NOVARTIS

Cinical Development

ABL001/Asciminib/Scemblix®

CABL001J12301

A phase III, multi-center, open-label, randomized study of oral asciminib versus Investigator selected TKI in patients with newly diagnosed Philadelphia Chromosome Positive Chronic Myelogenous Leukemia in Chronic Phase

Statistical Analysis Plan (SAP)

Document type:	SAP Documentation
Document status:	Amendment 2
Release date:	15-Sep-2023
Number of pages:	83

Property of Novartis
Confidential
May not be used, divulged, published or otherwise disclosed
without the consent of Novartis

Template Version 6.0, Effective from 23-Nov-2022

Document History – Changes compared to previous final version of SAP

Date	Time point	Reason for update	Outcome for update	Section and title impacted (Current)
25-Oct-2021	Prior to the randomization of first patient	NA	N/A - First version	NA
13-Mar-2023	Prior to dry-run 1	See details below	Amendment 1	See details below
15-Sep-2023	Prior to dry-run 2	See details below	Amendment 2	See details below

Amendment 1 (prior to dry-run 1) changes and sections impacted:

1. "BCR-ABL" revised to "BCR::ABL"
2. "Patients" revised to "participants"
3. Section 1: an introduction to the current amendment added
4. Section 1.1: ELTS risk groups defined based on ELTS score
5. Section 1.1: treatment group(s) defined as a synonym as treatment arm(s)
6. Table 1-2: wording revised to align with the protocol
7. Section 2.1.1: data used for each CSR clarified
8. [redacted]
9. Section 2.1.3.10: safety follow-up and survival follow-up clarified
10. Section 2.1.3.11: assessment target dates and windows revised to align with the protocol
11. Section 2.1.3.13: total follow-up time defined
12. Section 2.2: All Subjects Set defined
13. Section 2.2.4.1: analysis sets used for subgroup analyses clarified
14. Section 2.2.4.4: subgroup analysis for China defined
15. Section 2.4: section numbering changed due to the removal of an incorrectly placed sub-section heading
16. Section 2.4.1.1: reporting of summary statistics by individual TKI removed to align with the protocol
17. Section 2.4.1.2: calculations of planned cumulative dose defined
18. Section 2.5.2 and Section 2.6.2: the equal signs corrected to less-than-or-equal-to in the null hypotheses as in the protocol
19. Section 2.5.6: supplementary analysis added for the primary endpoint
20. Section 2.6.6: supplementary analysis added for the secondary endpoint
21. Section 2.7.5: supplementary analysis added for the secondary safety endpoint as per protocol
22. Section 2.9.1.1: endpoint BCR::ABL≤1% at and by scheduled time points added per protocol amendment 1.
23. Section 2.9.1.1: endpoints for "by" timepoint analyses clarified to exclude baseline values
24. Section 2.9.1.2: events considered competing risks for time to first MMR/MR4.0/MR4.5 defined

25. Section 2.9.1.3: definition of TTF and FFS (Table 2-7) corrected to align with protocol amendment 1
26. Section 2.9.2.1: subsections added to distinguish different analyses for different response endpoints
27. Section 2.9.2.2.3: analyses of other time-to-event endpoints specified
28. Section 2.11.1.1: exposure adjusted incidence rates for AESI added
29. Section 2.11.4.2: details of Framingham Cardiovascular Risk scores and Charlson Comorbidity Index added
30. Section 2.13 Analyses of exploratory objectives added. These include:
    - Section 2.13.1: biomarker endpoints
    - Section 2.13.2: pharmacogenomics endpoints
    - [redacted]
    - Section 2.13.4: Health Care Resource Utilization endpoints
    - Section 2.13.5: exploratory analysis for the primary and the key secondary endpoints
31. Sections 5.3.3.3, 5.3.3.4, and 5.3.3.5: definitions of loss of MMR, loss of MR4.0, and loss of MR4.5 updated to align with the protocol
32. Section 5.3.4: imputation of CCyR removed due to the addition of BCR::ABL ≤1% as a secondary endpoint
33. Section 5.3.4.1: definition of loss of CCyR added
34. Section 5.3.5.2: definition of loss of CHR added
35. Section 5.4: additional details to statistical models and analyses added
36. Section 5.4.5: calculation of exposure adjusted incidence rate added
37. Section 5.5.1: derivations of Framingham cardiovascular risk scores added
38. Section 5.5.2: derivations of Charlson Comorbidity Index added

There may be other clarifications or editorial changes, e.g., corrections of typos. All changes can be found in the tracked version of this document.

Amendment 2 (prior to dry-run 2) changes and sections impacted:

39. Section 1.2.3: the summary measure for the secondary safety estimand TTDAE is changed from the cumulative incidence function to cause-specific hazard to align with the protocol
40. Section 2.1.3.9: assessments used for baseline demographic and background data defined
41. Section 2.1.3.9: baseline for safety evaluation updated to include assessments taken on first dose day to be consistent with the assessment schedule of Table 8-1 in the protocol.
42. Section 2.1.3.11: rules for selecting from multiple molecule assessments revised
43. Section 2.1.3.11: rules for selecting from multiple ePRO records collected on the same day added
44. Section 2.2.1: labels for MMR/MR4.0/MR4.5 responder sets corresponding to FAS_{ima} and FAS_{2GTKI} provided
45. Section 2.2.4.3: additional outputs added for Japan subgroup analysis
46. Section 2.2.4.4: subgroups for analyses related to China submission re-defined

