

Statistical Analysis Plan

A Phase 1/2, Open-Labeled Study to Assess the Safety, Tolerability, Pharmacokinetics, and Anti-tumor Activity of BBT-176 in Patients with Advanced Non-Small Cell Lung Cancer (NSCLC) who Progressed Following Prior Therapy with an Epidermal Growth Factor Receptor Tyrosine Kinase Inhibitor (EGFR TKI) Agent

Bridge Biotherapeutics, Inc

Protocol BBT176-ONC-001

Protocol Version ■■■

Statistical Analysis Plan Version 1.0
28 JAN 2021

Contents

Signature Page	4
LIST OF ABBREVIATIONS AND DEFINITIONS OF TERMS.....	5
1. INTRODUCTION.....	7
2. SUMMARY OF KEY PROTOCOL INFORMATION.....	7
2.1 Study Objectives and Endpoints	7
2.1.1 Primary Objectives	7
2.1.2 Secondary Objectives	7
2.1.3 Exploratory Objectives.....	8
2.2 Study Design.....	8
2.2.1 Part 1 Study Design: Dose-Escalation Study	9
2.2.2 Part 1 Study Design: RP2D Exploration study.....	11
2.2.3 ■■■■■	11
2.3 Schedule of Activities (SOA).....	11
3. DETERmination of sample size	15
3.1 Part 1.....	15
3.2 ■■■.....	15
4. GENERAL ANALYSIS CONSIDERATION.....	16
4.1 Data Summaries	16
4.2 Data Handling	16
4.2.1 Baseline Value.....	16
4.2.2 Conventions	17
4.2.3 Standard Calculations.....	17
5. ANALYSIS POPULATION.....	18
6. STUDY POPULATION.....	18
6.1 Patient Disposition	18
6.2 Demographic and Baseline Characteristics.....	19
6.3 Medical History.....	19
6.4 Prior and Concomitant Medication	19
7. Efficacy Analyses.....	19
7.1 Tumor Response	20
7.1.1 Overall Time Point Response	20
7.1.2 Best Response.....	21
7.1.3 Best Overall Response.....	23
7.2 Objective Response Rate (ORR)	23
7.3 Duration of Response (DoR)	23
7.4 Progression-Free Survival (PFS).....	23
7.5 Handling Missing, Unused, and Spurious Data	24
7.6 ■■■■■	24
7.7 Methods of Efficacy Analysis.....	24
8. Pharmacokinetic (PK) Analyses	24
8.1 PK Sampling Schedule	25
8.2 Handling of the Below the Lower Limit of Quantification (BLQ) and the No Reportable Concentration Values	25

8.3 Handling of the Difference between the Scheduled and the Actual Sampling Times	25
8.4 Plasma PK Endpoints.....	25
8.5 Statistical Analyses	27
<hr/>	
9. Pharmacodynamic (PD) Analyses	29
10. Safety Analyses	30
10.1 Extent of Exposure.....	30
10.2 Adverse Events (AEs)	30
10.3 Clinical Laboratory Evaluations	31
10.4 Vital Signs.....	31
10.5 Physical Examination	31
10.7 Eastern Cooperative Oncology Group (ECOG)	32
REFERENCES.....	33
Appendix 1. Partial Date Conventions	34
Appendix 2. Response Evaluation Criteria in Solid Tumors (RECIST Version 1.1)	36
Appendix 3. Eastern Cooperative Oncology Group (ECOG) Performance Status	44

1. *What is the primary purpose of the study?*

LIST OF ABBREVIATIONS AND DEFINITIONS OF TERMS

Abbreviation Definition

AE	Adverse Event
ANC	Absolute Neutrophil Count
ATC	Anatomical Therapeutic Chemical
AUC	Area Under the Concentration-Time Curve

BMI	Body Mass Index
CI	Confidence Interval
Cmax	Peak Concentration
CR	Complete Response
CTCAE	Common Terminology Criteria for Adverse Event
ctDNA	Circulating Tumor DNA
DLT	Dose Limiting Toxicity
DoR	Duration of Response
ECG	Electrocardiogram
ECOG	Eastern Cooperative Oncology Group
eCRF	Electronic Case Report Form
EGFR	Epidermal Growth Factor Receptor
EOT	End of Treatment

HCV	Hepatitis C Virus
HED	Human Equivalent Dose
HIV	Human Immunodeficiency Virus
ICH	International Conference on Harmonisation
IV	Intravenous(ly)

MTD	Maximum Tolerated Dose
MUGA	Multi-gated Acquisition
NCI	National Cancer Institute
NE	Not Evaluable
NSCLC	Non-Small Cell Lung Cancer
ORR	Objective Response Rate
PD	Progressive Disease
PFS	Progression-Free Survival
PK	Pharmacokinetics
PD	Pharmacodynamics
PP	Per Protocol
PR	Partial Response
PS	Performance Status
QTc	corrected QT (interval)
QTcF	QT Interval Corrected Using Fridericia's Formula
RECIST	Response Evaluation Criteria In Solid Tumors

RP2D	Recommended Phase 2 Dose
SAE	Serious Adverse Event
SD	Stable Disease
SMC	Safety Monitoring Committee
SoA	Schedule of Activities
T _{1/2}	Half-life
AESI	Adverse Event of Special Interest
TKI	Tyrosine Kinase Inhibitor
Tmax	Time to Peak Concentration
WHO	World Health Organization

1. INTRODUCTION

This document outlines the statistical methods to be implemented during the analyses of data collected within the scope of Bridge Biotherapeutics Protocol Number BBT176-ONC-001 entitled “A Phase 1 / 2, Open-Label Study to Assess the Safety, Tolerability, Pharmacokinetics, and Anti-tumor Activity of BBT-176 in Patients with Advanced Non-Small Cell Lung Cancer (NSCLC) who Progressed Following Prior Therapy with an Epidermal Growth Factor Receptor Tyrosine Kinase Inhibitor (EGFR TKI) Agent.” The statistical methods described herein are based on the analyses proposed in the Protocol version [REDACTED]. This study is designed in 2 parts designated Part 1 (dose-escalation study) and Part 2 [REDACTED].

2. SUMMARY OF KEY PROTOCOL INFORMATION

2.1 Study Objectives and Endpoints

2.1.1 Primary Objectives

Table 1 shows the primary objectives and endpoints of Part 1 (dose-escalation study).

Table 1. Part 1 Primary Objectives and Endpoints

Part 1 Primary Objectives	Endpoints
<ul style="list-style-type: none"> To investigate the safety and tolerability of BBT-176 when given orally alone or in combination with cetuximab to patients with locally advanced or metastatic NSCLC with EGFR mutation To determine the MTD and/or RP2D of BBT-176 	Incidence of adverse events [AEs] and clinical laboratory abnormalities defined as DLTs

Table 2 shows the primary objectives and endpoints of Part 2 [REDACTED].

Table 2. Part 2 Primary Objectives and Endpoints

Part 2 Primary Objectives	Endpoints
To evaluate the anti-tumor activity of BBT-176 in selected mutant types by evaluation of ORR using RECIST version 1.1	Objective Response Rate [ORR] per RECIST version 1.1

2.1.2 Secondary Objectives

Table 3 shows the secondary objectives and endpoints of Part 1 [REDACTED].

Table 3. Secondary Objectives and Endpoints

Part 1 Secondary Objectives	Endpoints
<ul style="list-style-type: none"> To characterize the PK profile of BBT-176 after a single oral dose and steady state after multiple doses when given orally alone and in combination with cetuximab 	<ul style="list-style-type: none"> Plasma concentrations of BBT-176 PK parameters ($t_{1/2}$, C_{max}, t_{max}, AUC, C_{trough})

<ul style="list-style-type: none"> To evaluate the anti-tumor activity of BBT-176 in selected mutant types by evaluation of the ORR using RECIST version 1.1 	<ul style="list-style-type: none"> ORR per RECIST version 1.1
---	--

Table 4 shows the secondary objectives and endpoints of Part 2 [REDACTED]

Table 4. Secondary Objectives and Endpoints

Part 2 Secondary Objectives	Endpoints
<ul style="list-style-type: none"> To evaluate the anti-tumor activity of BBT-176 in selected mutant types by evaluation of the Duration of Response (DoR) and the Progression-Free Survival (PFS) using RECIST version 1.1 	<ul style="list-style-type: none"> DoR per RECIST version 1.1 PFS per RECIST version 1.1
<ul style="list-style-type: none"> To evaluate the safety and tolerability of BBT-176 at the PR2D 	<ul style="list-style-type: none"> Incidence of adverse events [AEs] and clinical laboratory abnormalities defined as DLTs
<ul style="list-style-type: none"> To characterize the snapshot PK of BBT-176 at steady state after multiple oral doses 	<ul style="list-style-type: none"> Plasma BBT-176 concentrations at steady state PK parameters (C_{max}, t_{max})

2.1.3 Exploratory Objectives

Table 5 shows the exploratory objectives and endpoints of Part 1 (dose-escalation study).

Table 5. Exploratory Objectives and Endpoints

Part 1 Exploratory Objectives	Endpoints
<ul style="list-style-type: none"> To investigate the DoR and PFS using RECIST version 1.1 	<ul style="list-style-type: none"> DoR per RECIST version 1.1 PFS per RECIST version 1.1
<ul style="list-style-type: none"> To investigate the relationship between the anti-tumor activity and blood-borne biomarkers (i.e., circulating-tumor DNA [ctDNA]) 	<ul style="list-style-type: none"> Pharmacodynamic biomarker (ctDNA)

Table 6 shows the exploratory objectives and endpoints of Part 2 [REDACTED]

Table 6. Exploratory Objectives and Endpoints

Part 2 Exploratory Objectives	Endpoints
<ul style="list-style-type: none"> To investigate the relationship between the anti-tumor activity and blood-borne biomarkers (i.e., circulating-tumor DNA [ctDNA]) 	<ul style="list-style-type: none"> Pharmacodynamic biomarker (ctDNA)

2.2 Study Design



2.2.1 Part 1 Study Design: Dose-Escalation Study



○ [REDACTED]
[REDACTED]

[REDACTED]
[REDACTED]
[REDACTED]

2.2.2 Part 1 Study Design: RP2D Exploration study

[REDACTED]
[REDACTED]
[REDACTED]
[REDACTED]
[REDACTED]
[REDACTED]
[REDACTED]
[REDACTED]

