ClinicalTrials.gov Document Cover Page

Statistical Analysis Plan

Study BLU-285-1303

Study Title: An International, Multicenter, Open-label, Randomized, Phase 3 Study of BLU-285 vs Regorafenib in Patients With Locally Advanced Unresectable or Metastatic Gastrointestinal Stromal Tumor (GIST)

Document Date: 01 February 2020

NCT Number: NCT03465722

STATISTICAL ANALYSIS PLAN

An International, Multicenter, Open-label, Randomized, Phase 3 Study of BLU-285 vs Regorafenib in Patients with Locally Advanced Unresectable or Metastatic Gastrointestinal Stromal Tumor (GIST)

Protocol Number: BLU-285-1303 Amendment 3

SAP Version:	Date of Statistical Analysis Plan:
Final 1.0	01 FEB 2020

Approval Signatures:

13 MAR 2020
Date

[redacted]

12 MAR 2020
Date

[redacted]
Blueprint Medicines Corporation

10 MAR 2020
Date

[redacted]
Blueprint Medicines Corporation

10 Mar 2020
Date

[redacted]
Blueprint Medicines Corporation

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

LIST OF ABBREVIATIONS AND DEFINITIONS OF TERMS

Abbreviation	Term
AE	Adverse event
AESI	Adverse event of special interest
ANC	Absolute neutrophil count
ATC	Anatomical therapeutic chemical
β-hCG	Beta human chorionic gonadotropin
BMI	Body mass index
BP	Blood pressure
BSC	Best supportive care
C3D1	Cycle 3 Day 1
CBR	Clinical benefit rate
CI	Confidence interval
CR	Complete response
CS	Clinically significant
CSR	Clinical study report
CT	Computed tomography
CTCAE	Common Terminology Criteria for Adverse Events
ctDNA	Circulating tumor deoxyribonucleic acid
CXDX	Cycle X Day X
D842V	Aspartic acid to valine at amino acid 842
DCR	Disease control rate
DE	Dose escalation
DOR	Duration of response
ECG	Electrocardiogram
ECOG PS	Eastern Cooperative Oncology Group Performance Status
EORTC-QLQ-C30	European Organisation for Research and Treatment of Cancer Quality of Life
EQ-5D-5L	EuroQol 5 Dimension
eCRF	Electronic case report form
EMA	European Medicines Agency
EOT	End-of-treatment
FDA	Food and Drug Administration
GIST	Gastrointestinal stromal tumors
IDMC	Independent Data Monitoring Committee
ITT	Intent-to-Treat
IV	Intravenous
IWRS	Interactive web response system
KIT	V-Kit Hardy-Zuckerman 4 feline sarcoma viral oncogene homolog
KM	Kaplan-Meier
MAF	Mutant allele fractions
MedDRA	Medical Dictionary for Regulatory Activities
MRI	Magnetic resonance imaging

Abbreviation	Term
mRECIST 1.1	Modified Response Evaluation Criteria in Solid Tumors version 1.1
NCI	National Cancer Institute
NCS	Not clinically significant
NE	Not evaluable
ORR	Overall response rate
OS	Overall survival
PD	Progressive Disease
PFS	Progression free survival
PDGFRA	Platelet-derived growth factor receptor alpha
PK	Pharmacokinetic
PR	Partial response
PRO	Patient reported outcome
PT	Preferred term
QD	Once daily
QoL	Quality of life
QTcF	QT interval corrected using Fridericia's formula
RDI	Relative dose intensity
RE	Response-evaluable
R/R	Relapsed or refractory
RPSFT	Rank preserving structural failure time
SAE	Serious adverse event
SAP	Statistical analysis plan
SD	Stable disease
SOC	System organ class
StdDev	Standard deviation
TEAE	Treatment-emergent adverse event
TKI	Tyrosine kinase inhibitor
TNM	Tumor/lymph nodes/metastasis
TTR	Time to response
ULN	Upper limit of normal
US	United States (of America)
VAS	Visual analog scale
WHO	World Health Organization

1. INTRODUCTION

This statistical analysis plan (SAP) describes the statistical analyses and data presentations to be performed for Blueprint Medicines, Protocol No. BLU-285-1303 (amendment 3 dated 20 JUN 2019) 'An International, Multicenter, Open-label, Randomized, Phase 3 Study of Avapritinib (BLU-285) vs Regorafenib in Patients with Locally Advanced Unresectable or Metastatic Gastrointestinal Stromal Tumor (GIST)'. It contains definitions of analysis populations, derived variables, and statistical methods for the analysis of efficacy and safety. It is to ensure that the data listings, summary tables, and figures which will be produced, and the statistical methodologies that will be used, are complete and appropriate to allow valid conclusions regarding the study objectives.

The SAP will be finalized and approved prior to the clinical database lock for the final analysis, when approximately 264 progression free survival (PFS) events are reached. The analyses of safety, efficacy and pharmacokinetics will include all data collected in the database through the data cutoff date.

A follow-up analysis for overall survival will be conducted when approximately 264 deaths have been observed. Additional data collected after the database lock from the primary analysis of the study will be prepared as an addendum to the study report per regulatory or scientific need.

Selected analyses from this SAP will be performed for the Independent Data Monitoring Committee (IDMC) to review accumulated data from the study. Further details are described separately in the IDMC Charter.

Analysis of pharmacokinetics (PK) is outside the scope of this SAP and will be addressed separately.

1.1 Study Design

This is an open-label, randomized, Phase 3 study in patients with locally advanced unresectable or metastatic GIST previously treated with imatinib and 1 or 2 other tyrosine kinase inhibitors (TKIs).

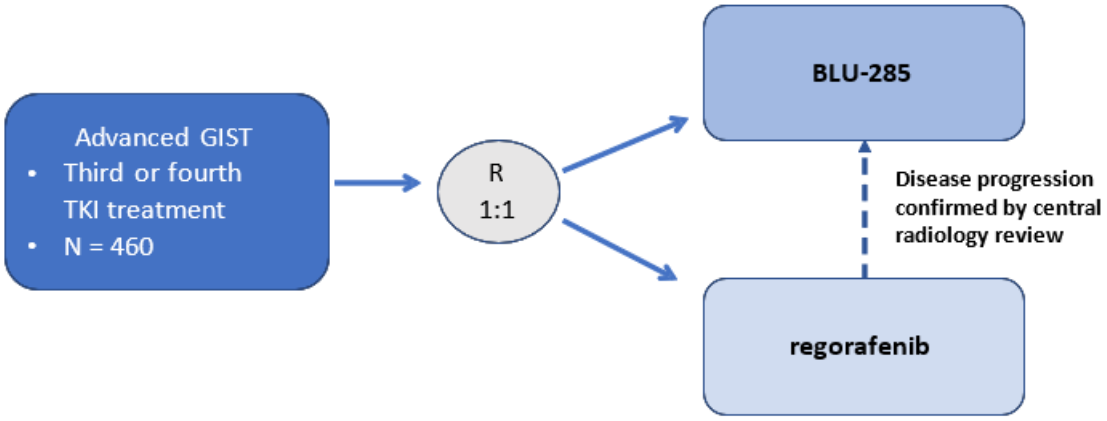
All study visits are intended to be conducted on an outpatient basis. After provision of written informed consent, patients will be evaluated for study eligibility during the screening period within 4 weeks (28 days) before study drug administration on Cycle 1 Day 1 (C1D1). During the screening period, eligibility will be confirmed; management of baseline concomitant conditions will be recorded and stabilized; baseline symptoms will be assessed; hematology, blood chemistry, mutation status, brain imaging (computed tomography (CT) scan or magnetic resonance imaging (MRI)), and baseline tumor assessments (CT scan or MRI) will be performed.

Approximately 460 patients are to be randomly assigned, in a 1:1 ratio, to 1 of 2 treatment arms: Arm A (avapritinib) or Arm B (regorafenib) stratified by TKI treatment (third versus fourth), geographic region (Asia versus rest of the world), and PDGFRA D842V mutation status measured in ctDNA or a tumor sample (PDGFRA D842V mutation present versus absent). PDGFRA D842V mutation present is defined if a PDGFRA D842V mutation is detected per local or/and central testing. At least 70% of the patients enrolled should be receiving the study drug as their third TKI treatment for GIST.

Patients randomized to Arm A will receive avapritinib 300 mg orally (PO) once daily (QD) with the option to escalate to 400 mg PO QD at the Investigator's discretion (the escalation option was removed in protocol amendment 3), if specific criteria are met as described in the protocol. Patients randomized to Arm B will receive regorafenib 160 mg PO QD for 3 weeks out of every 4 week (28 days) cycle (i.e., 3 weeks on/1 week off).