47. Section 2.3: updated to display summary statistics by treatment arms and not by total
48. Section 2.3.2: summary of "reasons for pre-randomization choice of TKI" to be produced for the CSR
49. Section 2.3.3: cross-table for ELTS risk groups recorded in the IRT and eCRF to be produced for the CSR
50. Section 2.3.5: wordings on separate listings for medical history removed since they are provided in other listings.
51. Section 2.7.1: definition of TTDAE updated to include death due to AE as event of interest, and death due to other causes as competing risks
52. Section 2.7.2: the main analysis for TTDAE updated to use cause-specific hazard to align with the protocol
53. Section 2.9.1.2: treatment failure added as one of the end dates for the durations of molecular responses to align with the protocol
54. Sections 2.9.1.2 and 2.9.3.1: censoring date for time to first MMR, MR4.0, or MR4.5 revised to use last PCR assessment date while on treatment to be consistent with other endpoints
55. Section 2.9.1.3 Table 2-7: definitions of FFS and EFS corrected to include competing events as specified in the protocol
56. Section 2.9.1.3: details added for the censoring time for TTS, FFS, EFS, and PFS as per protocol
57. Section 2.9.2: subsections headers 2.9.2.2 and 2.9.2.3 added for clarity. No changes to existing contents.
58. Section 2.9.2.3: analysis for FFS and EFS corrected to use cumulative incidence as specified in the protocol
59. Section 2.9.3.1: imputation rule for other secondary response endpoints removed as it only applies to the primary and key secondary endpoints.
60. Section 2.9.4: supplemental analysis for time to first MMR, MR4.0, and MR4.5 added to estimate median survival time
61. Section 2.11.1: wording clarified for summarizing first occurrences of AE and SAE
62. Section 2.11.1.1: competing events for the CIF of AESI specified
63. Section 2.11.1.1: for AESI plot of CIF and summary of first event removed
64. Section 2.11.3: handling of multiple lab assessments taken on the same day added
65. Section 2.11.4.3: text below Table 2-9 removed as it is kept by mistake.
66. Section 2.12: source of ePRO compliance summary changed from eCRF to the questionnaires
67. Section 2.13.1 and Section 5.3.3: a separate category added for subjects with no evidence of typical transcript for summarizing BCR::ABL ratio categories

[redacted]

69. Section 2.13.4: analysis sets for HCRU updated to align with the protocol
70. Section 5.1.3: imputation rule added for missing CML disease diagnosis date

71. Section 5.3.3: imputation of 0 value of BCR::ABL copy number for checking loss of molecular responses updated to include MR4.0 and MR4.5
72. Section 5.3.3: text on derivation of BCR::ABL ratio (IS%) simplified since central lab results are used.
73. Section 5.3.3.1.1: definition of responders of BCR::ABL≤ 1% added
74. Sections 5.3.3.2, 5.3.3.4, and 5.3.3.5: definition of MMR, MR4.0, and MR4.5, respectively, updated to exclude subjects with no evidence of typical transcript
75. Sections 5.3.3.3, 5.3.3.4, and 5.3.3.5: confirmed loss of CHR revised to loss of CHR in the definitions of confirmed loss of MMR, loss of MR4.0, and loss of MR4.5.
76. Section 5.3.6: details to the reporting of treatment failure per ELN criteria added
77. Section 5.3.7: handling of progression to CML-AP/BC during the first four weeks of treatment added per Protocol Amendment 2
78. Section 5.4.1.1: details to testing the hypothesis in SAS added
79. Section 5.4.1.1: details to deriving one-sided p-value in SAS added
80. Section 5.4.1.2: details to estimating response rates and their confidence intervals in SAS added
81. Section 5.4.3: details of estimating cause-specific hazard in SAS added
82. Section 5.4.4: SAS code for hazard ratio removed as it is not planned for any endpoint.
83. Section 5.5.1: conversion factor between mmol/L and mg/dl for cholesterol for calculating Framingham CVD risk scores corrected

There may be other clarifications or editorial changes, e.g., corrections of typos. All changes can be found in the tracked version of this document.

Table of contents

List of tables

List of figures

List of abbreviations

1G	First generation
2G	Second generation
AE	Adverse event
AESI	Adverse event of special interest
ALP	Alkaline phosphatase
ALT	Alanine aminotransferase
AP	Accelerated phase
AST	Aspartate aminotransferase
ATC	Anatomical therapeutic classification
AUC	Area under the curve
BC	Blast crisis
BCR ABL	BCR ABL fusion gene (also called the Philadelphia chromosome)
BLQ	Below the limit of quantification
CBC	Complete blood count
CVD	Cardiovascular disease
CCyR	Complete cytogenetic response
CHR	Complete hematologic response
CI	Confidence Interval
CIF	Cumulative incidence function
CMH	Cochrane-Mantel-Haenszel
CML	Chronic myelogenous leukemia
CML-CP	Chronic myelogenous leukemia in chronic phase
CRO	Contract research organization
CSP	Clinical study protocol
CSR	Clinical study report
CTCAE	Common terminology criteria for adverse events
CV	Coefficient of variation
DAR	Dosage administration record
DI	Dose intensity
DMC	Data monitoring committee
DRL	Drug reference listing
DSUR	Development safety update report
ECG	Electrocardiogram
eCRF	Electronic case report form
EFS	Event Free Survival
ELTS	EUTOS Long Term Survival
EOS	End of Study (study level)
EOT	End of treatment in the treatment phase (patient level)
EUTOS	The **EU**ropean **T**reatment **O**utcome **S**tudy
FAS	Full analysis set
FAS_{2GTKI}	Full analysis set for the stratum of patients with a 2G TKI as their pre-randomization selection of TKI
FAS_{IMA}	Full analysis set for the stratum of patients with an imatinib as their pre-randomization selection of TKI

FD	First dose date of study treatment during the treatment phase
FFS	Failure Free Survival
HCRU	Health care resource utilization
HLGT	High level group term
HLT	High level term
ICF	Informed consent form
IMA	Imatinib
IRT	Interactive response technology
IS	International scale
IS-TKI	Investigator selected TKI
KM	Kaplan Meier
LD	Last dose date of study treatment during the treatment phase
LPFT	Last patient first treatment
MCyR	Major cytogenetic response
mCyR	Minor cytogenetic response
MedDRA	Medical dictionary for regulatory activities
MMR	Major molecular response
NGS	Next generation sequencing
NMPA	National Medical Products Administration (China)
NMQ	Novartis MedDRA Query
OS	Overall survival
PAS	Pharmacokinetic analysis set
PCR	Polymerase chain reaction
PCyR	Partial cytogenetic response
PD	Pharmacodynamic
PDI	Planned dose intensity
PFS	Progression-free survival
Ph+	Philadelphia chromosome positive
PK	Pharmacokinetics
PMDA	Pharmaceuticals and Medical Devices Agency (Japan)
PPS	Per-protocol set
PRO	Patient-reported outcomes
PRS-TKI	Pre-randomization selection of TKI
PSUR	Periodic safety update report
PT	Preferred term
q.d.	Qua'que di'e / once a day
RDI	Relative dose intensity
SAE	Serious adverse event
SAP	Statistical analysis plan
SD	Standard deviation
SMQ	Standardized MedDRA query
SOC	System organ class
TBL	Total bilirubin
TKI	Tyrosine kinase inhibitor
TTDAE	Time to discontinuation of study treatment due to AE