2.2.3 Part 2 Study Design: [REDACTED]

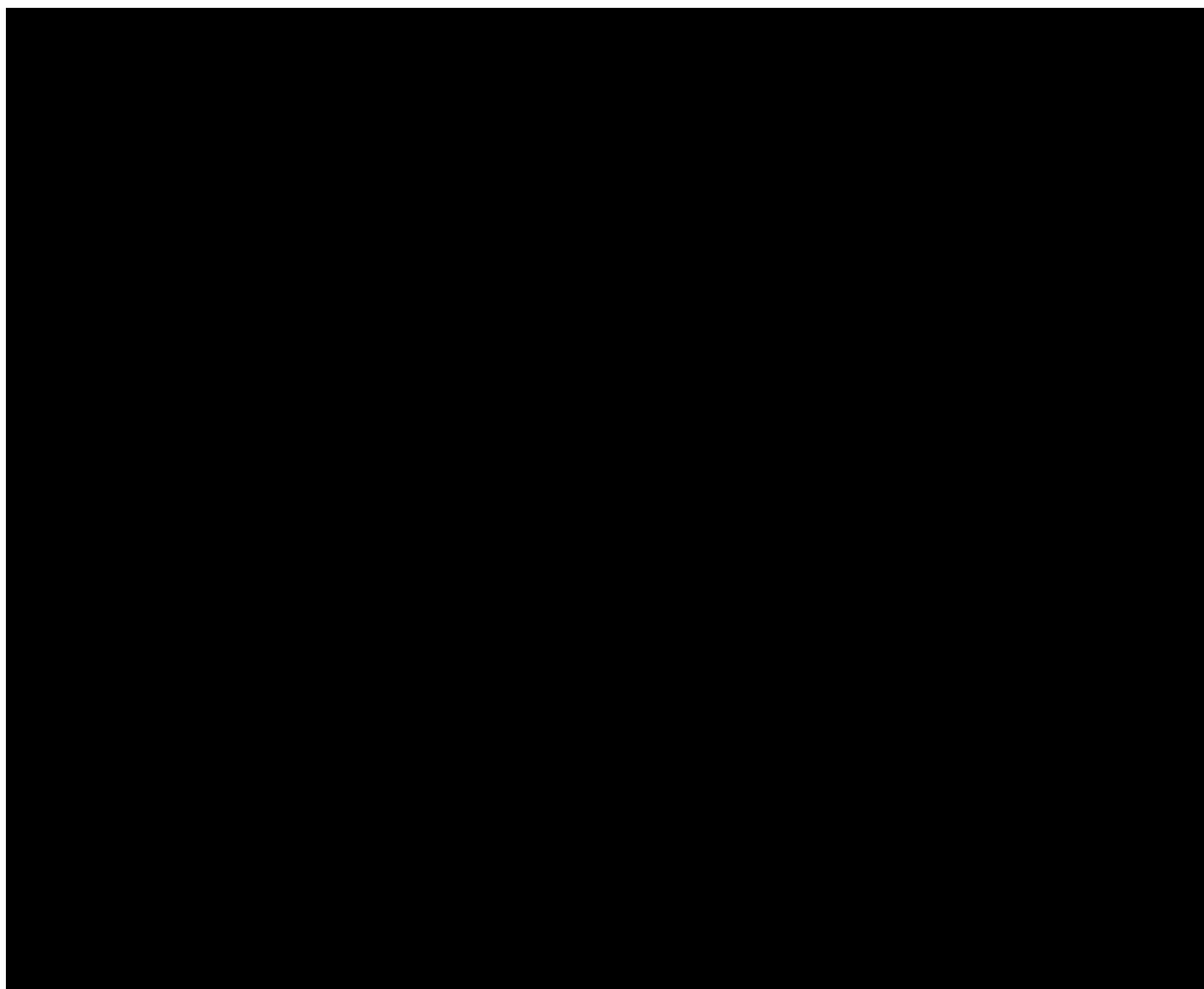
[REDACTED]
[REDACTED]
[REDACTED]
[REDACTED]
[REDACTED]

[REDACTED]	[REDACTED]	[REDACTED]

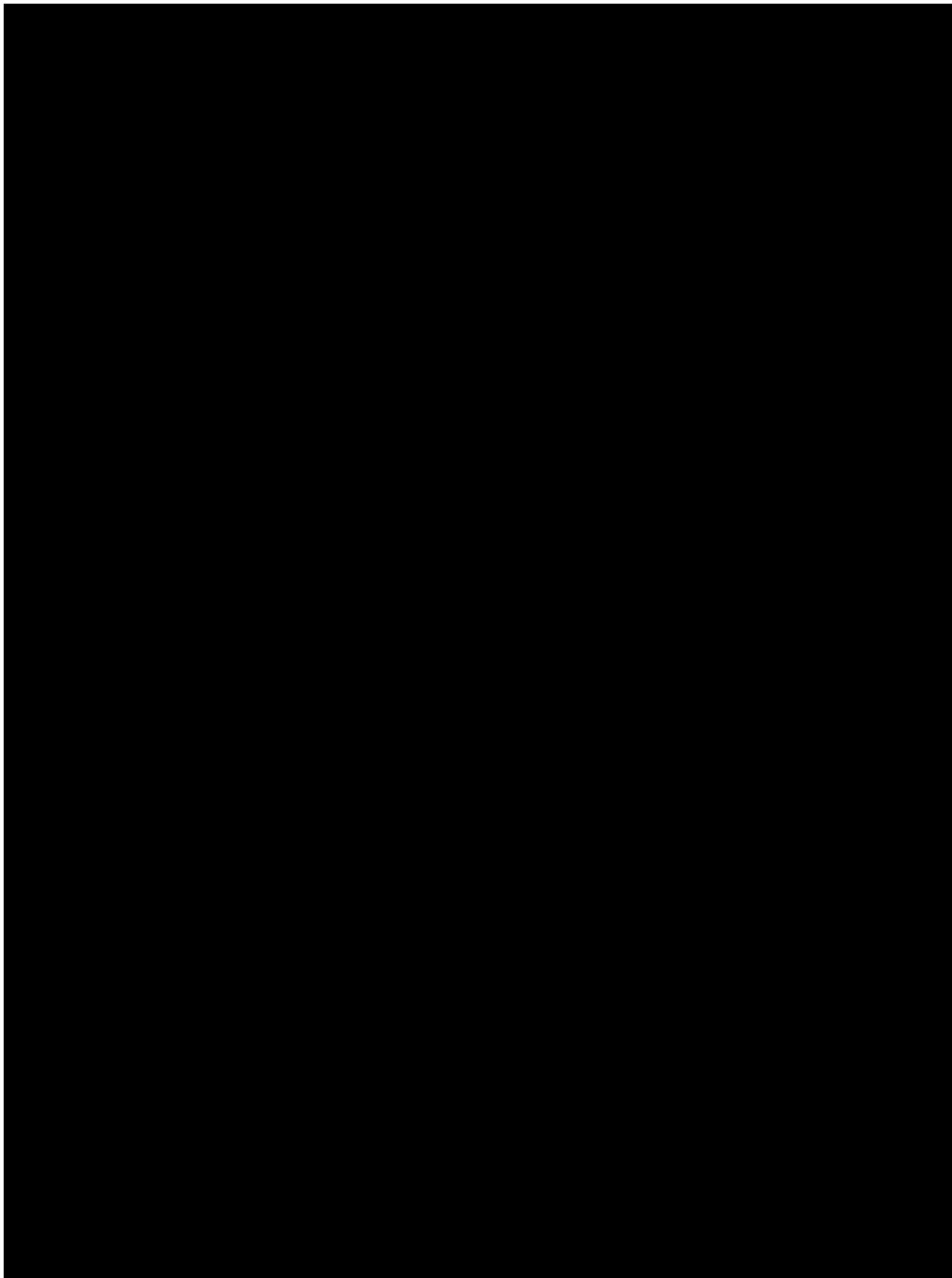
2.3 Schedule of Activities (SOA)

[REDACTED]
[REDACTED]

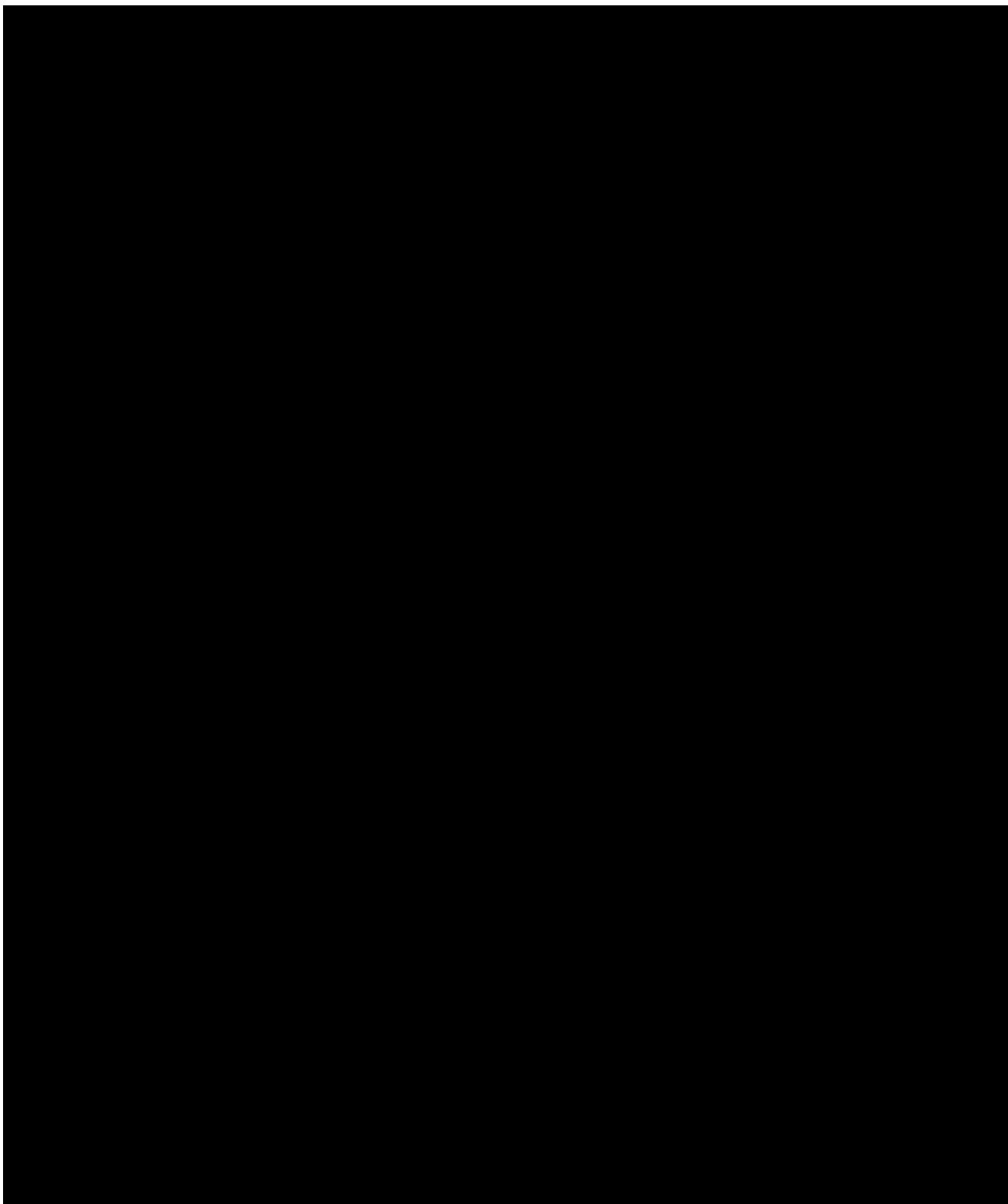
[REDACTED]							
------------	------------	------------	------------	------------	------------	------------	------------



STATISTICAL ANALYSIS PLAN
BBT176-ONC-001



STATISTICAL ANALYSIS PLAN
BBT176-ONC-001



3. DETERMINATION OF SAMPLE SIZE

3.1 Part 1

Dose-Escalation: up to █ patients with locally advanced or metastatic NSCLS with EGFR mutation are planned for enrollment. █
█
█
█

3.2 Part 2

█
█
█
█
█

█
█
█
█
█

█
█
█
█
█
█
█
█
█
█
█
█
█

█
█
█

4. GENERAL ANALYSIS CONSIDERATION

4.1 Data Summaries

The statistical analyses will be reported using summary tables, figures, and data listings. The International Conference on Harmonization (ICH) numbering convention will be used for all tables, listings, and figures. Unless otherwise specified, all statistical testing will be two-sided and will be performed at the 0.05 significance level. Unless otherwise specified, all confidence intervals (CIs) will be constructed at the 95% confidence level.

The results of the statistical analyses will be reported using summary tables, figures, and data listings. The International Conference on Harmonization (ICH) numbering convention will be used for all tables, listings, and figures. [REDACTED]

[REDACTED] For other estimates, unless otherwise specified, 95% CIs will be provided.

Continuous variables will be summarized with means, standard deviations, medians, minimums, and maximums. Categorical variables will be summarized by counts and by percentages of patients in corresponding categories. Time-to-event variables will be summarized using the Kaplan-Meier estimates. All summary tables will be presented by part and cohort. Baseline summaries will also include a total summary column.

Individual subject data obtained from the electronic case report forms (eCRFs), and any derived data will be presented by subject in data listings. Data from all assessments, whether scheduled or unscheduled, will be included in the listings. Summary tables will be presented for scheduled assessments only.

