Protocol B1871040

AN OPEN-LABEL BOSUTINIB TREATMENT EXTENSION STUDY FOR SUBJECTS WITH CHRONIC MYELOID LEUKEMIA (CML) WHO HAVE PREVIOUSLY PARTICIPATED IN BOSUTINIB STUDIES B1871006 OR B1871008

Statistical Analysis Plan
(SAP)

Version: 3.0

Author: PPD PhD, PPD Clinical Statistician

Date: 11-AUG-2020

TABLE OF CONTENTS

APPENDICES

1. AMENDMENTS FROM PREVIOUS VERSION(S)

The SAP version 3.0 amended the original SAP v2.0 dated November 16, 2016. The main changes are COVID-19 related and additional exploratory analyses which included:

- Impact on analyses from COVID-19.
- Analysis of molecular data.

The SAP version 2.0 amended the original SAP v1.0 dated April 24, 2013. The main changes to align with protocol amendment #1 included:

- Add pharmacokinetic (PK) endpoint.
- Include imputation of cytogenetic response from molecular monitoring data if cytogenetics is not performed per standard of care.
- Update appendices for cytogenetic and hematologic response criteria.

2. INTRODUCTION

This study is a treatment extension protocol to allow the opportunity of long term treatment with bosutinib for subjects who received bosutinib in previous Pfizer sponsored CML studies B1871006 and B1871008 and who are thought to have the potential, as judged by the investigator, to derive clinical benefit from continued treatment with bosutinib.

In addition this study:

- *Will allow collection of long term survival data for all bosutinib subjects who have discontinued bosutinib treatment in parent studies B1871006 and B1871008 but are known to be alive as of their last long term follow-up visit;*
- *Will provide the opportunity to gather additional long term safety and selected efficacy data for bosutinib in all CML subjects, including 1st line subjects in chronic phase, as well as second, third and even fourth line subjects in chronic, accelerated and blast CML phases;*
- *Will satisfy the EMA post- approval commitment to provide safety data about diarrhea incidence after the subjects switch from clinical study to commercial bosutinib formulation.*

2.1. Study Design

This is an open-label bosutinib treatment extension protocol. This protocol will be offered to those bosutinib patients who were previously enrolled in one of the two parent CML bosutinib studies (B1871006 or B1871008).

Patients to be enrolled will include those who – at the time of this protocol approval - are still receiving bosutinib in either one of the parent studies and are benefiting from bosutinib treatment as judged by the investigator, as well as those patients who have already discontinued bosutinib as part of the parent studies and are in long-term follow-up (LTFU) for survival. The former group will continue to receive bosutinib as part of the extension study; the latter group will only enter into the long-term survival follow-up part of the extension study. Data on any subsequent therapy with TKIs or other anti-cancer therapy and response following bosutinib will be collected on LTFU patients, if available.

Each patient will remain in the extension study, either on bosutinib treatment or in long-term survival follow-up phase, until the last patient has reached 10 years of follow-up, as calculated from the date of his/her first dose of bosutinib administered in the parent study. When this milestone is reached, the present study will be closed. At that time patients still benefiting from bosutinib will switch to the most appropriate therapy available at that time.

Patients are allocated to the cohort based on the derived assignment from the parent study at the time of database lock. Patients will be summarized separately or pooled based on these cohorts depending on the specific analysis. All B1871008 patients are chronic phase CML first line (CP1L). B1871006 patients are either (1) chronic phase CML second line (CP2L) resistant or intolerant to imatinib, (2) chronic phase CML third/fourth line (CP3L/CP4L) resistant or intolerant to imatinib and resistant or intolerant to dasatinib and/or nilotinib, (3) accelerated phase CML resistant or intolerant to imatinib only (AP2L) or resistant or intolerant to imatinib and at least one additional TKI including dasatinib and/or nilotinib (AP3L/AP4L), (4) blast phase CML resistant or intolerant to imatinib only (BP2L) or resistant or intolerant to imatinib and at least one additional TKI including dasatinib and/or nilotinib (BP3L/BP4L), or (5) Ph+ ALL resistant or intolerant to imatinib only or resistant or intolerant to imatinib and at least one additional TKI including dasatinib and/or nilotinib. Accelerated phase, blast phase, and Ph+ ALL combined are known as advanced patients.

All data from the 2 parent studies will be combined with data from B1871040 for analysis. Only baseline data from the 2 parent studies previously reported in the clinical study reports will not be summarized.

