are collected on the same day, the maximum grade value will be selected as the value for that day.

Detailed listings of data for subjects experiencing NCI CTCAE grade 3 to 4 blood chemistry and hematology values will be provided. All measurements collected, regardless of the number of days after last dose of study drug will be included in these listings.

For shifts related to National Cancer Institute Common Toxicity Criteria (NCI CTC), baseline and post-baseline laboratory observations will be categorized as grade 0, grade 1, grade 2, grade 3, or grade 4, or given no grade according to NCI CTC grade version 4.0. For each parameter, shift tables will be generated that cross tabulate the number of subject with baseline values of no grade, grade 1, grade 2, grade 3, or grade 4 versus final observations of no grade, grade 1, grade 2, grade 3, or grade 4. Additionally, for each parameter, the number and percentage of subjects that have a baseline observation with no grade or a baseline observation that is categorized as a grade 0, grade 1, grade 2, and that also have a grade 3 or 4 final observation will be presented. For a laboratory parameter for which grading is performed separately for low and high values, a separate summary will be presented for each grading scale. In addition to final observation, a similar set of summaries will be produced for the minimum or maximum post-baseline laboratory observations, as appropriate for the grading scale.

9.3.3 Laboratory Search Strategy for TLS

To determine if a subject's laboratory values qualify for TLS, the Howard criteria will be assessed. The Howard criteria for TLS comprises ≥ 2 of the electrolyte abnormalities within 24 hours of each other and are specified in Table 7. The number and percentage of subject having laboratory values qualify for TLS will be summarized. A listing of such subjects will be generated.

9.3.4 Assessment of Drug-Induced Liver Injury

The number and percentage of subjects who have at least one observed post-baseline value meeting the following criteria will be tabulated by dose and overall: ALT > 3 × ULN, AST > 3 × ULN, AST or ALT > 3 × ULN, total bilirubin value > 2 × ULN, ALT or AST > 3 × ULN and total bilirubin > 2 × ULN within 24 hours. A listing of all observed ALT, AST, and total bilirubin values will be generated for subjects with an observed value meeting any of these criteria.

9.4 Analysis of Vital Signs

9.4.1 Potentially Clinically Significant Values

For selected vital signs variables (systolic blood pressure, diastolic blood pressure, heart rate, and temperature), a listing of all observations collected will be generated for subjects that had at least one post-baseline observation meeting pre-defined criteria for potentially clinically significant values. The number and percentage of subjects who have at least one post-baseline observation meeting the pre-defined criteria for potentially clinically significant values will be provided for each variable by dose and overall.

Pre-defined criteria for potentially clinically significant vital signs values are given in Table 8.

9.4.2 Mean Change from Baseline

Changes from baseline will be analyzed for each scheduled post-baseline visit as well as for the final visit for vital signs variables (systolic blood pressure, diastolic blood

pressure, and heart rate). If more than one measurement exists for a subject on a particular day, then an arithmetic average will be calculated. This average will be considered to be that subject's measurement for that day. Post-baseline measurements more than 30 days after last study drug will not be included. Subjects that do not have a baseline measurement or do not have any post-baseline measurements will not be included.

Mean changes from baseline at each scheduled post-baseline visit as well as at Final Visit will be summarized with the baseline mean, visit mean, change from baseline mean, standard deviation, and median.

9.5 Analysis of ECG Parameters

A listing will be prepared that will include details about all abnormal ECG findings.

10.0 Efficacy Analysis

10.1 Response Rates

The distribution of best overall IMWG response (per investigator assessment) will be reported as the number and percentage of subjects who achieved each response as their "best" response. In other words, each subject will only be counted in one of the following categories:

- Stringent complete response (sCR)
- Complete response (CR)
- Very good partial response (VGPR)
- Partial Response (PR)
- Minimal response (MR)
- Stable disease (SD)
- Disease Progression (PD)
- Missing

All responses must be confirmed by a consecutive assessment in which that subject achieved as good or better of a response, with the exception of SD, which is confirmed by the baseline. Section 4.3.2.1 details how PD is confirmed.

In addition to the aforementioned distributional breakdown of best overall response, the number, percentage, and 95% confidence interval (exact binomial distribution; Clopper-Pearson methodology) will be reported for the following:

- ORR (PR or better: PR + VGPR + CR + sCR)
- ≥ VGPR rate (VGPR or better: VGPR + CR + sCR)

10.2 Time to Event Endpoints

Analysis of the time to event endpoints (TTP and DOR) will include the following:

- Number and % of subjects experiencing an event
- Number and % of censored subjects
- Point estimates and 95% confidence intervals (log-log; the SAS default) for survival time (months) for the 25th, 50th (median), and 75th percentiles, and for the survival estimates at months 6, 12, and 24 (% survivors remaining)
- Kaplan-Meier plot w/curves and risk table

Analysis of TTR (PR or better) will include the mean and 95% CI, standard deviation, median, and range (i.e., minimum and maximum).

