pharmacyclics®
An AbbVie Company

Statistical Analysis Plan

A Multi-Center Open-Label Study of the Bruton's Tyrosine Kinase (BTK) Inhibitor, Ibrutinib, in Combination with MEDI4736, in Subjects with Relapsed or Refractory Lymphomas

Protocol PCYC-1136-CA; Phase 1b/2

Prepared by: Shanhong Guan

10 January 2018

Version 1.0

Confidentiality Statement

TABLE OF CONTENTS

LIST OF TABLES

LIST OF ABBREVIATIONS

ABC	Activated B-cell
AE(s)	adverse event(s)
BTK	Bruton's tyrosine kinase
CNS	central nervous system
CR	complete remission (response)
CrCl	creatinine clearance
CRF	case report form
CT	computed tomography
CTCAE	Common Terminology Criteria for Adverse Events
CYP	cytochrome P450
DLBCL	diffuse large B-cell lymphoma
DLT	dose-limiting toxicity
DOR	duration of response
ECG	Electrocardiogram
ECOG	Eastern Cooperative Oncology Group
FDA	Food and Drug Administration
FL	follicular lymphoma
GCB	germinal center B cell-like
GCP	Good Clinical Practice
GEP	gene expression profile
HDT/ASCT	high-dose chemotherapy with autologous stem cell support
HIV	human immunodeficiency virus
IB	Investigator's Brochure
ICF	informed consent form
ICH	International Conference on Harmonisation
Ig	Immunoglobulin
IHC	immunohistochemisty
IND	Investigational New Drug
INR	international normalized ratio
IV	Intravenous
LLN	low limit of normal
MAD	maximum administered dose
MedDRA	Medical Dictionary for Regulatory Activities
MEDI4736	durvalumab
MRD	minimal residual disease
MRI	magnetic resonance imaging

MTD	maximum tolerated dose
NHL	non-Hodgkin's lymphoma
Non-GCB	Non-Germinal Center B-cell like
ORR	objective response rate
OS	overall survival
PD	progressive disease
PFS	Progression-free survival
PK	pharmacokinetics
PO	per os (oral)
PR	partial remission (response)
RP2D	recommended phase 2 dose
PT	prothrombin time
QTc	corrected QT interval
SAE(s)	serious adverse event(s)
SAP	statistical analysis plan
SD	stable disease
SI	standard international units
SOC	system organ class
ULN	upper limit of normal

1. INTRODUCTION

This statistical analysis plan (SAP) lays out key elements including definitions and statistical methods for analysis of data in evaluation of efficacy and safety for the PCYC-1136-CA study. Analyses of pharmacokinetics data and biomarkers will be addressed in separate documents.

1.1. Stud y Design

This is a Phase 1b/2, open-label multi-center study to assess the safety and efficacy of ibrutinib in combination with MEDI4736 in subjects with relapsed or refractory lymphomas. A 6+3 de-escalation design will be employed in Phase 1b to assess doses of ibrutinib in combination with MEDI4736 to determine the recommended Phase 2 dose (RP2D) for this study.

Phase 1b:
In the safety portion of the study, a starting dose of 560 mg of ibrutinib and 10 mg/kg of MEDI4736 will be explored in Cohort 1 and will follow a 6+3 dose de-escalation design including a sentinel subject with at least a 3-day observation period in each Phase 1b dosing cohort prior to dosing additional subjects to monitor for any unexpected acute or overlapping toxicities.

Subjects with one of the following two tumor types will be eligible for enrollment in Phase 1b:

- Follicular lymphoma (FL) (grade 1, 2, or 3A)
- Diffuse large B cell lymphoma (DLBCL)

In cohort 1, ibrutinib will be administered orally (PO) daily in 6 subjects at a dose of 560 mg in combination with MEDI4736 at a dose of 10 mg/kg intravenously (IV) every 2 weeks in 28-day cycles for 12 months or until unacceptable toxicity or disease progression occurs.