2.2. Study Objectives

- *To allow long term bosutinib treatment in patients with chronic or advanced phases Ph+ CML who received bosutinib in a previous Pfizer sponsored CML study (ie, studies B1871006 and B1871008) and who have the potential, as judged by the investigator, to derive clinical benefit from continued treatment with bosutinib;*

- *To collect long term safety and efficacy data for bosutinib;*

- *To assess the duration of clinical benefit for Ph+ CML patients treated with bosutinib;*

- *To fulfill the European Medicines Agency (EMA) post approval requirement for the collection and analysis of safety data about diarrhea incidence after switch from clinical study to commercial bosutinib formulation;*

- *To fulfill the EMA post-approval requirement for the analysis of the PK of bosutinib administered once daily.*

3. INTERIM ANALYSES, FINAL ANALYSES AND UNBLINDING

There are no interim analyses planned for this study. The interim analyses were already conducted as part of the parent studies, along with the unblinding for patients in B1871008.

The final analysis will be performed when all patients have finished the 10 year follow-up period, calculated as 10 years from first dose for B1871008 and B1871006 patients.

4. HYPOTHESES AND DECISION RULES

4.1. Statistical Hypotheses

There are no statistical hypotheses in this study.

4.2. Statistical Decision Rules

There are no statistical decision rules for this study and no adjustments for multiple comparisons. Confidence intervals (CI) are intended for descriptive purposes only.

5. ANALYSIS SETS

5.1. Full Analysis Set

The full analysis set is all patients randomized to the bosutinib arm from B1871008 and all dosed patients from B1871006.

5.2. 'Per Protocol' Analysis Set

Per Protocol analyses will not be performed in this study. The Per Protocol analyses were conducted as part of the parent studies for the primary (B1871008, B1871006) and secondary (B1871006) endpoints.

5.3. Safety Analysis Set

The safety analysis set are those patients who received at least one dose of bosutinib. For B1871006 patients, this is the same as the full analysis set.

5.4. Other Analysis Sets

The evaluable analysis set are those dosed patients from B1871006 with a valid baseline efficacy assessment from B1871006 for the respective endpoint (ie, [1] at least 20 metaphases or at least 1 Ph+ metaphase from the baseline bone marrow cytogenetic assessment or [2] a valid baseline hematologic assessment as defined in Appendix 1 or [3] are not from China, South Africa, Russian Federation, or India for molecular analyses as molecular samples were not evaluated). This analysis set will be used as deemed appropriate for cytogenetic, hematologic, and molecular analyses for B1871006 patients. There is no potential bias expected from utilizing this analysis set as patients are included based on baseline information collected prior to the first dose of study medication from B1871006.

5.5. Treatment Misallocations

All B1871008 patients were correctly randomized to the appropriate treatment group. Three bosutinib patients were randomized but not dosed. These 3 patients will be included in the full analysis set but not the safety analysis set. B1871006 was a single arm study so there were no treatment misallocations. One patient was enrolled but not dosed and will be excluded from the full and safety analysis sets.

5.6. Protocol Deviations

Per Protocol analyses will not be performed for this study. The full list of protocol deviations for the study report will be compiled prior to database closure. A listing of COVID-19 related protocol deviations will be provided.

6. ENDPOINTS AND COVARIATES

There is no formal hypothesis testing therefore endpoints are not designated as primary or secondary.

In general, the baseline value will be defined as the last non-missing value prior to first dose from B1871008 and B1871006, unless specifically stated otherwise. Assessments performed on Week 1 day 1 will be assumed to be pre-dose per the schedule of activities in the Protocol.

An on-treatment efficacy evaluation is defined as any evaluation done during bosutinib therapy or within 30 days of last dose for B1871008 patients and up to date of treatment discontinuation for B1871006 patients. An on-treatment safety evaluation is defined as any evaluation done during bosutinib therapy or within 30 days of last dose.

6.1. Efficacy Endpoint(s)

- Duration of major/complete cytogenetic response (MCyR/CCyR) for B1871006 patients only;
- Duration of major/deep molecular response (MMR/MR4) for B1871006 patients only;
- Duration of complete hematologic response (CHR) for B1871006 patients only;
- Duration of overall hematologic response (OHR) for B1871006 advanced patients only;
- Progression-free survival (PFS) for B1871006 patients only;
- Transformation to accelerated or blast phase for B1871006 patients only;
- BCR-ABL mutations present at treatment discontinuation;
- Overall Survival (OS).