TTF	Time to treatment failure
ULN	Upper limit of norm
VAS	Visual analogue scale
WBC	White blood cells
WoC	Withdrawal of consent
WHO-DD	World Health Organization Drug Dictionary

1 Introduction

Introduction in SAP version Final 0.0

The current Statistical Analysis Plan (SAP) describes the planned primary analyses for the primary objectives, the key secondary objectives, and other secondary objectives as defined in the CABL001J12301 Clinical Study Protocol (CSP) version 00, as well as for the postings for ClinTrial.gov and EudraCT. These analyses are defined prior to the randomization of the first participant in the study.

The SAP provides details to statistical models and data analyses to supplement the information in the CSP.

Introduction to SAP Amendment 1

The content of this SAP is based on the amended protocol CABL001J12301 v01 (dated 09-May-2022). In addition, planned analyses for exploratory endpoints are defined.

Introduction to SAP Amendment 2

The revision of the SAP is based on the amended protocol CABL001J12301 v02 (dated 15-May-2023).

All decisions regarding the primary analysis, as defined in the SAP document, have been made prior to database lock and unblinding of the study data.

The output shells accompanying this document can be found in the Tables, Figures, and Listings (TFL) shells document. The specifications for derived variables and datasets can be found in the Programming Datasets Specifications (PDS) document.

All changes to the planned analysis described in this document required before or after database lock will be made through an amendment or addendum, respectively.

1.1 Study design

Study CABL001J12301 is a phase III, multi-center, open-label, randomized study in adult participants with newly diagnosed Philadelphia Chromosome Positive Chronic Myelogenous Leukemia in Chronic Phase (Ph+ CML-CP). The two treatment arms (or treatment groups) are:

1. Investigational treatment arm: oral asciminib 80 mg q.d.
2. Control (or comparator) treatment arm: Investigator selected TKI (IS-TKI)

Unless otherwise specified, in this document asciminib refers to asciminib 80 mg q.d. The IS-TKI is one of the following treatment options approved by major health authorities (for example, FDA or EMA) for first-line treatment of CML-CP*:

- Imatinib (starting dose 400 mg q.d.)
- Nilotinib (starting dose 300 mg b.i.d.)
- Dasatinib (starting dose 100 mg q.d.)
- Bosutinib (starting dose 400 mg q.d.)

** Dose recommendations (including starting dose) as per local label should be followed for participants with hepatic and renal impairment.*

Imatinib is considered as a first generation TKI (1G), the latter three—nilotinib, dasatinib, and bosutinib—are second generation (2G) TKIs.

It is planned that 402 participants are to be randomized in a 1:1 ratio to the two treatment arms. The final number of randomized participants may be larger than 402. The randomization is stratified based on the following two factors:

- ELTS risk groups defined based on ELTS score:
  - Low: ELTS $\leq$ 1.5680
  - Intermediate: 1.5680 < ELTS $\leq$ 2.2185
  - High: ELTS > 2.2185
- Pre-randomization selection of TKI (PRS-TKI): 1G (imatinib) or 2G TKI.

(The difference between PRS-TKI and IS-TKI is that the former refers to the second stratification factor, whereas the latter refers to the control arm treatment.)

The IRT system allocates participants equally to the treatment arms within a stratum, and allocates approximately 50% of participants in each of the two PRS-TKI strata to reflect the distribution of 1G and 2G TKIs in clinical practice.

Participants will continuously receive the assigned treatment until the End of Study (EoS), or until premature treatment discontinuation. The definition of EoS can be found in Section 2.1.3.3 of the current document.

No formal interim analysis is planned in the study. The study includes a data monitoring committee (DMC), which evaluates safety data at protocol defined intervals (i.e., beginning approximately 6 months after the first randomized participant has started treatment and approximately every 6 months thereafter) and provide recommendations to the sponsor to continue, modify, or stop the study early.

1.2 Study objectives, endpoints and estimands

1.2.1 Study objectives

Table 1-1 contains the study objectives and the associated endpoints that will be formally tested using the pre-specified testing strategy (see Section 2.7 of the current document). Table 1-2 contains other objects and endpoints for which descriptive statistics will be provided.

Table 1-1 Objectives and related endpoints to be tested

Primary objectives	Endpoints
1. To compare the efficacy of asciminib versus IS-TKI 2. To compare the efficacy of asciminib versus IS-TKI within the stratum of participants with imatinib as their PRS-TKI	Major Molecular response (MMR) at Week 48 (Yes/No)
Key Secondary objectives	**Endpoints**
1. To compare the efficacy of asciminib versus IS-TKI 2. To compare the efficacy of asciminib versus IS-TKI within the stratum of participants with imatinib as their PRS-TKI	MMR at Week 96 (Yes/No)
Secondary safety objective	**Endpoint**
To characterize the safety and tolerability profile of asciminib versus 2G TKIs during the course of study.	Time to discontinuation of study treatment due to AE

Table 1-2 Other objectives and endpoints

Secondary efficacy objectives	Endpoints
To estimate the efficacy of asciminib versus IS-TKI within the strata of participants with a 2G TKI as their PRS-TKI	• MMR at Week 48 • MMR at Week 96
To compare the efficacy of asciminib versus (1) IS-TKI arm, (2) within the stratum of participants with imatinib as their PRS-TKI, and (3) within the stratum of participants with 2G TKI as their PRS-TKI	• MMR at all scheduled data collection time points (except at Week 48 and at Week 96) • MMR by all scheduled data collection time points
To estimate the efficacy of asciminib versus (1) IS-TKI arm, (2) within the stratum of participants with imatinib as their PRS-TKI, and (3) within the stratum of participants with 2G TKI as their PRS-TKI	• MR4.0 and MR4.5 at and by all scheduled data collection time points • Complete Hematological response (CHR) at and by all scheduled data collection time points • BCR::ABL1 ≤1% at and by all scheduled data collection time points. • Complete Cytogenetic response (CCyR) by Week 48 and by Week 96 • Time to first MMR, MR4.0, MR4.5 • Duration of MMR, MR4.0, MR4.5 • Time to treatment failure (TTF) • Progression free survival (PFS) • Event free survival (EFS) • Failure free survival (FFS) • Overall survival (OS)
Secondary PK objective	**Endpoints**
To characterize the PK of asciminib	• Trough plasma concentrations • For the full PK group: Cmax, Tmax, AUCtau, AUClast, CL/F