Patients who experience disease progression on avapritinib as determined by central radiology may be offered the opportunity to continue treatment with avapritinib if there is no clinical evidence of disease progression or laboratory abnormalities attributable to disease progression, no rapid progression of disease or a progressive tumor requiring urgent alternative medical intervention at critical anatomical sites, and no decline in Eastern Cooperative Oncology Group Performance Status (ECOG PS). Patients who experience disease progression on regorafenib, as determined by central radiology, may be offered the opportunity to cross over to avapritinib.

Figure 1: Study Schematic



Abbreviations: GIST = gastrointestinal stromal tumor; N = number of patients; R = randomized; TKI = tyrosine kinase inhibitor.

1.2 Study Objectives

1.2.1 Primary Objectives

The primary objective is to demonstrate the efficacy of avapritinib based on PFS as determined by central radiological assessment per mRECIST 1.1 in patients with advanced GIST following 2 or 3 prior TKI therapies, including imatinib, compared to patients treated with regorafenib.

1.2.2 Secondary Objectives

The key secondary objectives are:

- To evaluate objective response rate (ORR) as determined by central radiological assessment per mRECIST 1.1 in patients with advanced GIST treated with avapritinib compared to patients treated with regorafenib.
- To evaluate overall survival (OS) in patients with advanced GIST treated with avapritinib compared to patients treated with regorafenib.

Additional secondary objectives are:

- To evaluate the European Organisation for Research and Treatment of Cancer Quality of Life (EORTC-QLQ-C30) individual scores in patients with advanced GIST treated with avapritinib compared to patients treated with regorafenib.
- To evaluate the safety and tolerability of avapritinib compared to regorafenib.
- To evaluate objective response rate as assessed by the Investigator per mRECIST 1.1 in patients with advanced GIST treated with avapritinib compared to patients treated with regorafenib.
- To evaluate objective response rate as determined by central radiological assessment per Choi criteria in patients with advanced GIST treated with avapritinib compared to patients treated with regorafenib.
- To evaluate disease control rate (DCR) per mRECIST 1.1 in patients with advanced GIST treated with avapritinib compared to patients treated with regorafenib.
- To evaluate duration of response (DOR) per mRECIST 1.1 in patients with advanced GIST treated with avapritinib compared to patients treated with regorafenib.
- To determine steady state systemic exposure of avapritinib (out of scope for this SAP).
- To assess the patient-reported perception of abdominal pain.

1.2.3 Exploratory Objectives

[redacted]

[redacted]

1.3 Sample Size Justification

The sample size calculation was based upon the assumption that the median PFS for regorafenib is approximately 5 months (Demetri 2013). Assuming avapritinib can reduce the risk of PFS events by 33%, i.e., a hazard ratio (HR) of 0.67 (avapritinib compared to regorafenib), a minimum of 264 PFS events are needed to provide 90% power at a 2-sided alpha of 0.05 for the study. With an 18-month accrual period and a 6-month follow-up period, the total sample size needed for the study is approximately 460 patients (230 patients per arm). This sample size takes a 2% dropout rate into consideration and assumes accrual will follow a truncated exponential distribution with a scaled power parameter of -9.

The sample size calculated for the primary endpoint, PFS, is also sufficient to detect a difference in ORR between the two arms. Assuming a null ORR of 5% and an alternative ORR of 15%, a sample size of 460 patients provides approximately 95% power to detect a difference in ORR using a test of two-proportions at a 2-sided alpha of 0.05.

Assuming the same effect size for OS as PFS (HR = 0.67), the same number of events (264) is required to power the OS analysis. The timing of the OS analysis will occur later than the PFS analysis as it will be at such time when 264 deaths occur.

1.4 Endpoints

1.4.1 Primary Efficacy Endpoint

The primary endpoint is PFS, based on central radiological assessment per mRECIST 1.1, in patients with advanced GIST. Progression-free survival is defined as time from randomization to the earlier of disease progression or death due to any cause.

1.4.2 Secondary Efficacy Endpoints

The key secondary endpoints are:

- ORR defined as the percentage of patients whose best response is complete response (CR) or partial response (PR) as assessed by central radiology using mRECIST 1.1.
- OS defined as the time from date of randomization to death due to any cause.

Additional secondary endpoints include:

- All individual EORTC-QLQ-C30 scores, including symptom and function scales, and the global health status score.
- ORR as assessed by Investigator per mRECIST 1.1.
- ORR as assessed by central radiology per Choi criteria.
- DCR is defined as the rate of CR or PR of any duration or stable disease (SD) lasting at least 16 weeks from the date of randomization.
- Duration of response defined as the time from first documentation of tumor response to disease progression or death due to any cause.
- Adverse events, serious AEs (SAEs), and changes in safety laboratory parameters, 12-lead ECG evaluations, and ECOG PS. The intensity of AEs will be assessed by the NCI CTCAE, version 4.03 or 5.0.
- Abdominal pain as measured by a numeric rating scale (0-10).

1.4.3 Exploratory Endpoints

2. POPULATIONS FOR ANALYSIS

2.1 Intent-to-Treat (ITT) Population

The ITT Population includes all randomized patients, independent of whether they received the study medication or not. All primary efficacy analyses will be based on the ITT Population. All patients in the ITT Population will be analyzed according to the treatment they were randomized to receive.

2.2 Safety Population

The Safety Population includes all patients who have received at least 1 dose of study drug. All patients in the Safety Population will be analyzed according to the treatment actually received.

2.3 Per Protocol (PP) Population

The PP Population includes all patients in the ITT Population who have no major violations of the inclusion/exclusion criteria or other major deviations as specified in the "*Protocol Deviations and Non-Compliance Management Plan*". Patients who actually receive incorrect treatment compared to what they were randomized to receive will be excluded from the PP Population. The PP Population will be used for selected efficacy analyses as supportive analyses.

2.4 Response-Evaluable (RE) Population

The RE Population is defined as all patients in the ITT Population who received at least 1 dose of avapritinib or regorafenib, have at least 1 target lesion at baseline as determined by central radiology per mRECIST 1.1, and have at least 1 post-baseline disease assessment by central radiology per mRECIST 1.1. Selected efficacy analysis will be performed using the RE Population. Results in this population will be analyzed according to the treatment to which patients were randomized to receive.

3. GENERAL STATISTICAL CONSIDERATIONS

3.1 Randomization, Stratification, and Blinding

Patients who meet all study eligibility criteria will be randomized by an interactive web response system (IWRS).

The randomization will be stratified by the following 3 variables (stratification factors):

- TKI treatment (third versus fourth)
- Geographic region (Asia versus rest of the world)
- PDGFRA D842V mutation (present versus absent).

If a patient is randomized using incorrect baseline information from the stratification factors, that randomization will be accepted but the analysis will be based on the correct baseline information.

If fewer than 5 patients are in a stratification group of all three factors, that group will be combined with another stratification group. If the PDGFRA D842V is present in fewer than 24 patients overall (i.e., on average, each combination of the other two stratification under each treatment arm contains fewer than 3 patients), then all stratified analyses in this SAP will be based only on TKI treatment and geographic region stratification factors.

To control bias during the study conduct and analyses, and to ensure proper type I error control, patient treatment assignments are blinded to pre-defined study team members (refer to "*Plan for Providing Blinded Data to Biostatistics*" for details). An unblinded team will conduct analyses by treatment, e.g., IDMC until the final database lock, upon which the study team will be unblinded to conduct the primary analysis.

3.2 Data Analysis General Information and Definition

Summary statistics for continuous variables will include n (non-missing observations), mean, standard deviation (StdDev), minimum, median, and maximum. Summary statistics for categorical variables will be presented in terms of frequencies and percentages. Time to event data will be summarized using the Kaplan-Meier (KM) method, which will include the estimated median with 95% confidence interval (CI) and the 25th and 75th percentiles. All relevant patient data will be included in listings.

3.2.1 Study Drug

Study drug or study treatment refers to avapritinib or regorafenib.

3.2.2 Day 1 and Other Days

Date of first dose of study drug (or Day 1) is defined as the day of the first administration of study drug after enrollment. First dose refers to first dose in study unless otherwise specified.

Date of first dose of study drug during extended study treatment period (or CC1D1) is defined as the day of the first administration of study drug after cross-over or continuation post disease progression per central radiology.

Date of last dose of study drug is defined as the date of last administration of study drug (last dose) in the study. For patients who have not ended treatment and whose last dose date is missing, the cutoff date will be used. The date of last dose of study drug will be calculated separately for the main study treatment period, the extended study treatment period, and both periods combined.

The **extended study treatment period** (or extended study) is defined as any time on or after CC1D1 in the study.

The **main study treatment period** (or main study) is defined as any time before CC1D1 in the study.

Day -1 is defined as the day before the first administration of study drug in the study (Day 1).

3.2.3 Calculations Using Dates

Calculations using dates will adhere to the following conventions:

Study day for a date of interest (TARGET DATE) is calculated as

STUDY DAY = TARGET DATE – Day 1 + 1 if TARGET DATE is on or after Day 1;

STUDY DAY = TARGET DATE – Day 1 otherwise.