Duration of response (DOR) is defined as the time from first response to confirmed loss, progression of disease, or on-treatment death due to any cause ([date of loss – date of first response]/7). Response is unconfirmed for cytogenetic and molecular and confirmed for hematologic with 2 consecutive responses at least 28 days apart. Confirmed loss is defined as 2 consecutive non-responses at least 28 days apart for MCyR and 14 days apart for hematologic, 1 non-response confirmed by progression of disease or death, or no loss of response and progression of disease or death. For CCyR, confirmed loss is defined as 2 consecutive assessments with >0 Ph+ metaphases or ≥1% positive cells from fluorescence in situ hybridization (FISH) at least 28 days apart or progression or death. For MMR, confirmed loss is defined as 2 consecutive assessments at least 28 days apart with a <3-log (>0.1% from B1871006 or MR^2 or worse denoted on the B187040 case report form [CRF]) reduction in transcripts from standardized baseline one of which corresponds to a ≤2-log reduction (≥1% from B1871006 or MR^1 or worse denoted on the B1871040 CRF) or progression or death. For MR^4, confirmed loss is defined as 2 consecutive assessments at least 28 days apart with a <4-log (>0.01% from B1871006 or MR^3 or worse denoted on the B187040 CRF) reduction in transcripts from standardized baseline one of which corresponds to a ≤3-log reduction (≥0.1% from B1871006 or MR^2 or worse denoted on the B1871040 CRF) or progression or death. Patients without confirmed loss are censored at the last valid assessment where response could be assessed. Cytogenetic and hematologic responses are defined in Appendix 2 and Appendix 1. MCyR is defined as a CCyR or partial cytogenetic response (PCyR). OHR is defined as CHR or return to chronic phase (RCP). To be considered a responder, the patient must have an improvement from baseline or maintenance of baseline response for ≥5 weeks for hematologic response or ≥4 weeks for cytogenetic response. Patients with a non-valid cytogenetic assessment (<20 metaphases) or missing baseline assessment must demonstrate the best response, CCyR or CHR, to be considered a responder for the respective endpoint.

Molecular response in B1871006 was based on the difference between the subject's change in Bcr-Abl/Abl ratio and a laboratory reference baseline. The Quest (vendor) laboratory reference Bcr-Abl/Abl ratio of 4.1325 derived from a pool of 120 previously untreated CML subjects was used as the standardized (or reference) baseline. Molecular response was not analyzed for subjects in China, India, the Russian Federation, and South Africa due to logistical constraints in transporting samples to the central laboratory.

To be considered a responder for MMR in B1871006, a subject must have had at least a 3-log reduction from the Quest standardized baseline, a detectable Bcr-Abl transcript at baseline or any time post-baseline, and must have maintained or attained a CCyR. Additionally, as a conservative approach given the known variability with standard MMR analyses, a subject who had a non-CCyR on the same day as the MMR assessment was not considered to have a MMR at that time point. MR^4 was defined similar to MMR except at least a 4-log reduction was required. In B1871040, responders had MMR/MR^4 or better denoted on the CRF as assessed by a local laboratory. On-treatment response is independent of a patient's baseline molecular response.

PFS is defined as the time from first dose to progression as assessed by the investigator and noted as the reason for permanent treatment discontinuation on the CRF or on-treatment death due to any cause ([date of progression/death – first dose date]/30.4), whichever occurs first. Patients without events are censored at the last cytogenetic, hematologic, or laboratory assessment where progression could be assessed. Progression for B1871006 patients is defined as:

- *Evolution from chronic phase (or from return to chronic phase for advanced subjects) to accelerated phase or blast crisis (on two consecutive assessments at least a week apart).*
- *Evolution from accelerated phase to blast crisis (on two consecutive assessments at least a week apart).*
- *One of the following conditions occurring after dose escalation or presence of adverse events prohibiting dose escalation:*
    - *For 2nd or later line subjects, loss of major cytogenetic response (need at least 30% increase).*
    - *For all lines of treatment, loss of complete hematologic response confirmed by 2 assessments at least 2 weeks apart.*
    - *For all lines of treatment, increasing WBC defined as doubling of WBC, occurring over a period of ≥1 month, with the second WBC >20 x 10^{9}/L and confirmed at least 1 week later.*

For DOR and PFS, if there is an unacceptable gap between the date of progression or death and the date of the most recent prior disease assessment (>48 weeks), the event will not be used and censorship will be at the most recent prior assessment.

Transformation, defined by the first two progression bullets above, is the time from first dose to confirmed transformation ([date of transformation – first dose date + 1]/7). Confirmed transformation is defined as 2 consecutive assessments at least 1 week apart or 1 assessment confirmed by progression of disease or death. Patients without events are censored at the last hematologic or extramedullary assessment where transformation could be assessed.

BCR-ABL mutational status at treatment discontinuation will be assessed. Emergent mutations are defined as those mutations which were not present at baseline. B1871006 patients who were not assessed for baseline mutations are not assessable for emergent mutations. All B1871008 patients are assessable for emergent mutations.

OS is defined as the time from randomization for B1871008 patients and from first dose B1871006 patients to the date of death ([date of death – randomization/first dose date]/30.4). Patients who have not died will be censored at the last known alive date.

