NCT #: NCT03580655

STATISTICAL ANALYSIS PLAN

An Open-label, Single Arm, Phase 2 Study to Evaluate Efficacy and Safety of Avapritinib (BLU-285), A Selective KIT Mutation-targeted Tyrosine Kinase Inhibitor, in Patients with Advanced Systemic Mastocytosis

Protocol Number: BLU-285-2202

SAP Version:	Date of Statistical Analysis Plan:
Final 1.0	29 June2020

Approval Signatures:

See eSignature
Date

[redacted]
Prepared by:
[redacted]
Blueprint Medicines Corporation

See eSignature
Date

[redacted]
Blueprint Medicines Corporation

See eSignature
Date

[redacted]
Blueprint Medicines Corporation

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

LIST OF ABBREVIATIONS AND DEFINITIONS OF TERMS

Abbreviation	Term
AdvSM	Advanced systemic mastocytosis including subtypes of aggressive systemic mastocytosis (ASM), systemic mastocytosis with an associated hematologic neoplasm (SM-AHN), and mast cell leukemia (MCL)
AdvSM-SAF	AdvSM Symptom Assessment Form
AE	Adverse event
AESI	Adverse events of special interest
AHN	Associated hematologic neoplasm
ALB	Albumin
ALP	Alkaline phosphatase
ALT	Alanine aminotransferase
AML	Acute myeloid leukemia
ANC	Absolute neutrophil count
AP	Appetite loss
ASM	Aggressive systemic mastocytosis
AST	Aspartate aminotransferase
ATC	Anatomical therapeutic chemical
BM	Bone marrow
BMB	Bone marrow biopsy
BMI	Body mass index
C	Cycle
CD	cluster designation or cluster of differentiation
CDF	Cumulative density function
CEL	Chronic eosinophilic leukemia
CF	Cognitive functional domain
CI	Clinical improvement
CMML	Chronic myelomonocytic leukemia
CO	Constipation
CR	Complete remission
CRh	Complete remission with partial recovery of peripheral blood counts
CRF	Case report form
CSR	Clinical study report
CT	Computed tomography
CTCAE	Common Terminology Criteria for AEs
D	Day
ddPCR	Droplet Digital polymerase chain reaction
DI	Diarrhea
DNA	Deoxyribonucleic acid
DOB	Date of birth
DOR	Duration of response
DY	Dyspnea

Abbreviation	Term
ECG	Electrocardiogram
ECOG	Eastern Cooperative Oncology Group
EF	Emotional functional domain
EFS	Event free survival
[redacted]	[redacted]
EOT	End of treatment
FA	Fatigue symptom scale
FDA	Food and Drug Administration
FI	Financial difficulties
GI	Gastrointestinal
HGB	Hemoglobin
IWG-MRT-ECNM	International working group – myeloproliferative neoplasms research and treatment & European competence network on mastocytosis consensus response criteria
KIT	V-Kit Hardy-Zuckerman 4 feline sarcoma viral oncogene homolog
KM	Kaplan-Meier
LoR	Loss of response
MAF	Mutant allele fractions
MC	Mast cells
MCL	Mast cell leukemia
mCR	Morphologic CR
mCRh	Morphologic CRh
mDOR	Morphologic DOR
MDS	Myelodysplastic syndrome
MedDRA	Medical Dictionary for Regulatory Activities
mIWG	Modified international working group – myeloproliferative neoplasms research and treatment & European competence network on mastocytosis consensus response criteria
moCR	Molecular CR
moCRh	Molecular CRh
MPN	Myeloproliferative neoplasms
mPR	Morphologic PR
MRI	Magnetic resonance imaging
NA	Not Applicable
NCI	National Cancer Institute
NE	Not evaluable
NV	Nausea and vomiting symptom scale
ORR	Overall response rate
OS	Overall survival
PA	Pain symptom scale
PD	Progressive Disease
PF	Physical functional domain

Abbreviation	Term
PFS	Progression free survival
[redacted]	[redacted]
PK	Pharmacokinetic
PLT	Platelet (count)
PPR	Pure pathologic response
PPRE	Pure pathologic response-evaluable
PR	Partial remission
PRO	Patient reported outcome
PT	Preferred term
QD	Once daily
QoL	Quality of life
RBC	Red blood cell
RDI	Relative dose intensity
RE	Response-evaluable
RF	Role functional domain
RNA	Ribonucleic acid
RS	Raw score
S	Score
SAE	Serious adverse event
SAP	Statistical analysis plan
SAS	Statistical Analysis System
SD	Stable disease
SF	Social functional domain
SL	Insomnia
SM	Systemic mastocytosis
SM-AHN	Systemic mastocytosis with an associated hematologic neoplasm
SOC	System organ class
SSC	Study Steering Committee
TF	Treatment failure
TKI	Tyrosine kinase inhibitor
TSS	Total symptom score
TTR	Time to response
TTmR	Time to morphologic response
TTmCR	Time to morphologic CR
US	United States (of America)
WHO	World health organization

1. INTRODUCTION

This statistical analysis plan (SAP) describes the statistical analysis and data presentations to be performed for Blueprint Medicines Corporation, Protocol Number. BLU-285-2202 (amendment 7.1 dated 21 April 2020) "An Open-label, Single Arm, Phase 2 Study to Evaluate Efficacy and Safety of Avapritinib (BLU-285), A Selective KIT Mutation-targeted Tyrosine Kinase Inhibitor, in Patients with Advanced Systemic Mastocytosis" It contains definitions of analysis populations, derived variables, and statistical methods for the analysis of efficacy and safety.

This study consists of 2 cohorts:

- Cohort 1: Patients who are evaluable by modified International Working Group-Myeloproliferative Neoplasms Research and Treatment and European Competence Network on Mastocytosis (IWG-MRT-ECNM) criteria (whose with evaluable C-findings at baseline or MCL, regardless of C-findings)
- Cohort 2: Patients who are not evaluable by modified IWG-MRT-ECNM criteria (lacking C-findings at baseline, except for MCL)

One interim analysis (IA) for superiority will be performed at approximately 50% of the information fraction (ie, when 32 patients enrolled in Cohort 1 who are evaluable for response).

If the testing for primary endpoint is statistically significant at the interim analysis, the key secondary endpoint, other secondary endpoints, and exploratory endpoints will be tested and summarized in the corresponding patient populations. If the testing for primary endpoint is not statistically significant at the interim analysis, the final analysis will be conducted when 63 patients have been enrolled in cohort 1 and are evaluable for response.

The study data will be summarized descriptively by disease subtype, and overall for the clinical study report (CSR). Database will be locked after completion of medical review of the data, identification of protocol deviations, and data cleaning. SAP will be finalized and approved prior to the clinical database lock for interim analysis.

At the time of the data cutoff date for IA, the planned enrollment in Cohort 1 will be close to completion, however enrollment to Cohort 2 will be ongoing. An addendum to the CSR analysis will be conducted when all patients in the study have been enrolled and had sufficient treatment duration.

All statistical analyses will be conducted using Statistical Analysis System (SAS) version 9.4 or higher.

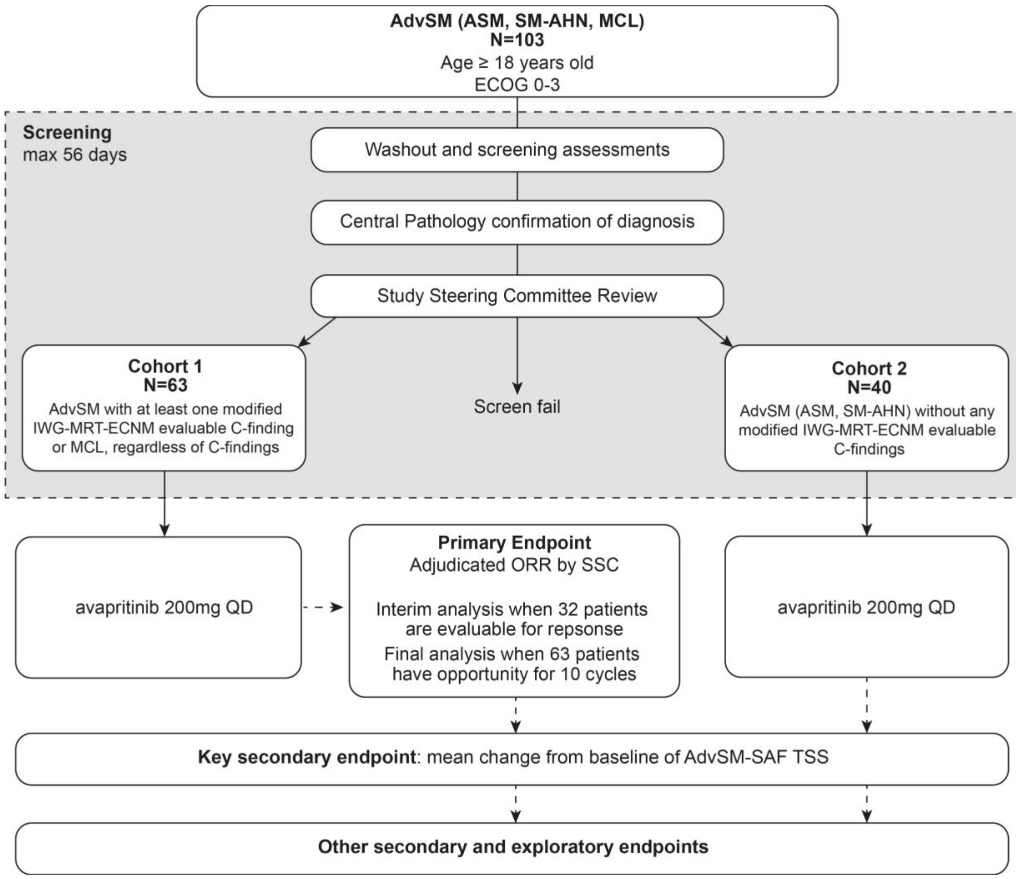
Analyses of Pharmacokinetic (PK) will be addressed separately in a standalone analysis plan.

1.1 Study Design

This is an open-label, single-arm, Phase 2 study evaluating the efficacy and safety of avapritinib in patients with a WHO diagnosis of AdvSM, including patients with ASM, SM-AHN, and MCL. Patients with a centrally confirmed WHO diagnosis of AdvSM will be enrolled into 1 of 2 cohorts. Cohort 1 will enroll patients, confirmed by a Study Steering Committee (SSC), to meet modified IWG-MRT-ECNM criteria for evaluable disease (have severe and quantifiable organ damage (an evaluable C-finding) or have MCL (regardless of C-findings). Cohort 1 patients will support the primary objective of determining adjudicated ORR by modified IWG-MRT-ECNM criteria.

Cohort 2 will enroll patients with AdvSM who are not considered eligible for an adjudicated response, those who were confirmed centrally to have ASM or SM-AHN but are lacking an evaluable C-finding (as determined by the SSC). Cohort 1 and Cohort 2 will be included in the key secondary objective of assessing efficacy on TSS improvement. Both cohorts will also support additional secondary and exploratory endpoints such as investigator determined response and safety. A study schematic is in Figure 1.

Figure 1: Study Schematic



Abbreviations: AdvSM = Advanced Systemic Mastocytosis; ASM = aggressive systemic mastocytosis; ECOG = Eastern Cooperative Oncology Group; IWG-MRT-ECNM = International Working Group-Myeloproliferative Neoplasms Research and Treatment and European Competence Network on Mastocytosis; MCL = mast cell leukemia; ORR = overall response rate;QD = once daily; SAF = Symptom Assessment Form; SM-AHN = systemic mastocytosis with an associated hematologic neoplasm; SSC = Study Steering Committee; TSS = total symptom score.