All analyses and tabulations will be performed using SAS® v9.4 on a SAS server platform. Tables and listings will be presented in rich text format. Upon completion, all SAS programs will be validated by an independent programmer. In addition, all program output will undergo a senior level statistical review. The validation process will be used to confirm that statistically valid methods have been implemented and that all data manipulations and calculations are accurate. Checks will be made to ensure accuracy, consistency with this plan, consistency within tables, and consistency between tables and corresponding data listings. Upon completion of validation and quality review procedures, all documentation will be collected and filed by the project statistician or designee.

4.2 Data Handling

4.2.1 Baseline Value

Unless otherwise specified, the baseline value is defined as the last value obtained before the date and time of the first BBT-176 dose. Post-baseline values are defined as value obtained after the first dose of BBT-176. Change from baseline is defined as a post-baseline value minus the baseline value.

4.2.2 Conventions

Percentages are showed as a rate relative to the total number of subjects with a percent sign (%). For frequency counts of categorical variables, categories whose counts are zero will be displayed for the sake of completeness. For categories with zero counts no percentage will be displayed.

For continuous variables, the precision of the original measurements will be maintained in summaries and listings, when possible. Generally, means, medians and standard deviations will be presented with an increased level of precision; means and medians will be presented to one more decimal place than the raw data, and the standard deviations will be presented to two more decimal places than the raw data.

Summaries of continuous variables that have some values recorded using approximate values (e.g. < or >) will use imputed values. That is, approximate values will be imputed using the closest exact value for that measurement. Listing will present the data in the original format. For tables where rounding is required, rounding will be done to the nearest round-off unit. For example, if the round-off unit is the ones place (i.e., integers), values \geq XX.5 will be rounded up to XX+1 while values < XX.5 will be rounded down to XX.

Listings will present the data in the original format. Durations of AEs and concomitant medications will not be derived using imputed dates but, rather, be set to missing when an event has a partially or completely missing start date or end date.

4.2.3 Standard Calculations

Variables requiring calculation will be derived using the following formula.

- **Study Day** – Study Day 1 is defined as the date of the first treatment. For a given event date, Study Day is calculated relative to the date of first dose of study drug.
$$\text{Study Day} = [\text{Event Date} - \text{First Dose Date}] + 1 \text{ (in days)},$$
where the event date is on or after the first treatment date.
$$= [\text{Event Date} - \text{First Dose Date}] \text{ (in days)},$$
where the event date is before the first treatment date.
- **Duration (Days)** – A duration in days is calculated as the number of days between one date (Date1) and another later date (Date2), including both the start and end days.
$$\text{Duration (days)} = [\text{Date2} - \text{Date1}] + 1 \text{ (in days)}$$
- **Duration (Months)** – A duration in months is calculated as the duration in days divided by 30.4375 (~365.25/12), rounded to one decimal place.
- **Duration (Years)** – A duration in years is calculated as the duration in days divided by 365.25, rounded to one decimal place.
- **Age (Years)** – Age is calculated as the number of years from the date of birth to the date of informed consent using the following formula.
$$\text{Age (years)} = [\text{Date informed consent signed} - \text{Date of birth (in days)}] / 365.25,$$
rounded down to the closest year.

- **Height** – Height entries made in inches (in) are converted to centimeters using the following formula:
$$\text{Height (cm)} = \text{Height (in)} * 2.54$$
, rounded to the one decimal place.
- **Weight** – Weight entries made in pounds (lbs) are converted to kilograms (kgs) using the following formula:
$$\text{Weight (kg)} = \text{Weight (lb)} / 2.2046$$
, rounded to the one decimal place.
- **Body Mass Index (BMI)** – BMI is derived using following formula.
$$\text{BMI (kg/m}^2\text{)} = \text{Weight (kg)} / [(\text{Height (cm)} / 100)^2]$$
, rounded to one decimal place.
- **Temperature** – Temperature entries in degrees Centigrade are converted to degrees Fahrenheit using the following formula:
$$\text{Temperature (degrees Fahrenheit)} = 1.8 * \text{Temperature (degrees Centigrade)} + 32$$
, rounded to one decimal place.

5. ANALYSIS POPULATION

6. STUDY POPULATION

6.1 Patient Disposition

Patient disposition information will be summarized by each dose escalation cohort, overall patients in Part 1 and each expansion cohort, overall patients in Part 2. Summaries will include the number of enrolled patients, the number of patients in each analysis population, and the number of patients completing the study, and the primary reason for discontinuation.

6.2 Demographic and Baseline Characteristics

Demographic variables include: age, sex, ethnicity, and race. Age will be calculated in years relative to the informed consent date. Other baseline characteristics include: medical history, previous cancer treatment, previous tumor tissues collection, weight, height, and BMI.

Descriptive statistics will be presented for medical history, age, weight height, BMI. Frequency counts and percentages will be presented for sex, ethnicity and race. Demographic and baseline characteristics will be summarized by each dose escalation cohort, over all patients in Part 1 and each expansion cohort, overall patients in Part 2. All demographic and baseline characteristics will be presented in a data listing.

6.3 Medical History

The number and percentage of subjects will be summarized by cohort and overall for each body system.

6.4 Prior and Concomitant Medication

Verbatim terms on eCRFs will be mapped to Anatomical Therapeutic Chemical (ATC) class and Generic Drug Names using [REDACTED]

All medications taken within 28 days prior to the administration of study drug and during the course of the study (includes over-the-counter and prescription drugs) and any non-drug treatment/procedures must be recorded on the eCRF, including the reason for treatment, generic name of the drug, dosage, route, and date of administration.

Prior medications are those that start taken prior to the administration of study drug, even if they continue after dosing. Concomitant medications are those that start on or after the date of the first dose of study medication. See Appendix 1 for handling of partial dates for medications; in cases where it is not possible to define a medication as prior or concomitant, the medication will be classified by the worst case—ie, concomitant.

Prior and concomitant medications will be summarized by WHO ATC class include the therapeutic (i.e., the second level of Anatomical Therapeutic Chemical [ATC] classification system code, that is, corresponding to the first 3 figures), preferred name (the fourth level of the ATC code, that is, corresponding to the first 5 figures) and verbatim text and generic drug name for the Safety population. Summaries will be presented by each dose escalation cohort, over all patients in Part 1 and each dose expansion cohort, over all patients in Part 2. The number and percentage of patients using each medication will be presented. Patients may have more than one medication per ATC class and generic drug name. At each level of patient summarization, a patient is counted once if he/she reported one or more medications at that level. Each summary will be ordered in descending order of incidence of ATC class and generic drug name within each ATC class.

7. EFFICACY ANALYSES

All efficacy analysis will be based on the PP population. This study will utilize the definitions of tumor response as described in Response Evaluation Criteria in Solid Tumors (RECIST) version 1.1 (Appendix 2).

7.1 Tumor Response

Tumor assessment (RECIST version 1.1) will be performed at Screening and every 6 weeks from Cycle 1 Day 1 until progression of disease, EOT, or until the patient commences further anti-cancer treatment, whichever comes first. All efficacy data will be summarized [REDACTED]

[REDACTED] Baseline as well as change from baseline and percentage change from baseline measurements of the sum of the longest diameters for the target lesions will be summarized at each scheduled evaluation using descriptive statistics including mean, standard deviation, median, minimum and maximum values.

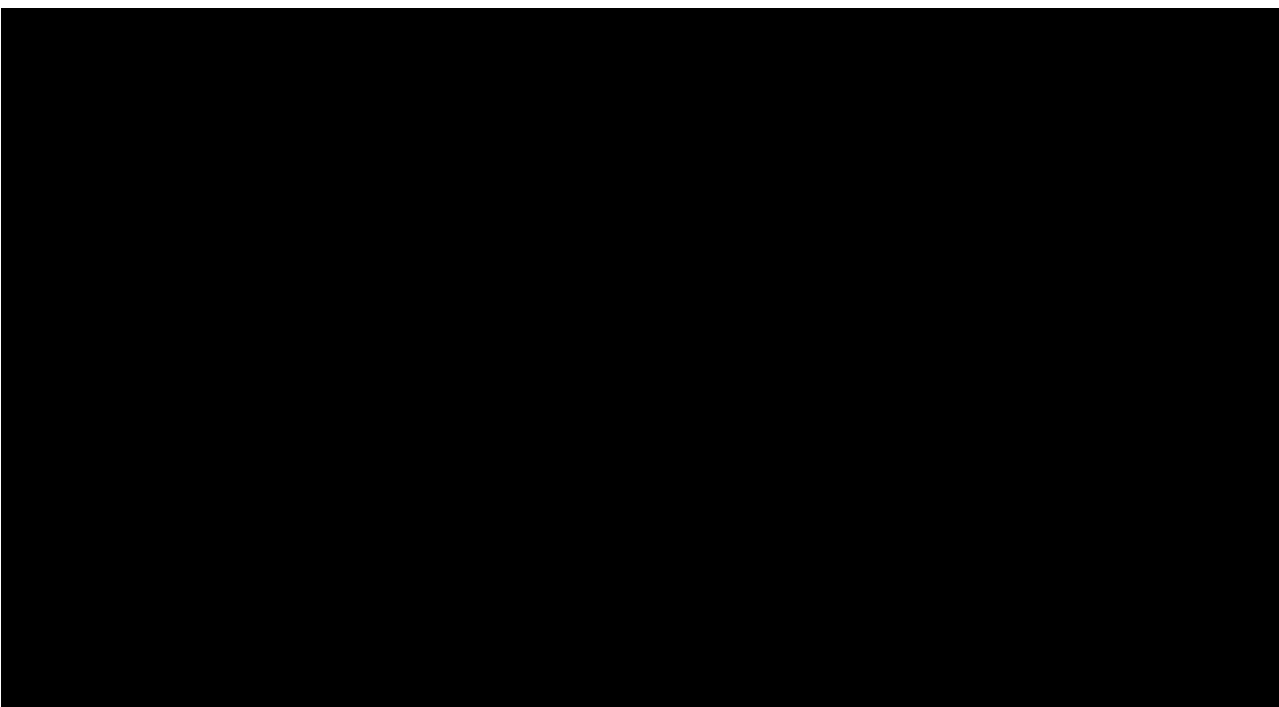
7.1.1 Overall Time Point Response

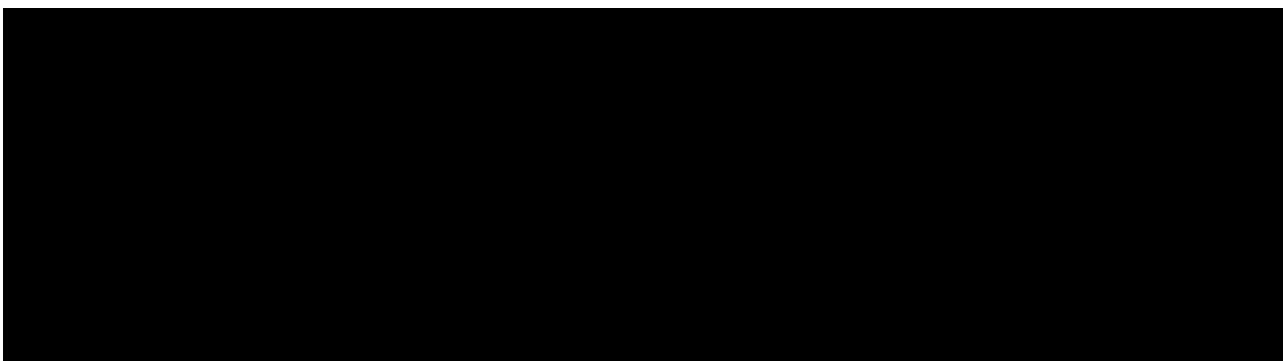
The investigator's assessment of a patient's target lesion response, non-target lesion response, and appearance of new lesions (collected from the eCRF) will be used to determine the overall tumor response at each time point as shown [REDACTED]

The possible overall tumor responses at a time point, from best to worst, are as follows.