Note that negative study days are reflective of observations obtained during the screening period.

Age (in days) is calculated as

AGE = CONSENT DATE – BIRTH DATE (DOB) + 1.

Age (in years) is calculated as

Age (in years) = (year of consent) - (year of DOB) if the day of consent ≥ birthday;

Age (in years) = (year of consent) - (year of DOB) - 1 otherwise.

This is equivalent to the following:

[(year of consent) - (year of DOB)] – [(month of consent) ≤ (month of DOB)] + [(month of consent) = (month of DOB) and (day of consent) ≥ (day of DOB)]

3.2.4 Duration Derivation

Unless otherwise specified for a specific panel or variable, duration variables will be derived according to the following rules:

Duration (in days) = [end date – start date +1]

Duration (in weeks) = [end date – start date +1] / 7

Duration (in months) = [end date – start date +1] / 30.4375

Duration (in years) = [end date – start date +1] / 365.25.

Duration variables expressed in units greater than day will be rounded to 1 decimal place.

3.2.5 Baseline and Rebaseline Values

Baseline is defined as the last observation prior to first dose of study drug, i.e., the Day 1 pre-dose value or the last available (including unscheduled) value before Day 1 if a Day 1 pre-dose value is unavailable. For summaries using the ITT Population, patients who were randomized but did not receive study drug will be baselined at their last available observation.

Rebaseline is defined as the last observation prior to first dose of avapritinib in the extended study treatment period. For summaries using the ITT Population, patients who signed informed consent to crossover to avapritinib post-disease progression but did not subsequently receive study drug will be re-baselined at their last available observation. Patients who are initially randomized to avapritinib will not be rebaselined for continuation to avapritinib post-disease progression.

3.2.6 Last Contact

Last contact or last date known alive is defined as the last non-imputed date of any patient record prior to or on the data cutoff date in the clinical database.

3.2.7 Table Headings

All analyses will be tabulated by treatment group and include a total if appropriate. The treatment groups are defined as below:

- avapritinib (“ava”) – Data collected during the main study treatment period only from patients randomized* to avapritinib.
- Regorafenib (“rego”) – Data collected during the main study treatment period only from patients randomized* to regorafenib.
- avapritinib overall (“ava overall”) – Data collected during both the main and extended study treatment periods from patients randomized* to avapritinib.

- regorafenib to avapritinib ("rego to ava") – Data collected during the extended study treatment period only from patients randomized* to regorafenib. During this period, patients are treated with avapritinib after crossing over from regorafenib.

*For analyses using the Safety Population, patients will be grouped based on the study drug actually received during the main study treatment period rather than the study drug they were randomized to.

Safety, laboratory, exposure, compliance, dose modification, concomitant medications, and major protocol deviation analyses will be presented using the 4 groups above.

Efficacy analyses will be presented using 2 groups (avapritinib, regorafenib).

Analysis populations, patient dispositions, baseline or demographic characteristics, patients by country by site, medical history, and prior therapies for the underlying malignancy will be presented using 3 groups (avapritinib, regorafenib, total).

3.3 Methods for Handling Missing Data

Refer to Section 7.1 for detailed date imputation guidelines.

3.4 Windowing of Visits

Data will be summarized by visit following the visit windows in Table 1 when appropriate (e.g., lab parameters). If multiple assessments occur within a visit window, the assessment closest to the target day will be used. In case of ties, with equal number of days on either side of target day, the value from the day before target day shall be used.

The extended study treatment period will be windowed similarly. The extended study treatment period visits will have a prefix "C", e.g., C Week 4 (CC2D1).

Table 1: Visit Windows

Visit	Start day	Target Day	End day
Week 4 (C2D1)	2	29 ± 2	43
Week 8 (C3D1)	44	57 ± 3	71
Week 12 (C4D1)	72	85 ± 3	99
Week 16 (C5D1)	100	113 ± 3	127
Week 20 (C6D1)	128	141 ± 3	155
Week 24 (C7D1)	156	169 ± 7	183
Week 28 (C8D1)	184	197 ± 7	211
Week 4*X (C(X+1)D1)	28*X–12	28*X+1 ± 7	28*X+15

Note: Visit window for Week 4 will not include Day 1 if the Day 1 observation is marked as baseline; X can be 4, 5, 6… until all visit windows are covered. CXDX in parentheses are for calculation purposes and will not be presented in outputs.

3.5 Withdrawals, Dropouts, Lost to Follow-Up

Patients may withdraw or be withdrawn from study treatments at any time for any of the following reasons:

- Withdrawal of consent
- Adverse event (AE)
- Disease progression
- Death
- Investigator decision
- Protocol deviation
- Pregnancy
- Lost to follow-up.

3.6 Protocol Deviations

All protocol deviations will be entered into the RAVE database by site staff, and their classifications will be entered by designated CRO or sponsor study team members. Details on reporting of protocol deviations can be found in the *Protocol Deviations and Non-Compliance Management Plan*.

Prior to database lock, protocol deviations will be reviewed and categorized as major or minor for summarization.

All major protocol deviations will be tabulated for the ITT Population by treatment.

A listing of patients with protocol deviations will be provided. A separate listing of patients with major protocol deviations that lead to exclusion from the PP population or any analysis will be provided.

4. STATISTICAL METHODOLOGY

4.1 Analysis Populations

A summary of analysis populations by treatment will be presented along with percentages relative to the ITT Population. Derived stratification factors will be tabulated for each factor and all combinations.

- Prior Lines of TKI (2, 3)
- PDGFRA D842V (present, absent)
- Region (Asia, Rest of World)

4.2 Patient Disposition

The number and percent of patients will be tabulated for the ITT, Safety, PP, and RE Populations, respectively, for the following categories:

- Patients in the population
- Patients who receive study treatment
- Patients who end main study treatment without continuation or cross-over to avapritinib
- Patients who continue or cross-over to avapritinib
- Patients who end treatment after continuation or cross-over to avapritinib
- Patients who receive subsequent anti-cancer therapy

Reasons for treatment discontinuation will be summarized with the following categories collected on the electronic case report form (eCRF):

- Patient refuses treatment
- Withdrawal of consent
- Pregnancy
- Adverse event and related adverse event
- Disease progression (confirmed by Central Review)
- Clinical progression
- Protocol deviation

- Investigator decision
- Lost to follow-up
- Administrative/Other
- Sponsor decision

Listings will be provided for patients who discontinue treatment with the reason for treatment discontinuation.

Reasons for treatment discontinuation after continuation or cross-over to avapritinib will be similarly summarized.

Reasons for study discontinuation will be summarized with the following categories:

- Withdrawal of consent
- Death
- Lost to follow-up
- Administrative/Other
- Sponsor decision

Listings will be provided for patients who discontinue from study with the reason for study discontinuation.

4.3 Demographics and Baseline Disease Characteristics

Demographic and baseline disease characteristic data will be summarized for the ITT, Safety, PP, and RE Populations. Medical history data will be summarized for the ITT Population.

4.3.1 Demographics

The number and percentage of patients in each of the following categories will be presented:

- Age group (<65 years or ≥65 years)
- Sex
- Region
- Ethnicity
- Race.

Patients' age (years), height (cm), weight (kg), body mass index (BMI) (kg/m^2), and other continuous demographic variables will be summarized descriptively.

BMI will be calculated as: BMI (kg/m^2) = (weight in kg) / (height in $m)^2$.

4.3.2 Baseline Disease Characteristics

The number and percentage of patients in each of the following categories will be presented:

- Mutation Status
  - KIT V654A or T670I (Present, Absent)
  - KIT mutation status in KIT V654A and T670I mutation absent patients (Exon 17 mutation detected, KIT mutation other than exon 17 detected, no KIT mutation detected)
  - Specific Exon 17 mutations (D816X, D820X, N822X, Y823X) in KIT V654A and T670I mutation absent patients
  - PDGFRA D842V
  - PDGFRA non-D842V
- The Eastern Cooperative Oncology Group (ECOG) performance status (0, 1, 2)
- Metastatic disease (Yes, No)
- Largest target lesion size (≤5 cm, >5 to ≤10 cm, >10 cm) by central radiology assessment
- Patient staged at screening by tumor/lymph nodes/metastasis (TNM) system (Yes, No)
- Current stage at screening by TNM
- Prior imatinib (Yes, No)
- Prior sunitinib (Yes, No)
- Primary tumor site of GIST at the time of diagnosis
- Site(s) of metastatic disease
- Prior surgical resection (Yes, No), and type of resection (total, partial, other).

Continuous baseline disease characteristics will be summarized using descriptive statistics.