1.2 Study Objectives

1.2.1 Primary Objectives

The primary objective is to determine adjudicated ORR of CR/CR with partial recovery of peripheral blood (PB) counts (CRh) + PR + CI based on modified IWG-MRT-ECNM consensus response criteria in patients with AdvSM treated with avapritinib and enrolled in Cohort 1.

1.2.2 Secondary Objectives

1.2.2.1 Key Secondary Objective

The key secondary objective for patients in Cohorts 1 and 2 is to assess mean change from baseline in AdvSM-SAF TSS.

1.2.2.2 Additional Secondary Objectives

Additional secondary objectives for patients in Cohorts 1 and 2 are to:

- Determine local investigator determined ORR based on modified IWG-MRT-ECNM consensus response criteria, using C-findings when present at baseline.
- Determine time-to-event outcomes including time-to-response (TTR), DOR, PFS, and OS.
- Determine morphologic response (CR/CRh+ PR) based on Pure Pathological Response criteria and as defined in protocol Appendix 8.
- Determine ORR and other clinical outcome measures (DOR, PFS, and OS) analyzed by prior therapy and by genotype.
- Assess changes in the following individual measures of MC burden:
  - BM MCs.
  - Serum tryptase.
  - *KIT* mutation burden (eg, D816V) in PB and BM.
  - Liver and spleen volume by imaging.
- Assess additional PROs using the AdvSM-SAF, including change from baseline in:
  - Symptom domain scores.
  - Individual symptom scores.

- Assess changes in QoL measures, including the [redacted]) and [redacted]
- Assess the safety of avapritinib by AEs, vital signs, electrocardiograms (ECGs), and laboratory tests.
- Assess the PK of avapritinib.
- Correlate avapritinib exposure with safety and efficacy endpoints.

1.2.3 [redacted]

[redacted]

1.3 Study Endpoints

1.3.1 Primary Endpoints

The primary endpoint is adjudicated ORR (CR/CRh + PR + CI) based on modified IWG-MRT-ECNM criteria, confirmed 12 weeks after initial response in patients in Cohort 1 only.

1.3.2 Secondary Endpoints

1.3.2.1 Key Secondary Endpoints

The key secondary endpoint is the mean change from baseline in AdvSM-SAF TSS in patients in Cohorts 1 and 2.

1.3.2.2 Other Secondary Endpoints

Additional secondary endpoints for patients in Cohorts 1 and 2 are:

- Local Investigator assessed ORR (CR/CRh + PR + CI) based on modified IWG-MRT-ECNM criteria, using C-findings when present, confirmed 12 weeks after initial response.
- Objective response rate (CR/CRh + PR) based on Pure Pathological Response criteria.
- Time-to-event outcomes including: TTR, DOR, PFS, and OS.
- CR/CRh + PR and clinical benefit (CR/CRh + PR + CI + SD) based on modified IWG-MRT-ECNM criteria.
- Overall response rate and other clinical outcome measures (DOR, PFS, OS) analyzed by prior therapy and by genotype.
- Changes in BM MCs, serum tryptase, KIT mutation burden (eg, D816V) in PB and BM, and liver and spleen volume by imaging.
- Mean change from baseline in AdvSM-SAF domain and individual symptom scores.
- Changes in [redacted]
- Safety of avapritinib, as assessed by AEs, changes in vital signs, ECGs, and laboratory testing.
- PK of avapritinib.
- Correlations between avapritinib exposure and safety and efficacy endpoints.

1.3.3 [redacted]

[redacted]

[redacted]

1.4 Sample Size Justification

The primary objective of this study is to determine adjudicated ORR (CR/CRh + PR + CI) based on modified IWG-MRT-ECNM criteria in patients with AdvSM treated with avapritinib and enrolled in Cohort 1. Assuming the null hypothesis ORR of 28% versus the alternative hypothesis ORR of 50%, and a 1-sided type I error rate of 0.025, a sample size of 63 patients in the modified IWG-MRT-ECNM evaluable cohort (Cohort 1) will have 93.5% power using the exact 1-sample binomial test. Enrollment of the SM-AHN subgroup will be capped at approximately 70% of the 63 Cohort 1 patients (ie, maximum of 45 patients) to ensure the study population reflects the general AdvSM patient population.

This sample size will also allow statistical testing of the key secondary objective of patients with symptom reduction based on AdvSM-SAF TSS. Assuming the null hypothesis is the mean change of TSS ≥ 0, versus the alternative hypothesis of the mean change of TSS is ≤ -10 with avapritinib treatment, the standard deviation of mean change of TSS is assumed to be 20, and a 1-sided type I error rate of 0.025, all treated patients in the modified IWG-MRT-ECNM evaluable cohort and the non-modified IWG-MRT-ECNM evaluable cohort will be included in the analysis. A sample size of 63 patients will have > 90% power using a 1 sample t-test. Testing for this key secondary endpoint will be sequential to ensure control of the study-wise type I error rate (ie, it will only be performed when the null hypothesis for the primary objective is rejected).

The non-modified IWG MRT-ECNM evaluable cohort will enroll approximately additional 40 patients, for a total of 103. With 103 patients, there is about 88% probability of observing at least 1 AE at 2% frequency, instead of 3.5% frequency with 60 patients.

2. POPULATIONS FOR ANALYSIS

2.1 Safety Population

The safety population will include all patients who receive at least 1 dose of study drug. The safety population will be used for the key secondary analysis and all safety analyses, PRO analyses, and efficacy analyses that are not based on modified IWG-MRT-ECNM response criteria.

2.2 Response-Evaluable Population

The response-evaluable (RE) population includes all patients as follows:

1) Received at least 1 dose of avapritinib,
2) And deemed evaluable per modified IWG-MRT-ECNM criteria at Baseline as assessed by SSC review,
3) And had one of the below conditions:
    a. at least 2 post baseline bone marrow assessments, and had been on study for at least 6 cycles (28x6=168 days),
    b. an end of study visit.

In addition to the primary efficacy analysis, the response-evaluable population will be used for all secondary efficacy analyses related to response, such as objective response, TTR, DOR, PFS, CI rate, and clinical benefit rate.

2.3 Pure Pathologic Response-Evaluable Population

The Pure Pathologic Response-Evaluable (PPRE) Population includes all patients as follows:

1) Received at least 1 dose of avapritinib
2) And had one of the below conditions:
    a. at least 2 post baseline bone marrow assessments, and had been on study for at least 6 cycles (28x6=168 days),
    b. an end of study visit.

The PPRE population will be used as analysis population for Pure Pathologic Response, such as objective response rate, TTR, DOR, PFS.

3. GENERAL STATISTICAL CONSIDERATIONS

3.1 Randomization, Stratification and Blinding

This is an open label, single arm study with no randomization or stratification.

3.2 Data Analysis General Information and Definition

Summary statistics for continuous variables will include n (non-missing observations), mean, standard deviation, 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 Kaplan-Meier (KM) method, which will include the number of events and censors, the estimated median with 95% confidence intervals, and 25th and 75th percentiles.

3.2.1 Study Drug

Study drug or study treatment used in the rest of the document refers to avapritinib (BLU-285).

3.2.2 Day 1 and Other Days

Date of first dose of avapritinib (or D1) is defined as the date of the first administration of study drug once enrolled in the study. First dose refers to first dose in study unless specified otherwise.

Date of last dose of avapritinib is defined as the date of last administration of study drug (last dose) in the study. For patients who have not ended treatment, the date of last dose is the last end of administration date if it is not missing. If the last end of administration date is missing, then the cutoff date will be used.

3.2.3 Calculation 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 – D1 + 1 if TARGET DATE is on or after D1;

  STUDY DAY = TARGET DATE – D1 otherwise.

  Note that negative study days are reflective of observations obtained during the screening period. Partial dates for the first study drug administration will not be imputed.

  Age at informed consent (in years) = (year of consent) - (year of date of birth (DOB)) if passed birthday;

  Age at informed consent (in years) = (year of consent) - (year of DOB) - 1 if 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 that are to be expressed in units greater than day will be rounded to 1 decimal place.

3.2.5 Baseline Values

In general, baseline is defined as the last observation before first dose date of study drug, including pre-dose assessments on the first dose date. Any deviation from this definition will be specified in the respected sections.

3.2.6 Last Contact

Last contact is the date of last measurement or assessment upon the patient.

3.3 Methods for Handling Missing Data

Refer to Section 7.1 for detailed date imputation guidelines.

3.4 Windowing of Visits

Visits during treatment (from C1D1 to End of Treatment (EOT), including unscheduled visits) will be windowed. Data are summarized by visit following the visit windows reported 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.

Table 1: Visit Windows

Visit	Start day (time)	Target Day (time)	End day (time)
Week 1 (C1D8)	1	8 ± 1	11
Week 2 (C1D15)	12	15 ± 2	18
Week 3 (C1D22)	19	22 ± 2	25
Week 4 (C2D1)	26	29 ± 3	36
Week 6 (C2D15)	37	43 ± 3	48
Week 8 (C3D1)	49	57 ± 7	71
Week 12 (C4D1)	72	85 ± 7	99
Week 16 (C5D1)	100	113 ± 7	127
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 consider D1 if D1 observation is marked as baseline; X can be 4, 5, 6… until cover all visit windows.

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:

- Disease progression (including "Confirmed SM Progressive Disease, per IWG criteria", "Confirmed AHN Progressive Disease requiring initiation of other Cytoreductive therapy", and "SM clinical progression")
- AE
- Death
- Lost to follow-up
- Protocol deviation
- Withdrawal of consent
- Pregnancy
- Investigator decision
- Non-compliance
- Sponsor decision
- Other

3.6 Protocol Deviations

Deviations from the protocol, as defined in the protocol and protocol deviation plan, will be documented in the study EDC system in Medidata RAVE by sites in an ongoing basis and reviewed by the study monitors and project manager throughout the study period.

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

Major protocol deviations will be summarized descriptively in the Safety Population.

A by-patient listing of patients with protocol deviations in the Safety Population will be provided.

4. STATISTICAL METHODOLOGY

All analyses in this section will be conducted and presented by 'starting daily dose 200 mg' and 'All Doses' for the Safety Population, and/or by disease subtype (grouped as 'ASM', 'SM-AHN', 'MCL', and 'All AdvSM') for the RE population and the PPRE population.

4.1 Analysis Populations

A summary of analysis populations by starting dose without grouping and 'All Doses' will be presented for the following:

- Safety Population
- RE Population
- PPRE Population

Disease diagnosis will be presented for the Safety, RE, and PPRE Populations using data collected from two sources: investigator reported diagnosis and SSC determined disease diagnosis separately for the following categories:

- ASM
- SM-AHN
    - Associated Hematologic neoplasm (AHN) subtype: chronic eosinophilic leukemia (CEL), chronic myelomonocytic leukemia (CMML), myeloproliferative neoplasms (MPN), myelodysplastic syndrome (MDS), MDS/MPN-U, Indolent lymphoid neoplasm, and Other
- MCL
    - AHN subtype: CEL, CMML, MPN, MDS, MDS/MPN-U, Indolent lymphoid neoplasm, and Other
- Other

Analysis populations, investigator reported diagnosis, and SSC determined diagnosis will be presented in a listing.