As a lead-in to the first cycle, ibrutinib will be given for 7 days prior to starting MEDI4736 in order to obtain single agent PK and biomarker samples that will be compared to combination therapy. Per the protocol, the 7-day lead-in period of ibrutinib is not part of the dose-limiting toxicity (DLT) evaluation period. The DLT assessment will occur from Day 1 of Cycle 1 of the combination therapy until the laboratory evaluations on Day 1 of Cycle 2 (28 days). If the subject incidence of DLTs from Day 1 of Cycle 1 of combination therapy to the laboratory evaluations on Day 1 of Cycle 2 is <33.3% (ie, ≤1 of 6 or ≤2 of 9 subjects with a DLT), this dose level will be considered safe to proceed to Phase 2, and will be defined as the RP2D. If 2 subjects within the initial cohort of 6 subjects experience a DLT, an additional 3 subjects will be enrolled at the same dose level. If 3 or more of 6-9 subjects experience a DLT, dose de-escalation will occur. A similar 6+3 cohort design will be utilized in the de-escalation cohorts. De-escalation cohorts -1A and -1B will be opened simultaneously to determine which dosing schedule is most appropriate for the Phase 2 portion of the study (for cohort definitions, see

section 4.2: Study Treatment). Determination will be based on the safety profile of the 2 treatment regimens. If necessary, the dose de-escalation cohort -2 will be enrolled thereafter.

The Phase 1b DLT observation period will occur from Day 1 of Cycle 1 of ibrutinib in combination with MEDI4736 treatment until the laboratory evaluations on Day 1 of Cycle 2 which will occur before the MEDI4736 infusion on Day 1 of Cycle 2. A DLT is an AE at least possibly related to study drug(s) defined as:

Any Grade 3 or higher non-hematologic AE (with additional clarifications below)

- Grade 3 or higher nausea, vomiting, or diarrhea (non-colitis-related) that is uncontrolled by maximal supportive care and persists for >7 days would be considered a DLT
- Grade 3 fatigue that persists for >7 days would be considered a DLT
- Any Grade 3-immune related adverse event (irAE) that does not downgrade to Grade ≤2 within 3 days after onset of the event despite maximal supportive care including systemic corticosteroids, or does not downgrade to Grade ≤1 or baseline within 14 days would be considered a DLT
- Any Grade ≥2 pneumonitis that does not resolve to Grade ≤1 within 3 days of the initiation of maximal supportive care would be considered a DLT.
- The DLT definition excludes the following conditions:
  - Grade 3 endocrinopathy that is managed with or without systemic corticosteroid therapy and/or hormone replacement therapy and the subject is asymptomatic
  - Grade 3 inflammatory reaction attributed to a local antitumor response (eg, inflammatory reaction at sites of metastatic disease, lymph nodes, etc.)
  - Vitiligo or alopecia of any AE grade would not be considered a DLT

Clinically relevant hematologic AE such as:

- Grade 3 or 4 neutropenia (ANC<1,000/ mm^3) lasting >7 days
- Febrile neutropenia (Grade 3 or 4 neutropenia with fever ≥38.3ºC)
- Grade 3 or 4 thrombocytopenia (<25,000/ mm^3) lasting >7 days or requiring transfusion or associated with Grade ≥2 bleeding
- Any other Grade 4 hematologic AE, except lymphopenia

During the Phase 1b DLT observation period of ibrutinib in combination with MEDI4736, subjects will be replaced for any of the following reasons:

- Missed ≥5 consecutive doses of ibrutinib or missed 7 total doses (not due to DLT)
- Missed dose of MEDI4736 or received the MEDI4736 dose outside of the protocol-allowed window (±2 days) (not due to DLT)
- Study drug discontinuation for any reason other than DLT

After the RP2D of ibrutinib in combination with MEDI4736 is defined, enrollment in Phase 2 will commence.

Phase 2:

Two tumor types will be enrolled in separate cohorts as part of this protocol:

- Follicular lymphoma (FL) (grade 1, 2, or 3A)
- Diffuse large B cell lymphoma (DLBCL)

For FL: n = 39
The FL cohort will enroll 39 response-evaluable subjects. An interim analysis will be performed after 19 subjects are evaluable for tumor response after completing at least 1 response assessment. If only 6 or fewer responders (≤6/19) are observed, the Sponsor may consider discontinuation of this cohort; however, assessment of biomarkers (that may aid in prospective enrichment of responders) or tumor measurements (showing clinically relevant tumor reductions, ie, 40%, that fit the criteria for SD) may support continued enrollment.