- Complete Response (CR)
- Partial Response (PR)
- Stable Disease (SD)
- Progressive Disease (PD)

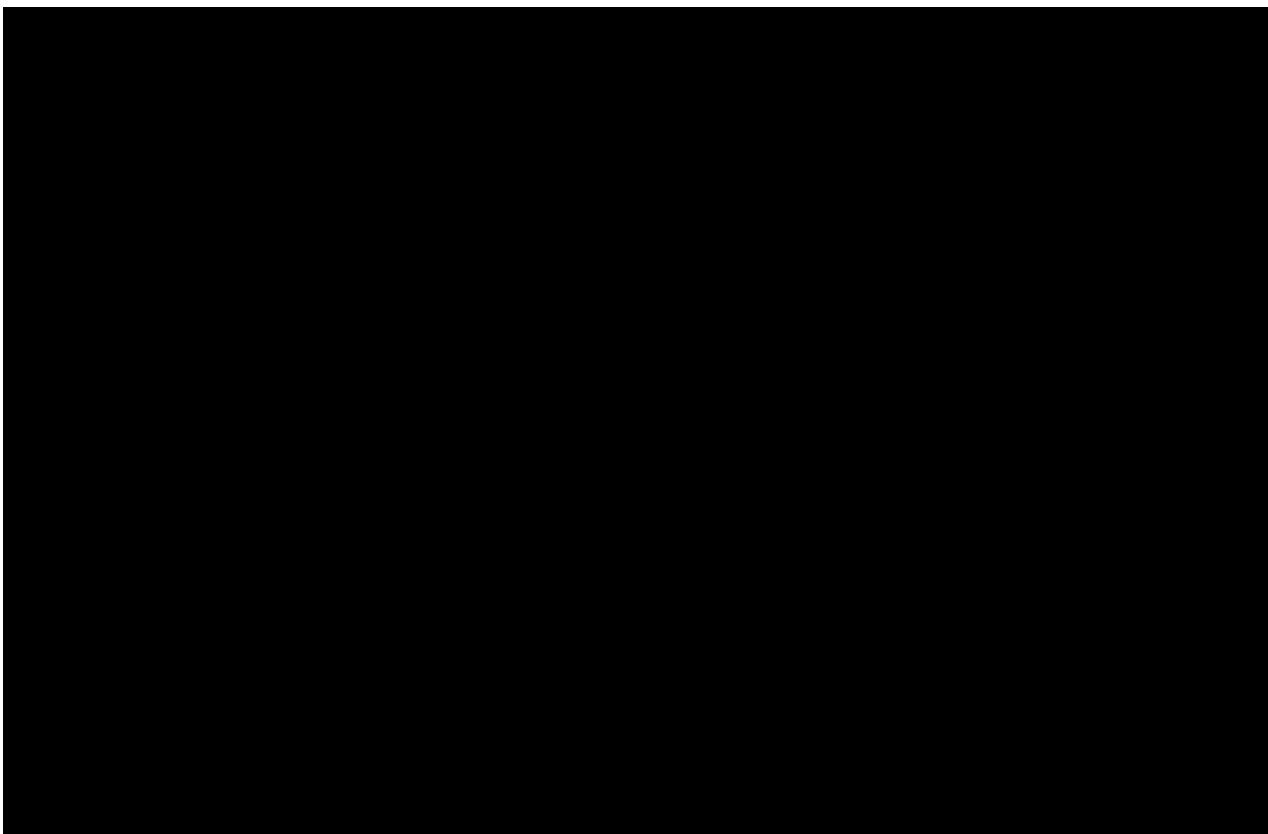
If no assessment is done or a response cannot be determined, it is designated as Not Evaluable (NE).



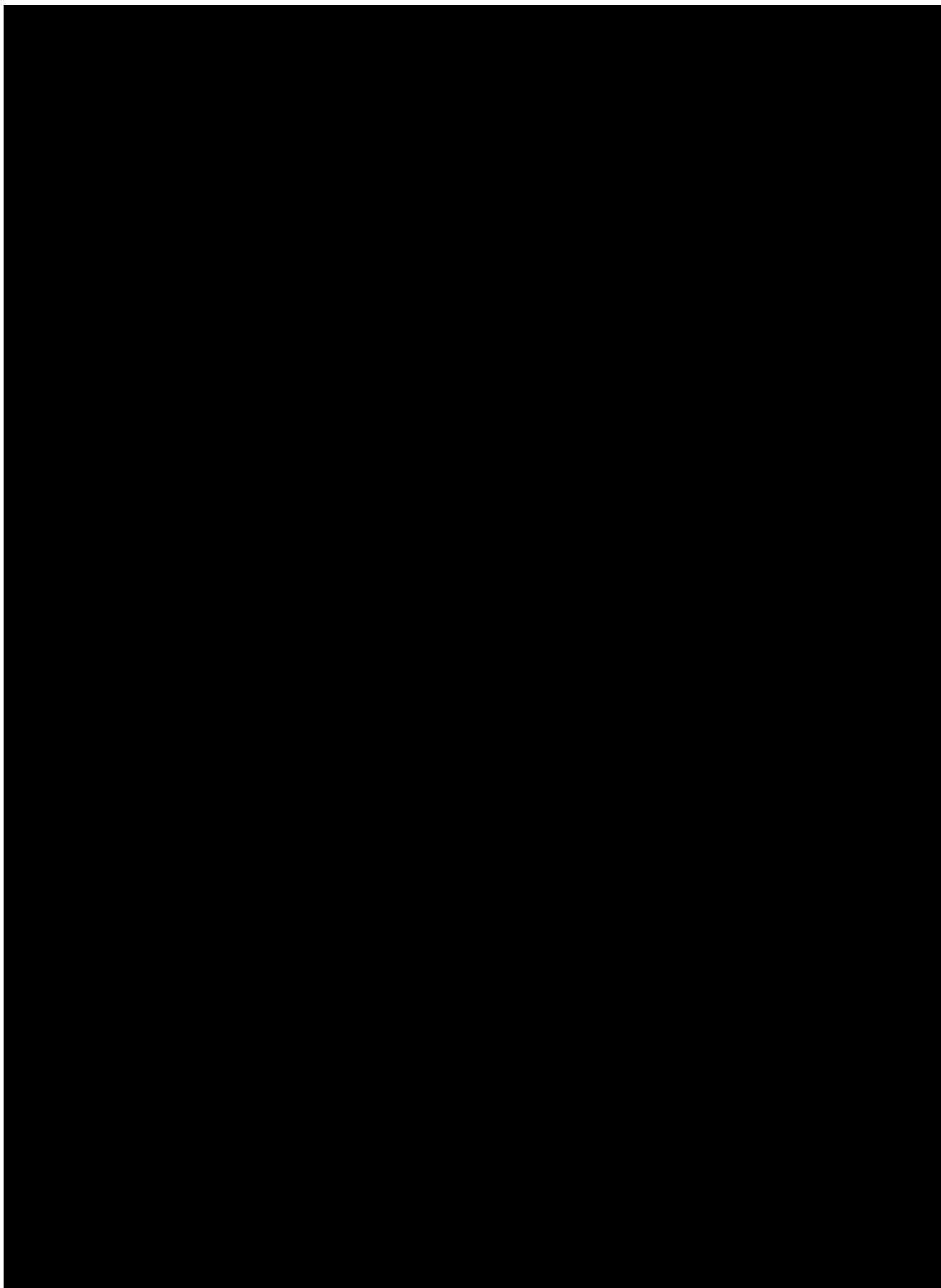


7.1.2 Best Response

The best responses for pairs of time point responses (when confirmation of CR and PR are required) are determined as shown



STATISTICAL ANALYSIS PLAN
BBT176-ONC-001



- [REDACTED]
- [REDACTED]
- [REDACTED]

7.1.3 Best Overall Response

A Best Response is determined for each sequential pair of post-treatment time points (and in some cases, triplets of sequential time points). Then, the Best Overall Response for a patient (when confirmation of CR and PR are required) is the best of these Best Responses.

7.2 Objective Response Rate (ORR)

The ORR will be estimated by the number of patients with a best overall response of CR or PR divided by the total number of patients who are evaluable for efficacy. The best overall response is the best response recorded from the start of the treatment until progression of disease, EOT, or until the patient commences further anti-cancer treatment, whichever comes first.

7.3 Duration of Response (DoR)

DoR is calculated for every patient with a response to therapy (PR and CR) and is defined as the number of days from the date of initial response to the date of the first documented disease progression/relapse (including clinical progression) or death, whichever occurs first. The exact date of a response is based on the tumor response assessment dates as defined in Section 7.1.1. If disease progression/relapse does not occur, DoR is censored as of the date of their last imaging exam of target or non-target lesions. Patients without an assessment of disease progression/relapse who discontinue the study and indicate progressive disease or clinical progression as the primary reason for ending treatment will be counted as having progressed on their last dose date of BBT-176—whichever date is earlier.

7.4 Progression-Free Survival (PFS)

PFS will be calculated for each patient as the number of days from the first day of treatment to the date of the first documented disease progression/relapse or date of death, whichever occurs first. The exact date of disease progression is based on the tumor response assessment dates as indicated in Section 7.1.1. If disease progression/relapse does not occur, PFS is censored as of the date of their last imaging exam of target or non-target lesions. However, for patients without an assessment of disease progression/relapse who escalate to a higher dose, PFS is censored as of the date of their first dose of study medication in the later dose cohort. Patients without an assessment of disease progression/relapse who discontinue the study and indicate progressive disease or clinical progression as the primary reason for ending treatment will be counted as having progressed on their last dose date of BBT-176—whichever date is earlier.

7.5 Handling Missing, Unused, and Spurious Data

All available efficacy data will be included in the data listings and tabulations. No imputation of values for missing data will be performed.

7.6 [REDACTED]

[REDACTED]

[REDACTED]

[REDACTED]

[REDACTED]

7.7 Methods of Efficacy Analysis

The extent of disease will be assessed using the sum of the longest diameter from the target lesions. Change from baseline and percent change from baseline measurements will be summarized at each scheduled evaluation using descriptive statistics.

The investigators' assessments of patients' target lesion response, non-target lesion response, and appearance of new lesions (collected from the eCRF) are used to determine the overall [tumor] response at each time point as shown in Section 7.1.1. The investigators' assessments will be shown in listings.

Next, as indicated in Section 7.1.2, the best [tumor] response for sequential pairs (or triplets) of time point responses (recorded from the start of study treatment until the end of treatment) are determined as shown in Table 14 and in the subsequent discussion. Finally, as indicated in Section 7.1.3, the Best Overall Response for a patient—when confirmation of CR and PR are required—is the best of these Best Responses. The number and percent of patients in each category of Best Overall Response (CR, PR, SD, and PD) will be summarized by cohort.

The ORR will be presented with frequencies and percentages with 80% confidence interval. The frequency and percentages of patients with a best overall response of CR, PR, with stable disease (SD), or with PD will be summarized by cohort, additionally. 80% confidence intervals will be provided.

DoR will be summarized using the Kaplan-Meier estimate including median estimate, first quartile, 3rd quartile.

Progression-free survival in days will be summarized using the Kaplan-Meier to estimate the median PFS. The quartiles, including the median time-to-event and its two-sided 95% CI will be presented for each endpoint. Summaries will be presented for each cohort and overall.

8. PHARMACOKINETIC (PK) ANALYSES

All pharmacokinetic analyses will be based on the PK populations. This SAP describes the method of deriving PK using non-compartmental analysis applying appropriate software, i.e., Phoenix™

WinNonlin® (Version 8.0 or higher, Certara Corporation), and all other analyses and summaries using SAS® (Version 9.4 or higher, SAS Institute Inc.).

The PK analyses will be conducted for BBT-176 in plasma, will include a listing, summary, and figure of concentrations, and derived PK parameters as described below.

8.1 PK Sampling Schedule

Plasma samplings for PK analyses will be done according to [REDACTED]

8.2 Handling of the Below the Lower Limit of Quantification (BLQ) and the No Reportable Concentration Values

[REDACTED]
[REDACTED]
[REDACTED]
[REDACTED]
[REDACTED]
[REDACTED]
[REDACTED].

Subjects for whom there are insufficient data to calculate the PK parameters will be included in the concentration tables only and will be excluded from the PK and statistical analysis.