Other secondary safety objectives	Endpoints
To characterize the safety and tolerability profile of asciminib versus IS-TKI during the course of study.	Type, frequency and severity of adverse events, changes in laboratory values that fall outside the pre-determined ranges and clinically notable ECG, and other safety data (vital signs, physical examination).
Secondary objectives for PROs	**Endpoints**
To assess the effect of asciminib versus IS-TKI on patient-reported disease-related symptoms, functioning, and health-related quality of life (HRQoL)	Change from baseline in overall scores and individual scales of the EORTC QLQ-C30 v3.0 and EORTC QLQ-CML24
Exploratory biomarker objectives	**Endpoints**
To characterize mutations in the [redacted] and their association with molecular response.	Proportion of participants who develop any [redacted]
To conduct gene expression analysis in peripheral blood to predict treatment response	Correlation/association between expression profiles changes from baseline and on treatment with response as an effect of asciminib.
To explore the impact of immune landscape of peripheral blood on treatment response.	Baseline and changes from baseline of immune markers and their correlation with treatment molecular response (MMR and MR4.0 when applicable).
Exploratory objective for Pharmacogenetics	**Endpoints**
To explore [redacted] activity variation on asciminib exposure	Baseline Genetic variant analysis of the [redacted] to investigate the effect of [redacted] activity variation on asciminib exposure.
Exploratory objective for healthcare resource utilization	**Endpoints**
To compare the impact of treatment on health care resource utilization between treatment arms in all participants	Health care resource burden over time
Exploratory objectives for PROs	**Endpoints**
To evaluate health-related quality of life and other PROs in each treatment arm	[redacted]
To explore participant self-reported treatment related symptomatic adverse events between treatment arms	[redacted]
To explore participant self-reported overall impact of side effects of treatment in each treatment arm on PROs from baseline during the course of the study	[redacted]

1.2.2 Primary and key secondary efficacy estimands

The primary and the key secondary efficacy estimands are defined in CSP Section 2.1 and Section 2.2.1, respectively. For completeness they are reproduced below with these minor modifications: (1) the use of notation T (defined below), (2) the use of abbreviations IS-TKI and PRS-TKI, (3) the omission of "80 mg q.d." when referring to asciminib, and (4) the parenthetical explanation to the first intercurrent event is modified.

Let T = 48 or 96 be the time points (Week) associated with primary or the key secondary efficacy estimand, respectively.

Combined Section 2.1 and Section 2.2.1 of CSP:

The primary and key secondary estimands will be based on the treatment policy and composite approaches as defined in this section.

The clinical questions of interest are:

1. What is the efficacy of asciminib compared to IS-TKI with respect to MMR at Week T,
2. What is the efficacy of asciminib compared to IS-TKI within the stratum of participants that have imatinib as their PRS-TKI with respect to MMR at Week T,

without meeting any treatment failure criteria [as per European Leukemia Network (ELN) criteria, (Hochhaus et al., 2020)] and without treatment discontinuation, prior to Week T, in newly diagnosed CML-CP patients; *regardless of* dose interruptions/reductions/allowed dose escalations; *regardless of* dosing errors, changes of concomitant medication, intake of prohibited medication; and *regardless of* taking a TKI different from their PRS-TKI in the comparator arm.

The justification for the primary and key secondary estimands are that these will capture the effects of the study drugs in a manner that reflects their use in current clinical practice.

The primary and key secondary estimands are described by the following attributes:

Population: Newly diagnosed adult Ph+ CML-CP patients, as defined by Inclusion/ Exclusion criteria, (1) overall and (2) within the stratum with imatinib as their PRS-TKI.

Endpoint: the composite endpoint of Major Molecular Response (MMR) at Week T, without meeting any treatment failure criteria prior to Week T and without discontinuation due to any reasons prior to Week T. A participant will be counted as being in MMR at Week T if he/she meets the MMR criterion (BCR::ABL ratio (IS) $\leq$ 0.1%) at Week T. If the participant meets any treatment failure criteria prior to Week T or discontinues treatment due to any reason prior to Week T, the participant is counted as not being in MMR at Week T.

Intercurrent events (IE):

- Taking a TKI different from their PRS-TKI stratum (i.e., taking imatinib as IS-TKI when a 2G has been the pre-randomisation selected TKI and vice versa): *treatment policy strategy*
- Change of study treatment per protocol (dose reduction/interruption/allowed dose escalations): *treatment policy strategy*
- Dosing errors (e.g., missed dose): *treatment policy strategy*
- Deviation in any intake of concomitant medications: *treatment policy strategy*
- Intake of prohibited medications: *treatment policy strategy*
- Meeting any treatment failure criteria prior to Week T or treatment discontinuation due to any reason prior to Week T: *composite*
- Handling of remaining IEs: no other IE foreseen.

Treatment:

1. The randomized treatment arm (the investigational treatment asciminib or the IS-TKI); with or without dose modifications (reductions/interruptions/allowed dose escalations);

regardless of dosing errors, deviation in any intake of concomitant medications, intake of prohibited medication; and regardless of taking or not taking a TKI different from their PRS-TKI stratum in the comparator arm.

2. The randomized treatment arm (the investigational treatment asciminib or the IS-TKI), within the stratum of participants that had imatinib as their PRS-TKI; with or without dose modifications (reductions/interruptions/allowed dose escalations); regardless of dosing errors, deviation in any intake of concomitant medications, intake of prohibited medication; and regardless of taking or not taking a TKI different from their PRS-TKI stratum in the comparator arm.