For DLBCL: n = 34
Subjects with germinal center B-cell (GCB) and non-GCB subtypes will be enrolled at a 1:1 ratio approximately. The DLBCL cohort will enroll 34 response-evaluable subjects. An interim analysis will be performed after 9 subjects are evaluable for tumor response after completing at least 1 response assessment. If only 1 or fewer responder (≤1/9) is observed, the Sponsor may consider discontinuation of this cohort; however, assessment of biomarkers (that may aid in prospective enrichment of responders) or tumor measurements (showing clinically relevant tumor reductions, ie, 40%, that fit the criteria for SD) may support continued enrollment.

In the DLBCL cohort, additional prospectively identified non-GCB subjects may be enrolled if a 22% ORR (2/9) is not achieved at the interim analysis in the unselected DLBCL population so that at least a total of 16 subjects with the non-GCB subtype are evaluated.

Note: The Phase 2 subjects include all of those response-evaluable subjects treated at R2PD in Phase 1b but with tumor type as defined for Phase 2.

During the Phase 2 portion of the study, if a subject discontinues study drug prior to the first tumor response assessment for reasons other than progressive disease, the subject may be replaced.

The primary analysis will occur at least 12 months after the last FL subject is enrolled and at least 9 months after the last DLBCL subject is enrolled.

Dosing of ibrutinib in combination with MEDI4736 will continue for a total of 12 cycles of therapy if the subject is deriving clinical benefit (CR, PR, or SD) and the subject is not experiencing unacceptable toxicity. After 12 cycles, dosing of ibrutinib alone will continue as

long as clinical benefit is seen without safety concerns or disease progression. Response evaluations will occur on Day 1 of Cycle 3 and then after every 3rd cycle through the 1st year and after every 6th cycle thereafter. Assessment will follow the Revised Criteria for Response Assessment of Malignant Lymphoma (Cheson 2014) using CT criteria.

In order to accommodate the potential for immune flare (pseudoprogression), treatment with MEDI4736 and ibrutinib may continue between the initial assessment of suspected progression and confirmation of progression. Subjects with presumed progressive disease (PD) who, in the Investigator's opinion, continue to receive clinical benefit from their treatment may continue to receive MEDI4736 and ibrutinib as dictated in the protocol after consultation with the Sponsor and at the Investigator's discretion. In the absence of clinically significant deterioration, a CT scan or biopsy may be performed at least 4 weeks later to confirm PD. If PD is confirmed at the later time point, PD should be assigned to the prior time point at which PD criteria were also met. MEDI4736 and ibrutinib should be discontinued if there is confirmed PD per Revised Criteria for Response Assessment of Malignant Lymphoma (Cheson 2014) or other clinical data suggest clear evidence of progression.

Subjects are required to undergo a pre-treatment tumor biopsy for all phases of this study; a post-progression biopsy is recommended but will be optional.

1.2. Endpoints

Primary Endpoints

Phase 1b:

- Recommended Phase 2 Dose (RP2D) of ibrutinib in combination with MEDI4736
- Safety and tolerability of ibrutinib in combination with MEDI4736

Phase 2:

- Overall Response Rate (ORR)

Secondary Endpoints

Phase 1b:

- Overall response rate (ORR)
- Duration of response (DOR)

Phase 2:

- DOR
- Progression-free survival (PFS)
- Overall survival (OS)

Safety Assessments

Safety and tolerability assessments include adverse events, clinical laboratory tests, and vital signs.

1.3. Statistical Hypotheses

There is no hypothesis testing for Phase1b. The primary objective of Phase 2 is to evaluate overall response rate (ORR) for the combination therapy at RP2D. Hypothesis testing will be conducted to address this objective as follows:

Null hypothesis (H_0): P <= x%; where P denotes overall response rate.
Alternative hypothesis (H_a): P >= y%; where P denotes overall response rate.

For FL cohort, a true response rate of 30% (H_0) versus 50% (H_a) will be tested. For DLBCL cohort, a true response rate of 15% (H_0) versus 35% (H_a) will be tested.

1.4. Sample Size Determination

Phase 1b:

6-36 subjects for DLT assessment (6-9 subjects in up to 4 dose de-escalation cohorts, regardless of tumor type).