4.3.3 Medical History

Ongoing medical history data will be summarized by system organ class (SOC) and preferred term (PT) per Medical Dictionary for Regulatory Activities (MedDRA) version 18.1. Patients with multiple events coded to the same system organ class or preferred term will only be counted once in the summary for that category.

All medical history data will be listed.

4.4 Prior and Concomitant Medications and Prior Therapies

Prior and concomitant medications and prior therapies will be summarized for the ITT population.

4.4.1 Prior and Concomitant Medications

Prior medications are all medications which started and stopped before exposure to study drug. Concomitant medications are all medications taken at any time from the first dose date of study drug to last dose date of study drug + 30 days.

When summarizing data collected only from the main study treatment period, concomitant medications include medications taken at any time from the first dose date of study drug to the earlier of the last dose date of study drug + 30 days or the first dose date of study drug during the extended study treatment period.

When summarizing data collected only from the extended study treatment period, concomitant medications include medications taken at any time from the first dose date of avapritinib during the extended study treatment period to the last dose date of avapritinib + 30 days.

If the last dose date of study drug is not available, then the cutoff date is used in place of the last dose date.

The number and percentage of patients taking concomitant medications and significant non-drug therapies will be summarized and listed by therapeutic area (ATCText3) and preferred drug name. A patient taking the same medication multiple times is counted only once under that preferred drug name.

Prior (with clear flag) and concomitant medications will be listed.

Medications are coded using the World Health Organization Drug dictionary B3 enhanced, version Mar 2019.

4.4.2 Prior Therapies and Surgery Procedures for the Underlying Malignancy

Prior therapy is defined as all treatment that started and ended on or before first dose date of study drug.

The number and percentage of patients for each category below will be summarized by PT:

- Prior non-TKI Therapy (Yes, No)
- Prior radiation therapy (Yes, No; type (external, brachytherapy or IORT))
- Prior lines of TKIs (2 lines, >2 lines)
- Prior cancer related surgery-procedures (Yes, No)
- Best response to any prior TKI (CR, PR, SD, progressive disease, NE)

- Best response to last prior TKI (CR, PR, SD, progressive disease, NE)
- Type of prior surgery procedure (total resection, partial resection or other), and the site of the procedure.

Continuous variables to be summarized include:

- Time between diagnosis and randomization date
- Time between the end of last prior line of anticancer therapy and randomization date
- Duration of previous exposure to TKI

The duration of previous exposure to TKI will be the sum of exposure from each prior TKI therapy, which is number of days from start date to end date captured on the prior cancer therapy eCRF page.

No overlap between the exposure to prior therapies and avapritinib will be allowed and any overlap of exposure will be queried at the data review stage. The end date of the last prior therapy will be imputed as 'Day 1 – 15' if the imputation of partial dates is within 14 days of Day 1. However, if Day 1 of study drug is within the first half of the month, and last prior therapy ended on the same month as study drug start, the day component of last prior therapy end date will be imputed as 1.

4.5 Study Treatments and Extent of Exposure

Study drug exposure, including the number of doses administered, cumulative dose, dose intensity, relative dose intensity (RDI), duration of treatment, and the proportion of patients with dose modifications will be summarized descriptively by time (every 4 weeks and the overall duration of the data collection period) for the Safety Population.

Day 1 + 28*(X–1) is the start date of each 4 week period. Day 28 + 28*(X–1) is the end date of each 4 week period. The end date of last 4 week period is the date of last dose of study drug.

4.5.1 Treatment Duration

Duration of treatment (days) = (treatment end date – treatment start date + 1).

Duration of treatment (weeks) = (treatment end date – treatment start date + 1) / 7.

The treatment start date is the first dose date of study drug.

For patients receiving avapritinib, the treatment end date is the last dose date of study drug. For patients receiving regorafenib, the treatment end date is the earliest of the last dose date of study drug + 7 days, the data cutoff date, or the death date, to account for the dosing schedule of regorafenib.

For patients with last dose date missing, the earlier of the data cutoff date or the death date will be used.

Descriptive statistics will be provided for the duration of treatment in weeks.

4.5.2 Cumulative Dose

Cumulative dose (mg) is defined as sum of all doses taken. The dosage will be counted as 0 for days when the study drug is not taken.

4.5.3 Average Daily Dose

Average daily dose (mg) is defined as the cumulative dose divided by the number of days dosed.

4.5.4 Dose Intensity

Dose intensity (mg/day) is defined as the cumulative dose divided by the duration of treatment (in days).

4.5.5 Relative Dose Intensity

Relative dose intensity is defined as the ratio: dose intensity / planned dose intensity. Planned dose intensity is based on initial assigned daily dose.

Relative dose intensity will be summarized both descriptively as a continuous variable and categorically in the following groups: <75%, 75% to <90%, 90% to 120%, and ≥120%.

4.5.6 Dose Modification

The number and percentage of patients with the following dose modifications will be summarized:

- Dose reduction due to AE (0, 1, 2, or >2)
- Dose interruption/missing due to AE (0, 1, 2, or >2)
- Dose discontinuation due to AE
- Dose reduction (0, 1, 2, or >2)
- Dose interruption/missing (0, 1, 2, or >2)
- Dose discontinuation
- Dose increase (Yes, No)

Time to first dose reduction among patients with at least one dose reduction is the time elapsed, in weeks, from the date of first dose of study drug to the date when the first dose reduction occurred. Time to first dose reduction will be summarized as a continuous variable.

Details of dose modifications will be provided in listings.

4.5.7 Treatment Compliance

The number of pills dispensed, returned, and taken (defined as the difference between the number of pills dispensed and returned), and compliance (defined as the percentage of pills taken relative to pills dispensed) will be summarized descriptively for the Safety Population.

4.6 Efficacy Analyses

All efficacy analyses will be conducted by treatment for the ITT Population. Selected analyses may be conducted in the PP and RE Populations. Endpoints involving response assessment will be based on central radiology per mRECIST 1.1. Investigator assessment per mRECIST 1.1, or central radiology assessment per Choi criteria will only be used supportively.

4.6.1 Multiplicity

A gate-keeping method will be implemented to control Type I error. This strategy will apply to the primary and key secondary efficacy endpoints: PFS, ORR, and OS.

PFS will be analyzed first when approximately 264 events are reached. In line with the gate keeping strategy, ORR will only be assessed for superiority should PFS have demonstrated superiority.

A formal OS interim analysis will be performed at the time of the final analysis of PFS and ORR for which superiority is demonstrated. The alpha for the interim analysis of the OS will follow the O'Brien-Fleming boundary method. The amount of alpha spent on OS at the time of the final PFS analysis will be specified according to the actual proportion of deaths observed at the time of the database lock regardless of whether the exact targeted number of PFS events has occurred at the lock. The exact amount of alpha spending at the interim analysis will be calculated using the alpha spending function developed by Lan and DeMet. OS will be analyzed when superiority is demonstrated for both PFS and ORR. The OS will also be analyzed at the follow-up when 264 events (deaths) occur, using the remaining unspent alpha.

4.6.2 Analysis of Primary Efficacy Endpoint

The primary endpoint of PFS is defined as the time from the date of randomization to the earlier of the date of first documented disease progression or death due to any cause.

The primary analysis for PFS is based on the central radiological assessment of PD per mRECIST 1.1.

Patients who are still in response at time of data cutoff will be censored at their last valid assessment. Complete censoring rules (both European Medicines Agency [EMA] and United States [US] Food and Drug Administration [FDA]) are specified in Table 2. The analysis will be primarily based on FDA Guidance for Cancer Trial Endpoints (FDA, 2007). The

censoring rules based on the EMA Guideline on the evaluation of anticancer medicinal products in man will be used as a sensitivity analysis (EMA, 2017).

The Kaplan-Meier (KM) method will be used to estimate the survival distribution function. The median PFS and its two-sided 95% CI and the 25th and 75th percentiles will be estimated. In addition, the event rates (or event-free) at specific time-points (e.g., 3-, 6-, and 12-month, *etc.*) will be computed, along with their standard errors using Greenwood's formula (Klein, 2003). Plots of survival curves by treatment using the KM method will be presented. Additional survival plots will be provided by treatment for each of the three stratification factors (defined in Section 3.1).

The primary treatment comparison is based on a stratified log-rank test.

The hazard ratio (HR) and its 95% CI will be estimated based on a stratified Cox model with treatment as the explanatory variable and strata for the stratification factors.

If patients were randomized based on incorrect strata, then the corrected strata will be used for stratified analyses rather than the incorrect strata recorded at randomization.

As a sensitivity analysis, the stratified analyses of PFS based on the central radiological assessment of PD per mRECIST 1.1 will be repeated for the PP Population. Unstratified analyses based on the ITT population may also be conducted.

PFS derived from the Investigator's assessment per mRECIST 1.1 will be used as a supportive analysis of the primary endpoint.