4.2 Patient Disposition

The number of patients discontinued from treatment, continuing on treatment, and discontinued from study will be summarized for the Safety, RE and PPRE Populations.

Reasons for treatment and study discontinuation listed in section 3.5 will be summarized. Disease progression to acute myeloid leukemia (AML) and discontinuation due to a treatment-related AE will also be summarized.

Patient disposition information will also be provided in a listing including reasons for treatment and study discontinuation.

The number of patients treated with avapritinib in the Safety Population will be summarized by region (North America including United States and Canada vs. Europe), by country, and by investigational site that participated in this study.

4.3 Demographics, Baseline Disease Characteristics and Medical History

Demographic and baseline disease characteristic data will be summarized in the Safety, RE and PPRE Populations. Medical history data will be summarized in the Safety 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
- Ethnicity
- Race

Patients' age at informed consent (years), height (cm), weight (kg), and body mass index (BMI) (kg/m^2) 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:

- Major and Minor WHO diagnosis criteria for SM by investigator and by SSC assessment
    - Major: multifocal dense infiltrations of Mast Cells (MC)
    - Minor: >25% MC atypical/ spindle shaped in biopsy or immature/atypical in aspirate
    - Minor: expression of cluster designation (CD) markers CD2 and/or CD25
    - Minor: *KIT* mutation at codon 816
    - Minor: Baseline serum tryptase >20 ng/mL
- WHO diagnosis criteria for AdvSM subclassification by investigator and by SSC assessment
    - MCL only: ≥20% immature MC in bone marrow (BM) aspirate smears
    - C-findings: Cytopenia(s), including absolute neutrophil count (ANC) <1.0 x 10^9/L or hemoglobin (HGB) <10 g/dL or platelet count (PLT) <100 x 10^9/L

  - C-findings: Hepatomegaly with impaired liver function, ie, elevated transaminases and/or bilirubin levels and/or hypoalbuminemia (with or without ascites or portal hypertension)
  - C-findings: Palpable splenomegaly with hypersplenism (eg, as documented by thrombocytopenia; ie, PLT <100 x 10^9/L)
  - C-findings: Malabsorption with hypoalbuminemia and/or significant weight loss defined as >10 % weight loss over the last 6 months
  - C-findings: Skeletal involvement with large osteolytic lesions with pathologic fractures
  - C-findings: Life-threatening organ damage in other organ systems caused by local mast cell infiltration in tissues
- Extracutaneous organ involvement (Yes, No)
- Mutation status
  - KIT mutation (as reported on case report form (CRF)) (Yes, No)
    - Source sample type (BM, Blood, Other)
    - Exon 17 mutation (D816V, D816Y, Wild Type (WT))
  - Non-KIT mutation (as reported on CRF) (Yes, No)
    - Source sample type (BM, Blood, Other)
    - ASXL1, CBL, DNMT3A, EZH2, JAK2, KRAS, NRAS, RUNX1, SF3B1, SRSF2, TET2, U2AF1, NOTCH1, SMC1A, and Other
  - KIT Exon 17 mutation (central assay)
    - Exon 17 mutation (D816V, D816Y, Wild Type (WT), Other)
  - SRSF2/ASXL1/RUNX1 mutation (S/A/R) (Positive, Negative) by central assay
  - Number of co-mutations (excluding KIT) by central assay (0, 1, 2-3, 4-5, >=6)
- The Eastern Cooperative Oncology Group (ECOG) performance status (0, 1, 2, 3)
- Corticosteroid therapy use for SM
- Spleen palpation
  - not palpable, palpable <5 cm, palpable =5 cm, palpable >5 cm)
- Liver palpation (not palpable, palpable)
- Albumin (ALB) Common Terminology Criteria for AEs (CTCAE) grade (0, 1, 2, 3)
- ANC CTCAE grade (0, 1, 2, 3, 4)

- HGB CTCAE grade (0, 1, 2, 3)
- PLT CTCAE grade (0, 1, 2, 3, 4)
- Direct bilirubin CTCAE grade (0, 1, 2, 3, 4)
- Alanine aminotransferase (ALT) CTCAE grade (0, 1, 2, 3, 4)
- Aspartate aminotransferase (AST) CTCAE grade (0, 1, 2, 3, 4)
- Alkaline phosphatase (ALP) CTCAE grade (0, 1, 2, 3, 4)
- SM related organ damage:
  - Ascites and Pleural Effusions
    - Ascites (No, Grade 1, Grade 2, Grade 3)
    - Pleural effusion (No, Grade 1, Grade 2, Grade 3)
    - >=2 Paracentesis or thoracentesis at least 28 days apart over 12 weeks before C1D-8 with 1 procedure performed during the 6 weeks before C1D-8 (Yes, No)
  - Liver function abnormalities
    - Other cause for abnormal liver function not related to SM (Yes, No)
    - Clinical-relevant portal hypertension (Yes, No)
    - Biopsy-proven liver MC infiltration (Yes, No)
  - Red blood cell (RBC) transfusion
    - RBC transfusion of ≥ 6 units within 12 weeks before C1D-8
      - No
      - Yes – related to cytopenia attributable to SM
      - Yes – related to bleeding
      - Yes – related to hemolysis
      - Yes – related to treatment-induced myelosuppression
      - Yes – with most recent transfusion within 4 weeks before C1D-8
  - PLTs transfusion
    - PLT transfusion of ≥ 6 units within 12 weeks before C1D-8 (Yes, No)
    - PLT transfusion of ≥ 2 units within 4 weeks before C1D-8 (Yes, No)

- Cutaneous Disease (Yes, No)
- Serum tryptase (<11.4 ng/mL, >=11.4 and <20 ng/mL, >=20 and <40 ng/mL, >=40 and <200 ng/mL, >=200 ng/mL)
- BM MC burden assessed by Central Pathology Review (aggregates, no aggregates)
- *KIT* D816V mutation allele burden as measured by mutation allele fraction (MAF) using Droplet Digital polymerase chain reaction (ddPCR) from BM (central assay) (<0.02%, >=0.02% to <1%, >=1%)
- *KIT* D816V mutation allele burden as measured by mutation allele fraction (MAF) using Droplet Digital polymerase chain reaction (ddPCR) from blood (central assay) (<0.02%, >=0.02% to <1%, >=1%)

Continuous baseline disease characteristics will be summarized using descriptive statistics for the following parameters:

- Serum tryptase (ng/mL) (central assay)
- BM MC burden assessed by Central Pathology Review as measured by % MC in BMB that are aggregates or interstitial.
- *KIT* D816V mutation allele burden as measured by MAF using ddPCR from BM (central assay)
- *KIT* D816V mutation allele burden as measured by MAF using ddPCR from blood (central assay)
- Spleen volume (mL) (central radiology review)
- Liver volume (central radiology review)
- ALB (g/L)
- ANC (x 10^9/L)
- HGB (g/L)
- PLT (x 10^9/L)
- Direct bilirubin (umol/L)
- ALT (U/L)
- AST (U/L)
- ALP (U/L)
- Absolute monocyte count (x 10^9/L)
- Number of co-mutations (excluding KIT) by central assay

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. All medical history data will be listed.

4.4 Prior and Concomitant Medications and Prior Therapies

4.4.1 Prior and Concomitant Medications

Prior medications are those that started and stopped before exposure to study drug. Concomitant medications are all medications taken at any time from first dose date of study drug to last dose date of study drug + 30 days. If the last dose date of study drug is not available, 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 (Anatomical Therapeutic Chemical (ATC) Text Level 3) and preferred drug name for the Safety Population. 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 for the Safety Population.

Corticosteroid use are common for SM treatment. Prior and concomitant systemic use of prednisone (and equivalent, 5 mg prednisone = 5 mg prednisolone = 20 mg hydrocortisone = 25 mg cortisone = 4 mg triamcinolone = 4 mg methylprednisolone = 0.8 mg betamethasone = 0.8 mg dexamethasone) will be medically reviewed (ATC Text Level 3 ‘CORTICOSTEROIDS FOR SYSTEMIC USE, PLAIN’). The number of patients with corticosteroid treatment at baseline, and the number of patients with decrease, increase, and completely discontinued use of corticosteroid during avapritinib treatment will be summarized in the Safety Population. Patient will be assigned to only one of the mutually exclusive group of ‘increase’, ‘decrease’, and ‘completely discontinued’ use of corticosteroid. If a patient had any increase use of corticosteroid during the treatment, the patient will not be considered as ‘decrease’ or ‘completely discontinued’ even if those happened.

Medications are coded using the WHO Drug Global B3 dictionary, version March 2019.

4.4.2 Prior Therapies for the Underlying Malignancy

Summary of prior therapy will be presented for the Safety, RE and PPRE Populations. Prior therapy is defined as all treatment that started and ended on or before first dose date of avapritinib.

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

- Prior antineoplastic therapy (Yes, No)

- By preferred drug name. A patient taking the same medication multiple times is counted only once under that agent name.

- Prior tyrosine kinase inhibitor (TKI) therapy (Yes, No)
- Prior radiation therapy (Yes, No)
- Prior cancer related surgery-procedures (Yes, No)
- Best response to any prior antineoplastic therapy CR, PR, CI, Stable Disease (SD), Loss of Response (LoR), Progressive Disease (PD), Not Applicable (NA), Not evaluable (NE))
- Prior midostaurin (Yes, No)
- Best response to prior midostaurin (CR, PR, CI, SD, LoR, PD, NA, NE)
- Reason for discontinuation of prior midostaurin (PD/Relapse, Toxicity, Other)

Continuous variables to be summarized include:

- Duration of treatment on midostaurin (months)

No overlap between the exposure to prior therapies and avapritinib will be allowed and any overlap of exposure will be queried at data review stage. See section 7.1.3 for imputation of partial dates.

By-patient listings of prior cancer therapy and prior cancer related surgical procedures will be presented for the Safety Population.

4.5 Study Treatments and Extent of Exposure

A summary of study drug exposure, including number of doses administered, cumulative dose, dose intensity, relative dose intensity (RDI), duration of treatment (months), and the proportion of patients with dose modifications will be produced for the Safety Population.

4.5.1 Treatment Duration and Exposure

Duration of treatment (months) = (last dose date – first dose date + 1) / 30.4375.

First and last dose dates are as defined in section 3.2.2.

Descriptive statistics will be provided for treatment duration in months.

4.5.2 Cumulative Dose

Cumulative dose (mg) is defined as sum of all dose 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. Each patient has an average daily dose.

4.5.4 Dose Intensity

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

4.5.5 Relative Dose Intensity

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

RDI is will be summarized descriptively as a continuous variable and be categorized into <75%, 75% to <90%, 90% to 120%, 120 to 150%, and ≥150%.

Treatment exposure described in sections 4.5.1 - 4.5.5 will be summarized descriptively by time (every 4 weeks and entire duration) in the Safety Population.

Duration in every 4 weeks is defined as the time from first day of the 4 weeks to the day prior to the first day of the subsequent 4 weeks; for the last 4 weeks, the end date is the date of last dose.

A by-patient listing of study drug including treatment duration, cumulative dose, average daily dose, dose intensity, planned initial dose, RDI will be provided.

4.5.6 Dose Modification

A summary of the frequency of patient with dose modification and reasons for dose modifications will be provided for the Safety Population.

Dose reduction due to AE will be summarized using the number and percent of patients with 0, 1, 2, or >2 dose reduction due to AE during the entire treatment period.