11.0 Pharmacokinetic Analysis

11.1 Tabulations and Summary Statistics

Plasma concentrations and PK parameter values of venetoclax will be tabulated for each subject and each dose level, and summary statistics will be computed for each sampling time and parameter. Significant pharmacokinetic sample time deviations will be identified and listed.

11.2 Model and Test

Dose Proportionality and Covariate Selection of Venetoclax on Cycle 2 Day 1

Pharmacokinetic parameters of venetoclax from the dose escalation subjects assessed on Cycle 2 Day 1 will be analyzed as follows. An analysis will be performed for T_{max}, dose normalized C_{max}, dose-normalized AUC_{24}, and β, provided that they can be adequately determined from the data. The model used for the statistical analyses will include dose level of venetoclax as a fixed effect. This may be done by classifying subjects by dose level or, if appropriate, using dose level as a continuous variable. Baseline covariates such as age, body weight, body surface area, calculated creatinine clearance, and gender will be included in the model initially. A covariate will be dropped from the model if the regression coefficient is not significant at alpha level 0.10. The natural logarithmic transformation will be employed for C_{max} and the AUCs unless the data clearly indicate that another transformation or the untransformed variable provides more nearly symmetric probability distributions and/or more nearly homogenous variances across dose levels. Within the framework of the model, tests that have good power for a trend with dose will be performed on the effect of dose level.

Formula for calculated creatinine clearance is defined as follows:

$$eC_{Cr} = \frac{(140 - \text{baseline Age}) \bullet \text{baseline weight (kg)} \bullet [1.23 \text{ if Male, } 1.04 \text{ if Female}]}{\text{baseline Serum Creatinine } (\mu\text{mol/L})}$$

Steady State Trough Hour 0 Concentrations

Exploratory analyses of trough (pre-dose) concentrations of venetoclax on Day 1 of Cycle 2, 3, 5, 7 and 9 for Phase 1 portion subjects will be done to characterize the achievement of steady-state pharmacokinetics. Cycle will be included in the model as a categorical variable (to generate pairwise comparisons of various time points). The term for dose will also be included in the model, and the potential interaction between cycle and dose will be evaluated.

11.3 Missing Values and Model Violations

The possibility of bias from missing data of subjects who prematurely discontinue due to an adverse event will be addressed. Normally values of PK variables (C_{max}, AUC, etc.) will be determined without replacing missing individual concentration values, simply using the available data, and if necessary doing the analysis with some missing values for a PK variable. However, missing concentration values for isolated individual blood samples may be replaced (imputed) if such might affect study conclusions or meaningfully affect point estimates.

12.0 References

1. Durie BG, Harousseau JL, Miguel JS, et al. International uniform response criteria for multiple myeloma. Leukemia. 2006;20(9):1467-73.
2. Durie BG, Harousseau JL, Miguel JS, et al. International uniform response criteria for multiple myeloma. Leukemia. 2006;20(9):1467-73. Erratum in: Leukemia. 2007;21(5):1134.
3. Rajkumar SV, Harousseau JL, Durie B, et al. Consensus recommendations for the uniform reporting of clinical trials: report of the International Myeloma Workshop Consensus Panel 1. Blood. 2011;117(18):4691-5.

13.0 Appendix

13.1 IMWG 2011[1-3] Response Criteria

Response Subcategory	Response Criteria[a]
Stringent complete response (sCR)*	• Negative immunofixation on the serum and urine (regardless of whether disease at baseline was measurable on serum, urine, both, or neither) and • Disappearance of any soft tissue plasmacytomas and • < 5% plasma cells in bone marrow[b] and • Normal FLC (free light chain) ratio** and • Absence of clonal plasma cells in bone marrow[b] by immunohistochemistry or immunofluorescence[c] or 2 to 4 color flow cytometry
Complete response (CR)*	• Negative immunofixation on the serum and urine (regardless of whether disease at baseline was measurable on serum, urine, both, or neither) and • Disappearance of any soft tissue plasmacytomas and • < 5% plasma cells in bone marrow[b] and • For subjects in whom the only measurable disease is by serum FLC levels, a normal FLC ratio** is also required
Very good partial response (VGPR)*	• Serum and urine M-component detectable by immunofixation but not on electrophoresis or • ≥ 90% reduction in serum M-component plus urine M-component < 100 mg per 24 hr For subjects in whom the only measurable disease is by serum FLC levels, VGPR is defined as: • ≥ 90% decrease in the difference between involved and uninvolved FLC levels
Partial response (PR)	• ≥ 50% reduction of serum M-protein and • ≥ 90% reduction in 24-hr urinary M-protein or to < 200 mg per 24 hr ○ If the serum and urine M-protein are unmeasurable[d], a ≥ 50% decrease in the difference between involved and uninvolved FLC levels is required in place of the M-protein criteria • ≥ 50% reduction in the size of soft tissue plasmacytomas, if present at baseline