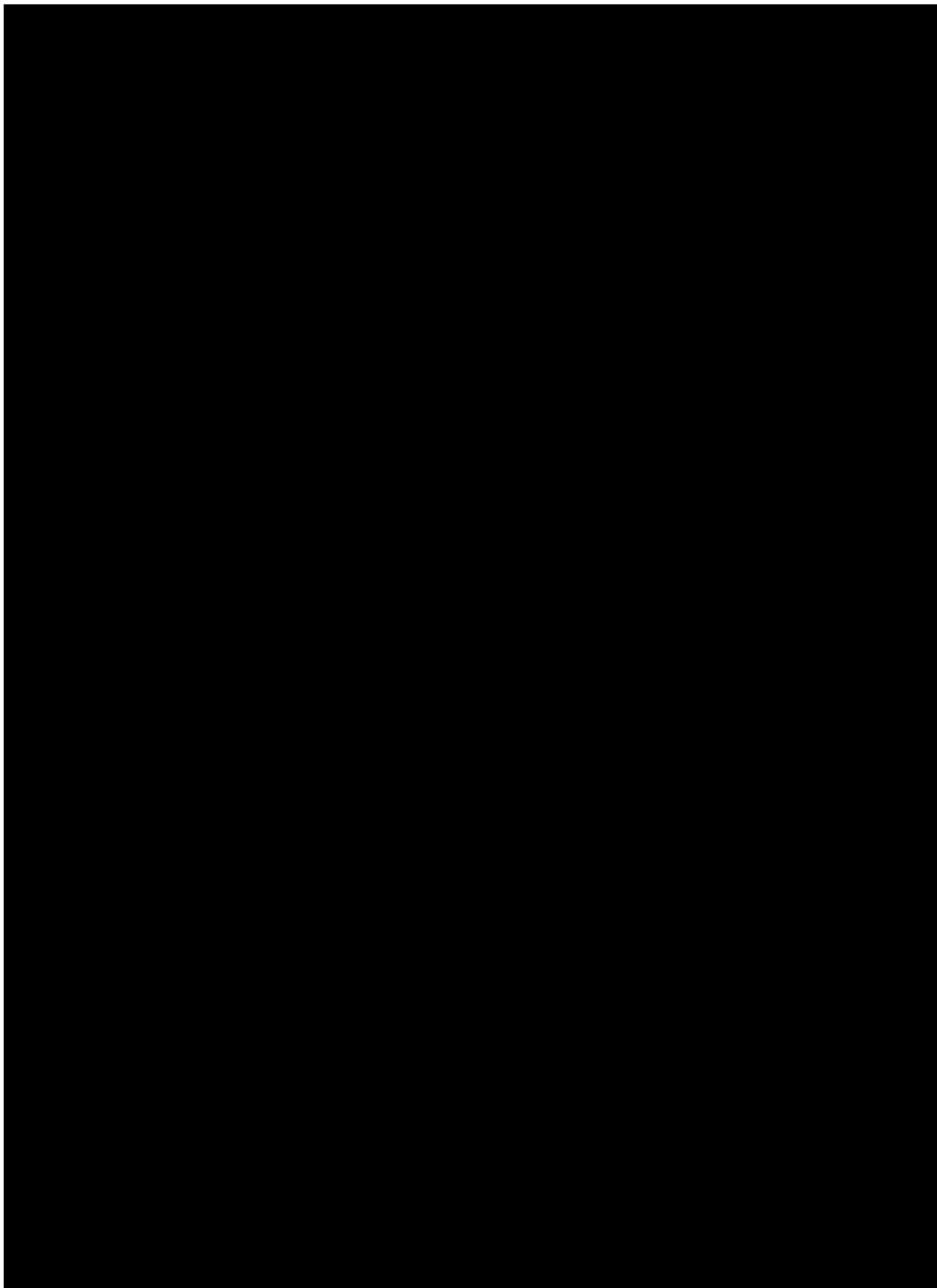
8.3 Handling of the Difference between the Scheduled and the Actual Sampling Times

The actual clock time for dosing and each collection time for the PK samples will be recorded. For all sampling times, the actual sampling times relative to dosing will be calculated as the difference between the actual clock time of sampling and the actual clock time of dosing. The actual post-dose sampling times relative to dosing expressed in hours and rounded off to three decimal digits will be used to calculate the PK parameters, except for pre-dose samples occurring prior to dosing, which will always be reported as zero (0.000), regardless of the time difference. Scheduled sampling times will be presented in concentration tables and mean graphs, while actual sampling times will be presented in the individual graphs.

8.4 Plasma PK Endpoints

[REDACTED]
[REDACTED]

STATISTICAL ANALYSIS PLAN
BBT176-ONC-001



[REDACTED]

[REDACTED]

[REDACTED]
[REDACTED]
[REDACTED]
[REDACTED]

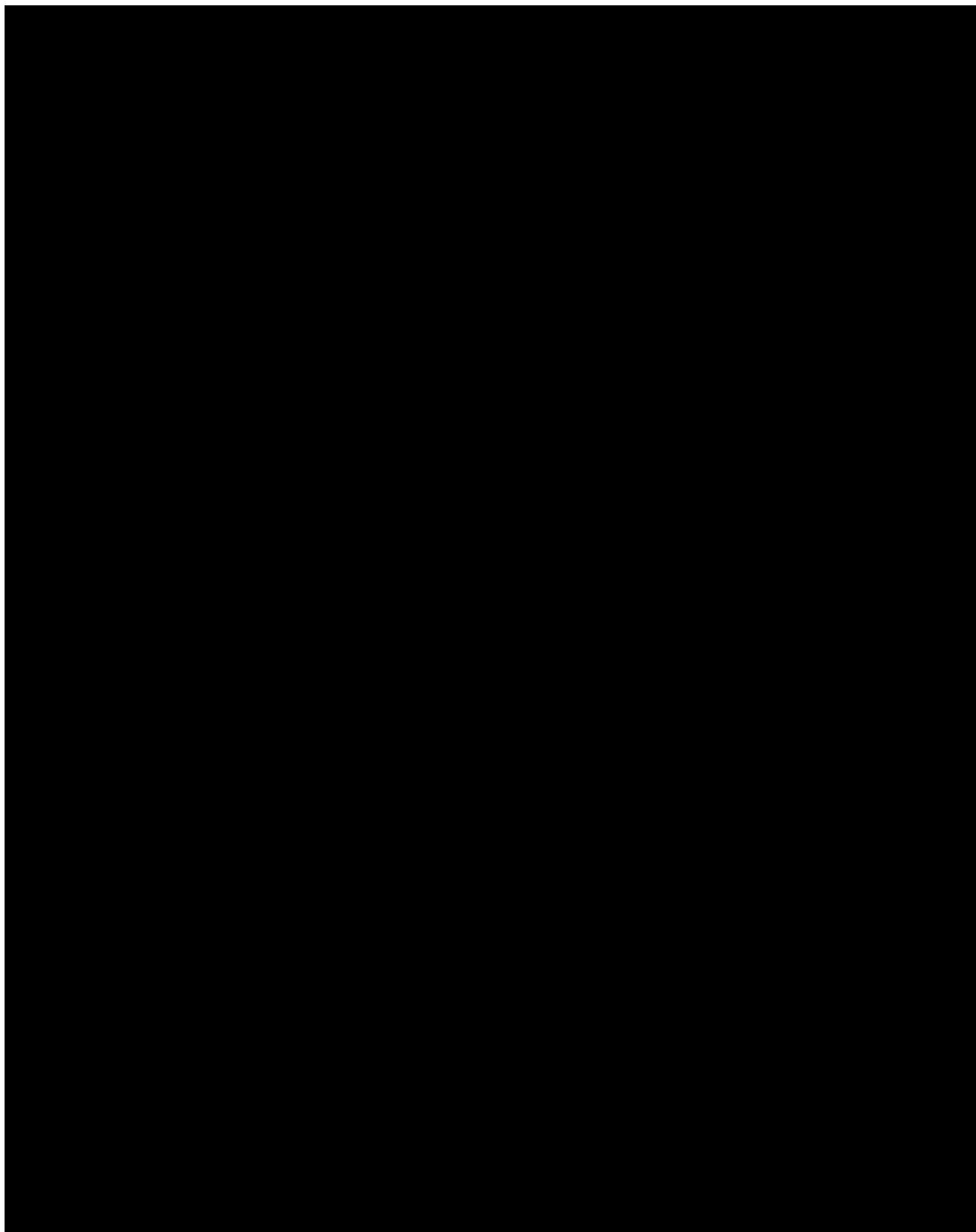
[REDACTED]

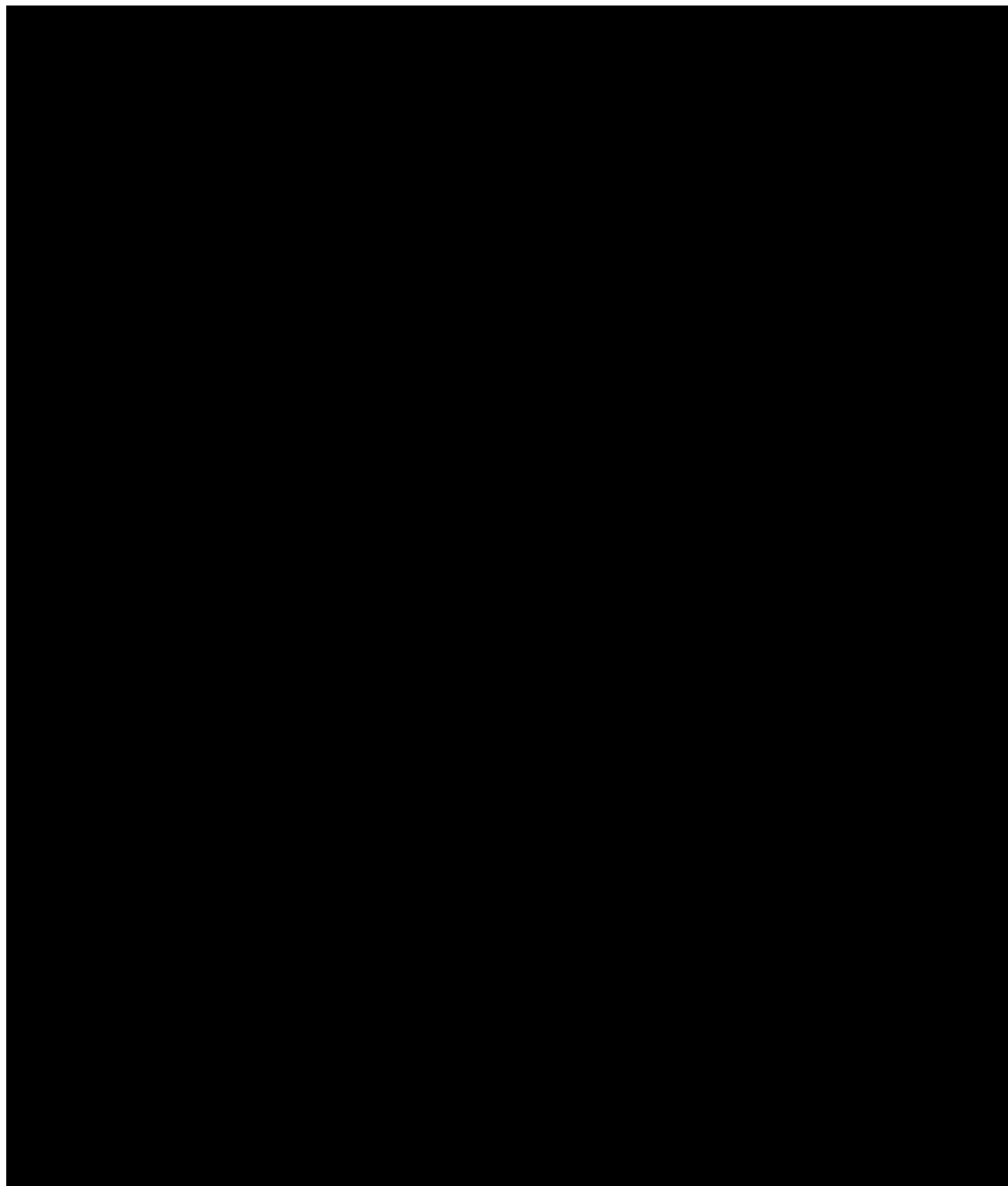
Additional PK parameters may be computed if deemed appropriate.