Table 2: Progression Free Survival Censoring Rules

Situation	Date of Progression or Censoring	FDA Censoring Rules	EMA Censoring Rules
No baseline assessments and alive after 2 scheduled assessments (defined by 130 days)	Date of randomization	Censored	Censored
Progression documented between scheduled visits	Date of radiological assessment showing new lesions or an increase in measured lesions	Event	Event
No progression	Date of last radiological assessment with evidence of no progression (or randomization date if no assessment)	Censored	Censored
New anticancer/ non-protocol treatment started prior to progression	Date of last radiological assessment with evidence of no progression prior to the start of new anticancer treatment	Censored	Event at date of disease progression/death
Death before the second scheduled post-baseline assessment (defined by within 130 days of date of randomization)	Date of death	Event	Event
Death between scheduled assessments	Date of death	Event	Event
Death or progression immediately after an extended lost to follow-up time (2 more missed scheduled assessments defined by at least x[1] days)	Date of last radiological assessment with evidence of no progression (randomization date if first 2 scheduled assessments are missing)	Censored	Event at date of disease progression/death

[1] x = 130 days if the death or progression date is ≤336 days from first dose date; x = 210 days if the death or progression date is ≥337 days from randomization date.

4.6.3 Analysis of Secondary Efficacy Endpoints

4.6.3.1 *Objective Response Rate*

The ORR is defined as the proportion of patients with a best response of CR or PR. The primary analysis of ORR will be assessed by central radiology per mRECIST 1.1 for the ITT population.

The ORR and its two-sided 95% CI will be estimated using the exact binomial distribution (Clopper-Pearson).

Additionally, the best overall response following the hierarchical order of CR, PR, SD, PD, and not evaluable (NE) will be tabulated. Non-CR/Non-PD will be treated as SD.

A Cochran-Mantel-Haenszel (CMH) test will be performed to test for treatment difference. Stratification will be by the randomization stratification factors. An unstratified CMH test will be performed as a sensitivity analysis.

The following additional sensitivity analyses will be conducted to assess the robustness of the primary analysis:

- ORR assessed by central radiology per mRECIST 1.1 in the PP and RE Populations.
- ORR assessed by central radiology per mRECIST 1.1 in the ITT Population by each stratification factor.
- ORR as assessed by the Investigator in the ITT Population will be summarized for best overall response.

4.6.3.2 *Overall Survival*

Overall survival is defined as the time from randomization to the date of death. Patients who die before or on the data cutoff date will be considered to have had an OS event. All patients who do not have a death record prior to or on the cutoff date will be censored at the last date known alive.

The primary treatment comparison will be based on a stratified log-rank test using the ITT Population. The HR and its 95% CI will be estimated using a stratified Cox model with treatment as the explanatory variable and strata as the stratification factors. An unstratified log-rank test will be performed as a sensitivity analysis.

A rank preserving structural failure time (RPSFT) model will be applied as a supportive analysis to account for treatment crossover effects from regorafenib to avapritinib. The survival time gained/lost by receiving avapritinib after crossover in the regorafenib group will be estimated. The RPSFT reconstructs the survival duration of crossover patients as if they had never received avapritinib, assuming treatment acts by multiplying survival time by a given factor once a patient starts receiving avapritinib (Korhonen et al, 2012). See Appendix 2 (Section 7.4) for further details of implementing this method.

Kaplan-Meier plots will be presented for overall survival for both the primary and sensitivity analyses.

4.6.3.3 *Duration of Response*

Duration of response is defined as the time from first documented response (CR/PR) to earlier of the date of first documented disease progression or death due to any cause.

Patients without CR or PR will be excluded from this analysis. Patients who are still in response at time of data cutoff will be censored at their last valid assessment. Complete censoring rules (both European Medicines Agency [EMA] and United States [US] Food and Drug Administration [FDA]) are specified in Table 2. The analysis will be primarily based on FDA Guidance for Cancer Trial Endpoints (FDA, 2007). The censoring rules based on the EMA Guideline on the evaluation of anticancer medicinal products in man will be used as a sensitivity analysis (EMA, 2017).

Duration of response will be analyzed for the responders (CR/PR) in the ITT Population using KM methods and will include the estimated median with two-sided 95% CI and the 25th and 75th percentiles. DOR at specific time-points (e.g., 3-, 6-, and 12-month, etc.) will be computed with respective standard errors using Greenwood's formula (Klein, 2003).

Sensitivity analyses will be conducted for DOR based on Investigator assessment per mRECIST 1.1. Both FDA and EMA censoring rules will be applied.

4.6.3.4 *Disease Control Rate*

Disease control rate (DCR) is defined as the proportion of patients with a CR/PR of any duration, or SD lasting for 16 weeks from date of randomization. The response will be assessed per mRECIST 1.1. DCR and its two-sided 95% CI will be estimated for the ITT Population using the exact binomial distribution (Clopper-Pearson).

4.6.3.5 *Response Rate by Choi Criteria*

The ORR per Choi response criteria (Weng, 2013) and its two-sided 95% CI will be estimated for the ITT Population using the exact binomial distribution (Clopper-Pearson).

The best overall response per Choi criteria following the hierarchical order of CR, PR, SD, PD, and NE will be tabulated.

4.6.3.6 *EORTC-QLQ-C30*

All questionnaire analyses will be based on the ITT population, and only includes those patients with baseline assessments for each instrument.

The EORTC QLQ-C30 (version 3.0) is a 30-item questionnaire used to evaluate QoL, and includes five functional domains (physical (PF), cognitive (CF), role (RF), emotional (EF), and social (SF)), three symptom scales (fatigue (FA), nausea and vomiting (NV), and pain (PA)), a global health status / QoL scale (QL), and six single items (dyspnea (DY), insomnia (SL), appetite loss (AP), constipation (CO), diarrhea (DI), and financial difficulties (FI)).

All of the scales and single-item measures range in score from 0 to 100. A high score for a functional scale represents a high/healthy level of functioning, a high score for the global health status represents a high QoL, but a high score for a symptom scale / item represents a high level of symptomatology / problems. The scoring method is outlined in Table 3.

Table 3: Scoring the QLQ-C30 version 3.0

	Scale	Number	Item	Version	Function
Global health Status /QoL	QL	2	6	29, 30	
Functional Scales					
Physical functioning	PF	5	3	1 to 5	F
Role functioning	RF	2	3	6,7	F
Emotional functioning	EF	4	3	21 to 24	F
Cognitive functioning	CF	2	3	20, 25	F
Social function	SF	2	3	26, 27	F
Symptom scales/items					
Fatigue	FA	3	3	10, 12, 18	
Nausea and vomiting	NV	2	3	14,15	
Pain	PA	2	3	9, 19	
Dyspnea	DY	1	3	8	
Insomnia	SL	1	3	11	
Appetite loss	AP	1	3	13	
Constipation	CO	1	3	16	
Diarrhea	DI	1	3	17	
Financial difficulties	FI	1	3	28	

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

Raw score (RS) is calculated as the average of item score when at least half of the items are not missing. A linear transformation to 0-100 will then be applied to get the score (S) use ranges provided in Table 3.

For functional scales $S = (1 – (RS – 1) / \text{range}) \times 100$

For symptom scales / items and global health status $S = ((RS – 1) / \text{range}) \times 100$

Calculated scale scores will be descriptively summarized at baseline and Day 1 of each assessed cycle. Change and percentage change from baseline will also be summarized for post-baseline assessments.

Box plots of the change from baseline will be presented.

As a sensitivity analysis, cumulative density function (CDF) plots of QoL will be produced at Day 1 of each assessed cycle: the horizontal axis displays the absolute score change from baseline, and the vertical axis shows the proportion of patients whose absolute score change from baseline is less than or equal to the value indicated on the x-axis. Similar CDF plots will be produced for all 6 functional scales at Day 1 of each assessed cycle.

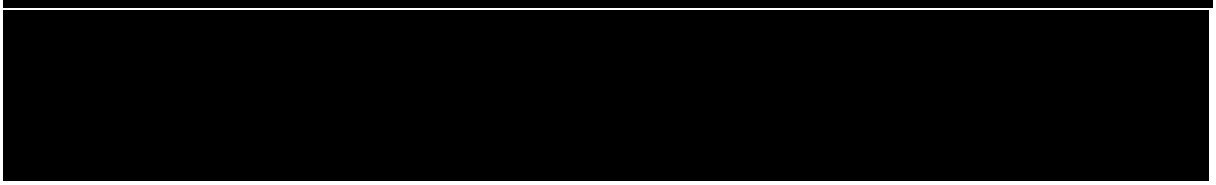
For physical functioning, pain, role functioning and appetite loss, the mean change from baseline to Week 12 will be compared between treatment groups using a t-test.

4.6.3.7 *Abdominal Pain*

Abdominal pain is measured as a numeric rating scale (0-10).