6.2. Safety Endpoints

Treatment-emergent adverse events (TEAE) were defined as any event increasing in severity from baseline or any new event starting during bosutinib therapy or within 30 days of the last dose of study drug.

6.2.1. Exposure and Compliance

Exposure and compliance data will be collected during this study. The primary exposure endpoints while the patient is on treatment are duration of treatment, cumulative dose received, and reasons for dose modifications (reduction, delay, or escalation). Duration of treatment is the time from first non-zero dose to last non-zero dose ([date of last non-zero dose – first dose date + 1]/30.4). Planned dose is based on the initial dose for B1871006 patients and randomized dose for B1871008 patients.

6.2.2. Laboratory Evaluations, Vital Signs, and Cardiac Evaluations

Laboratory values, vital signs, electrocardiogram (ECG), and left ventricular ejection fraction (LVEF) data from echocardiogram (ECHO) or multiple gated acquisition scan (MUGA) will be collected during this study. Collection time points for these measurements are specified in the protocol. Laboratory values, vital signs, ECG, and LVEF data will be flagged according to the potentially clinically important (PCI) criteria for abnormal values. The PCI criteria for laboratory and LVEF values are those with grade 3 or 4 from NCI CTC version 3.0. The PCI criteria for vital signs and ECG are specified in Appendix 3.

For continuous ECG parameters, the average of all valid measurements among the last performed “triplicate" ECG observed prior to taking the first dose of test article is defined as the baseline value. For Overall Evaluation, the most serious of all valid measurements among the last performed “triplicate" ECG observed prior to taking the first dose of test article is defined as the baseline value.

If multiple post-baseline observations occur at the same visit, the values used for summaries and analyses are defined as the following:

- For PCI laboratory evaluations, PCI vital signs (and body weight), and PCI LVEF, the most serious (highest grade) evaluation will be used for all summaries and analyses.

- For PCI ECG, the mean of the “triplicate” ECGs at each visit will be calculated and then the most serious (highest) mean will be used for all summaries and analyses. For Overall Evaluation, the most serious evaluation will be used for all summaries and analyses.

6.2.3. Laboratory Evaluations of Special Interest

Laboratory parameters including alanine aminotransferase (ALT), aspartate aminotransferase (AST), platelets, neutrophils, and hemoglobin will be summarized. Endpoints for these parameters include distribution of maximum toxicity, time to first event (time from first dose to date of first event including only non-partial dates), duration of any grade event (time from start date to stop date including only non-partial dates), and cumulative duration of grade

3/4 events (sum of time from start date to stop date including only non-partial dates for all events).

A listing will be provided for patients who potentially meet the Hy's law criteria. Potential is defined as meeting these 3 criteria any time during the study (non-concurrent): AST or ALT level of ≥3 x upper limit of normal (ULN), total bilirubin level of >2 x ULN, and alkaline phosphatase level <2 x ULN. Total Bilirubin, AST, ALT, and alkaline phosphatase will be listed.

6.2.4. Non-study Medications

Non-study medications are coded using the World Health Organization (WHO) Drug Dictionary. The Prior, Concomitant, After (PCA) flag for non-study medications will be derived: prior non-study medications are defined as any non-study medications taken before the first dose of test article administration; concomitant non-study medications are defined as any medications taken during the treatment period (includes the 30 days after the last dose of the test article); after non-study medications are defined as any non-study medications taken during the follow-up period.

6.2.5. Death

Death data will be collected during this study. Mandatory variables captured are date of death and cause of death. Deaths within 30 days of last dose and death due to AEs are safety endpoints.

6.2.6. Adverse Events

A 3-tier approach will not be used to summarize AEs since only the bosutinib patients from the randomized B1871008 are eligible for this study; the imatinib patients were not offered enrollment to this extension protocol. Furthermore, the B1871006 study was a single-arm non-comparative study.

To ensure accuracy and consistency in data analysis and reporting, all adverse events must be classified using the Medical Dictionary for Regulatory Activities (MedDRA) before database lock (DBL). The severity of AEs will be assessed using the National Cancer Institute Common Terminology Criteria for Adverse Events (NCI CTCAE) version 3.0. The causal relationship between test article and an AE will be assessed by investigator's judgment.

AEs will be classified as follows:

- TEAEs;
- AEs leading to treatment discontinuations;
- TEAEs leading to dose reductions;
- TEAEs leading to dose delays;
- TEAEs by maximum toxicity grade (grades 1 to 5);

- TEAEs of higher toxicity grades (grades 3 or 4);
- TEAEs related to test article;
- Serious adverse events (SAEs).