The summary measure: stratum adjusted difference in the proportions of participants that are in MMR at Week T and corresponding 95% confidence interval (CI), between the

1. Randomized treatments (asciminib versus IS-TKI).
2. Randomized treatments (asciminib versus IS-TKI), within the stratum of participants that have imatinib as their PRS-TKI.

1.2.3 Secondary safety estimand

A secondary safety estimand related to the endpoint time to discontinuation of study treatment due to AE (TTDAE) will be based on the treatment policy and while on treatment approaches as defined in this section.

A secondary safety related clinical question of interest is: what is the safety and tolerability of asciminib compared to 2G TKIs with respect to the TTDAE, where prior treatment discontinuation due to other reasons is considered a competing risk event, in newly diagnosed CML-CP patients regardless of dose interruptions, reductions, or allowed dose escalations, regardless of dosing errors, changes on concomitant medication, intake of prohibited medication.

The justification for this secondary estimand is that it will capture the safety and tolerability of asciminib compared to 2G TKIs and the effect of additional medications, mirroring the conditions in current clinical practice.

This estimand is described by the following attributes:

Population: Newly diagnosed adult Ph+ CML-CP patients, satisfying the study inclusion and exclusion criteria and are treated with asciminib or a 2G TKI as their first starting dose.

Endpoint: TTDAE

Intercurrent events (IE):

- Discontinuation from study treatment due to other reasons: *while on treatment strategy*
- Change on study treatment per protocol (dose reduction, interruption, or allowed dose escalations): *treatment policy strategy*
- Dosing errors (e.g., missed dose): *treatment policy strategy*
- Deviation in any intake of concomitant medications: *treatment policy strategy*
- Intake of prohibited medications: *treatment policy strategy*

Handling of remaining IEs: no other intercurrent events foreseen.

Treatment:

- The actual treatment received (asciminib or a 2G TKI) with or without dose modifications (reductions, interruptions, or allowed dose escalations) regardless of dosing errors, deviation in any intake of concomitant medications, or intake of prohibited medications.

The summary measure: the cause-specific hazard for TTDAE.

2 Statistical methods

2.1 Data analysis general information

The planned analyses will be performed by Novartis or a designated CRO. SAS version 9.4 or later will be used to perform the data analyses and to generate tables, figures, and listings. The multiplicity adjusted p-value will be generated using the gMCP package version 0.8-15 or a higher version. RStudio Server Pro with R version 3.6.1 (2019-07-05) or above in the Novartis DaVinci Production Computing Environment will be used to run the gMCP package.

For between-treatment comparisons of efficacy endpoints, randomization stratification factors, ELTS and PRS-TKI, as recorded in the IRT system, will be included in respective stratified statistical tests.

2.1.1 Data included in the analyses

The analysis data cut-off dates for the planned analyses are:

- 48-week primary analysis: after all randomized participants have completed their 48-week assessments or discontinued earlier.
- 96-week analysis: after all randomized participants have completed their 96-week assessments or discontinued earlier.
- Final analysis: approximately 5 years from the date when the last randomized participant receives the first study treatment dose.

Additional data updates for regulatory purposes, e.g., Day 120 Safety update, may be performed as required.

All data with an assessment date or event start date (e.g., vital sign assessment date or start date of an adverse event) prior to or on the cut-off date will be included in the analysis, with one exception: for efficacy endpoints that are defined as response or no response, e.g., MMR at or MMR by time points, data will be presented for up to week 48 (or week 96) from each participant for the 48-week (or 96-week) analysis. For time-to-event efficacy endpoints and for safety endpoints, all data collected up to the data cut-off date will be presented. Any data collected beyond the cut-off date will not be included in the analysis and will not be used for any derivations.

All events with start date before or on the respective cut-off date and end date after the respective cut-off date will be reported as ongoing. The same rule will be applied to events starting before or on the respective cut-off date but without a documented end date. This

approach applies, in particular, to adverse event and concomitant medication reporting. For these events, the end date will not be imputed and therefore do not appear in the listings.

2.1.2 General analysis conventions

Unless specified otherwise, data from all study sites and centers will be pooled for the analysis. Due to expected small number of participants enrolled at centers, no center effect will be assessed.

Qualitative (or categorical) data (e.g., gender, race) will be summarized by means of contingency tables by treatment arm; a missing category will be included as applicable. Frequency counts and percentages will be calculated using the number of participants in the relevant population or subgroup as the denominator.

Quantitative (or continuous) data (e.g., age, body weight) will be summarized by appropriate descriptive statistics (i.e., mean, standard deviation, median, minimum, and maximum) by treatment arm. For pharmacokinetics (PK) concentration and parameters descriptive statistics may also include coefficient of variation (CV), geometric-mean, and geometric-CV.

2.1.3 General definitions

2.1.3.1 Investigational drug and study treatment

The investigational drug refers to asciminib only. The study treatment refers to either asciminib or the control arm treatment of any of the IS-TKI received during the treatment phase.

2.1.3.2 Treatment arms

Treatment arms are defined in Section 1.1 of the current SAP. No crossover of study treatment across arms, and no change of study treatment within the IS-TKI will be allowed.

2.1.3.3 Date of end of study (EoS)

The date of EoS occurs 5 years from the date of last participant's first dose in the study. Participants who discontinue the study treatment prematurely due to any reason are followed up for survival and progression to AP or BC until EoS. Note that EoS is a study level concept, i.e., there is only one EoS date.

2.1.3.4 Date of end of study treatment (EOT) for each participant

For each participant, the EOT date is the date that the study treatment is ended for him/her. On this date, the participant is treated for at least 5 years unless he/she has discontinued study treatment earlier. The date of EOT is the date the participant takes the last dose of study treatment (as recorded in the Dosage Administration Record (DAR) page). The participant should complete his/her EOT assessments (Visit number 220 in Table 8-1 of CSP) following EOT. It is possible that there is a gap between the date of EOT and the date of EOT assessment visit.

The date of EOT is a participant level concept.

2.1.3.5 Date of first administration of randomized study treatment for each participant

The date of the first administration of a randomized study treatment, or first dose (FD), is derived as the first date when a non-zero dose of study treatment is administered to a participant as per the DAR page. The date of FD will also be referred as the start of study treatment. The date of FD is a participant level concept.