For each scheduled assessment a summary will be given of the observed data. In addition, for post-baseline assessments, the change and percentage change from baseline will also be summarized.

4.6.4 Analysis of Exploratory Efficacy Endpoints



[redacted]

4.6.4.6 [redacted]

[redacted]

[redacted]

4.6.5 Subgroup Analysis

For ORR, DOR, and PFS as assessed by central radiology, and OS, the following subgroup analyses will be conducted. Corresponding forest plots will be provided based on the odds ratio or hazard ratio by treatment for each subgroup. Analyses will only be performed if the subgroup contains ≥ 5 patients.

- Age (<65 years, ≥65 years)
- Gender (Male, Female)
- Region (US, Europe, Asia, Other) *Australia to be combined with Europe
- Race (White, Asian, Other)
- Largest target lesion (≤10 cm, >10 cm)
- ECOG Performance Score at Baseline (0, 1)
- V654A or T670I Mutation Status (Present, Absent)
- PDGFRA D842V Mutation Status (Present, Absent)
- Stratification factors (Respective stratification levels).
- The following subgroup analyses will be performed in patients who are V654A mutation negative and T670I mutation negative:
    - D816X Mutation Status (Present, Absent)
    - D820X Mutation Status (Present, Absent)
    - N822X Mutation Status (Present, Absent)
    - Y823X Mutation Status (Present, Absent)

4.7 Safety Analyses

Safety will be evaluated by the incidence of AEs, causality, intensity, seriousness, and outcome of AEs, the patient's vital signs, ECOG PS, clinical laboratory test results, and ECG data.

All safety data will be summarized by treatment group for the Safety Population.

4.7.1 Adverse Events

Adverse Events will be analyzed using treatment-emergent adverse events (TEAEs). The TEAE definition is dependent on the study period (e.g., main study treatment period) reported. An AE is considered a TEAE for a specified study period if the AE occurs during or after administration of the first dose of study drug from the same study period through up to the earlier of the last dose of study drug in that period + 30 days or the day before the start of crossover/extended treatment.

Adverse Event refers to TEAEs unless otherwise specified throughout this document. All AEs will be coded using MedDRA version 18.1. The severity will be graded using the National Cancer Institute (NCI) Common Terminology Criteria for Adverse Events (CTCAE), version 4.03 (patients enrolled before protocol amendment 2) and 5.0.

A patient experiencing multiple AEs under the same PT or SOC will be counted only once for that PT or SOC by maximum severity. If a patient experiences the same AE more than once with more than one causal relationship to study drug, the strongest causal relationship to study drug will be given precedence. Any missing severity, causality, or outcome will not be imputed and classified as 'Missing'. Detailed imputation rules for missing AE dates are in Section 7.1.

The number of patients with at least one AE will be summarized (Table 4). The following tables summarize incidences of various aspects of AEs and serious adverse events (SAEs):

1. AE Summary (including all subsequent items)
2. AE by PT
3. AE by SOC and PT
4. AE related to study drug by PT
5. AE related to study drug by SOC and PT
6. SAE by PT
7. SAE by SOC and PT
8. SAE related to study drug by PT
9. SAE related to study drug by SOC and PT
10. Grade 3/4/5 AE by PT
11. Grade 3/4/5 AE by SOC and PT
12. Grade 3/4/5 AE related to study drug by PT
13. Grade 3/4/5 AE related to study drug by SOC and PT
14. AE by PT and NCI CTCAE Grade
15. AE by SOC, PT, and NCI CTCAE Grade

16. AE related to study drug by PT and NCI CTCAE Grade
17. AE related to study drug by SOC, PT, and NCI CTCAE Grade
18. SAE by PT and NCI CTCAE Grade
19. SAE by SOC, PT, and NCI CTCAE Grade
20. SAE related to study drug by PT and NCI CTCAE Grade
21. AE leading to interruption of study drug by SOC and PT
22. AE leading to interruption of study drug by PT
23. AE leading to dose reduction of study drug by SOC and PT
24. AE leading to dose reduction of study drug by PT
25. AE leading to permanent discontinuation of study drug by SOC and PT
26. AE leading to permanent discontinuation of study drug by PT
27. AE related to study drug leading to interruption of study drug by PT
28. AE related to study drug leading to dose reduction of study drug by PT
29. AE related to study drug leading to permanent discontinuation of study drug by PT

All AEs will be listed by patient.

Subgroup analyses will be conducted for items 1 – 13 above (details in Table 4) in the Safety Population.

4.7.2 Adverse Events of Special Interest

Adverse events of special interest (AESI) will be summarized by category (cognitive effects and intracranial bleeding) and relevant PTs.

- Cognitive effects consisting of 4 PTs: cognitive disorder, memory impairment, confusional state, encephalopathy
- Intracranial bleeding consisting of 3 PTs: haemorrhage intracranial, cerebral haemorrhage, subdural haematoma.

The general rules for summarizing AE specified in Section 4.7.1 apply for AESIs for avapritinib. The following AESI tables will be presented:

- AESI by AESI category and PT
- AESI related to study drug by AESI category and PT
- Serious AESI by AESI category and PT
- Serious AESI related to study drug by AESI category and PT
- Grade 3/4/5 AESI by AESI category and PT

- Grade 3/4/5 AESI related to study drug by AESI category and PT
- AESI by AESI category, PT, and NCI CTCAE Grade
- AESI leading to permanent discontinuation of study drug by PT
- AESI related to study drug leading to permanent discontinuation of study drug by PT

4.7.3 Dementia Adverse Events

Dementia adverse events will be summarized using MedDRA Dementia SMQ PTs. The general rules for summarizing AE specified in Section 4.7.1 apply for dementia AE for avapritinib. The following dementia AE tables will be presented:

- Dementia AE by PT
- Dementia AE related to study drug by PT
- Serious Dementia AE by PT
- Serious Dementia AE related to study drug by PT
- Grade 3/4/5 Dementia AE by PT
- Grade 3/4/5 Dementia AE related to study drug by PT
- Dementia AE by PT and NCI CTCAE Grade
- Dementia AE related to study drug by PT and NCI CTCAE Grade
- Serious Dementia AE by PT and NCI CTCAE Grade
- Serious Dementia AE related to study drug by PT and NCI CTCAE Grade

Table 4: Summary of Adverse Event Tables:

Safety Endpoint	Main Table	by PT		by SOC, PT		by PT, CTCAE	by SOC, PT, CTCAE
		Overall	Subgroup	Overall	Subgroup	Overall	Overall
Summary Table (each item below one row)	X						
AE		X	X	X	X	X	X
SAE		X	X	X	X	X	X
Grade 3+ AE		X	X	X	X		
Related AE		X	X	X	X	X	X
Related SAE		X	X	X	X	X	
Related Grade 3+ AE		X	X	X	X		
AE Leading to Permanent Study Discontinuation		X		X			
Related AE Leading to Permanent Study Discontinuation		X					
AE Leading to Dose Interruption		X		X			
Related AE Leading to Dose Interruption		X					
AE Leading to Dose Reduction		X		X			
Related AE Leading to Dose Reduction		X					
Listing (Death, Discontinuation, SAE)	X						
AESI by Category and		X				X	
Related AESI by Category and		X					
Serious AESI by Category and		X					
Related Serious AESI by Category and		X					
Grade 3+ AESI by Category and		X					
Related Grade 3+ AESI by Category and		X					
AESI Leading to Permanent Study Discontinuation		X					
Related AESI Leading to Permanent Study Discontinuation		X				X	
Dementia AE		X				X	
Related Dementia AE		X				X	

Serious Dementia AE		X				X	
Related Serious Dementia AE		X				X	
Grade 3+ Dementia AE		X					
Related Grade 3+ Dementia AE		X					

4.7.4 Deaths

On treatment deaths will be summarized by treatment group by SOC and PT.

A listing of patients who died will be provided specifying the date of death, the cause of death, and the date of last dose of study drug. On treatment deaths, defined as deaths which occur between the first dose date of study drug to last dose date of study drug during the relevant study treatment period + 30 days, will be clearly marked.

4.7.5 Clinical Laboratory Evaluations

The central laboratory parameters with both baseline and post-baseline assessments will be summarized by treatment group and scheduled visit for the Safety Population. If there are multiple values recorded for a specific visit, then the last reported value will be used for that visit. Baseline is defined as the last assessment prior to start of study treatment. Data will be presented in Standard International units.

Clinical laboratory values will be graded according to NCI CTCAE version 4.03 or 5.0 for applicable parameters. Shift tables of laboratory data from baseline to worst grade during treatment, including unscheduled visits, will be presented. For laboratory parameters (including hemoglobin, glucose, sodium, potassium, magnesium, glucose, bilirubin, and phosphate), shift tables in maximum CTCAE grades of worst high and worst low assessment will be presented separately.