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

Dose interruption/missing due to AE, and due to AESI will be summarized using the number and percentage of patients who have 0, 1, 2, or >2 dose interruptions/missing during the entire treatment period.

Details of study drug administration data and dose modifications will be provided in by-patient listings.

4.5.7 Prescribed Dose Level

Prescribed dose level at D1 is the assigned starting dose level for each patient. However, it changes over time due to dose reduction, re-escalation, and interruption. Prescribed dose level will be summarized at Week 1 D1, and every 8 weeks interval (eg, Week 9 D1, Week 17 D1, Week 25 D1, etc.) for the Safety Population. Percent of patients at a specific prescribed dose level at each time point will be calculated among patients who are treated on or beyond that time point.

4.6 Efficacy Analyses

Response on treatment in cohort 1 will be assessed using IWG-MRT-ECNM (IWG) criteria (Gotlib, 2013) with modification (mIWG) which are specified in protocol appendix 6, and 7 and were presented and discussed with FDA and CHMP.

Response assessments performed by investigators are recorded at the C1D15, beginning of C2, 3, 7, 11, and every 6 cycles thereafter. Unscheduled assessments may also be performed. Investigators assess responses based on information collected at the investigational site (local assessment), including, BM MCs, spleen and liver, hematology and blood chemistry, and other assessments associated with organ damage due to SM. Serum tryptase results will come from Central Laboratory testing.

For all patients, BM samples collected at the investigational sites are reviewed centrally by the independent pathologist at baseline for diagnosis, and post baseline for response assessment. Images of spleen and liver by CT or MRI are reviewed centrally by independent radiologists. Serum tryptase levels are measured at a central laboratory. In order to provide consistent and unbiased assessment of response using the mIWG criteria, a SSC is formed with a subset of investigators and the independent pathologist, chaired by Dr. Jason Gotlib, the primary author of the IWG (Gotlib, 2013). Disease diagnosis and response adjudicated by SSC using centrally collected data (BM, tryptase, spleen volume), spleen and liver palpations, hematology and blood chemistry, and other assessments of organ damage due to SM from local evaluations will be documented in the electronic data collection system. The mIWG used by SSC adjudication are listed in appendix 7.2 and 7.3. Based on feedback from FDA, medically documented weight loss will be dropped from the modification list and excluded as an evaluable c-finding.

In addition, a computer algorithm will be implemented using SAS to derive response per the IWG without modifications. The difference between IWG and mIWG as implemented by SSC can be found in appendix 7.4. This computer algorithm will use the same source data used by SSC including adjudicated diagnosis by SSC. Response produced by this computer algorithm will be denoted as Algorithm-IWG response.

Secondly, attribution of C-findings to SM is inherently subjective and challenging because multiple reasons may exist for each C-finding, such as the AHN component for cytopenia and splenomegaly. For this reason, a Pure Pathologic Response (PPR) criteria proposed by SSC (appendix section 7.5) based on MC burden, tryptase, and blood counts is more objective and may allow clear demonstration of reduction in MC burden, which lead to the symptom of SM. Because PPR does not include the C-findings, it can be applied to all patients with measurable MC burden in both cohorts. PPR will be derived by a computer algorithm using the same assessments that are used for mIWG derivation (BMB, tryptase, ANC, HGB, and PLT) and additionally KIT D816V MAF by central assay. PPRE population will be used for analysis of Pure Pathologic responses.

4.6.1 Analysis of Primary Efficacy Endpoints

4.6.1.1 Overall Response Rate

ORR is defined as the proportion of patients with a confirmed best response of CR, CRh, PR or CI by mIWG criteria.

Primary ORR analysis will be based on SSC adjudicated responses in the RE Population. ORR will be estimated using frequency, percentage, and two-sided 95% confidence intervals based on the exact binomial distribution (Clopper-Pearson). Statistical test on binomial proportion against a null of 28% will be performed using one sided $\alpha=0.025$. Wald test p-value will be presented. There is one interim analysis when 32 patients (with the SM-AHN subgroup capped at approximately 70%) enrolled in the cohort 1 are evaluable for response. If the 1-sided p-value is < 0.00625, the null hypothesis will be rejected, and avapritinib will be deemed effective in treating patients with AdvSM and the interim analysis will be used to support a marketing application. If the 1-sided p-value is ≥ 0.00625 at the interim analysis, the final analysis will be used to support a marketing application. The final analysis of the primary efficacy endpoint will occur after all 63 patients (with the SM-AHN subgroup capped at approximately 70%) are enrolled in the cohort 1 and have had the opportunity to receive avapritinib treatment for at least 10 cycles or have discontinued treatment earlier. The final primary efficacy analysis will be tested at 1-sided alpha level of 0.02178 (See Section 5).

CR + CRh + PR rate will be summarized similarly as ORR. Statistical test on binomial proportion against a null of 17% will be performed using one sided $\alpha=0.025$. Wald test p-value will be presented.

CR + CRh rate and clinical benefit rate (CR/CRh + PR + CI + SD) will be summarized similarly as ORR. No statistical testing will be performed.

Additionally, the best overall response following the hierarchical order of CR, CRh, PR, CI, SD, PD, and NE will be tabulated for RE Population.

Sensitivity analysis will be conducted using Algorithm-IWG responses in the RE Population.

Pathologic ORR is defined as the proportion of patients with a confirmed best response of morphologic CR (mCR), morphologic CRh (mCRh), or morphologic PR (mPR) by PPR criteria.

Pathologic ORR will be analyzed in the PPRE Population and RE Population. Pathologic ORR will be estimated using frequency, percentage, and two-sided 95% confidence intervals based on the exact binomial distribution (Clopper-Pearson).

mCR + mCRh will be summarized similarly as Pathologic ORR.

Additionally, the best overall response following the hierarchical order of mCR, mCRh, mPR, Treatment Failure (TF), PD, and NE will be tabulated for PPRE Population. Among patients with mCR/mCRh, the proportion who achieve molecular CR (moCR) or molecular CRh (moCRh) will also be summarized.

Furthermore, to assess the treatment effect in all treated patients, the investigator assessed responses in RE Population will be summarized.

4.6.2 Analysis of Secondary Efficacy Endpoints

4.6.2.1 Analysis of Key Secondary Endpoint

4.6.2.1.1 AdvSM Symptom Assessment Form Total Symptom Score (TSS)

Patients will complete AdvSM symptom assessment form (AdvSM-SAF) daily from 7 days prior to first dose of study drug through C17D28.

The AdvSM-SAF is a 10-item questionnaire to assess symptoms (abdominal pain, diarrhea, nausea, vomiting, spots, itching, flushing, fatigue, each scored 0-10), and functional domains (gastrointestinal (GI) domain including abdominal pain, diarrhea, nausea, vomiting, scored 0-40; Skin domain including spots, itching, flushing, scored 0-30) specific to AdvSM. Total Symptom Score (TSS) is the sum of all eight symptoms (scored 0-80).

A symptom/domain/TSS score for any given day is determined by the 7-day average prior to that day. For example, Baseline score is determined by the 7-day average from C1D-7 to C1D-1; score on C3D1 is determined by the 7-day average from C2D22 to C2D28, so on so forth. Domain scores and TSS on any given day will only be calculated if no symptom score within the domain or the TSS is missing. If a symptom/domain/TSS score is available on at least 4 or more days during the 7-day period (no consecutive requirement), the average score is calculated by the sum of the available scores divided by the number of days with available scores.

Change in AdvSM-SAF TSS is calculated from Baseline to C11D1 for patients on treatment for more than 10 cycles. The last observation carried forward method will be used for patients who end treatment before completing 10 cycles. Mean change will be calculated for all patients in the safety population and tested against the null hypothesis of ≥ 0 in mean change of TSS. If the one-sided one sample t-test p-value is < 0.025, the null hypothesis will be rejected, and avapritinib will be deemed effective in reducing symptoms in patients with AdvSM.

A sensitivity analysis for change in AdvSM-SAF TSS is calculated from Baseline to C7D1 for patients on treatment for more than 6 cycles. The last observation carried forward method will be used for patients who end treatment before completing 6 cycles. One-sided one sample t-test p-value for test against the null hypothesis of ≥ 0 in mean change of TSS will be provided.

AdvSM-SAF individual and domain score will be summarized similarly as TSS. No statistical testing will be performed.

Two set of box plots of change from Baseline to D1 of each cycle will be presented, one set including all patients in the Safety Population with Baseline value for the symptom/domain/TSS scores, the other set will be limited to the ones with Baseline value above median of the specific symptom/domain/TSS scores (higher symptom).

In order to assess the compliance of completing AdvSM-SAF by patients. The proportion of patients who have 7-days average symptom/domain/TSS scores, will be calculated using the number of patients who received avapritinib treatment on or beyond a given time point as denominator. Time points to be summarized include D1 of C1 to C18.

4.6.2.2 Analysis of Other Secondary Endpoints

4.6.2.2.1 Time to Response

Time to response (TTR) is defined as the time from the start of treatment to the time a response (CR/CRh/PR/CI) by mIWG is first met. Patients without confirmed response will be excluded from this analysis.

Time to CR/CRh/PR is defined as the time from the start of treatment to the time a CR/CRh/PR by mIWG is first met. Patients without confirmed CR/CRh/PR will be excluded from this analysis.

Time to CR is defined as the time from the start of treatment to the time a CR/CRh by mIWG is first met. Patients without confirmed CR/CRh will be excluded from this analysis.

Time to Morphologic Response (TTmR) is defined as the time from the start of treatment to the time a response (mCR/mCRh/mPR) by PPR is first met. Patients without confirmed response will be excluded from this analysis.

Time to morphologic CR (TTmCR) is defined as the time from the start of treatment to the time a mCR/mCRh by PPR is first met. Patients without confirmed mCR/mCRh will be excluded from this analysis.

Summary statistics of TTR/TTmR (months) will be presented for the following:

- Patients achieved CR/CRh/PR/CI in RE Population using SSC adjudicated response
- Patients achieved CR/CRh/PR/CI in RE Population using Algorithm-IWG response
- Patients achieved mCR/mCRh/mPR in PPRE Population using PPR

- Patients achieved CR/CRh/PR/CI in RE Population using investigator assessed response

Summary statistics of Time to CR/CRh/PR (months) will be presented for the Patients achieved CR/CRh/PR in RE Population using SSC adjudicated response.

Summary statistics of Time to CR or TTmCR (months) will be presented for the following:

- Patients achieved CR/CRh in RE Population using SSC adjudicated response
- Patients achieved CR in RE Population using Algorithm-IWG response
- Patients achieved mCR/mCRh in PPRE Population using PPR
- Patients achieved CR/CRh in RE Population using investigator assessed response

4.6.2.2.2 Duration of Response

DOR is defined as the time from first documented response (CR/CRh/PR/CI) to the date of first documented PD/LoR or death due to any cause, whichever occurs first. Responses are determined by mIWG criteria.

Morphologic DOR (mDOR) is defined as the time from first documented PPR (mCR/mCRh/mPR) to the date of first documented PD/LoR or death due to any cause, whichever occurs first. Responses are determined by PPR criteria.

DOR/mDOR will be analyzed using KM methods and will include number of events and censors, the estimated median with two-sided 95% confidence intervals, and 25th and 75th percentiles. DOR/mDOR at specific time-points (e.g. 3-, 6-, 9-, 12-, 18-, 24- month, etc.) will be computed, along with the standard errors using Greenwood's formula (Klein 2003).