8.5 Statistical Analyses

[REDACTED]
[REDACTED]
[REDACTED]
[REDACTED]
[REDACTED]

STATISTICAL ANALYSIS PLAN
BBT176-ONC-001





9. PHARMACODYNAMIC (PD) ANALYSES



No formal statistical analysis of pharmacodynamics endpoints will be performed. Data listings will be created for each pharmacodynamics biomarker by cohort and patient.

Summary tables of values (number of observations, mean, median, standard error of the mean, standard deviation) will be created for each pharmacodynamics biomarker.

10. SAFETY ANALYSES

All safety analyses will be based on the Safety population. Safety variables to be assessed will include AEs, laboratory test results (hematology, clinical chemistry, and urinalysis), ECGs, weight, vital signs.

10.1 Extent of Exposure

Study drug exposure will be summarized for each dose escalation cohort and all patients in the dose escalation phase. Duration of treatment is defined as the last dose date minus the first dose date plus 1. Drug adjustments and interruptions are listed and summarized for each cohort.

10.2 Adverse Events (AEs)

An AE is any untoward, undesired, or unplanned event in the form of signs, symptoms, disease, or laboratory or physiologic observations occurring in a person given a test article (study drug) in a clinical study. The description found in the revised NCI CTCAE version 5.0 will be utilized for AE reporting.

All AEs will have their relationship to study treatment assessed as follows:

- Definite – The AE is *clearly related* to the study treatment
- Probable – The AE is *likely related* to the study treatment
- Possible – The AE *may be related* to the study treatment
- Unlikely – The AE is *doubtfully related* to the study treatment
- Unrelated – The AE is *clearly NOT related* to the study treatment

A serious AE (SAE) is an AE that:

- Results in death.
- Is life-threatening.
- Requires inpatient hospitalization or prolongation of an existing hospitalization.
- Results in a persistent or significant disability or incapacity.
- Results in a congenital anomaly or birth defect.
- Additionally, important medical events that may not result in death, be life-threatening, or require hospitalization may be considered SAEs when, based on appropriate medical judgment, they may jeopardize the patient and may require medical or surgical intervention to prevent one of the outcomes listed in this definition.

A Treatment-emergent AE (TEAE) is an AE with an onset day on or after the day of the initial dose of study drug.

AEs of special interest (AESI) is an AE that:

- Gastrointestinal, skin (>Grade 2), pulmonary (\geq Grade 2), and/or serious events requiring reporting to the Sponsor as an SAE.
- Any laboratory hepatic toxicity meeting Hy's law criteria for suspected severe drug induced liver injury, defined as a rise in serum aminotransferase (AST or ALT) of $>3\times$ ULN and total bilirubin $\geq 2 \times$ ULN, in the absence of any alternative reason (e.g., liver metastasis, liver disease, or concomitant medication)

AE terms recorded on the eCRF will be mapped to preferred terms using [REDACTED]

[REDACTED] An overall summary will be provided of the number and percentage of patients reporting AEs, TEAEs, treatment-related TEAEs, serious TEAEs, treatment-related serious TEAEs, AEs of special interest (AESIs), TEAEs leading to study drug withdrawal and TEAEs leading to death. In the overall summary, additionally, the number and percentage of patients with at least one severe TEAE, with at least one moderate TEAE but no severe TEAE and with only mild TEAEs will be provided.

The number and percentage of patients with AEs will be summarized for each dose level and overall in Part 1, and by cohort and overall in Part 2 according to the system organ class and preferred term. AEs (including AESI) will be summarized for overall frequency (number and percentage of patients), worst reported severity, and relationship to study treatment for each preferred term per patient. SAEs will be similarly summarized. Listings of death, SAEs, and AEs leading to early termination of study treatment or premature withdrawal from study will be also provided.

10.3 Clinical Laboratory Evaluations

Laboratory parameters (serum chemistry, hematology, coagulation, urinalysis, and serology test) will be presented in data listings. Abnormal values will be flagged as high or low relative to the local lab normal ranges, where applicable. Values that are deemed as abnormal, clinically significant (as collected on the eCRF) will also be flagged in the data listing.

Laboratory variables will be summarized using actual values and changes from baseline to scheduled time points. Laboratory values will also be categorized according to their NCI-CTCAE version 5.0 toxicity grade and tabulated by worst on-study toxicity grade. The baseline value of a variable is defined as the last value obtained on or before the date and time of the first BBT-176 dose.

10.4 Vital Signs

Vital signs (pulse, respiration, blood pressure, oral temperature, and weight) will be summarized by dose cohort for the dose escalation phase and for each expansion cohort using descriptive statistics at baseline and at each post-baseline time point. Actual values and changes from baseline to scheduled time points will also be summarized.

10.5 Physical Examination

Physical examination results will be summarized by shift tables from baseline compared to scheduled post-baseline time-points presenting the number and percentage of patients with normal /abnormal NCS / abnormal CS and not evaluable results.

10.6 Electrocardiogram (ECG)

The ECG data analysis will be conducted based on methodology recommended in the ICH E14 guidance, The Clinical Evaluation of QT/QTc Interval Prolongation and Proarrhythmic Potential for Non-antiarrhythmic Drugs. Descriptive statistics at baseline and at each scheduled post-baseline time point as well as changes from baseline will be summarized for each ECG parameter (heart rate, RR interval, PR interval, QRS interval, QT interval, QT interval corrected using Frederica's formula (QTcF)) by dose escalation cohort, overall patients in Part 1 and each expansion cohort and overall patients in Part 2. The number and percentage of patients with normal, abnormal and abnormal clinically significant categories will be provided for the overall ECG evaluation for each visit by dose escalation cohort.

The RR interval will be calculated from heart rate using the following formula:

- RR (milliseconds) = $60*1000 / [\text{heart rate in beats per minute}]$

Frederica's formula, which corrects the QT interval for the effect of heart rate, is

- QTcF (milliseconds) = $QT / [(RR/1000)^{0.33}]$ or, equivalently, $QT / [(60/\text{heart rate})^{0.33}]$

The RR value derived from heart rate will be used in any derivations (eg, QTcF) rather than the value collected on the eCRF. Further, the derived QTcF values rather than the QTcF collected in the eCRF will be used for all analyses. Both the collected and derived RR and QTcF values will be shown in the listings.

All ECG information as well as change from baseline will be listed by dose escalation cohort and over all patients in Part 1 and each expansion cohort and overall patients in Part 2.

In addition, a categorical summary of abnormal QTcF values will be presented for each dose escalation cohort and overall patients in Part 1 and each expansion cohort and overall patients in Part 2. Additionally, a frequency table with number and percentage of patients with noteworthy QTcF values (> 450 , > 480 and > 500 ms) during the study and of patients noteworthy QTcF changes from baseline (≥ 30 but < 60 ; ≥ 60 ms) will be provided for each dose escalation cohort,

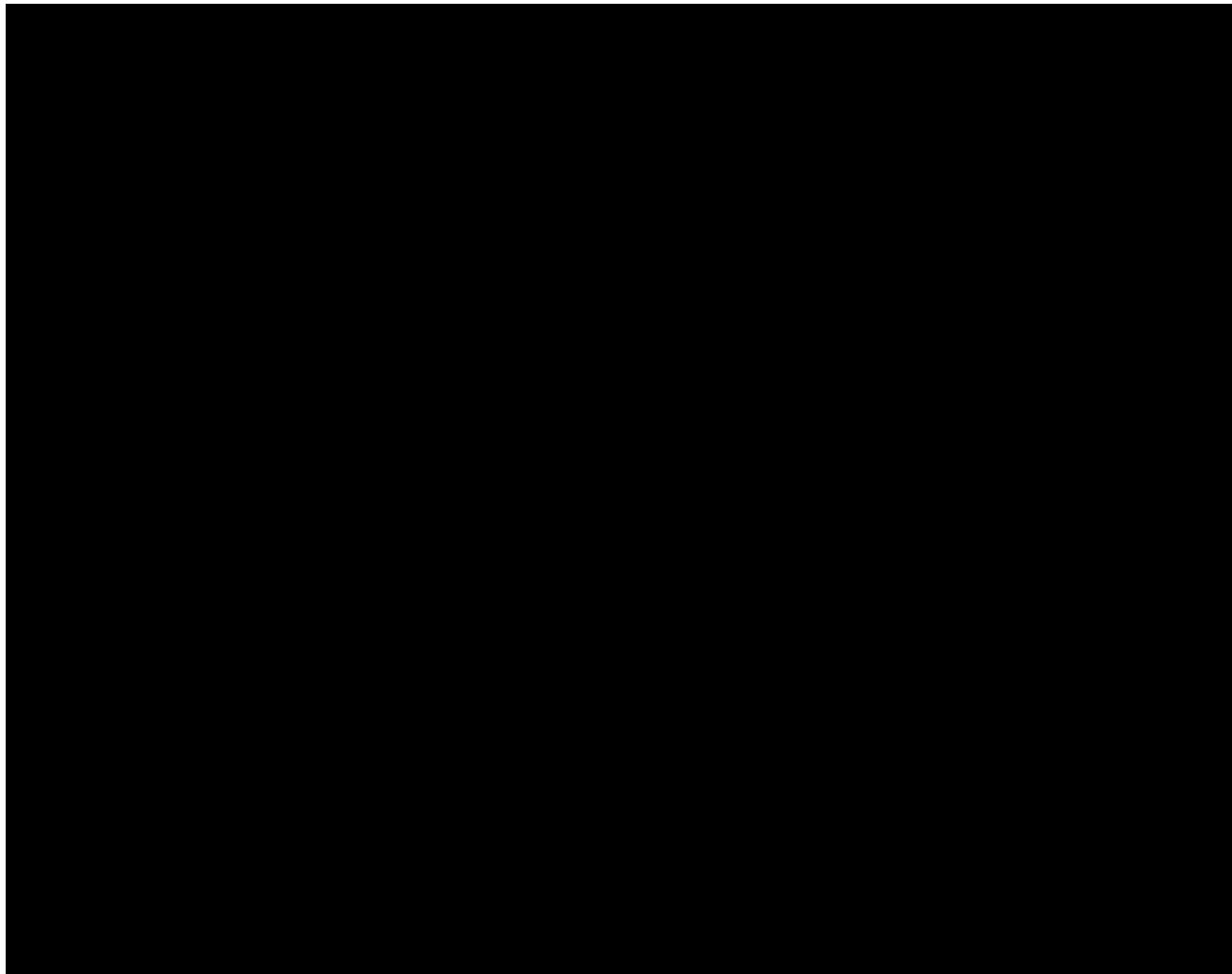
10.7 Eastern Cooperative Oncology Group (ECOG)

Eastern Cooperative Oncology Group (ECOG) performance status scores will be included in a data listing; shift tables assessing the worst post-baseline value will be presented (see Appendix 3).