Overall summary and changes from baseline will be presented for hematology and serum chemistry (including coagulation) laboratory values.

Listings will be provided for each laboratory parameter. All results outside predefined normal ranges and CTCAE grade will be flagged in the data listings.

Urinalysis will be presented in listing only.

Boxplots of selected lab tests (including leukocytes, neutrophils, platelets, hemoglobin, alanine aminotransferase, aspartate aminotransferase, alkaline phosphatase, total bilirubin, and phosphate) values will be presented by visit.

Local laboratory tests will be listed only, with the exception of the shift tables, which will include both central and local results when deriving the worst grade.

4.8 Vital Signs

Vital signs (weight, systolic blood pressure, diastolic blood pressure, heart rate, and temperature) will be summarized at each scheduled assessment for the Safety Population. The change and percentage change from baseline will also be summarized for post-baseline assessments.

4.9 Electrocardiograms

Electrocardiogram overall interpretations (normal, abnormal not clinically significant [NCS], abnormal clinically significant [CS], and NE) and changes from baseline [expressed as improvement, no change, and deterioration] will be summarized by treatment group and scheduled visit for the Safety Population.

- Improvement = abnormal CS to abnormal NCS or normal, abnormal NCS to normal
- Deterioration = normal to abnormal NCS or CS, abnormal NCS to abnormal CS
- No change = normal to normal, abnormal NCS to abnormal NCS, abnormal CS to abnormal CS.

If an interpretation is missing for any patient, then an 'Unknown' category will be presented.

ECG parameters (heart rate, PR interval, QT interval, QRS interval, and QTcF) will be descriptively summarized by scheduled visit. The change and percentage change from baseline will be summarized for post-baseline assessments.

The proportion of patients with a QT interval and QTcF interval >450, >480 and >500 msec will be summarized. In addition, the proportion of patients with a change from baseline in QT interval and QTcF interval >30 and >60 msec will also be summarized.

Note, ECGs are not assessed following either continuation or crossover of treatment to avapritinib, following disease progression based on central review.

4.10 Physical Examination

Any abnormal, clinically significant physical examination results captured at screening will be treated as either medical history or adverse event data, as appropriate. Data will be listed for patients who have physical examinations performed.

4.11 Basic Neurological Assessment

Basic neurological assessment results at screening will be summarized.

4.12 ECOG Performance Status

ECOG PS will be summarized by scheduled visit. The number and percentage of patients with each response will be summarized.

Note, ECOG PS is also scheduled to be captured following either continuation or crossover of treatment to avapritinib, following disease progression based on central radiological assessment. These scheduled assessments will also be summarized.

5. INTERIM ANALYSES

No interim analysis is planned for the primary endpoint of this study. An interim analysis of OS is planned at the time of final analysis when superiority of the PFS and ORR are established.

The OS interim analysis will be performed at the time of the final analysis of PFS and ORR for which superiority is demonstrated. The alpha for the interim analysis of the OS will follow the O'Brien-Fleming boundary method. The amount of alpha spent on OS at the time of the final PFS analysis will be specified according to the actual proportion of deaths observed at the time of the database lock regardless of whether the exact targeted number of PFS events has occurred at the lock. The exact amount of alpha spending at the interim analysis will be calculated using the alpha spending function developed by Lan and DeMet. OS will be analyzed when superiority is demonstrated for both PFS and ORR. The OS will also be analyzed at the follow-up when 264 events (deaths) occur, using the remaining unspent alpha.

6. CHANGES TO PLANNED ANALYSES FROM THE PROTOCOL

There is no change to planned analyses from the protocol.

7. APPENDICES

7.1 Data Imputation Guidelines

No imputation will be made for completely missing date unless otherwise specified. General imputation rules mentioned below apply to partially missing or impossible dates:

- If the stop date is not missing, and the imputed start date is after the stop date, the start date will be imputed by the stop date
- If the start date is not missing, and the imputed stop date is before the start date, then the imputed stop date will be equal to the start date
- Any imputed dates need to be logical. For example, last dose date should not be later than death date.

When imputation rules in subsequent sections contradicts the general rule, always follow the general rule.

7.1.1 Adverse Event Date Imputation

Follow the general rule specified in Section 7.1.

Incomplete Start Date:

Missing day, month, and year

- No imputation will be made; the corresponding AE will be included.

Missing day and month

- If the year is the **same** as the year of the first dose date, then impute day and month as the day and month of the first dose date
- If the year is **prior to** the year of the first dose date, then impute day and month as 31 Dec
- If the year is **after** the year of the first dose date, then impute day and month as 01 Jan.

Missing day only

- If the month and year are the **same** as those of the first dose date, then impute day as the day of the first dose date
- If either the year of partial date is **before** the year of the first dose date, or the years are the same, but the month of partial date is **before** the month of the first dose date, then impute day as last day of the month
- If either the year of partial date is **after** the year of the first dose date, or the years are the same, but the month of partial date is **after** the month of the first dose date, then impute day as first day of the month.

Incomplete Stop Date:

Missing day, month, and year

- No imputation will be made.

Missing day and month

- If the year is the **same** as the year of the last dose date, then impute day and month as the day and month of the last dose date
- If the year is **prior to** the year of the last dose date, then impute day and month as 31 Dec
- If the year is **after** the year of the last dose date, then impute day and month as 01 Jan.

Missing day only

- If the month and year are the **same** as those of the last dose date, then impute day as the day of the last dose date;
- If either the year of partial date is **not the same as** the year of the last dose date, or the years are the same, but the month of partial date is **not the same as** the month of the last dose date, then impute day as last day of the month.

7.1.2 Concomitant Medication Date Imputation

Follow the general rules specified in Section 7.1 and rules in Section 7.1.1.

7.1.3 Prior Therapies Date Imputation

Follow the general rule specified in Section 7.1.

For start date partial as month and year are available, then impute day as '01'. *E.g.* impute partial date of 'DEC2013' as '01DEC2013'.

For start date partial as year only is available, then impute day and month as '01JAN'. *E.g.* impute partial date of '2013' as '01JAN2013'.

For end date partial as month and year are available, then impute day as last day of the month. *E.g.* impute partial date of 'JUN2013' as '30JUN2013'.

For end date partial as year only is available, then impute day and month as the last day of the year. *E.g.* impute partial date of '2013' as '31DEC2013'.

If the imputed starting date is earlier than initial diagnosis date, it should be set as the initial diagnosis date. No overlap between the exposure to prior therapies and study drug will be allowed and any overlap of exposure will be queried at data review stage. The end date of prior therapies will be imputed to first dose date of study drug – 15 if there is overlap due to imputation of partial dates. When there are multiple lines of prior therapies, the end date of

prior therapies will be imputed to first dose date of the subsequent line of therapy -1 if there is overlap due to imputation of partial dates. If the end year for one prior therapy is the same as the start year for the next prior therapy, impute Jun 30 and Jul 1 of that year for the end and start days if missing months and days. If the end year and month for one prior therapy is the same as the start year and month for the next prior therapy, impute 15th and 16th of that year and month for the end and start days if missing days.

7.1.4 Death Date Imputation

- If death date is completely missing, will use latest alive date + 1
- If both month and day are missing, then use the first date (01 JAN) of the year, or latest alive date + 1, whichever is later
- If only day is missing, then use the first day of the month, or latest alive date + 1, whichever is later.

7.1.5 Post-Therapies Date Imputation

Follow the general rule specified in Section 7.1.

7.1.6 Other Imputations

Follow the general rule specified in Section 7.1.

7.2 Table, Listing, and Figure Format

In the top left portion of each table/listing, a *table/listing number* followed by the *title* of the table/listing will be presented. After the title line, optional *sub-title* or *population* information can be presented. Horizontal lines will appear before and after the column heading of the table/listing. *Footnotes* will be put under the main body of text at the bottom of the page.

The *sponsor name*, *protocol number,* programmer user identification, status of the table/listing (i.e. draft or final) and *SAS program name* will appear bottom left in a string and the *page number* will appear on the bottom right corner of each table/listing. The *date and time of creation* of table/listing will appear bottom left under the sponsor name. The source listing number will appear bottom left.

A *landscape layout* is for both table and listing presentations for post-text tables. A *portrait* layout is for in text table.

The *left* and *right margins* of all tables and listings will be a minimum of 2.1 cm from the left and 1.9 cm from the right. The *top and bottom margins* will be a minimum 2.92 cm. *Header and footer* will be both 1.27 cm.

There is no special requirement of *font type* and *size*, but an *8-point* font size for tables and *7*or *8-point* for listings is proposed using *Courier New* font. A maximum SAS line size = 141 and page size = 44 for *8-point* font size, and line size = 161 and page size = 50 for *7-point* will be used to fit on both United Kingdom and US paper sizes.

In a listing, in the case that a patient's record has been continued to the next page, an appropriate identification (e.g. the patient identification number) must be presented at the beginning of that page.