Primary DOR analysis will be based on SSC adjudicated responses in the RE Population. Patients without confirmed response (CR/CRh/PR/CI) 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.

A sensitivity analysis for duration of CR/CRh/PR based on SSC adjudicated responses in the RE Population will be summarized similarly as DOR.

The following analysis will be conducted on the DOR/mDOR to assess the robustness of the primary analysis by including broader subset and all treated patients:

- DOR of Algorithm-IWG responders in the RE Population
- mDOR of PPR responders in the PPRE Population
- DOR in responders by investigator assessment in the RE Population

The plots of DOR/mDOR curves using the KM method will be presented.

4.6.2.2.3 Progression Free Survival

PFS is defined as the time from the start of treatment to the date of first documented PD or death due to any cause, whichever occurs first. If a patient has not had an event, PFS is censored at the date of last valid assessment that is LoR or better. As specified in the IWG, development of AML will be treated as disease progression.

The KM method will be used to estimate the PFS distribution function. The median PFS along with its two-sided 95% confidence intervals and 25th and 75th percentiles will be estimated. In addition, the event rates (or event-free) at specific time-points (e.g. 3-, 6-, 9-, 12-, 18-, 24-month, etc.) will be computed, along with the standard errors using Greenwood's formula (Klein 2003).

The following PFS analysis will be presented:

- PFS in RE Population using SSC adjudicated response and progression
- PFS in RE Population using Algorithm-IWG response and progression
- PFS in PPRE Population using PPR response and progression
- PFS in RE Population using investigator assessed response and progression

The plots of PFS curves using the KM method will be presented.

4.6.2.2.4 Overall Survival

OS is defined as the time from the start of treatment to the date of death. Patients who die before or on the data cutoff date will 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.

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. It can be the last visit date or last contact date that the patient is known to be alive.

The survival distribution of OS will be estimated using the KM method. The median OS along with its two-sided 95% confidence intervals and 25th and 75th percentiles will be estimated. In addition, the survival rate at specific time-points (e.g. 3-, 6-, 9-, 12-, 18-, 24- month, etc.) will be computed, along with the standard errors using Greenwood's formula (Klein 2003).

The following OS analysis will be presented:

- OS in Safety Population, and by baseline S/A/R genotype, and by prior treatment with midostaurin
- OS in RE Population, and by baseline S/A/R genotype, and by prior treatment with midostaurin

The plots of survival curves using the KM method will be presented.

Duration of follow up (months) for the Safety Population will be summarized by KM method using OS data with reversed event/censor flag. Descriptive statistics will also be provided among patients who are still alive.

4.6.2.2.5 Bone Marrow Mast Cells

Bone marrow biopsy and aspirate and a PB smear will be obtained at screening (baseline), C3D1, C7D1, and C11D1 and then every 6 cycles, as well as at EOT and at the time of PD for SM response assessment according to mIWG and for AHN response assessment. Patients with a confirmed CR/CRh based on Central Pathology Review (confirmed 12 weeks after the first CR/CRh; ie, 2 consecutive BM biopsies demonstrating absence of neoplastic MCs) are not required to have subsequent BM biopsy unless loss of response is suspected.

Percentage of total mast cells present (out of total nucleated cells) from bone marrow core biopsy will be summarized. Descriptive statistics and percent change from baseline will be presented by visits in the Safety Population.

In Safety Population, the number of patients achieved total clearance of neoplastic MC aggregates among patients with baseline neoplastic MCs (IWG CR) vs. the number of patients achieved ≥50 % reduction (IWG PR) from baseline vs. the number of patients with <50 % reduction from baseline will be plotted in a stacked bar chart by visits.

Waterfall plot that represent the maximum reduction in percent of neoplastic MCs from baseline for each patient will be presented for the Safety Population. Distinct bar color will be used for patients with ≥50 % reduction from baseline that lasted for at least 2 consecutive bone marrow biopsies in the graph.

Percent of patients achieved total clearance of neoplastic MC aggregates who has neoplastic MC at baseline among all patients (IWG CR) and percent of patients achieved ≥50 % reduction from baseline (IWG PR) will be tabulated.

4.6.2.2.6 Serum Tryptase

Serum Tryptase will be measured at screening (baseline), C1D15, C2D1, C3D1, C5D1, C7D1, C11D1 and every 6 cycles thereafter, and EOT at a central laboratory. Also performed 12 weeks (± 4 weeks) after documentation of CR or PR to confirm response and 4 weeks after SM and/or AHN PD or clinical progression to confirm PD. Descriptive statistics and percent change from baseline will be presented by visits in the Safety Population.

In Safety Population, the number of patients achieved serum tryptase < 20 ng/mL who has baseline serum tryptase ≥ 40 ng/mL (IWG CR) vs. the number of patients achieved ≥50 % reduction from baseline (IWG PR) vs. the number of patients with <50 % reduction from baseline will be plotted in a stacked bar chart by visits.

Waterfall plot that represent the maximum reduction in percent of serum tryptase from baseline for each patient will be presented for the Safety Population. Distinct bar color will be used for patients with ≥50 % reduction from baseline that lasted for at least 2 cycles (56 days) in the graph.

Percent of patients achieved serum tryptase < 20 ng/mL (IWG CR) who has baseline serum tryptase ≥ 40 ng/mL (IWG CR) among all patients and percent of patients achieved ≥50 % reduction from baseline (IWG PR) will be tabulated.

4.6.2.2.7 *KIT* D816V Mutation Allele Fraction

KIT D816V MAF (%) in blood will be measured at screening, C1D1 (baseline), C1D15, C2D1, C3D1, C7D1, C11D1, and every 6 cycles thereafter, and EOT at a central laboratory by ddPCR. *KIT* D816V MAF in BM samples will be measured at Screening, C3D1, C7D1, C11D1, and every 6 cycles thereafter, and EOT at the same central laboratory by ddPCR. Descriptive statistics and percent change from baseline will be presented by in the Safety Population for blood *KIT* D816V MAF and for BM *KIT* D816V MAF separately.

In Safety Population, among patients with baseline *KIT* D816V MAF, the number of patients achieved ≥50 % reduction from baseline vs. the number of patients with <50 % reduction from baseline will be plotted in a stacked bar chart by visits. For the subset of patients with baseline MAF ≥1 % stacked bar chart will be plotted for the number of patients with MAF <0.02% (assay detectable level), MAF <1 %, MAF ≥1 % by visits.

Waterfall plot that represent the maximum reduction in percent of MAF from baseline for each patient will be presented for the Safety Population with baseline MAF, and the subset of patients with baseline MAF ≥1 %. Distinct bar color will be used for patients with ≥50 % reduction from baseline that lasted for at least 2 cycles (56 days) in the graph. Patients with MAF <0.02% and MAF <1 % will be flagged.

Percent of patients with MAF <0.02% who has baseline MAF ≥1 % among all patients, percent of patient with <1% who has baseline MAF ≥1 % among all patients and percent of patients with ≥50 % reduction from baseline among patients with baseline KIT D816V MAF will be tabulated.

4.6.2.2.8 Spleen and Liver Measurements

Imaging of the spleen and liver will be performed at screening (baseline), C3D1, C7D1, C11D1, and every 6 cycles thereafter, and EOT at a central laboratory. Images by CT or MRI are centrally assessed by independent radiologist. Volume (mL) and the longest vertical diameter (cm) are reported. Descriptive statistics and percent change from baseline of the spleen and liver volumes will be presented by visits in the Safety Population.

In Safety Population, among patients with baseline spleen volume assessment, the number of patients achieved ≥35 % reduction from baseline vs. the number of patients with <35 % reduction from baseline will be plotted in a stacked bar chart by visits.

Waterfall plot that represent the maximum reduction in percent of spleen volume from baseline for each patient will be presented for the Safety Population, and the subset of patients with baseline palpation ≥5 cm.

Waterfall plot that represent the maximum reduction in percent of liver volume from baseline for each patient will be presented for the Safety Population, and the subset of patients with baseline palpable liver.

Percent of patients achieved ≥35 % reduction from baseline among patients with baseline spleen volume assessment will be tabulated.

4.6.2.2.9 [redacted]

[redacted]

4.6.2.2.10 [redacted]

[redacted]

Table 2: [redacted]

[redacted]

[redacted]

4.6.3 [redacted]

4.6.3.1 [redacted]

[redacted]

4.6.3.2 [redacted]

[redacted]

4.6.3.3 [redacted]

[redacted]

4.6.3.4 [redacted]

[redacted]

4.6.3.5 [redacted]

[redacted]

4.6.3.6 [redacted]

[redacted]

4.6.3.7 [redacted]

[redacted]

[redacted]

4.6.4 Event Free Survival

Event Free Survival (EFS) is defined as the time from the start of treatment to the first onset date of an event. In addition to the PFS events (PD, development of AML, or death), clinical progression of AHN or SM that requiring EOT or start of new therapy, or EOT due to toxicity will also be considered as events. If a patient has not had an event, EFS is censored at the date of last valid assessment that is LoR or better.

The KM method will be used to estimate the EFS distribution function. The median EFS along with its two-sided 95% confidence intervals and 25th and 75th percentiles will be estimated. In addition, the event rates (or event-free) at specific time-points (e.g. 3-, 6-, 9-, 12-, 18-, 24-month, etc.) will be computed, along with the standard errors using Greenwood's formula (Klein 2003).

The following EFS analysis will be presented:

- EFS in RE Population using RAC adjudicated response and progression
- EFS in RE Population using Algorithm-IWG response and progression
- EFS in RE Population using investigator assessed response and progression

The plots of EFS curves using the KM method will be presented.

4.6.5 Subgroup Analysis

ORR analysis in the RE Population will be presented for the following subgroups.

- Age (<65 years, ≥65 years)
- Gender (male, female)
- Region (North America, Europe)
- Baseline S/A/R genotype (positive, negative)
- Prior treatment with midostaurin (yes, no)
- Prior antineoplastic therapy (yes, no)

DOR, PFS in the RE population and OS analysis in the safety population will be presented for the following subgroups.

- Baseline S/A/R genotype (positive, negative)
- Prior treatment with midostaurin (yes, no)
- Prior antineoplastic therapy (yes, no)

4.7 Safety Analyses

4.7.1 Adverse Events

AEs will be analyzed in terms of treatment-emergent AEs which are defined as any AE that occurred between the first dose of a study drug through 30 days after the last dose of any study drug.

AEs refers to treatment-emergent AEs unless otherwise specified throughout this document. All AEs will be coded using MedDRA version 18.1. The severity will be graded per the National Cancer Institute (NCI) CTCAE, version 5.0.

Relationship of AE to treatment is defined as AE that is considered to be related to the study drug by the Investigator according to the following criteria:

- Not Related: Exposure to the study treatment did not occur, or the occurrence of the AE is not reasonably related in time, or the AE is clearly caused by something other than the study treatment.
- Related: The AE does not meet the criteria for "Not Related."

A patient experiencing multiple AEs under the same PT (SOC) will be counted only once for that PT (SOC) by maximum severity. A patient experiences the same AE more than once with more than one 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 rule for missing AE dates are in Section 7.1.