Phase 2:

Approximately 73 response-evaluable subjects (including those subjects treated at the RP2D in Phase 1b but with tumor type as defined for Phase 2).

For FL:
A maximum of 39 response-evaluable subjects (including those subjects treated at the RP2D in Phase 1b and with tumor type of FL) will be enrolled. An interim analysis will be performed after 19 subjects are evaluable for tumor response after completing at least 1 response assessment. If only 6 or fewer responders (≤6/19) are observed, the Sponsor may consider discontinuation of this cohort.

The null hypothesis that the true ORR is 30% will be tested against a one-sided alternative that the ORR is 50%. The null hypothesis will be rejected if 17 or more responses are observed in the 39 subjects. This design yields a 1-sided type I error rate of 0.05 and power of 80% when the true ORR is 50%. This statistical design including number of subjects and number of responders follows the statistical framework of Simon's minimax two-stage design (Simon, 1989). Enrollment will continue while the interim analysis is performed.

For DLBCL:
A maximum of 34 response-evaluable subjects with GCB and non-GCB (inclusive of subjects treated at the RP2D in Phase 1b in the unselected DLBCL population) will be enrolled at a 1:1 ratio approximately.
An interim analysis will be performed after 9 subjects are evaluable for tumor response after completing at least 1 response assessment. If only 1 or fewer responder (≤1/9) is observed, the Sponsor may consider discontinuation of this cohort.

The null hypothesis that the true ORR is 15% will be tested against a one-sided alternative that the ORR is 35%. The null hypothesis will be rejected if 9 or more responses are observed in the 34 subjects. This design yields a 1-sided type I error rate of 0.05 and power of 80% when the true ORR is 35%. This statistical design including number of subjects and number of responders follows the statistical framework of Simon's optimal two-stage design (Simon 1989). Enrollment will continue while the interim analysis is performed.

1.5. Planned Analyses

1.5.1. Phase 1b Analyses

A Dose Level Review Committee will be established to evaluate the safety data from each cohort of the Phase 1b on an ongoing basis. Dose de-escalation will follow the 6+3 design described in Section *3.1.1* and up to 4 cohorts including 3 dose de-escalation cohorts (1, -1A, -1B, and -2) will be enrolled if needed to determine the RP2D for the combination therapy.

1.5.2. Phase 2 Analyses

An interim analysis is planned for each tumor type:

An interim analysis for FL will be performed after 19 subjects are evaluable for tumor response after completing at least 1 response assessment. If only 6 or fewer responders (≤6/19) are observed, the Sponsor may consider discontinuation of this cohort; however, assessment of biomarkers (that may aid in prospective enrichment of responders) or tumor measurements (showing clinically relevant tumor reductions, ie, 40%, that fit the criteria for SD) may support continued enrollment.

An interim analysis for DLBCL will be performed after 9 subjects are evaluable for tumor response after completing at least 1 response assessment. If only 1 or fewer responder (≤1/9) is observed, the Sponsor may consider discontinuation of this cohort; however, assessment of biomarkers (that may aid in prospective enrichment of responders) or tumor measurements (showing clinically relevant tumor reductions, ie, 40%, that fit the criteria for SD) may support continued enrollment.

The primary efficacy analysis will occur at least 12 months after the last FL subject is enrolled and at least 9 months after the last DLBCL subject is enrolled.

1.6. Testing Procedure and Level of Significance

The 2-sided significance level for the final analysis of primary endpoint (ORR) will be 0.05. Analyses of secondary endpoints will be performed using the same significance level.

2. GENERAL ANALYSIS CONSIDERATION

Time to event or duration of event endpoints will be based on the actual event date (or censoring date) rather than visit number or visit label. Missing efficacy or safety data will not be imputed unless otherwise specified.

In general, the baseline value is defined as the last valid measurement on or prior to the first dose of study treatment. For by-visit analysis, visit windows will be used to associate assessment with a scheduled visit and will be created in reference to the date of first dose of study treatment to assign visit number based on assessment date.

2.1. Analysis Sets

Response-evaluable Population

The Response-evaluable Population is defined as all enrolled subjects who received at least one dose of study treatment (ibrutinib or MEDI4736) and provided at least one post-baseline response assessment. This population includes all the qualified subjects from both Phase 1b and Phase 2. The Response-evaluable Population will be used as the primary population for analyses based on overall response rate. In the analysis, subjects who had clinical progression per investigator will be included in the response-evaluable population.