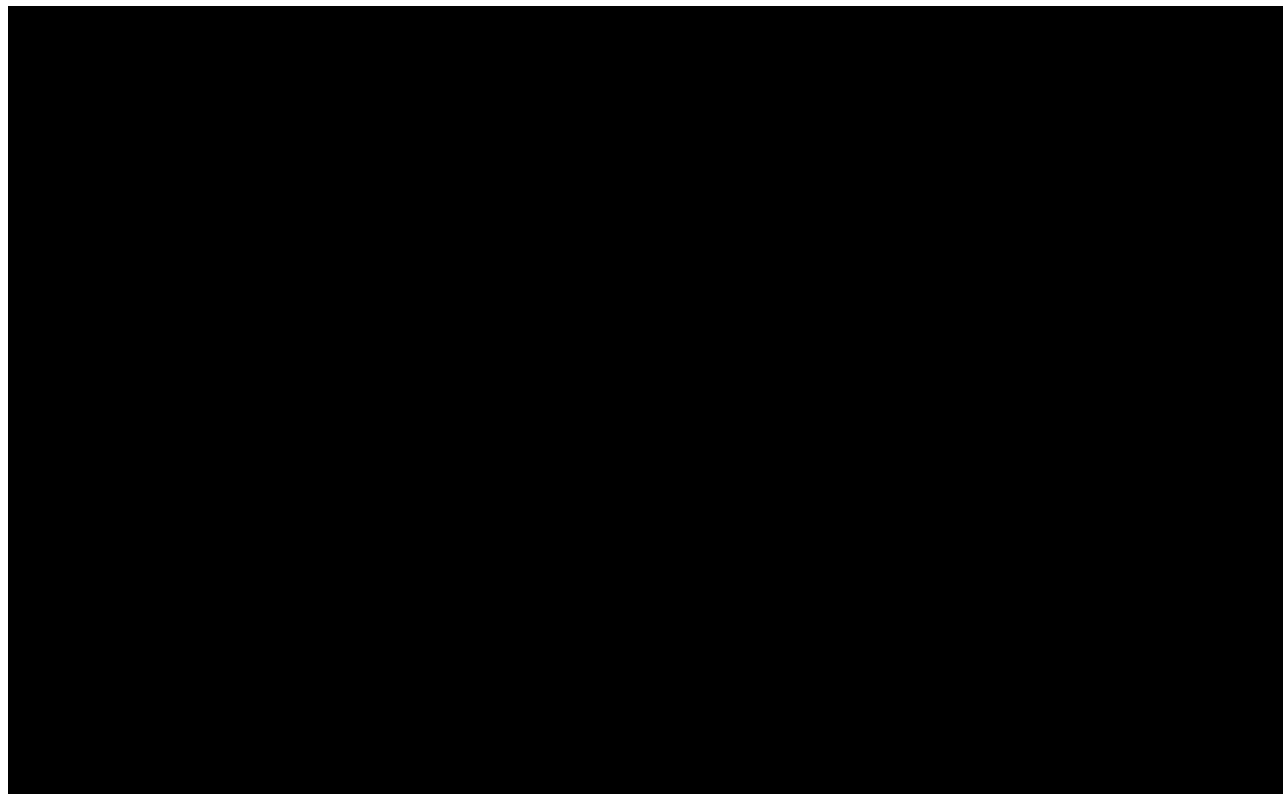
REFERENCES

1. Gettinger, Scott N., et al. Activity and safety of brigatinib in ALK-rearranged non-small-cell lung cancer and other malignancies: a single-arm, open-label, phase 1/2 trial. *The Lancet Oncology* 2016; 17.12: 1683-1696
2. Guidance for Industry E14 The Clinical Evaluation of QT/QTc Interval Prolongation and Proarrhythmic Potential for Non-antiarrhythmic Drugs, U.S. Department of Health and Human Services, Food and Drug Administration, Center for Drug Evaluation and Research (CDER), October 2005.
3. Gough K, Hutchinson M, Keene O, Byrom B, Ellis S, Lacey L, McKellar J. Assessment of Dose Proportionality: Report from the Statisticians in The Pharmaceutical Industry/Pharmacokinetics UK joint working party. *Drug Inf J.* 1995; 29:1039-1048.
4. Smith BP, Vandenhende FR, DeSante KA, Farid NA, Welch PA, Callaghan JT, Forgue ST. CI criteria for assessment of dose proportionality, *Pharmaceut. Res.* 2000; 17: 1278-1283.
5. Hummel J, McKendrick S, Brindley C and French, R. Exploratory assessment of dose proportionality: review of current approaches and proposal for a practical criterion. *Pharmaceut. Statist.* 2009; 8: 38-49.

APPENDIX 1. PARTIAL DATE CONVENTIONS



STATISTICAL ANALYSIS PLAN
BBT176-ONC-001



APPENDIX 2. RESPONSE EVALUATION CRITERIA IN SOLID TUMORS (RECIST VERSION 1.1)

The text below was obtained from the following reference:

Eisenhauer EA, Therasse P, Bogaerts J, Schwartz LH, Sargent D, Ford R, et al. New response evaluation criteria in solid tumours: revised RECIST guideline (version 1.1). European Journal of Cancer.2009;45:228-247.

DEFINITIONS

Response and progression will be evaluated in this trial using the international criteria proposed by the Response Evaluation Criteria in Solid Tumors (RECIST) Committee (version 1.1). Changes in only the largest diameter (uni-dimensional measurement) of the tumor lesions are used in the RECIST criteria. Note: Lesions are either measurable or non-measurable using the criteria provided below. The term “evaluable” in reference to measurability will not be used because it does not provide additional meaning or accuracy.

Measurable Disease

Tumor lesions: Must be accurately measured in at least 1 dimension (longest diameter in the plane of measurement is to be recorded) with a minimum size of:

- 10 mm by CT scan (irrespective of scanner type) and MRI (no less than double the slice thickness and a minimum of 10 mm).
- 20 mm by chest X ray (if clearly defined and surrounded by aerated lung).

Malignant lymph nodes: To be considered pathologically enlarged and measurable, a lymph node must be ≥ 15 mm in short axis when assessed by CT scan (CT scan slice thickness recommended to be no greater than 5 mm). At baseline and in follow up, only the short axis will be measured and followed.

Non-measurable Disease

All other lesions (or sites of disease), including small lesions (longest diameter ≥ 10 to < 15 mm with conventional techniques or < 10 mm using spiral CT scan), are considered non-measurable disease. Leptomeningeal disease, ascites, pleural, or pericardial effusion, inflammatory breast disease, lymphangitic involvement of skin or lung, abdominal masses/abdominal organomegaly identified by physical examination that is not measurable by reproducible imaging techniques are all non-measurable.

Bone lesions:

- Bone scan, PET scan, or plain films are not considered adequate imaging techniques to measure bone lesions. However, these techniques can be used to confirm the presence or disappearance of bone lesions
- Lytic bone lesions or mixed lytic-blastic lesions, with identifiable soft tissue components, that can be evaluated by cross sectional imaging techniques such as CT or MRI can be considered as measurable lesions if the soft tissue component meets the definition of measurability described above.
- Blastic bone lesions are non-measurable.

Cystic lesions:

- Lesions that meet the criteria for radiographically defined simple cysts should not be considered as malignant lesions (neither measurable nor non measurable) since they are, by definition, simple cysts.
- ‘Cystic lesions’ thought to represent cystic metastases can be considered as measurable lesions, if they meet the definition of measurability described above. However, if non cystic lesions are present in the same patient, these are preferred for selection as target lesions.

Lesions with prior local treatment:

- Tumor lesions situated in a previously irradiated area, or in an area subjected to other locoregional therapy, are usually not considered measurable unless there has been demonstrated progression in the lesion. Trial protocols should detail the conditions under which such lesions would be considered measurable.

Target Lesions

All measurable lesions up to a maximum of 2 lesions per organ and 5 lesions in total, should be identified as target lesions and recorded and measured at baseline. Target lesions should be selected on the basis of their size (lesions with the longest diameter), be representative of all involved organ, but in addition should be those that lend themselves to reproducible repeated measurements.

Lymph nodes merit special mention since they are normal anatomical structures which may be visible by imaging even if not involved by tumor. Pathological nodes which are defined as measurable and may be identified as target lesions must meet the criterion of a short axis of ≥ 15 mm by CT scan. Only the short axis of these nodes will contribute to the baseline sum. The short axis of the node is the diameter normally used by radiologists to judge if a node is involved by solid tumor. Nodal size is normally reported as 2 dimensions in the plane in which the image is obtained (for CT scan this is almost always the axial plane; for MRI the plane of acquisition may be axial, sagittal, or coronal). The smaller of these measures is the short axis. For example, an abdominal node which is reported as being 20 mm \times 30 mm has a short axis of 20 mm and qualifies as a malignant, measurable node. In this example, 20 mm should be recorded as the node measurement. All other pathological nodes (those with short axis ≥ 10 mm but < 15 mm) should be considered non target lesions. Nodes that have a short axis < 10 mm are considered non pathological and should not be recorded or followed.

A sum of the diameters (longest for non-nodal lesions, short axis for nodal lesions) for all target lesions will be calculated and reported as the baseline sum diameters. If lymph nodes are to be included in the sum, then as noted above, only the short axis is added into the sum. The baseline sum diameters will be used as reference to further characterize any objective tumor regression in the measurable dimension of the disease.

Non-target Lesions

All other lesions (or sites of disease) including pathological lymph nodes should be identified as non-target lesions and should also be recorded at baseline. Measurements are not required and these lesions should be followed as “present”, “absent”, or in rare cases “unequivocal progression” (more details to follow). In addition, it is possible to record multiple non-target lesions involving the same organ as a single item on the case record form (e.g., “multiple enlarged pelvic lymph node” or “multiple liver metastases”).

GUIDELINES FOR EVALUATION OF MEASURABLE DISEASE

All measurements should be recorded in metric notation, using calipers if clinically assessed. All baseline evaluations should be performed as close as possible to the treatment start and never more than 4 weeks before the beginning of the treatment.

The same method of assessment and the same technique should be used to characterize each identified and reported lesion at baseline and during follow-up. Imaging based evaluation should always be done rather than clinical examination unless the lesion(s) being followed cannot be imaged but are assessable by clinical examination.

CT, MRI: CT is the best currently available and reproducible method to measure lesions selected for response assessment. This guideline has defined measurability of lesions on CT scan based on the assumption that CT slice thickness is 5 mm or less. As is described in Appendix II of the original source article cited above, when CT scans have slice thickness greater than 5 mm, the minimum size for a measurable lesion should be twice the slice thickness. MRI is also acceptable in certain situations (e.g., for body scans).

Endoscopy, laparoscopy: The utilization of these techniques for objective tumor evaluation is not advised. However, they can be useful to confirm complete pathological response when biopsies are obtained or to determine relapse in trials where recurrence following complete response or surgical resection is an endpoint.

Cytology, histology: These techniques can be used to differentiate between PR and CR in rare cases if required by protocol (for example, residual lesions in tumor types such as germ cell tumors, where known residual benign tumors can remain). When effusions are known to be a potential adverse effect of treatment (e.g., with certain taxane compounds or angiogenesis inhibitors), the cytological confirmation of the neoplastic origin of any effusion that appears or worsens during treatment can be considered if the measurable tumor has met criteria for response or stable disease in order to differentiate between response (or stable disease) and progressive disease.

RESPONSE CRITERIA

Evaluation of Target Lesions

Complete Response (CR): Disappearance of all target lesions. Any pathological lymph nodes (whether target or non-target) must have reduction in short axis to < 10 mm.

Partial Response (PR): At least a 30% decrease in the sum of diameters of target lesions, taking as reference the baseline sum diameters.

Progressive Disease (PD): At least a 20% increase in the sum of diameters of target lesions, taking as reference the smallest sum on study (this includes the baseline sum if that is the smallest on study). In addition to the relative increase of 20%, the sum must also demonstrate an absolute increase of at least 5 mm. (Note: the appearance of one or more new lesions is also considered progression).

Stable Disease (SD): Neither sufficient shrinkage to qualify for PR nor sufficient increase to qualify for PD, taking as reference the smallest sum diameters while on study.

Lymph nodes. Lymph nodes identified as target lesions should always have the actual short axis measurement recorded (measured in the same anatomical plane as the baseline examination), even if the nodes regress to below 10 mm on study. This means that when lymph nodes are included as target lesions, the “sum” of lesions may not be zero even if complete response criteria are met, since a normal lymph node is defined as having a short axis of < 10 mm. Case report recorded in a separate section where, in order to qualify for CR, each node must achieve a short axis < 10 mm. For PR, SD and PD, the actual short axis measurement of the nodes is to be included in the sum of target lesions.