Safety Population will be used for all AE summaries. Number of patients with at least one AE will be summarized (Table 3:). The following tables summarizing incidence of various aspects of AEs and serious AEs (SAEs) by PT, SOC and PT, and CTCAE grades:

1. AE Summary (including all subsequent items)
2. AE
3. AE related to study drug
4. SAE
5. SAE related to study drug
6. Grade 3/4/5 AE
7. Grade 3/4/5 AE related to study drug
8. AE leading to interruption of study drug regardless of causality

9. AE leading to reduction of study drug regardless of causality
10. AE leading to permanent discontinuation of study drug regardless of causality

All AEs will be listed by patient.

Subgroup analysis will be conducted for endpoints 2 – 7 above (details in Table 3:) for all patients in the Safety Population. Corresponding forest plots will be provided based on the odds ratio for each of the following subgroups:

- Age (<65 years, ≥65 years)
- Gender (male, female)
- Region (North America, Europe)
- Patient ≥50K nadir platelet prior to dosing (Yes, No)

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

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

- AESI
- Serious AESI
- AESI leading to permanent discontinuation of study drug regardless of causality

Time to first onset of AESI (any cognitive effects, any intracranial bleeding, and for each PT) is defined as from day 1 of avapritinib to the first AESI start date. For patients without AESI time to first onset will be censored at the minimum of (date of last study drug dose + 30 days, cutoff date, death date, end of study date). KM analysis will be presented for time to first onset of AESI (months) in the Safety Population.

Incidence and prevalence of AESI (any cognitive effects, any intracranial bleeding, and for each PT) will be summarized in the Safety Population by intervals of 4 weeks from D1. While incidence accounts for any patients with new onset of AESI during the 4-week interval, prevalence accounts for any patients with ongoing AESI during the 4-week interval. Denominator for the incidence and prevalence calculation will include all patients with any drug exposure at or beyond the 4-week intervals.

Table 3: Summary of AE Tables

Safety Endpoint	Main Table	by PT	by SOC, PT		by PT, CTCAE	by SOC, PT, CTCAE
		Overall	Overall	Subgroup[1]	Overall	Overall
Summary Table (each item below one row)	X					
AE		X	X	X	X	X
SAE		X	X	X	X	X
Grade 3/4/5 AE		X	X	X		
Related AE		X	X	X	X	X
Related SAE		X	X	X	X	
Related Grade 3/4/5 AE		X	X	X		
Death (Grade 5 AE)		X				
AE Leading to Permanent Treatment Discontinuation		X	X			
AEs Leading to Dose Interruption		X	X			
AEs Leading to Dose Reduction		X	X			
Listing (Death, Discontinuation, SAE)	X					
AESI by Category		X			X	
Serious AESI by Category		X			X	
AESI Leading to Permanent Study Discontinuation		X				
Time to Onset of AESI	X					
Incidence of AESI	X					
Prevalence of AESI	X					

Abbreviations: AE = adverse event; AESI = adverse event of special interest; CTCAE = Common Terminology Criteria for AEs; PT = preferred term; SAE = serious adverse event; SOC = system organ class.

1 Among all patients by Subgroup of Age at Informed Consent (<65 years, ≥65 years), Gender (male, female), and Region (North America, Europe).

4.7.3 Deaths

Death (Grade 5 AE) will be summarized during the active treatment phase and within 30 days after the last date of the study treatment by SOC and PT.

On treatment death will be provided in a listing.

4.7.4 Clinical Laboratory Evaluations

The laboratory parameters (hematology, serum chemistry, and coagulation) and change from Baseline will be summarized for the Safety Population by visit. If there are repeated values for a specific visit, the last reported value will be used for that visit.

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

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.

Boxplot of selected lab tests (including leucocytes, neutrophils, PLT, HGB, ALB, ALT, aspartate aminotransferase, ALP, direct bilirubin, total bilirubin) values will be presented by visit.

4.8 Vital Signs

Vital sign parameters (height, weight, BMI, systolic and diastolic blood pressure, heart rate, and body temperature) and change from Baseline will be summarized for the Safety Population by visit.

Listings will be provided for each vital sign parameter.

Boxplot of weight percent change from baseline will be presented by visit.

4.9 Electrocardiograms

ECGs assessment will be presented in a listing for the Safety Population.

4.10 Brain Imaging

Brain imaging assessment will be presented in a listing for the Safety Population.

4.11 ECOG Performance Status

Overall summary and changes from baseline will be presented for ECOG performance status for the Safety Population by visit.

4.12 Platelet aggregation studies at baseline and during treatment, including at the time of bleeding events.

At selected sites, platelet aggregation assay for newly enrolling patients with a platelet count > 150, 000/µL are performed locally on fresh blood samples at baseline, C2D1, C3D1, and C7D1 or at any applicable time point for patients on study with a platelet count >150,000/µL, if no baseline value was assessed.

Platelet aggregation result (normal/abnormal/NE) and the categories for abnormal result (Decreased aggregation to any agonist/Increased aggregation to any agonist/Other) will be tabulated for the subset of Safety population - patient with Platelet aggregation results.

Platelet aggregation result will be presented in a by-patient-listing.

5. INTERIM ANALYSES

An interim analysis (IA) for superiority will be conducted to determine whether avapritinib is effective in treating patients with AdvSM, enrollment will be continuing regardless of the result of IA. Primary endpoint ORR (CR/CRh, PR, or CI) based on modified IWG-MRT-ECNM consensus response criteria per SSC review will be used as the endpoint for the interim analysis. The interim analysis will be performed at approximately 50% of the information fraction; it will include the first 32 patients (with the SM-AHN subgroup capped at approximately 70%) enrolled in the modified IWG-MRT-ECNM- evaluable cohort who are considered evaluable for response.

To establish the statistical monitoring guideline to control the overall Type I error rate at 0.025 (1-sided), critical values for the interim and final analyses of ORR is determined using power error spending function with parameter $\rho=2$. The interim analysis will be considered positive if the observed p value is less than 0.00625 (1-sided) at the interim analysis. The final analysis will be tested at 1-sided alpha level of 0.02178.

If the testing for the primary endpoint ORR is statistically significant at the interim analysis, avapritinib will be deemed effective in treating patients with AdvSM. The key secondary endpoint will be tested in all treated patients with alpha level of 0.025 (1-sided). Other secondary endpoints, safety analyses, PRO analyses, and efficacy analyses that are not based on modified IWG-MRT-ECNM response criteria will be summarized as specified in above sections respectively.

If the testing for the primary endpoint ORR is not statistically significant at the interim analysis, the ORR testing will be conducted again at the final analysis when 63 patients (with the SMAHN subgroup capped at approximately 70%) have been enrolled in the modified IWGMRTECNM evaluable cohort and have had the opportunity to receive avapritinib treatment for at least 10 cycles or have discontinued treatment earlier.

6. CHANGES TO PLANNED ANALYSES FROM THE PROTOCOL

- Data for the following exploratory objective will not be available at the time of IA, CSR or CSR addendum, therefore the analysis method will not be addressed in this plan.
    - [redacted]
- Data for Unconfirmed ORR, and unconfirmed CR/CRh + PR are not collected in the SSC meeting, therefore will not be summarized.
- PFS is defined as the time from the start of treatment to the date of first documented PD or death due to any cause, whichever occurs first. If a patient has not had an event, PFS is censored at the date of last valid assessment that is **SD** or better. However after discussed with SSC member, If a patient has not had an event, PFS is censored at the date of last valid assessment that is **LoR** or better.
- EFS analysis is added, to capture a broader definition of events compared to PFS
- Update the definition of response-evaluable population, "…had at least 2 complete post baseline assessments (inclusive of bone marrow and/or other involved tissue biopsy and liver and spleen imaging)…" to "..had at least 2 post baseline bone marrow assessments…" to be consistent across studies.

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'. If the imputed starting date is earlier than initial diagnosis date, it should be set as the initial diagnosis date.

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'. If the imputed starting date is earlier than initial diagnosis date, it should be set as the initial diagnosis date.

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'. If the imputed ending date is later than day 1 of avapritinib, then the ending date will be set as the maximum of (first day of the known ending month and year, first dose date of study drug – 15)

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 ending date is later than day 1 of avapritinib, then the ending date will be set as the maximum of (01JAN of the known ending year, first dose date of study drug – 15)

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 mIWG Response Criteria

Response	Criteria for Response

Complete remission (CR)[a]	Requires all 4 of the following criteria, and response duration must be ≥ 12 weeks:
	• No presence of compact neoplastic MC aggregates in the BM or other biopsied extracutaneous organ
	• Serum tryptase level < 20 ng/mL[b]
	• Peripheral blood count remission defined as:
	  o ANC ≥ 1 × 10^9/L with normal differential (absence of neoplastic MCs and blasts < 1%) *and*
	  o Platelet count ≥ 100 × 10^9/L *and*
	  o Hgb level ≥ 11 g/dL
	• Complete resolution of palpable hepatosplenomegaly and all biopsy-proven or suspected SM-related organ damage (C-findings)[c]

CR with partial recovery of peripheral blood counts (CRh)[a]	Requires all criteria for CR be met and response duration must be ≥ 12 weeks; however, patient may have residual cytopenias. The following minimum recovery of peripheral blood counts is required: • ANC > 0.5 × 10^9/L with normal differential (absence of neoplastic MCs and blasts < 1%) *and* • Platelet count > 50 × 10^9/L *and* • Hgb level > 8.0 g/dL

Response	**Criteria for Response**	
Partial remission (PR)[a]	Requires all 3 of the following criteria, and response duration must be ≥ 12 weeks, in the absence of CR/CRh and PD: • Reduction by ≥ 50% in neoplastic MCs in the BM[d] *and/or* other extracutaneous organ at biopsy demonstrating eligible SM-related organ damage • Reduction of serum tryptase level by ≥ 50%[b] • Resolution of 1 or more biopsy-proven or suspected SM-related organ damage (C-finding[s])[c]	
Clinical improvement (CI)[a]	Response duration must be ≥ 12 weeks Requires 1 or more of the nonhematologic and/or hematologic response criteria to be fulfilled (refer to CI criteria below) in the absence of CR, CRh, PR, or PD	
Stable disease (SD)	Not meeting criteria for CR/CRh, PR, CI, or PD	
Progressive disease (PD)[e]	Requires at least 1 element from the criteria below; duration must be ≥ 4 weeks:	
	Baseline	**Postbaseline**
	Any Grade 2 nonhematologic organ damage	• Worsening by 1 grade *and* • Minimum 100% increase (doubling) of laboratory abnormality
	≥ Grade 2 albumin	• Worsening by 1 grade *and* • Decrease by ≥ 0.5 g/dL
	≥ Grade 3 nonhematologic organ damage	Minimum 100% increase (doubling) of laboratory abnormality
	≥ Grade 2 transfusion-independent anemia or thrombocytopenia	New transfusion dependence for an 8-week period of ≥ 4 units of PRBCs or platelets
	Transfusion-dependent anemia or thrombocytopenia	• ≥ 100% increase in the average transfusion frequency for an 8-week period compared with the 12-week pretreatment period
	≥ Grade 3 neutropenia	• > 50% decrease in neutrophil count *and* • Absolute decrease of neutrophil count of ≥ 0.25 x 10^9/L *and* • Grade 4 (< 0.5 x 10^9/L)
	Baseline spleen size of not palpable or < 5 cm	• Development of at least 10 cm palpable symptomatic splenomegaly *or* • Increase in spleen volume ≥ 25%
	Splenomegaly ≥ 5 cm	• > 50% worsening *and*

Response	Criteria for Response	
		• Development of ≥ 10 cm of palpable symptomatic splenomegaly compared with the baseline value *or* • Increase in spleen volume ≥ 25%
Loss of response (LOR)	Loss of a documented CR/CRh, PR, or CI that must be for ≥ 8 weeks. Downgrading of CR/CRh to PR or PR to CI is considered as such but is not considered a loss of response unless CI is also lost for a minimum of 8 weeks. The baseline value for LOR is the pretreatment measurement(s) and not the nadir values during response.	