2.1.3.6 Date of last administration of randomized study treatment for each participant

Similarly, the date of the last administration of a randomized study treatment, or last dose (LD), is defined as the last date when a non-zero dose of study treatment is administered to a participant as per DAR eCRF. The date of LD is also a participant level concept.

2.1.3.7 Study day

For each participant, the date of randomization is defined as Day 1. The day before randomization is defined as Day -1.

The study day describes the day of an event or an assessment relative to the reference start date.

The study day is defined as:

- The date of the event (visit date, onset date of an event, assessment date, etc.) − reference start date + 1 if event is on or after the reference start date.
- The date of the event (visit date, onset date of an event, assessment date, etc.) − reference start date if event precedes the reference start date.

The reference start date for each participant:

- For safety assessments (e.g., adverse event onset, laboratory abnormality occurrence, vital sign measurement, dose interruption, PK) is the start of the study treatment.
- For all other, non-safety assessments (e.g., molecular response, survival time, disease progression, ECOG performance status, patient reported outcomes (PRO)) is the date of the randomization.

The study day will be displayed in the data listings. If an event starts before the reference start date, the study day displayed on the listing will be negative.

2.1.3.8 Time unit

A year is defined as 365.25 days.

A month is 30.4375 (=365.25/12) days. If duration is reported in months, duration in days will be divided by 30.4375. If duration is reported in years, duration in days will be divided by 365.25.

A week is defined as 7 days. If duration is reported in weeks, duration in days is divided by 7.

2.1.3.9 Baseline for the treatment period

For efficacy evaluations, the last non-missing assessment, including unscheduled assessments on or before the date of randomization is taken as baseline value or baseline assessment. This also applies to the evaluation of PRO endpoints.

For safety evaluations and for baseline demographic and background data, the last available assessment, including unscheduled assessments on or before FD is taken as baseline assessment. When the exact time of an assessment is known, the last one before dosing is used.

For pre-dose electrocardiogram (ECG), the last available assessment before the treatment start time is used for baseline. When multiple replicates are available, the average will be used as baseline.

For BCR::ABL1 mutation related analyses, the earliest assessment taken on or before Day 84 is taken as baseline. Day 84 is the mid-point between randomization and Week 24 target day (Day 168, see Table 2-1).

In the rare case where multiple laboratory measurements meet the baseline definition, and no further flag or label can identify the chronological order, and if values are from both the central and the local laboratories, the value from the central assessment will be considered as baseline. Otherwise, if the measurements come from a single laboratory, the average will be used as baseline.

If no measurements meet the above definition, the baseline value will be considered missing. Missing baseline values will not be imputed.

2.1.3.10 On-treatment assessment/event and observation periods

The following three mutually exclusive segments of each participant's overall observation period are defined for adverse event (AE) reporting:

- Pre-treatment period: from the day of the participant's informed consent to the day before his/her FD.
- On-treatment period: from the date of the participant's FD to 30 days after the date of LD (including start and stop date). The 30 days post LD is also referred to as the safety follow-up.
- Post-treatment period: starting at Day 31 after the participant's date of LD. The post-treatment period is also referred to as the survival follow-up. This period ends at approximately 5 years after the date of the LPFT.

Efficacy summaries on the FAS (and respectively FAS_{IMA} and FAS_{2GTKI}), apart from OS, PFS, FFS, and EFS (defined in Section 2.9.1.3), will include data from baseline up to either the last assessment on or before the EOT assessment visit, or before or on treatment failure, whichever is the earliest.

The efficacy assessments, if any, collected post-treatment failure, or post-EOT visit will not be included in any efficacy analyses, except for OS, PFS, FFS, and EFS analyses. However, they will be listed and flagged as appropriate.

2.1.3.11 Windows for multiple assessments

Data such as molecular response collected over time (including unscheduled visits) will be summarized by scheduled time point. As participants do not always adhere to the visit schedule, visits will be remapped according to visit windows defined in Table 2-1 to [redacted] of this document to enable at- or by-visit analysis.

For molecular and cytogenetic responses, only those protocol defined visits will have the visit window defined (Table 2-1and Table 2-2, respectively). Each assessment (including the EOT assessment), either scheduled or unscheduled, will have a mapped visit assigned, if study day is available, according to the defined visit window up to the date with data included.

If more than one molecular assessment are assigned to the same time window, the assessment performed closest to the target date will be used for at- or by-visit analyses. If multiple assessments within a visit window are equidistant from the target date, then the assessment associated with the lowest value is used; if multiple assessments have the same lowest value, the earliest is used. Data from all assessments (scheduled and unscheduled), including multiple assessments, will be listed.

Table 2-1 Time windows for molecular response

Assessment	Target day of assessment	Time Interval
Baseline	1	≤ Day 1#
Week 4	28	Day 2 to Day 42
Week 8	56	Day 43 to Day 70
Week 12	84	Day 71 to Day 126
Week 24	168	Day 127 to Day 210
Week 36	252	Day 211 to Day 294
Week 48	336	Day 295 to Day 378
Week k (k=60,72,84,96, …, EOT)	7 * k	Day (7*k-41) to Day (7*k+42)

Day 1 = Date of randomization
EOT assessments are mapped to the time points as needed.

Table 2-2 Time windows for cytogenetic response (as clinically indicated) and complete hematological response

Assessment	Target day of assessment	Time Interval
Baseline	1	≤ Day 1#
Week 2	14	Day 2 to Day 21
Week 4	28	Day 22 to Day 35
Week 6	42	Day 36 to Day 49
Week 8	56	Day 50 to Day 63
Week 10	70	Day 64 to Day 77
Week 12	84	Day 78 to Day 126
Week 24	168	Day 127 to Day 210
Week 36	252	Day 211 to Day 294
Week 48	336	Day 295 to Day 378

Assessment	Target day of assessment	Time Interval
Week k (k=60,72,84,96, …, EOT)	7 * k	Day (7*k-41) to Day (7*k+42)

Day 1 = Date of randomization
EOT assessments are mapped to the time points as needed.