6.2.7. Adverse Events of Special Interest

Adverse events including ALT, AST, thrombocytopenia (including MedDRA preferred term [PT] Platelet Count Decreased), neutropenia (including PT Neutrophil Count Decreased), anemia (including PT Hemoglobin Decreased), diarrhea, nausea, vomiting, cardiac toxicity (definition to be determined before DBL), vascular toxicity (definition to be determined before DBL), and renal toxicity (definition to be determined before DBL) will be summarized. Additional AEs or AE clusters may be summarized if deemed appropriate. AE characteristics to be summarized include maximum toxicity, time to first event, duration of any grade event, cumulative duration of grade 3/4 events, and successful rechallenge (those patients who were re-dosed and did not discontinue bosutinib due to this AE after having a dose delay).

The diarrhea characteristics will also be summarized separately for data from B1871008 and B1871006 compared to B1871040 in order to assess the impact of changing to the clinical formulation of bosutinib from the parent studies to the commercial formation in this study. For ongoing events at the time of the switch to commercial formulation, the stop date for the event attributed to the clinical formulation will be set to date of the start of commercial formulation minus one and the start date for the event attributed to the commercial formulation will be set to the start date of commercial formulation.

6.3. Other Endpoints

6.3.1. PK Endpoints

Steady-state trough concentrations (C_{trough}) of bosutinib will be determined. A total of 1 PK samples per subject will be drawn. All patients will provide a pre-dose blood sample following at least 2 weeks of uninterrupted dosing with bosutinib at the same dose level. This sample will be collected at their first scheduled visit following all appropriate approvals and implementation of protocol amendment # 1. Biological specimen samples for pharmacokinetic analysis must be identified with labels provided by the sponsor.

Concentrations of bosutinib will be determined in plasma using a Pfizer -approved and validated liquid chromatography/tandem mass spectrometry (LC/MS/MS) assay.

The objective of the pharmacokinetic analysis is to compare the C_{trough} of bosutinib in this study to C_{trough} of previous studies. Definitions of and analyses for PK parameters will be outlined in a formal PK analysis plan and will not be included in this SAP.

6.3.2. PD Endpoints

There are no PD data being collected in this study.

6.3.3. Outcomes Research Endpoints

There are no outcomes research data being collected in this study.

6.4. Covariates

Subgroup analyses by gender, age (<65, ≥65), and race (Asian, Black, Other, White) will be explored as deemed appropriate. Additional supportive analyses may be explored if deemed appropriate.

The randomization in B1871008 was stratified by region and Sokal score; however, since only bosutinib patients from B1871008 are eligible for B1871040, none of the CP1L analyses will be stratified.

7. HANDLING OF MISSING VALUES

In general, no imputation such as last observation carried forward (LOCF) will be used. For the time-to-event endpoints, such as DOR, time to transformation, PFS, and OS, the primary missing data handling method will be censoring as described in Section 6.1.

In compliance with Pfizer standards for handling partial dates, if the month and year are present but day is missing, start date is set to 1st of month, and stop date is set to last day of month. If year is present but day and month are missing, start date is set to January 1st, and stop date is set to December 31st.

Handling of missing/incomplete patient cytogenetic data from parent study B1871006 was done by the following procedure: for B1871006 baseline visits, only bone marrow information is used. If <20 metaphases at baseline, patients with ≥8 Ph+ will be assigned a minor response and those with 1 to 7 Ph+ will be assigned a partial cytogenetic response (PCyR). For post-baseline visits in B1871006 and B1871040, bone marrow information with 20 to 100 metaphases will be used over FISH information with at least 200 nuclei at a specific visit. No conventional bone marrow cytogenetic assessment will be derived when >100 cells are reported as analyzed for Philadelphia chromosome as this will be considered an error in data entry. If on-treatment molecular monitoring is performed for visits in B1871040 in lieu of conventional cytogenetics or FISH, CCyR may be imputed on a specific date if an MMR or better is achieved and denoted on the CRF on that date. The results of conventional cytogenetic testing, banding with ≥20 metaphases, will be used in preference to molecular testing, provided they are available and performed on the same date however molecular results will be used over FISH with ≥200 nuclei if results are discrepant. Handling of missing/mismatched patient hematologic efficacy data for from parent study B1871006 was done using the following procedure: for each patient, the observations were sorted by visit and the relative day to first dose from B1871006.

First, for each observation, peripheral blood and/or bone marrow differentials should add up to 100%. Some sites may not record 0% to reflect uncounted cell types as instructed. Instead, the corresponding field may appear as missing or “ND” even though data collection procedures have been vigorously followed. Also, the differentials may not exactly add up to 100%. To correct this problem, each domain of the differentials was summed together.

If the total was between 98.5% and 101%, the missing values were assigned 0%. Otherwise, the observation was left as is and the record was also to be reported as a data issue.