Safety Population (Treated Population)

Safety population (Treated Population) includes all subjects who were enrolled in the study and received at least 1 dose of study treatment (ibrutinib or MEDI4736). The safety population will

be used to summarize the safety (including dosing) data and for the analysis of PFS, OS and ORR sensitivity analysis.

DLT Evaluable Population for Phase 1b

DLT evaluable population includes Phase 1b subjects who received study treatment and did not meet the criteria for subject replacement during the DLT observation period.

2.2. Definition of Subgroups

For the DLBCL cohort, a subgroup analysis will also be conducted to calculate ORR and its 95% CI for

- GCB and non-GCB DLBCL types per IHC test.
- GCB, ABC and Unclassified DLBCL types per GEP test.

3. SUBJECT INFORMATION

3.1. Subject Disposition

Subject disposition will be summarized by each tumor type cohort. Subject enrollment will be summarized by region, country, and investigator.

Time on study is defined in the same way as overall survival with reversed censoring, i.e., subjects who died will be censored at death date. The Kaplan-Meier method will be used to estimate the median time on study.

3.2. Demographics and Baseline Characteristics

Baseline characteristics and demographic information at baseline will be summarized with descriptive statistics by tumor type cohort.

3.3. Prior and Concomitant Medications

Medications will be coded to a generic name and an Anatomical Therapeutic Chemical (ATC) class per the World Health Organization (WHO) drug dictionary. Concomitant medications will be summarized by therapeutic class and preferred term and by tumor type cohort. Concomitant medications are defined as medications that were taken at any time on treatment (i.e. from the date of the first dose of study treatment through the date of the last dose of study treatment). The following concomitant medications will be summarized separately: growth factors, blood supportive products and immunoglobulin, CYP3A inhibitors/inducers, anticoagulants and/or antiplatelets.

3.4. Extent of Exposure to Study Treatment

Exposure to study treatment will be summarized by tumor type cohort. Descriptive statistics will be provided for treatment duration and dosing information (e.g. total cumulative dose administered, relative dose intensity, dose reduction due to adverse events) for all study treatments.

3.5. Previous Treatment History and Subsequent Antineoplastic Therapies

Previous treatment history and subsequent antineoplastic agents will be summarized separately.

4. EFFICACY AND SAFETY ANALYSES

4.1. Efficacy Analyses

Efficacy endpoints and analysis methods are summarized in Table 1.

Table 1: Summary of Efficacy Analyses

Endpoint	Definition	Analysis Method
Primary Endpoint		
Overall Response Rate (ORR)	Proportion of subjects achieving the best overall responses of complete response (CR) or partial response (PR) with confirmation per the investigator using the Revised Criteria for Response Assessment of Malignant Lymphoma (Cheson 2014)	Primary (Response-evaluable population) • Proportion and 2-sided 95% CI based on exact Binomial distribution (Clopper-Pearson) Sensitivity (Treated population) • Proportion and 2-sided 95% CI based on exact Binomial distribution (Clopper-Pearson) • Waterfall plot of best change from baseline of tumor load from target lesion(s) Subgroup (Response-evaluable and treated population): For the DLBCL cohort, a subgroup analysis will also be conducted to calculate ORR and its 95% CI for • GCB and non-GCB DLBCL types per IHC test • GCB, ABC and Unclassified DLBCL types per GEP test
Secondary Endpoints		
Duration of Response (DOR)	Duration of time from the date of initial response to the date of disease progression or the date of death due to any cause, whichever occurs first	Primary (Response-evaluable population) • Kaplan-Meier estimate
Progression Free Survival (PFS)	Duration of time from the first dose date of study treatment (ibrutinib or MEDI4736) to the first documentation of disease progression or death	Primary (Treated population) • All PD and death considered as events regardless of any anticancer therapy prior to PD • Kaplan-Meier curve

Endpoint	Definition	Analysis Method
	from any cause, whichever occurs first.	
Overall Survival (OS)	Duration of time from the first dose date of study treatment (ibrutinib or MEDI4736) to the date of death due to any cause	Primary (Treated population) • All death considered as events regardless any anticancer therapy prior to death • Kaplan-Meier curve

4.2. Safety Analyses

Safety data will be summarized by tumor type cohorts. Table 2 summarizes the safety analyses to be performed for all tumor type cohorts.