Since some of the ePRO's are completed by participants at home outside the protocol defined site visits, these assessments will appear as unscheduled in the database. All ePRO data are mapped to timepoints given in Table 2-3 or [redacted] as appropriate. If more than one assessment is available in the same time window, the assessment closest to the planned date will be considered. If two assessments are obtained with the same time difference compared to the scheduled visit day or target assessment day, the assessment obtained prior to the visit or target day will be considered. If multiple assessments are obtained on the same day, only one will be used in analysis. When no assessments are mapped to a timepoint, missing is assumed.

Table 2-3 Time windows for PRO: EORTC QLQ-C30, QLQ-CML24, [redacted]

Assessment	Target day of assessment	Time Interval
Baseline	1	≤ Day 1#
Week 4	28	Day 2 to Day 42
Week 8	56	Day 43 to Day 70
Week 12	84	Day 71 to Day 126
Week 24	168	Day 127 to Day 252
Week 48	336	Day 253 to Day 504
Week 96	672	Day 505 to Day 714
Week j (EOT)	7 * j	Day [1+upper limit of previous interval] to Day 7*j+14
Week k (k=EOT+4, EOT+8, EOT+12)	7 * k	Day (7*k-13) to Day (7*k+14)

Day 1 = Date of randomization
EOT assessments are mapped to the time points as needed.

[redacted]

[redacted]

2.1.3.12 Last contact date

The last contact date will be derived for participants not known to have died at the respective analysis data cut-off date using the last complete date among the following:

Table 2-5 Last contact date data sources

Source data	Conditions
Date of randomization	No condition
Last contact date/last date participant was known to be alive from Survival Follow-up page	Participant status is reported to be alive, lost to follow-up or unknown
Start/End dates from further antineoplastic therapy	Non-missing medication/procedure term
Start/End dates from drug administration record	Non-missing dose. Doses of 0 are allowed
End of treatment date from end of treatment page	No condition
Any specific efficacy (molecular or cytogenetic) assessment date if available	Evaluation is marked as 'done'
Laboratory/PK collection dates	Sample collection marked as 'done'
Vital signs date	At least one non-missing parameter value
Performance status date	Non-missing performance status
Start/End dates of AE	Non-missing verbatim term

The last contact date is defined as the latest complete date from the above list on or before the respective data cut-off date. The cut-off date will not be used for last contact date, unless the participant has been seen or contacted on that date. No date after the cut-off date will be used. Completely imputed dates (e.g., the analysis data cut-off date programmatically imputed to replace the missing end date of a dose administration record) will not be used to derive the last contact date.

The last contact date will be used for censoring participants in the analysis of the time-to-event endpoints.

2.1.3.13 Total follow-up time

The total follow-up for each participant is defined as the time between randomization and the earliest of the following four dates:

1. the date of the data cut-off for each analysis, or the date of EoS for the last analysis, for all on-going participants, or
2. date of WoC, or

3. the time of last contact (cf. Section 2.1.3.12) for those who are no longer in the study but have not WoC and are not known to have died, or
4. date of death.

2.2 Analysis sets

The following analysis sets are all subsets of the All Subjects Set, which consists of all participants who have signed the study informed consent form.

2.2.1 Analysis sets

The analysis sets are defined in Section 12.1 of the CSP. With two minor changes they are reproduced verbatim below for completeness. The two changes are (1) the omission of the two paragraphs on the main and supplement primary and key secondary estimands, and (2) the paragraph on the secondary safety estimand.

Section 12.1 of the CSP:

The **Full Analysis Set** (FAS) comprises all participants to whom study treatment has been assigned by randomization.

The **IMA Full Analysis Set** (FAS_{IMA}) comprises all participants from the FAS, whose pre-randomization selection of TKI is imatinib.

The **2G TKI Full Analysis Set** (FAS_{2GTKI}) comprises all participants from the FAS, whose pre-randomization selection of TKI is a 2G TKI (nilotinib, dasatinib or bosutinib).

According to the intent to treat principle (reflecting the treatment policy estimands approach), participants in the FAS, FAS_{IMA}, FAS_{2GTKI} will be analyzed according to the treatment and stratum they have been assigned to during the randomization procedure.

The **Safety Set** comprises all participants who receive at least one dose of any study treatment. Participants will be analyzed according to the *actual study treatment received*, where actual treatment received is defined as the randomized treatment if the participant took at least one dose of that treatment, or the first treatment received if the randomized treatment was never received.

The **Pharmacokinetic analysis set** (PAS) includes all participants who provide at least one evaluable asciminib PK concentration. For a concentration to be evaluable, participants are required to:

- Take a dose of asciminib prior to sampling,
- Take the same dose of asciminib for at least 3 consecutive days without dose interruption or dose modification prior to sampling,
- For post-dose samples, do not vomit within 4 hours after the dosing of asciminib (this is the current dose); for pre-dose samples do not vomit within 4 hours after the dosing of asciminib prior to sampling (this is the previous dose),
- Have the pre-dose sample collected before the next dose administration.

Other analysis sets

For duration of MMR, the **MMR Responder Set, IMA MMR Responder Set, and 2G TKI MMR Responder Set** will be used, which comprises the subset of participants from the FAS, FAS_{IMA}, or FAS_{2GTKI}, respectively, who achieve MMR at any time.

For duration of MR4.0, the **MR4.0 Responder Set, IMA MR4.0 Responder Set, and 2G TKI MR4.0 Responder Set** will be used, which comprises the subset of participants from the FAS, FAS_{IMA}, or FAS_{2GTKI}, respectively, who achieve MR4.0 at any time.

For duration of MR4.5, the **MR4.5 Responder Set, IMA MR4.5 Responder Set, and 2G TKI MR4.5 Responder Set** will be used, which comprises the subset of participants from the FAS, FAS_{IMA}, or FAS_{2GTKI}, respectively, who achieve MR4.5 at any time.

2.2.2 Participant classification

Participants may be excluded from the analysis populations defined above based on the protocol deviations entered in the database and/or on specific subject classification rules defined in Table 2-6.