Second, for the set of variables collected to derive peripheral blood or bone marrow differentials at each post-baseline visit through week 12 on B1871006, if a given set were still incomplete, a non-missing set of lab variables obtained ±3 days from the scheduled visit was utilized to complete the derivation. For visits after week 12 (relative to first B1871006 dose) from B1871006 and B1871040, a non-missing set of lab variables obtained ±7 days of the visit will be utilized to complete that post-baseline assessment. For B1871006, the last non-missing peripheral blood or bone marrow differentials collected at week 1 day 1 or up to 28 days prior to week 1 day 1 was used for baseline assessment. For example, if the bone marrow blast count was missing, the entire non-missing bone marrow assessment, not just the blast field, from another observation was used. Some data handling rules for data combination are specified below.

- Data combination is needed only when the full dataset is not available on a single day at a given visit. The full dataset includes peripheral blood differential for chronic phase, and both bone marrow and peripheral blood differentials for advanced phase. The extramedullary assessment when absent should not preclude the derivation, and when present should be taken into account per response criteria rules.

- During the post-baseline data combination, for visits up to and including week 12 from B1871006, the ±3 window will be applied to days 1 and 7 separately; for visits after week 12 (relative to first B1871006 dose) from B1871006 and B1871040, the ±7 window will be based on day 7 only. The latest date will be used as the date of visit and the more conservative value will be taken if there is more than one measurement for a variable. For platelet counts, in case of multiple values, the lowest value will be taken. If any value falls above or below normal range, the out-of range value will be taken.

8. STATISTICAL METHODOLOGY AND STATISTICAL ANALYSES

8.1. Statistical Methods

8.1.1. Analyses of Continuous Data

For continuous variables, the descriptive statistics include n, mean, median, standard deviation (SD), and range (minimum, maximum).

8.1.2. Analyses of Categorical Data

For categorical variables, the descriptive statistics include the count in each category, the total n, and percentage.

8.1.3. Analyses of Binary Endpoints

Exact 2-sided 95% CIs will be provided for molecular, cytogenetic, and hematologic response rates.

8.1.4. Analyses of Time-To-Event Endpoints

For time-to-event endpoints such as DOR, time to transformation, PFS, and OS, the method of analysis will either be Kaplan-Meier (K-M) or cumulative incidence of the event adjusting for the competing risk of treatment discontinuation due to any reason without the event (Gray).[1]

When applying the K-M method, medians and quartiles with the associated 2-sided 95% CI will be provided. The CIs are based on the Brookmeyer-Crowley linear transformation method will be provided. The K-M yearly rates with the associated 2-sided 95% CI based on Greenwood's formula will also be provided.

When applying the cumulative incidence method, yearly rates with the associated 2-sided 95% CI based on the delta method using a log transformation will be provided.

8.2. Statistical Analyses

The efficacy endpoints will be analyzed according to the cohorts specified in Section 2.

8.2.1. Analysis of Efficacy Endpoints

The duration of cytogenetic, hematologic, and molecular response analyses are based on the full analysis set for the subset of B1871006 patients who respond for the respective endpoint. The K-M method, as described in Section 8.1, will be used provided most responding patients transferred to B1871040 remain on treatment until progression or loss of response. The percentage of MR^4, MMR, MCyR, CCyR, and PCyR and CHR and major hematologic response (MHR)/OHR (advanced patients only) for the respective endpoint, along with confirmed losses and censored patients still on-treatment will be summarized.

The time to transformation analysis is based on the full analysis set for the B1871006 patients. Cumulative incidence of transformation adjusting for the competing risk of treatment discontinuation due to any reason without transformation, as described in Section 8.1, will be used. The percentage of transformation without treatment discontinuation and the percentage of treatment discontinuation due to any reason without transformation will be summarized. The specific disease stage transformation (accelerated or blast phases) will also be listed.

The PFS analysis is based on the full analysis set for the B1871006 patients. The K-M method, with or without unsatisfactory response as assessed by the investigator as an event and cumulative incidence of on-treatment progression or death adjusting for the competing risk of treatment discontinuation due to any reason without the event, as described in Section 8.1, will be used. For K-M, the percentage of on-treatment death or progression, with or without unsatisfactory response as an event, will be summarized. For cumulative incidence, the percentage of on-treatment death or progression without treatment discontinuation and the percentage of treatment discontinuation due to any reason without on-treatment progression and death will be summarized.

The OS analysis is based on the full analysis set for all patients. The K-M method, as described in Section 8.1, will be used. The number of deaths and causes of death will be summarized. Time on study from date of randomization for B1871008 patients and date of first dose for B1871006 patients to date of last contact, and reasons for discontinuing participation will also be summarized.

The percentage of patients with emergent mutations will be summarized.