Target lesions that become “too small to measure”. While on study, all lesions (nodal and non-nodal) recorded at baseline should have their actual measurements recorded at each subsequent evaluation, even when very small (e.g., 2 mm). However, sometimes lesions or lymph nodes which are recorded as target lesions at baseline become so faint on CT scan that the radiologist may not feel comfortable assigning an exact measure and may report them as being “too small to measure”. When this occurs it is important that a value be recorded in the eCRF. If it is the opinion of the radiologist that the lesion has likely disappeared, the measurement should be recorded as 0 mm. If the lesion is believed to be present and is faintly seen but too small to measure, a default value of 5 mm should be assigned (Note: It is less likely that this rule will be used for lymph nodes since they usually have a definable size when normal and are frequently surrounded by fat such as in the retroperitoneum; however, if a lymph node is believed to be present and is faintly seen but too small to measure, a default value of 5 mm should be assigned in this circumstance as well). This default value is derived from the 5 mm CT slice thickness (but should not be changed with varying CT slice thickness). The measurement of these lesions is potentially non-reproducible, therefore providing this default value will prevent false responses or progressions based upon measurement error. To reiterate, however, if the radiologist is able to provide an actual measure, that should be recorded, even if it is below 5 mm.

Lesions that split or coalesce on treatment. When non-nodal lesions “fragment”, the longest diameters of the fragmented portions should be added together to calculate the target lesion sum. Similarly, as lesions coalesce, a plane between them may be maintained that would aid in obtaining maximal diameter measurements of each individual lesion. If the lesions have truly coalesced such that they are no longer separable, the vector of the longest diameter in this instance should be the maximal longest diameter for the “coalesced lesion”.

Evaluation of Non-target Lesions

While some non-target lesions may actually be measurable, they need not be measured and instead should be assessed only qualitatively at the time points specified in the protocol.

Complete Response (CR): Disappearance of all non-target lesions and normalization of tumor marker level. All lymph nodes must be non-pathological in size (< 10 mm short axis).

Non-CR/Non-PD: Persistence of one or more non-target lesion(s) and/or maintenance of tumor marker level above the normal limits.

Progressive Disease (PD): Unequivocal progression (see comments below) of existing non-target lesions. (Note: the appearance of one or more new lesions is also considered progression).

When the patient also has measurable disease. In this setting, to achieve “unequivocal progression” on the basis of the non-target disease, there must be an overall level of substantial worsening in non-target disease such that, even in presence of SD or PR in target disease, the overall tumor burden has increased sufficiently to merit discontinuation of therapy. A modest “increase” in the size of one or more non-target lesions is usually not sufficient to qualify for unequivocal progression status. The designation of overall progression solely on the basis of change in non-target disease in the face of SD or PR of target disease will therefore be extremely rare.

When the patient has only non-measurable disease. This circumstance arises in some phase 3 trials when it is not a criterion of trial entry to have measurable disease. The same general concept apply here as noted above, however, in this instance there is no measurable disease assessment to factor into the interpretation of an increase in non-measurable disease burden. Because worsening in non-target disease cannot be easily quantified (by definition: if all lesions are truly non-measurable) a useful test that can be applied when assessing patients for unequivocal progression is to consider if the increase in overall disease burden based on the change in non-measurable disease is comparable in magnitude to the increase that would be required to declare PD for measurable disease: i.e., an increase in tumor burden representing an additional 73% increase in “volume” (which is equivalent to a 20% increase diameter in a measurable lesion). Examples include an increase in a pleural effusion from “trace” to “large”, an increase in lymphangitic disease from localized to widespread, or may be described in protocols as “sufficient to require a change in therapy”. If “unequivocal progression” is seen, the patient should be considered to have had overall PD at that point. While it would be ideal to have objective criteria to apply to non-measurable disease, the very nature of that disease makes it impossible to do so, therefore the increase must be substantial.

New Lesions

The appearance of new malignant lesions denotes disease progression; therefore, some comments on detection of new lesions are important. There are no specific criteria for the identification of new radiographic lesions; however, the finding of a new lesion should be unequivocal: i.e., not attributable to differences in scanning technique, change in imaging modality or findings thought to represent something other than tumor (for example, some “new” bone lesions may be simply healing or flare of pre-existing lesions). This is particularly important when the patient’s baseline lesions show partial or complete response. For example, necrosis of a liver lesion may be reported on a CT scan report as a “new” cystic lesion, which it is not.

A lesion identified on a follow-up trial in an anatomical location that was not scanned at baseline is considered a new lesion and will indicate disease progression. An example of this is the patient who has visceral disease at baseline and while on trial has a CT or MRI brain ordered which reveals metastases. The patient’s brain metastases are considered to be evidence of PD even if he/she did not have brain imaging at baseline.

If a new lesion is equivocal, for example because of its small size, continued therapy and follow-up evaluation will clarify if it represents truly new disease. If repeat scans confirm there is definitely a new lesion, then progression should be declared using the date of the initial scan.

While FDG-PET response assessments need additional study, it is sometimes reasonable to incorporate the use of FDG-PET scanning to complement CT scanning in assessment of progression (particularly possible “new” disease). New lesions on the basis of FDG-PET imaging can be identified according to the following algorithm:

- Negative FDG-PET at baseline, with a positive FDG-PET at follow-up is a sign of PD based on a new lesion.
- No FDG-PET at baseline and a positive FDG-PET at follow-up: If the positive FDG-PET at follow-up corresponds to a new site of disease confirmed by CT, this is PD. If the positive FDG-PET at follow-up is not confirmed as a new site of disease on CT, additional follow-up CT scans are needed to determine if there is truly progression occurring at that site (if so, the date of PD will be the date of the initial abnormal FDG-PET scan). If the positive FDG-PET at follow-up corresponds to a pre-existing site of disease on CT that is not progressing on the basis of the anatomic images, this is not PD.

Evaluation of Best Overall Response

The best overall response is the best response recorded from the start of the study drug treatment until the end of treatment taking into account any requirement for confirmation. On occasion a response may not be documented until after the end of therapy so protocols should be clear if post-treatment assessments are to be considered in determination of best overall response. Protocols must specify how any new therapy introduced before progression will affect best response designation. The patient’s best overall response assignment will depend on the findings of both target and non-target disease and will also take into consideration the appearance of new lesions. Furthermore, depending on the nature of the trial and the protocol requirements, it may also require confirmatory measurement. Specifically, in non-randomized trials where response is the primary endpoint, confirmation of PR or CR is needed to deem either one the “best overall response”.

The best overall response is determined once all the data for the patient are known. Best response determination in trials where confirmation of complete or partial response IS NOT required: Best response in these trials is defined as the best response across all time points (for example, a patient who has SD at first assessment, PR at second assessment, and PD on last assessment has a best overall response of PR). When SD is believed to be best response, it must also meet the protocol specified minimum time from baseline. If the minimum time is not met when SD is otherwise the best time point response, the patient’s best response depends on the subsequent assessments. For example, a patient who has SD at first assessment, PD at second and does not meet minimum duration for SD, will have a best response of PD. The same patient lost to follow-up after the first SD assessment would be considered inevaluable.

Target Lesions	Non-target Lesions	New Lesions	Overall Response
CR*	CR	No	CR
CR	Non-CR/non-PD	No	PR
CR	Not evaluated	No	PR
PR	Non-PD or not all evaluated	No	PR
SD	Non-PD or not all evaluated	No	SD
Not all evaluated	Non-PD	No	NE
PD	Any	Yes or No	PD
Any	PD	Yes or No	PD
Any	Any	Yes	PD

* CR = complete response; PR = partial response; SD = stable disease; PD = progressive disease; and NE = invaluable. See text for more details.

Note: When nodal disease is included in the sum of target lesions and the nodes decrease to “normal” size (< 10 mm), they may still have a measurement reported on scans. This measurement should be recorded even though the nodes are normal in order not to overstate progression should it be based on increase in size of the nodes. As noted earlier, this means that patients with CR may not have a total sum of “zero” in the eCRF.

In trials where confirmation of response is required, repeated ‘NE’ time point assessments may complicate best response determination. The analysis plan for the trial must address how missing data/assessments will be addressed in determination of response and progression. For example, in most trials it is reasonable to consider a patient with time point responses of PR-NE-PR as a confirmed response.

Patients with a global deterioration of health status requiring discontinuation of treatment without objective evidence of disease progression at that time should be reported as “symptomatic deterioration”. Every effort should be made to document objective progression even after discontinuation of treatment. Symptomatic deterioration is not a descriptor of an objective response: it is a reason for stopping trial therapy.

In some circumstances it may be difficult to distinguish residual disease from normal tissue. When the evaluation of complete response depends upon this determination, it is recommended that the residual lesion be investigated (fine needle aspirate/biopsy) before assigning a status of complete response. FDG-PET may be used to upgrade a response to a CR in a manner similar to a biopsy in cases where a residual radiographic abnormality is thought to represent fibrosis or scarring. The use of FDG-PET in this circumstance should be prospectively described in the protocol and supported by disease-specific medical literature for the indication. However, it must be acknowledged that both approaches may lead to false positive CR due to limitations of FDG-PET and biopsy resolution/sensitivity.

For equivocal findings of progression (e.g., very small and uncertain new lesions; cystic changes, or necrosis in existing lesions), treatment may continue until the next scheduled assessment. If at the next scheduled assessment, progression is confirmed, the date of progression should be the earlier date when progression was suspected.

CONFIRMATORY MEASUREMENT/DURATION OF RESPONSE

Confirmation

In non-randomized trials where response is the primary endpoint, confirmation of PR and CR is required to ensure that the identified responses are not the result of measurement error. This will also permit appropriate interpretation of results in the context of historical data where response has traditionally required confirmation in such trials. However, in all other circumstances, i.e., in randomized trials (phase 2 or 3) or trials where stable disease or progression are the primary endpoints, confirmation of response is not required since it will not add value to the interpretation of trial results. However, elimination of the requirement for response confirmation may increase the importance of central review to protect against bias, in particular in trials which are not blinded.

In the case of SD, measurements must have met the SD criteria at least once after trial entry at a minimum interval (in general not less than 6 to 8 weeks) that is defined in the trial protocol.

Duration of Overall Response

The duration of overall response is measured from the time that measurement criteria are first met for CR/PR (whichever is first recorded) until the first date that recurrent or progressive disease is objectively documented (taking as reference for progressive disease the smallest measurements recorded on study).

The duration of overall complete response is measured from the time measurement criteria that are first met for CR until the first date that recurrent disease is objectively documented.

Duration of Stable Disease

Stable disease is measured from the start of the treatment (in randomized trials, from date of randomization) until the criteria for progression are met, taking as reference the smallest sum on study (if the baseline sum is the smallest, this is the reference for calculation of PD).

The clinical relevance of the duration of stable disease varies in different studies and diseases. If the proportion of patients achieving stable disease for a minimum period of time is an endpoint of importance in a particular trial, the protocol should specify the minimal time interval required between two measurements for determination of stable disease.

Note: The DoR and stable disease as well as the PFS are influenced by the frequency of follow-up after baseline evaluation. It is not in the scope of this guideline to define a standard follow-up frequency. The frequency should take into account many parameters including disease types and stages, treatment periodicity, and standard practice. However, these limitations of the precision of the measured endpoint should be taken into account if comparisons between trials are to be made.

**APPENDIX 3. EASTERN COOPERATIVE ONCOLOGY GROUP (ECOG)
PERFORMANCE STATUS**

Grade	Description
0	Fully active, able to carry on all pre-disease performance without restriction
1	Restricted in physically strenuous activity but ambulatory and able to carry out work of a light or sedentary nature, e.g., light house work, office work
2	Ambulatory and capable of all self-care but unable to carry out any work activities. Up and about more than 50% of waking hours
3	Capable of only limited self-care, confined to bed or chair more than 50% of waking hours
4	Completely disabled. Cannot carry on any self-care. Totally confined to bed or chair
5	Dead

Source: Oken MM, Creech RH, Tormey DC, Horton J, Davis TE, McFadden ET, et al. Toxicity and response criteria of the Eastern Cooperative Oncology Group. Am J Clin Oncol.1982;5:649-5