Abbreviations: ANC = absolute neutrophil count; BM = bone marrow; CI = clinical improvement; CR = complete remission; CRh = complete remission with partial recovery of peripheral blood counts; CT = computed tomography; Hgb = hemoglobin; IWG-MRT-ECNM = International Working Group-Myeloproliferative Neoplasms Research and Treatment; LOR = loss of response; MC = mast cells; MRI = magnetic resonance imaging; PBRCs = packed red blood cells; PD = progressive disease; PR = partial remission; SD = stable disease; SM = systemic mastocytosis.

Note: Guidelines for assessing response are as follows: (A) Only disease-related ≥ Grade 2 organ damage is evaluable as a primary endpoint. (B) Response assessments of CR, CRh, PR, SD, PD, and LOR should only be applied to these ≥ Grade 2 organ damage findings in the context of trials. (C) Disease status at the time of patient removal from the study singularly relates to the updated status of initial ≥ Grade 2 organ damage finding(s). (D) Exclusion of drug-related toxicity and/or other clinical issues (eg, gastrointestinal tract bleeding in the case of worsening anemia/transfusion-dependence) should be undertaken before assigning the designation PD or LOR in a patient with worsening of baseline ≥ Grade 2 organ damage.

[a] Responses that are not maintained for a period of at least 12 weeks do not fulfill criteria for CR/CRh, PR, or CI; however, both maintained and unmaintained (< 12 weeks duration) responses should be recorded in the electronic case report form each time they are observed in order to measure duration of response.

[b] Only valid as a response criterion if the pretreatment serum tryptase level is ≥ 40 ng/mL (ie, if pretreatment serum tryptase is < 40 ng/mL, it will not be considered as a criterion in evaluation of response).

[c] Biopsy of organ(s) in addition to the bone marrow to evaluate for SM-related organ damage may be considered.

[d] Only valid as a response criterion if the pretreatment BM MCs are ≥ 5% (ie, if pretreatment BM MCs are < 5%, BM MCs will not be considered as a criterion in evaluation of response).

[e] Preservation of at least 1 CI finding permits a patient to maintain the response of CI if 1 or more CI findings are lost but none meet criteria for PD. However, if 1 or more of the CI findings become PD, then the CI finding assignment is lost and the patient meets criteria for PD. The baseline value for evaluating PD is the pretreatment measurement(s). The PD findings must be considered related to the underlying disease and not to other clinical factors. Progression of an underlying chronic myeloid neoplasm to acute myeloid leukemia is also considered PD.

Source: Adapted from (Gotlib, et al., 2013).

CI criteria are listed below.

Criteria	**Baseline criteria**	**Criteria for response**
Nonhematologic		
Ascites or pleural effusions	(1) Symptomatic ascites or pleural effusion requiring medical intervention such as use of diuretics (Grade 2) or (2) ≥ 2 therapeutic paracenteses or thoracenteses at least 28 days apart over 12 weeks before C1D-8 (Grade 3) and 1 of the procedures is performed 6 weeks before C1D-8	(1) Complete resolution of symptomatic ascites or pleural effusion (including trace or minimal on radiographic imaging) and no longer in need of diuretics for ≥ 12 weeks and no longer in need of diuretics for ≥ 12 weeks or (2) No therapeutic paracenteses or thoracentesis for ≥ 12 weeks
Liver function test abnormalities	≥ Grade 2 abnormalities in direct bilirubin (> 1.5 × ULN), AST (> 3.0 × ULN), ALT (> 3.0 × ULN) or AP (> 2.5 × ULN) in the presence of: ascites and/or clinically relevant portal hypertension and/or liver MC infiltration that is biopsy-proven or other causes of abnormal liver function are not identified	Reversion of ≥ 1 LFTs to normal range for ≥ 12 weeks
Hypoalbuminemia	≥ Grade 2 hypoalbuminemia (< 3.0 g/dL)	Reversion of albumin to normal range for ≥ 12 weeks
Marked splenomegaly	A spleen that is palpable ≥ 5 cm below left costal margin	≥ 35% reduction in spleen volume based on 3D MRI or CT scan for ≥ 12 weeks
Hematologic		
ANC	≥ Grade 3 ANC (< 1.0×10^9/L)	≥ 100% increase and an absolute increase ≥ 0.5×10^9/L for ≥ 12 weeks
Anemia (transfusion-independent)	≥ Grade 2 Hgb (< 10 g/dL)	An increase in Hgb ≥ 2 g/dL that is maintained for ≥ 12 weeks
Anemia (transfusion-dependent)	Transfusion of ≥ 6 units PRBCs in the 12 weeks before C1D-8 *and* Most recent transfusion occurring in the 4 weeks before C1D-8 *and* Transfusions administered for Hgb ≤ 8.5 g/dL *and*	Transfusion independence for ≥ 12 weeks and maintenance of Hgb ≥ 8.5 g/dL at the end of the 12-week period of response duration

Criteria	Baseline criteria	Criteria for response
	Reason for transfusions is not bleeding, hemolysis or therapy-related	
Thrombocytopenia (transfusion-independent)	≥ Grade 2 thrombocytopenia (< 75×10^9/L)	≥ 100% increase and an absolute increase ≥ 50×10^9/L and no need for platelet transfusion for ≥ 12 weeks
Thrombocytopenia (transfusion-dependent)	(1) Transfusion of ≥ 6 units of apheresed (or ≥ 6 pools of random donor or buffy coat) platelets during the 12 weeks before C1D-8, and (2) ≥ 2 units transfused in the 4 weeks before C1D-8, and (3) Transfusions given for platelet count < 20×10^9/L	Transfusion independence for ≥ 12 weeks and maintenance a platelet count ≥ 20×10^9/L

Abbreviations: 3D = 3 dimension; ALT = alanine aminotransferase; ANC = absolute neutrophil count; AP = alkaline phosphatase; AST = aspartate aminotransferase; CT = computed tomography; Hgb = hemoglobin; IWG-MRT-ECNM = International Working Group-Myeloproliferative Neoplasms Research and Treatment and European Competence Network on Mastocytosis; MC = mast cell; MRI = magnetic resonance imaging; LFT = liver function test; PRBC = packed red blood cell; ULN = upper limit of normal.
Note: Toxicity grading is based on Common Terminology Criteria for Adverse Events, Version 5.0.

Source: Adapted from (Gotlib, et al., 2013).

7.3 mIWG Definition of Evaluable C-findings

Baseline eligibilities are listed below.

Nonhematologic C-findings	
Ascites or pleural effusions	Symptomatic ascites or pleural effusion requiring medical intervention such as: (1) Use of diuretics (Grade 2) or (2) ≥ 2 therapeutic paracenteses *or* thoracenteses (Grade 3) at least 28 days apart over 12 weeks before C1D-8 with 1 procedure performed during the 6 weeks before C1D-8
Liver function abnormalities	• ≥ Grade 2 abnormalities in direct bilirubin (> 1.5 × ULN), AST (> 3.0 × ULN), ALT (> 3.0 × ULN) or AP (> 2.5 × ULN) in the presence of: o Ascites *and/or* o Clinically relevant portal hypertension, *and/or* o Liver MC infiltration that is biopsy-proven *or* o No other identified cause of abnormal liver function
Hypoalbuminemia	≥ Grade 2 hypoalbuminemia (< 3.0 g/dL)
Marked splenomegaly	A spleen that is palpable ≥ 5 cm below the left costal margin

Hematologic C-findings	
ANC	≥ Grade 3 ANC (< 1.0 × 10^9/L)
Anemia (transfusion-independent)	≥ Grade 2 Hgb (< 10 g/dL)
Anemia (transfusion-dependent)	• Transfusion of ≥ 6 units PRBCs in the 12 weeks before C1D-8 *and* • Most recent transfusion occurring during the 4 weeks before C1D-8 *and* • Transfusions administered for Hgb ≤ 8.5 g/dL *and* • Reason for transfusions is not bleeding, hemolysis, or therapy-related
Thrombocytopenia (transfusion-independent)	≥ Grade 2 thrombocytopenia (< 75 × 10^9/L)
Thrombocytopenia (transfusion-dependent)	• Transfusion of ≥ 6 units of apheresed platelets (or ≥ 6 pools of random donor or buffy coat) 12 weeks before C1D-8 *and* • ≥ 2 units transfused 4 weeks before C1D-8 *and* • Transfusions administered for platelet count < 20 × 10^9/L

Abbreviations: ALT = alanine aminotransferase; ANC = absolute neutrophil count; AP = alkaline phosphatase; AST = aspartate aminotransferase; C1D1 = Cycle 1 Day 1; Hgb = hemoglobin; IWG-MRT-ECNM = International Working Group-Myeloproliferative Neoplasms Research and Treatment; MC = mast cell; PBRC = packed red blood cell; ULN = upper limit of normal.

Notes: One or more C-findings is required for eligibility. Grade is based on the Common Terminology Criteria for Adverse Events, Version 5.0.

Source: Adapted from (Gotlib, et al., 2013).

7.4 Response Derivation Algorithm

This section outlines the details on how to algorithmically apply IWG to derive response.

7.4.1 Diagnosis of AdvSM

Diagnosis of AdvSM adjudicated by SSC will be used for algorithmic derivation of responses.

Only patients with ASM, SM-AHN, or MCL are considered as having AdvSM. They will be included in determining evaluability by IWG.

7.4.2 Evaluability by Algorithmic-IWG

The following table outlines the rules on determining baseline evaluability by IWG. MCL by default is evaluable by IWG. ASM/SM-AHN need to meet at least one criteria listed below to be evaluable by IWG.

Baseline IWG Evaluable Criteria
(Ascities Grade 2) and (attribution to SM = Yes)
(Ascities Grade 3) and (paracentesis or thoracentesis >=2 in last 12 wk and at least 28 days apart) and (paracentesis or thoracentesis >=1 in last 6 wk) and (attribution to SM = Yes)
(Pleural Effusions Grade 2) and (attribution to SM = Yes)
(Pleural Effusions Grade 3) and (paracentesis >=2 or thoracentesis >=2 in last 12 wk and at least 28 days apart) and (paracentesis or thoracentesis >=1 in last 6 wk) and (attribution to SM = Yes)
(Direct bilirubin Grade >=2) and ('Ascities = Yes' or 'Portal Hypertension = Yes' or 'Liver biopsy-proven MC infiltration = Yes' or 'Other Causes for abnormal liver function = No') and (attribution to SM = Yes)
(AST Grade >=2) and ('Ascities = Yes' or 'Portal Hypertension = Yes' or 'Liver biopsy-proven MC infiltration = Yes' or 'Other Causes for abnormal liver function = No') and (attribution to SM = Yes)
(ALT Grade >=2) and ('Ascities = Yes' or 'Portal Hypertension = Yes' or 'Liver biopsy-proven MC infiltration = Yes' or 'Other Causes for abnormal liver function = No') and (attribution to SM = Yes)
(ALP Grade >=2) and ('Ascities = Yes' or 'Portal Hypertension = Yes' or 'Liver biopsy-proven MC infiltration = Yes' or 'Other Causes for abnormal liver function = No') and (attribution to SM = Yes)
(Albumin Grade >=2) and (attribution to SM = Yes)
Spleen Palpation >5 cm' and 'Symptomatic splenomegaly = Yes' and (attribution to SM = Yes)
(ANC Grade >=3) and (attribution to SM = Yes)
(RBC transfusion in last 12 wk >=6 units = Yes) and (RBC transfusion in last 4 wk = Yes) and (attribution of baseline anemia to SM = Yes)
(HGB transfusion dependent not Yes) and (HGB Grade >=2) and (attribution of baseline anemia to SM = Yes)
(PLT transfusion in last 12 wk >=6 units = Yes) and (PLT transfusion in last 4 wk >=2 units = Yes) and (attribution of baseline thrombocytopenia to SM = Yes)
(PLT transfusion dependent not Yes) and (PLT Grade >=2) and (attribution of baseline thrombocytopenia to SM = Yes)

7.4.3 Response at a Time Point

Response at a timepoint will be derived first. A response will be assigned in the order of PD, CR, CRh, PR, CI, and SD. Criteria for PD, CR, CRh, PR, and CI are shown below. Disease progression to AML will be assigned as PD. SD will be assigned when no other response can be assigned.