Missing values and incomplete data will be handled as described in Section 7 of this analysis plan.

Listings for all efficacy data will also be provided.

No COVID-19 related efficacy analyses will be performed as efficacy is only summarized for data up through 10 years after start of the parent study and all patients enrolled in B1871040 except one reached 10 years of follow-up prior to the March 11, 2020 COVID-19 anchor date and all China patients reached 10 years of follow-up prior to their January 9, 2020 anchor date. A summary of missing molecular and cytogenetic data before and after the anchor dates will be summarized for patients enrolled in B1871040.

A listing of COVID-19 related study discontinuations will be provided.

8.2.2. Analysis of Safety Endpoints

The safety analysis set will be used for all safety analyses unless specifically stated otherwise. The safety endpoints will be analyzed according to the cohorts specified in Section 2 and also analyzed as a pooled set of all treated patients from B1871006 and B1871008.

8.2.2.1. Baseline Characteristics

Demographic, medical history, and other baseline characteristics will not be summarized for B1871040 as this information resides solely in the parent study databases. This summary information will be taken directly from the B1871006 and B1871008 clinical study reports.

8.2.2.2. Treatment and Compliance

Descriptive statistics for continuous exposure variables treatment duration and cumulative received dose, and received dose over time (% receiving 600, 500, 400, 300, 200, or 0 mg at specific time points) will be calculated.

The number and percentage of patients with at least one dose reduction due to AE, dose delay due to AE, and escalation to 600 mg will be summarized. The number and percentage of patients who experienced more than 1 event will also be summarized.

The number and percentage for reasons for treatment discontinuation will be summarized and listed. A listing of COVID-19 related treatment discontinuations will be provided.

8.2.2.3. Laboratory Evaluations, Vital Signs, and Cardiac Evaluations

The incidence of clinical laboratory, QTcF/QTcB, and LVEF abnormalities by maximum toxicity categorized by NCI CTC V3.0 grade will be presented at baseline, during the treatment period, and post treatment period.

Laboratory values that are outside the normal range also be flagged in the data listings. Any out-of-range values that are identified as being clinically significant will be presented in a data listing. PCIs for ECG, LVEF, vital signs and body weight will also be flagged in listings.

A summary of missing safety data before and after the anchor dates will be summarized for patients enrolled in B1871040.

8.2.2.4. Laboratory Evaluations of Special Interest

The number and percentage of maximum on-treatment grade 0 through 4 for ALT, AST, platelets, neutrophils, and hemoglobin will be summarized. Descriptive statistics for time to first event, duration of any grade event, and cumulative duration of grade 3/4 events will be calculated for each of these parameters.

8.2.2.5. Non-study Medications

A listing of non-study medications data will be reported.

8.2.2.6. Death

The number of deaths within 30 days of last dose and cause of these deaths will be summarized.

8.2.2.7. Adverse Events

The number and percentage of AEs listed below will be summarized overall and by MedDRA body system and PT. Sorting will be done alphabetically by body system and then by decreasing frequency of PT.

- TEAEs;
- AEs leading to treatment discontinuations;
- TEAEs leading to dose reductions;
- TEAEs leading to dose delays;
- TEAEs by maximum toxicity grade (grades 1 to 5);
- TEAEs of higher toxicity grades (grades 3 or 4);
- TEAEs related to test article;
- Serious adverse events (SAEs);

- COVID-19 related TEAEs.

8.2.2.8. Adverse Events of Special Interest

The number and percentage of maximum treatment-emergent CTCAE grade events and successful rechallenge for ALT, AST, thrombocytopenia (including MedDRA PT Platelet Count Decreased), neutropenia (including PT Neutrophil Count Decreased), anemia (including PT Hemoglobin Decreased), diarrhea, nausea, and vomiting will be summarized. Descriptive statistics for time to first event, duration of any grade event, and cumulative duration of grade 3/4 events will be calculated for each of these parameters.

The number and percentage of maximum treatment-emergent CTCAE grade events over time for diarrhea will be summarized. Only patients receiving treatment within a given time period will be included in the denominator for this time period. The incidence relative to the first dose of both the parent studies and the first dose of B1871040 will be calculated separately. Furthermore, the incidence relative to the last dose of bosutinib from the parent studies will also be calculated in order to see the percentage of patients experiencing diarrhea during the last 6 months prior to enrolling in B1871040. Time to first event in B1871040, defined as any new event in B1871040 or any increase in severity for ongoing events from B1871006 or B1871008, and duration of B1871040 events will also be summarized.

8.2.2.9. Additional Safety

Listings of physical examination data, chest x-ray, and ECOG performance status will be provided.