Adverse events (AEs) will be coded in accordance with the Medical Dictionary for Regulatory Activities (MedDRA). Severity of AEs will be graded by the investigator according the NCI-CTCAE v4.03.

All laboratory values will be converted to and reported as international standard (SI) units. In general, only data from the central laboratory will be summarized and analyzed. Laboratory parameters will be graded using the NCI CTCAE v4.03.

Unless otherwise specified, only baseline and post-baseline values collected during the treatment-emergent period will be included in the safety analysis. A treatment-emergent adverse event is defined as any event that begins or worsens in either intensity or frequency on or after the first dose of study treatment (either ibrutinib or MEDI4736). Only treatment-emergent adverse events (TEAEs) occurring and reported during the study period will be included in the adverse event summaries. The TEAEs will be derived as those events that 1) occur after the first dose of study drug (either ibrutinib or MEDI4736), through the 30 days following the last dose of ibrutinib or 90 days following the last dose of MEDI4736 (whichever date is later); 2) any event with missing onset date and its resolution date during the treatment phase; 3) any event that is considered study drug-related regardless of the start date of the event; or 4) any event that is present at baseline but worsens in severity, seriousness, or becomes subsequently considered drug-related by the investigator.

Table 2: Summary of Safety Analyses

Safety Assessment	Definition	Analysis Methods
Safety and tolerability	AE: TEAEs, SAEs, grade 3 or worse TEAEs, related TEAEs, TEAEs leading to treatment discontinuation, TEAEs leading to dose reduction, TEAEs leading to death, protocol- defined events of special interest, and other safety observations Worst post-baseline toxicity grade for selected lab tests: Worst post-baseline toxicity grade, Hgb, creatinine clearance, abnormal uric acid, liver function abnormalities Vital signs (blood pressure, heart rate, temperature, respiratory rate), and weight	Descriptive summary statistics and/or listings

AE: adverse event; SAE: serious adverse event; TEAE: treatment-emergent adverse event

5. MODIFICATION OF ANALYSIS TO THE PROTOCOL

Below is the major change made to the analyses in the protocol:

- The timing of interim analysis for the FL cohort is changed from "An interim analysis will be performed after 19 subjects are evaluable for tumor response after completing at least 2 response assessments" to "An interim analysis will be performed after 19 subjects are evaluable for tumor response after completing at least 1 response assessment" due to the number of early treatment discontinuations prior to 2 response assessments.
- The timing of the interim analysis for the DLBCL cohort is changed from "An interim analysis will be performed after 9 subjects are evaluable for tumor response after completing at least 2 response assessments." to "An interim analysis will be performed after 9 subjects are evaluable for tumor response after completing at least 1 response assessment.
- Because of the substantial number of subjects who discontinued treatment due to PD after one post-baseline tumor assessment, the sponsors (Pharmacyclics and AstraZeneca) decided to hold the enrolment during the interim analysis.
- The terminology 'Safety Assessments' instead of 'Safety Endpoints' is used in Section 1.2 safety analysis part.

6. REFERENCES

Brookmeyer R., and Crowley J., A Confidence Interval for the Median Survival Time. Biometrics, March 1982: 29-41.

Cheson B.D, Pfistner B, Juweid ME, et al. Revised Response Criteria for Malignant Lymphoma, Journal of Clinical Oncology 2007; 25: 579-586.

Collett D. Modeling Survival Data in Medical Research, 2nd Edition, Chapman & Hall CRC, 2003

European Commission, Detailed Guidance on the Collection, Verification, and Presentation of Adverse Event/Reaction Reports Arising From Clinical Trials on Medicinal Products for Human Use, Official Journal of the European Union, C 172, 6 November 2011.

Fleiss J.L, et al., Statistical Methods for Rates and Proportions. Wiley Inter-science, Third Edition, 2003.

US Food & Drug Administration: Guidance for Industry: Clinical Trial Endpoints for the Approval of Cancer Drugs and Biologics. 2007.