Baseline Evaluable Criteria	PD (any one condition)	CR (require all condition meet)	CRh (require all condition meet)	PR (require MC and tryptase and >=1 of the other evaluable c-finding)	CI (>=1 of the evaluable c-finding)
% MC (multifocal dense infiltrates or interstitial)		central pathology assessment of CR = Yes	central pathology assessment of CR = Yes	central pathology assessment of PR = Yes	
Serum Tryptase >=40 ng/mL		absolute value <20 ng/mL	absolute value <20 ng/mL	if meet baseline evaluable criteria and >=50% reduction from baseline	
(Ascites Grade 2) and (attribution to SM = Yes)	if meet baseline evaluable criteria and (Ascites Grade >=3) and (PI response assessment for this C-finding = PD)	if meet baseline evaluable criteria and Ascites Grade=0	if meet baseline evaluable criteria and Ascites Grade=0	if meet baseline evaluable criteria and Ascites Grade=0	if meet baseline evaluable criteria and Ascites Grade=0
(Ascites Grade 3) and (paracentesis or thoracentesis >=2 in last 12 wk and at least 28 days apart) and (paracentesis or thoracentesis >=1 in last 6 wk) and (attribution to SM = Yes)	if meet baseline evaluable criteria and (Ascites Grade >=4) and (PI response assessment for this C-finding = PD)	if meet baseline evaluable criteria and (no paracentesis)	if meet baseline evaluable criteria and (no paracentesis)	if meet baseline evaluable criteria and (no paracentesis)	if meet baseline evaluable criteria and (no paracentesis)
(Pleural Effusions Grade 2) and (attribution to SM = Yes)	if meet baseline evaluable criteria and (Pleural Effusions Grade >=3) and (PI response assessment for this C-finding = PD)	if meet baseline evaluable criteria and Pleural Effusions Grade=0	if meet baseline evaluable criteria and Pleural Effusions Grade=0	if meet baseline evaluable criteria and Pleural Effusions Grade=0	if meet baseline evaluable criteria and Pleural Effusions Grade=0
(Pleural Effusions Grade 3) and (paracentesis or thoracentesis >=2 in last 12 wk and at least 28 days apart) and (paracentesis or thoracentesis >=1 in last 6 wk) and (attribution to SM = Yes)	if meet baseline evaluable criteria and (Pleural Effusions Grade >=4) and (PI response assessment for this C-finding = PD)	if meet baseline evaluable criteria and (no thoracentesis)	if meet baseline evaluable criteria and (no thoracentesis)	if meet baseline evaluable criteria and (no thoracentesis)	if meet baseline evaluable criteria and (no thoracentesis)

Liver Palpation		not palpable	not palpable		
(Direct bilirubin Grade >=2) and ('Ascites = Yes' or 'Portal Hypertension = Yes' or 'Liver biopsy-proven MC infiltration = Yes' or 'Other Causes for abnormal liver function = No') and (attribution to SM = Yes)	if meet baseline evaluable criteria and (worsening by >=1 grade and absolute value at least double) and (PI response assessment for this C-finding = PD)	if meet baseline evaluable criteria and Dbili Grade =0	if meet baseline evaluable criteria and Dbili Grade =0	if meet baseline evaluable criteria and Dbili Grade =0	if meet baseline evaluable criteria and Dbili Grade =0
(AST Grade >=2) and ('Ascites = Yes' or 'Portal Hypertension = Yes' or 'Liver biopsy-proven MC infiltration = Yes' or 'Other Causes for abnormal liver function = No') and (attribution to SM = Yes)	if meet baseline evaluable criteria and (worsening by >=1 grade and absolute value at least double) and (PI response assessment for this C-finding = PD)	if meet baseline evaluable criteria and AST Grade =0	if meet baseline evaluable criteria and AST Grade =0	if meet baseline evaluable criteria and AST Grade =0	if meet baseline evaluable criteria and AST Grade =0
(ALT Grade >=2) and ('Ascites = Yes' or 'Portal Hypertension = Yes' or 'Liver biopsy-proven MC infiltration = Yes' or 'Other Causes for abnormal liver function = No') and (attribution to SM = Yes)	if meet baseline evaluable criteria and (worsening by >=1 grade and absolute value at least double) and (PI response assessment for this C-finding = PD)	if meet baseline evaluable criteria and ALT Grade =0	if meet baseline evaluable criteria and ALT Grade =0	if meet baseline evaluable criteria and ALT Grade =0	if meet baseline evaluable criteria and ALT Grade =0
(ALP Grade >=2) and ('Ascites = Yes' or 'Portal Hypertension = Yes' or 'Liver biopsy-proven MC infiltration = Yes' or 'Other Causes for abnormal liver function = No') and (attribution to SM = Yes)	if meet baseline evaluable criteria and (worsening by >=1 grade and absolute value at least double) and (PI response assessment for this C-finding = PD)	if meet baseline evaluable criteria and ALP Grade =0	if meet baseline evaluable criteria and ALP Grade =0	if meet baseline evaluable criteria and ALP Grade =0	if meet baseline evaluable criteria and ALP Grade =0
(Albumin Grade >=2) and (attribution to SM = Yes)	if meet baseline evaluable criteria and (worsening by >=1 grade and absolute value decrease by >= 0.5 g/dL) and (PI response assessment for this C-finding = PD)	if meet baseline evaluable criteria and ALB Grade =0	if meet baseline evaluable criteria and ALB Grade =0	if meet baseline evaluable criteria and ALB Grade =0	if meet baseline evaluable criteria and ALB Grade =0

Spleen Palpation >5 cm' and 'Symptomatic splenomegaly = Yes' and (attribution to SM = Yes)	(if (baseline spleen size is not palpable or <=5 cm) and (palpation is >=10 cm)) or (if (meet baseline evaluable criteria) and (>50% increase in palpation size) and (palpation is >=10 cm)) and (PI response assessment for this C-finding = PD)	not palpable	not palpable		
Spleen volume (central radiology)	spleen volume increase ≥ 25%			(spleen volume decrease >= 35%)	(spleen volume decrease >= 35%)
(ANC Grade >=3) and (attribution to SM = Yes)	((if meet baseline evaluable criteria) and (>50% decrease in ANC) and (absolute decrease >= 0.25 x 10^9/L) and (Grade>=4)) and (PI response assessment for this C-finding = PDs)	ANC >= 1 x 10^9/L	ANC > 0.5 x 10^9/L	(if meet baseline evaluable criteria) and (>=100% increase in ANC) and (ANC > 0.5 x 10^9/L)	(if meet baseline evaluable criteria) and (>=100% increase in ANC) and (ANC > 0.5 x 10^9/L)
(RBC transfusion in last 12 wk >=6 units = Yes) and (RBC transfusion in last 4 wk = Yes) and (attribution of baseline anemia to SM = Yes)	((if meet baseline evaluable criteria) and (average 8-wk transfusion frequency increased >=100% compare to the 12-wk baseline frequency)) and (PI response assessment for this C-finding = PD)	HGB >= 11 g/dL and (no transfusion)	HGB >8 g/dL and (no transfusion)	(if meet baseline evaluable criteria) and (transfusion free) and (HGB value >=8.5 g/dL)	(if meet baseline evaluable criteria) and (transfusion free) and (HGB value >=8.5 g/dL)
(HGB transfusion dependent not Yes) and (HGB Grade >=2) and (attribution of baseline anemia to SM = Yes)	((if meet baseline evaluable criteria) and (average 8-wk transfusion frequency >= 0.5 unit of RBCs)) and (PI response assessment for this C-finding = PD)	HGB >= 11 g/dL and (no transfusion)	HGB >8 g/dL and (no transfusion)	(if meet baseline evaluable criteria) and (HGB increase >=2 g/dL)	(if meet baseline evaluable criteria) and (HGB increase >=2 g/dL)

(PLT transfusion in last 12 wk >=6 units = Yes) and (PLT transfusion in last 4 wk >=2 units = Yes) and (attribution of baseline thrombocytopenia to SM = Yes)	((if meet baseline evaluable criteria) and (average 8-wk transfusion frequency increased >=100% compare to the 12-wk baseline frequency)) and (PI response assessment for this C-finding = PD)	PLT >= 100 x 10^9/L and (no transfusion)	PLT > 50 x 10^9/L and (no transfusion)	(if meet baseline evaluable criteria) and (transfusion free) and (PLT >= 20 x 10^9/L)	(if meet baseline evaluable criteria) and (transfusion free) and (PLT >= 20 x 10^9/L)
(PLT transfusion dependent not Yes) and (PLT Grade >=2) and (attribution of baseline thrombocytopenia to SM = Yes)	((if meet baseline evaluable criteria) and (average 8-wk transfusion frequency >= 0.5 unit of PLTs)) and (PI response assessment for this C-finding = PD)	PLT >= 100 x 10^9/L and (no transfusion)	PLT > 50 x 10^9/L and (no transfusion)	(if meet baseline evaluable criteria) and (>=100% increase in PLT) and (PLT increase >= 50 x 10^9/L) and (transfusion free)	(if meet baseline evaluable criteria) and (>=100% increase in PLT) and (PLT increase >= 50 x 10^9/L) and (transfusion free)

7.4.4 Best Confirmed Response

Once response at a timepoint is assigned, confirmed response will be derived by enforcing duration requirements. A confirmed PD requires 8 weeks duration by IWG and 4 weeks duration by mIWG, with the exception of PD due to AML that will not need confirmation. A confirmed CR, CRh, PR, or CI require 12 weeks (+/- 7 days to accommodate visit scheduling) duration. CR, CRh or PR cannot be confirmed until at least 2 valid BMB assessments are performed post baseline.

LoR will be assigned as a response when a confirmed CR/CRh/PR/CI is interrupted by SD for at least 8 weeks.

7.5 [redacted]

[redacted]

[redacted]

7.6 [redacted]

[redacted]

[redacted]

8. REFERENCES

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

Gotlib et al. International Working Group-Myeloproliferative Neoplasms Research and Treatment (IWG-MRT) & European Competence Network on Mastocytosis (ECNM) consensus response criteria in advanced systemic mastocytosis. Blood 2013 121: 2393-2401.