8.2.3. Analyses of B1871040 Patient Disposition Endpoints

The listings of violations of B1871040 inclusion/exclusion criteria will be reported. This study will include the summaries of the numbers of patients who: sign the informed consent form (ICF), are screening failures, are enrolled, and are part of the full/safety/evaluable analysis sets. Descriptive statistics for the duration on-study (time from first dose for B1871006 patients and randomization for B1871008 patients to last contact and time from first dose on B1871040 to last contact) will also be calculated.

9. REFERENCES

1. A class of k-sample tests for comparing the cumulative incidence of a competing risk. Gray, RJ. 3, 1988, Annals of Statistics, Vol. 16, pp. 1141-1154.

APPENDICES

Appendix 1. HEMATOLOGIC RESPONSE DEFINITIONS

Hematologic Responses	***Definition***
Return to Chronic Phase (RCP) *(or CP for subjects with CP)* *Minor Hematologic Response (MiHR) (for Ph+ ALL subjects)* *Subjects had to meet all criteria*	*Disappearance of features defining accelerated and blast phases, but still in chronic phase as noted by:* • *<15% Blasts in both blood and marrow* • *<30% Blasts + promyelocytes in both blood and marrow* • *<20% Basophils in both blood and marrow (marrow not applicable for Ph+ ALL)* • *No extramedullary involvement other than liver/spleen*
Accelerated Phase *had to meet at least 1 criterion*	• *15-29% Blasts in blood or marrow* • *≥30% Blasts + promyelocytes in blood or marrow* • *≥20% Basophils in blood or marrow* <u>*only applicable to advanced phase*</u> • *Platelets <100x10^9/L (not related to therapy)*
Blast Phase *had to meet at least 1 criterion*	• *≥30% blasts in blood or bone marrow* • *Extramedullary involvement other than liver/spleen*
No Evidence of Leukemia (NEL) *had to meet all criteria*	• *WBC ≤ institutional ULN* • *450 x10^9/L > Platelets ≥20x10^9/L* • *ANC ≥0.5x10^9/L* • *<20% basophils in blood* • *No blasts or promyelocytes in blood* • *Myelocytes + metamyelocytes <5% in blood* • *No extramedullary involvement (incl. hepato- or splenomegaly)* <u>*only applicable for advanced phase*</u> • *≤5% bone marrow blasts*
Complete Hematologic Response (CHR) *had to meet all criteria*	• *WBC ≤ institutional ULN* • *Platelets <450 × 10^9/L* • *<20% basophils in blood* • *No blasts or promyelocytes in blood* • *Myelocytes + metamyelocytes <5% in blood* • *No extramedullary involvement (incl. hepato- or splenomegaly)* <u>*only applicable to advanced phase*</u> • *≤5% bone marrow blasts* • *ANC ≥1.0 × 10^9/L* • *Platelets ≥100 × 10^9/L*
Overall Hematologic Response (OHR) *had to meet at least 1 criterion*	*CHR +NEL +RCP + MiHR (if applicable)*
Major Hematologic Response (MHR) *(had to meet at least 1 criterion)*	*CHR +NEL*

Abbreviations: ANC=absolute neutrophil count, ULN=upper limit of normal range, WBC=white blood cell count

Appendix 2. CYTOGENETIC RESPONSE DEFINITIONS

*Cytogenetic Responses**	*% Philadelphia chromosome positive cells*
None	*>95%*
Minimal	*66-95%*
Minor	*36-65%*
Partial	*1-35%*
Complete	*0% from conventional cytogenetics or <1% from FISH*
Major	*Complete + Partial Rates*

*** Based on analysis of 200 nuclei if done from blood sample or based on analysis of 20 metaphases if done from bone marrow sample.*

Appendix 3. POTENTIALLY CLINCALLY IMPORTANT EVENTS

Potentially Clinically Important Criteria for Vital Signs and Body Weight

Variable	Criteria
Heart rate	Value <40 beats/min and value >150 beats/min
Systolic blood pressure	<80 or >210 mmHg
Diastolic blood pressure	<40 or >130 mmHg
Temperature	<32 or >40°C
Body weight	Change of greater than or equal to 10% in body weight
Respiratory Rate	<10 or >50 breaths/min

Potentially Clinically Important Criteria for Electrocardiogram Evaluations

Variable	Criteria
Heart rate	Increase >15 beats/min and value ≥120 beats/min
	Decrease >15 beats/min and value ≤45 beats/min
PR interval	Increase of ≥20 msec and value ≥220 msec
QRS interval	Value ≥120 msec
QTcF interval	Value >500 msec Increase >60 msec
QTcB interval	Value >500 msec Increase >60 msec

Abbreviations: min=minute; QTcB=QT interval corrected using the Bazett formula; QTcF=QT interval corrected using the Fridericia formula; VS=vital signs.