Response Subcategory	Response Criteria[a]
Minimal response (MR)	• 25% – 49% reduction of serum M-protein, and • 50% – 89% reduction in 24-hour urinary M-protein, and • 25% – 49% reduction in size of soft tissue plasmacytomas, if present at baseline, and • No increase in size or number of lytic bone lesions (development of compression fracture dose not exclude response)
Stable disease (SD)[e]	Not meeting criteria for sCR, CR, VGPR, PR, MR or progressive disease (PD)

a. All response categories require two consecutive assessments made at any time before the institution of any new therapy; all categories also require no known evidence of progressive or new bone lesions if radiographic studies were performed. Radiographic studies are not required to satisfy these response requirements.

b. Confirmation with repeat bone marrow biopsy not needed.

c. Presence/absence of clonal cells is based upon the k/λ ratio. An abnormal k/λ ratio by immunohistochemistry and/or immunofluorescence requires a minimum of 100 plasma cells for analysis. An abnormal ratio reflecting presence of an abnormal clone is k/λ of > 4:1 or < 1:2.

d. Measurable disease is defined as meeting at least one of the following four measurements: serum M-protein ≥ 1 g/dL (≥ 10 g/L) or urine M-protein ≥ 200 mg/24 hr or serum FLC assay with an involved FLC level ≥ 10 mg/dL (≥ 100 mg/L) provided serum FLC ration is abnormal or bone marrow plasma cells ≥ 30%.

e. Not recommended for use as an indicator of response; stability of disease is best described by providing the time to progression estimates.

* Serum and urine M-protein testing is required to fulfill requirements of VGPR and CR categories regardless of whether disease at baseline was measurable on serum, urine, both, or neither.

** Clarification to IMWG criteria for coding CR and VGPR in patients in whom the only measurable disease is by serum FLC levels: CR in such patients is defined as a normal FLC ratio of 0.26 – 1.65 in addition to CR criteria listed above. VGPR in such patients is defined as a > 90% decrease in the difference between involved and uninvolved free light chain FLC levels.

Relapse Subcategory	Relapse Criteria
Progressive disease (PD)*	Requires any one or more of the following: • Increase of ≥ 25% from lowest response level in any of the following: o Serum M-protein[a] absolute increase ≥ 0.5 g/dL, and/or o Urine M-protein absolute increase ≥ 200 mg/24 hr, and/or o In subjects without measurable serum and urine M-protein levels, the difference between involved and uninvolved FLC levels (the absolute increase must be > 10 mg/dL), and/or • Definite development of new bone lesions or soft tissue plasmacytomas, and/or • Definite increase in the size of existing bone lesions or soft tissue plasmacytomas, and/or • Development of hypercalcemia (corrected serum calcium > 11.5 mg/dL) that can be attributed solely to the plasma cell proliferative disorder
Clinical Relapse	Requires any one or more of the following direct indicators of increasing disease and/or end organ dysfunction (CRAB features). • Development of new soft tissue plasmacytomas or bone lesions on skeletal survey, MRI or other imaging, and/or • Definite increase in size of existing plasmacytomas or bone lesions. A definite increase is defined as a 50% (and at least 1 cm) increase as measured serially by the sum of the products of the cross-diameters of the measurable lesion, and/or • Hypercalcemia > 11.5 mg/dL (> 2.875 mmol/L), and/or • Decrease in hemoglobin of ≥ 2 g/dL (≥ 1.25 mmol/L) or < 10 g/dL, and/or • Rise in serum creatinine by > 2 mg/dL (≥ 177 mmol/L), and/or • Hyperviscosity

a. For progressive disease, serum M-component increases of ≥ 1 g/dL are sufficient to define relapse if starting M-component is ≥ 5 g/dL.

* The "25% increase" refers to M-protein, FLC, and bone marrow results and does not refer to bone lesions, soft tissue plasmacytomas, or hypercalcemia. The "lowest response value" does not need to be a confirmed value.