7.3 Conventions

Continuous variables will be summarized using the number of observations (n), mean, standard deviation (Std Dev), median, minimum, and maximum. The same number of decimal places as in the raw data will be presented when reporting minimum and maximum, 1 more decimal place than in the raw data will be presented when reporting mean and median, and 2 more decimal places than in the raw data will be presented when reporting standard deviation.

Time to event variables will be summarized using Kaplan-Meier (KM) method, which will include the estimated median with 95% CIs and 25th and 75th percentiles.

Unless otherwise specified frequency tabulations will be presented by number and percentage, where the percentage is presented in brackets to 1 decimal place (XX.X%).

P-values, if applicable, will be presented to 3 decimal places. If a p-value is less than 0.05 but is greater than or equal to 0.01, then an asterisk (*) will be added next to this value. If a p-value is less than 0.01 but is greater than or equal to 0.001, then two asterisks (**) will be

added next to this value. Finally, if the p-value is less than 0.001 then three asterisks (***) will be added next to this value and it will be presented as <0.001. If the rounded result is a value of 1.000, it will be displayed as >0.999.

Any date information in the listing will use the *date9.* format, for example, 07MAY2002. In the listing, a unit associated with a variable will be presented only once within parentheses either below or next to that variable in the heading portion. If a parameter has multiple units, each unit will be displayed only once, as applicable.

All tables will have their source listing referenced in a footnote. Listings should be sorted by analysis group, patient and visit and have the source data received by data management referenced in a footnote. All tables and listings will be converted into Microsoft Word documents and collated into two complete documents.

An Unknown or Missing category are added to each parameter for which information is not available for 1 or more patients.

Unless otherwise specified, the estimated mean and median for a set of values are printed out to 1 more significant digit than the original values, and standard deviations are printed out to 2 more significant digits than the original values. The minimum and maximum should report the same significant digits as the original values. For example, for systolic blood pressure:

N	XX
Mean (Std Dev)	XXX.X (X.XX)
Median	XXX.X
Min, Max	XXX, XXX

Tabular display of data for medical history, prior/concomitant medications, and all tabular displays of adverse event data are presented by the body system, treatment class, or SOC with the highest occurrence in the avapritinib treatment group in decreasing order, assuming all terms are coded. Within the body system, drug class and SOC, medical history (by preferred term), drugs (by ATC1 code), and adverse events (by preferred term) are displayed in decreasing order. If incidence for more than 1 term is identical, they should then be sorted alphabetically. Missing descriptive statistics or p-values which cannot be estimated are reported as “-”.

The percentage of patients is normally calculated as a proportion of the number of patients assessed in the relevant treatment group (or overall) for the analysis set presented.

For categorical summaries (number and percentage of patients) where a patient can be included in more than one category, describe in a footnote or programming note if the patient is included in the summary statistics for all relevant categories or just 1 category and the criteria for selecting the criteria.

7.4 Implementing the Rank-Preserving Structural Failure Time Model

A conventional analysis of overall survival, taking into account only the randomized treatment arm, does not account for the crossover from regorafenib to avapritinib, and the analysis can be biased and treatment effect underestimated. The rank-preserving structural failure time (RPSFT) model attempts to minimize this bias.

The model assumes that once a patient has switched from regorafenib to avapritinib, the overall survival that would have been seen, had the patient not switched, is multiplied by a given factor once the patient has switched. The observed survival for a patient who has switched is then transformed as if they had never received avapritinib and had continued on regorafenib.

The overall survival for a given patient, can be written as the sum of the time they were off avapritinib + the time they were on avapritinib. Here the time they were off avapritinib is when they were on regorafenib:

$T_i = T_{offi} + T_{oni}$

The above formulation is also valid for patients randomized to avapritinib ($T_{offi} = 0$) and those randomized to regorafenib who do not switch ($T_{oni} = 0$).

The aim of the RPSFT model is to estimate for patients who do switch treatment, the survival time that would have been seen if they had not switched treatment. This adjusted survival time, U_i, is derived as:

$U_i = T_{offi} + T_{oni}(\exp(\psi))$ [1]

$\exp(\psi)$, the acceleration factor, is the factor by which the time on avapritinib is multiplied in order to estimate the time that would have been seen if the patient had not switched.

In applying the model, the first step is to estimate ψ. This is done by applying formular [1] to all patients. At the correct value of ψ, an analysis of U_i by randomized treatment should result in a test statistic of 0. The log rank test will be used for this analysis, and the correct value of ψ is derived iteratively by running through a search grid in order to find the optimum value.

Once ψ has been derived, the survival times for patients who switch treatment are then transformed according to [1] above. Note, censoring needs to be applied as appropriate for transformed times that extend further than the data cutoff/end of study.

Once the transformation has been done, the data is then analyzed as usual by an appropriate method.

Further details of the method, and specifically the implementation in SAS, can be found in the PharmaSUG 2018 – Paper EP-04, Implementing the Rank-Preserving Structural Failure Time Model in SAS and R, Bradford J Danner and Indrani Sarkar.

8. REFERENCES

Cassier PA, Fumagalli E, Rutkowski P, Schoffski P, Van Glabbeke M, Debiec-Rychter M, et al. Outcome of patients with platelet-derived growth factor receptor alpha-mutated gastrointestinal stromal tumors in the tyrosine kinase inhibitor era. Clin Cancer Res. 2012 Aug 15;18(16):4458-64.

Demetri GD, Reichardt P, Kang YK, Blay JY, Rutkowski P, Gelderblom H, et al. Efficacy and safety of regorafenib for advanced gastrointestinal stromal tumours after failure of imatinib and sunitinib (GRID): an international, multicentre, randomised, placebo-controlled, phase 3 trial. Lancet. 2013a;381(9863):295-302.

Eisenhauer EA, Therasse P, Bogaerts J, Schwartz LH, Sargent D, Ford R, Dancey J, Arbuck S, Gwyther S, Mooney M, Rubinstein L, Shankar L, Dodd L, Kaplan R, Lacombe D, Verweij J. New response evaluation criteria in solid tumours: revised RECIST guideline (version 1.1). Eur J Cancer. 2009; 45:228–47.

EMA. Guideline on the evaluation of anticancer medicinal products in man [Online]. Available at: https://www.ema.europa.ed/docs/en_GB/document_library/Scientific_ guidelines /2017/11/WC500238764.pdf. Updated 22 Sep 2017.

Goodman VL, Rock EP, Dagher R, Ramchandani RP, Abraham S, Gobburu JV, et al. Approval summary: sunitinib for the treatment of imatinib refractory or intolerant gastrointestinal stromal tumors and advanced renal cell carcinoma. Clin Cancer Res. 2007 Mar 1;13(5):1367-73.

Heinrich MC, Maki RG, Corless CL, Antonescu CR, Harlow A, Griffith D, et al. Primary and secondary kinase genotypes correlate with the biological and clinical activity of sunitinib in imatinib-resistant gastrointestinal stromal tumor. J Clin Oncol. 2008 Nov 20;26(33):5352-9.

Kang YK, Ryu MH, Yoo, C, Ryoo BY, Kim HJ, Lee JJ, et al. Resumption of imatinib to control metastatic or unresectable gastrointestinal stromal tumours after failure of imatinib and sunitinib (RIGHT): a randomised, placebo-controlled, phase 3 trial. Lancet Oncol. 2013 Nov; 14(12):1175-82.

Klein JP, Moeschberger ML. Survival analysis technique for censored and truncated data (second edition). Springer-Verlag; 2003.

Stivarga® [Prescribing Information]. Whippany, NJ: Bayer HealthCare Pharmaceuticals Inc.; 2017 [Online]. Available at: http://labeling.bayerhealthcare.com/html/products/pi/ Stivarga_PI.pdf. Accessed 31 August 2017.

Sutent® [Prescribing Information]. New York, NY: Pfizer Labs; 2014 [Online]. Available at: http://labeling.pfizer.com/ShowLabeling.aspx?id=607. Accessed 09 March 2015.

US FDA. Guidance for Industry Clinical Trial Endpoints for the Approval of Cancer Drugs and Biologics [Online]. Available at: https://www.fda.gov/downloads/Drugs/Guidances/ ucm071590.pdf. Updated May 2007.

Weng Z, Ertle J, Zheng S, Lauenstein T, Mueller S, Bockisch A, et al. Choi criteria are superior in evaluating tumor response in patients treated with transarterial radioembolization for hepatocellular carcinoma. Oncol Lett. 2013 Dec; 6(6):1707–12.

Yoo C, Ryu MH, Jo J, Park I, Ryoo BY, Kang YK. Efficacy of imatinib in patients with platelet-derived growth factor receptor alpha-mutated gastrointestinal stromal tumors. Cancer Res Treat. 2016 Apr;48(2):546-52.