Table 4. Time Windows for Analysis of Laboratory and Vital Signs Parameters – Phase 1 Monotherapy Subjects

Scheduled Visit	Nominal Day	Time Window (Study Rx Day Range)
Screening	Screening	Screening Visit
Baseline	≤ 1	See the baseline definition (Section 6.2)
Lead-in Day 1	1	[1]
Lead-in Day 2	2	[2, 4]
Lead-in Day 8	8	[5, 11]
Cycle 1 Day 1	15	[12, 15]
Cycle 1 Day 2	16	[16]
Cycle 1 Day 3	17	[17, 19]
Cycle 1 Day 8	22	[20, 25]
Cycle 1 Day 15	29	[26, 32]
Cycle 2 Day 1	36	[33, 36]
Cycle 2 Day 2	37	[37, 39]
Cycle 2 Day 8	43	[40, 46]
Cycle 2 Day 15	50	[47, 53]
Cycle 3 Day 1	57	[54, 60]
Cycle 4 Day 1	78	[61, 95]
Cycle 5 Day 1	99	[96, 130]
Cycle 7 Day 1	141	[131, 174]
Cycle 9 Day 1	183	[175, 224]
Cycle 13 Day 1	267	[225, 309]
Cycle 17 Day 1	351	[310, 394]
Every 4th Cycle thereafter	Rx day of Cycle X Day 1	[Rx day of Cycle (X) Day 1 – 42 to Rx day of Cycle (X) Day 1 + 42]
Final Observation		Last value within 30 days of study drug discontinuation

Table 5. Time Windows for Analysis of Laboratory and Vital Signs Parameters – Phase 1 VenDex Combination Subjects

Scheduled Visit	Nominal Day	Time Window (Study Rx Day Range)
Screening	Screening	Screening Visit
Baseline	≤ 1	See the baseline definition (Section 6.1)
Cycle 1 Day 1	1	[1]
Cycle 1 Day 2	2	[2]
Cycle 1 Day 3	3	[3]
Cycle 1 Day 8	8	[4, 11]
Cycle 1 Day 15	15	[12, 18]
Cycle 2 Day 1	22	[19, 22]
Cycle 2 Day 2	23	[23, 26]
Cycle 2 Day 8	29	[27, 31]
Cycle 2 Day 15	36	[32, 40]
Cycle 3 Day 1	43	[41, 53]
Cycle 4 Day 1	64	[54, 75]
Cycle 5 Day 1	85	[76, 97]
Cycle 7 Day 1	127	[98, 140]
Cycle 9 Day 1	169	[141, 183]
Cycle 13 Day 1	211	[184, 226]
Cycle 17 Day 1	253	[227, 269]
Every 4th cycle thereafter	Rx day of Cycle X Day 1	[Rx day of Cycle (X) Day 1 – 42 to Rx day of Cycle (X) Day 1 + 42]
Final Observation		Last value within 30 days of study drug discontinuation

Table 6. Clinical Laboratory Tests – Phase 1 Portion

Hematology	Clinical Chemistry
Hemoglobin White Blood Cell (WBC) count Platelet count Neutrophils	Creatinine Total bilirubin Serum glutamic-oxaloacetic transaminase (SGOT/AST) Serum glutamic-pyruvic transaminase (SGPT/ALT) Potassium Calcium Inorganic phosphate Uric acid

a. Chemistry tests should be obtained under fasting conditions, if possible.

Table 7. Definitions of Laboratory Tumor Lysis Syndrome (TLS)

Metabolic Abnormality	Howard Criteria for Classification of Laboratory Tumor Lysis Syndrome (TLS)
Hyperruricemia	Uric acid > 8.0 mg/dL (475.8 μmol/L)
Hyperphosphatemia	Phosphorus > 4.5 mg/dL (1.5 mmol/L)
Hyperkalemia	Potassium > 6.0 mmol/lL
Hypocalcemia	Corrected calcium < 7.0 mg/dL (1.75 mmol/L) or ionized calcium < 1.12 mg/dL (0.3 mmol/L)[a]

a. The corrected calcium level in mg/dL = measured calcium level in mg/dL + 0.8 × (4-albumin in g/dL).

Note: In laboratory tumor lysis syndrome, 2 or more metabolic abnormalities must be present during the same 24-hour period within 3 days before the start of therapy or up to 7 afterward. Clinical tumor lysis syndrome requires the presence of laboratory tumor lysis syndrome plus an increased creatinine level, seizures, cardiac dysrhythmia, or death.

Source: Howard SC, Jones DP, Pui CH. The tumor lysis syndrome. N Engl J Med. 2011;364(19):1844-54.

Table 8. Criteria for Potentially Clinically Significant Vital Signs Parameters

Vital Signs Variables	Criterion	Definition of Potentially Clinically Significant
Systolic Blood Pressure	High	Value ≥ 160 mmHg
Diastolic Blood Pressure	High	Value ≥ 100 mmHg
Heart Rate	Low	Value < 50 bpm
	High	Value ≥ 120 bpm
Temperature	Low	Value < 36°C
	High	Value ≥ 38.5°C