Table 2-6 Subject classification based on protocol deviations and non-protocol deviation criteria

Analysis set	Protocol deviations leading to exclusion	Non protocol deviation leading to exclusion
FAS, FAS_{IMA}, FAS_{2GTKI}	No written inform consent	Not applicable
Safety set	No written inform consent	No dose of study medication
PK analysis set (PAS)	No written inform consent	See definition of PAS
MMR, MR4.0, or MR4.5 Responder Sets	No written inform consent	See definition of Responder Set for MMR, MR4.0, or MR4.5, respectively

2.2.3 Withdrawal of Informed Consent

Data collected in the database (via clinical database or third-party vendor data transfer) after a subject withdraws informed consent from all further participation in the trial will not be included in the analysis. The date on which a participant withdraws full consent is recorded in the eCRF. Data records containing confirmed cases of biological sample analyzed after WoC when not allowed per ICF or local regulations will be flagged and excluded from all analyses including listings.

2.2.4 Subgroup of interest

Subgroup analyses will use the same method as for the respective overall analysis sets.

Except for the subgroup analyses for Japan and for China (see Section 2.2.4.3 and Section 2.2.4.4, respectively), the objective for these analyses is to identify any potential patterns, trends, or issues that may be limited to a subgroup of participants. Unless otherwise

specified, summary tables and figures will be generated only for subgroups with at least 15 participants.

2.2.4.1 Subgroup analysis for efficacy endpoints

Subgroup analysis will be conducted for the primary and key secondary efficacy endpoints MMR at Week 48 and MMR at Week 96 based on FAS, FAS_{ima}, and FAS_{2GTKI} to examine the homogeneity of treatment effect, provided that the respective primary or key secondary efficacy analysis based on the FAS or FAS_{IMA} is statistically significant:

- Stratification factor: ELTS (low vs. intermediate vs. high) based on randomization data from IRT
- Stratification factor: ELTS (low vs. intermediate vs. high) based on data collected on the eCRF
- Sex: Female or male
- Race: Asian, Caucasian, or others
- Age groups
  - 18 years ≤ age < 65 years
  - 65 years ≤ age < 75 years
  - age ≥ 75 years

No formal statistical test of hypotheses will be performed for the subgroups, only point estimate of the treatment effect and 95% CI will be provided both in tables and in forest plots.

2.2.4.2 Subgroup analysis for safety endpoints

Subgroup analyses for selected safety endpoints will be defined when required.

2.2.4.3 Subgroup analysis for Japan

For submission to PMDA in Japan, two subgroups will be formed based on geographic regions:

- Japan
- Other regions, i.e., outside Japan

These subgroup analyses will only be used for submission to Japan health authority. Even if the number of participants in each subgroup is less than 15, these subgroup analyses will be conducted.

Summary tables and figures will be presented for the two subgroups for the following outcome measures:

- Baseline characteristics: Tables of demographics, diagnosis and extent of cancer, participant disposition, analysis sets by stratum.
- Exposure: Tables of duration of exposure, dose received.
- Tables of concomitant medications as well as surgical and medical procedures.
- PK (only in asciminib arm): Table and figure of asciminib concentration by time, table of asciminib PK parameters (participants with full PK sampling).

- AEs: Tables of all AEs, treatment-related AEs, AE leading to treatment discontinuation, treatment related AE leading to treatment discontinuation, AEs requiring dose adjustment or interruption, AEs requiring additional therapies, serious adverse events (SAEs), treatment related SAE, adverse events of special interest (AESIs), overview table of AEs and death.
- Safety: dose adjustments and discontinuation of study drug, notable changes in vital signs.
- ECG: Tables of Notable ECG values, change from baseline in ECG parameters values.
- Lab: Hematology shift table, biochemistry shift table based on CTC grades.
- Efficacy: Tables of MMR/MR4.0/MR4.5 rate at and by each time point, CHR at and by each time point, CCyR by Week 48 and by Week 96, time to MMR/MR4.0/MR4.5, duration of MMR/MR4.0/MR4.5, TTF, EFS, FFS, PFS, OS.
- Figures of cumulative incidence of time to MMR/MR4.0/MR4.5.
- Figures of cumulative incidence of time to treatment discontinuation due to AE.

The primary and the key secondary efficacy endpoints (MMR at Week 48 and MMR at Week 96) and MMR at scheduled time points based on FAS_{ima} and FAS_{2GTKI} will also be summarized for Japan subgroup to examine the homogeneity of treatment effect, provided that the respective primary or key secondary efficacy analysis based on the overall comparison in the FAS is statistically significant. All other analyses are based on FAS only.

2.2.4.4 Subgroup Analysis for China

For submission to NMPA in China, the following subgroups will be formed:

- China Mainland: all participants from sites within China (site numbers between 6000 and 6023, inclusive)
- Other regions, i.e., outside China Mainland.
- East Asian: all participants from China, Japan, South Korea, and Taiwan (province of China).

These subgroup analyses will only be used for submission to China health authority.

Summary tables and figures will be presented for the two subgroups for the outcome measures listed for Japan subgroup analysis (see Section 2.2.4.3). In addition, the follow analyses will be provided:

- Baseline characteristics: Framingham CVD risks and CCI
- Efficacy:
  - Odds ratio of MMR rate at 48 weeks adjusted for the stratification factors
  - Summary and change from baseline in EORTC QLQ-C30 total and sub-scores by time points
  - Summary and change from baseline in EORTC QLQ-CML24 total and sub-scores by time points
- Safety:
  - Treatment related AE leading to dose reduction or interruption

As with Japan subgroup analysis, the primary and the key secondary efficacy endpoints (MMR at Week 48 and MMR at Week 96) based on FASima and FAS2GTKI will also be summarized for China subgroups to examine the homogeneity of treatment effect, provided that the respective primary or key secondary efficacy analysis based on the overall comparison in the FAS is statistically significant. All other efficacy analyses are based on FAS only.

2.3 Participant disposition, demographics, and other baseline characteristics

Demographic and other baseline data including disease characteristics will be summarized descriptively by treatment arm for the FAS, FAS_{IMA} and FAS_{2GTKI}, unless otherwise specified. Summaries will be reported by treatment arm, and listings will be reported by treatment arm.

Categorical data will be presented as frequencies and percentages. For continuous data, mean, standard deviation, median, minimum, and maximum will be presented. For selected parameters, 25th and 75th percentiles will also be presented. No inferential statistics will be provided.

2.3.1 Participants disposition

Enrollment by country and center will be summarized for all screened participants and by treatment arm using the FAS, FAS_{IMA} and FAS_{2GTKI}. The number and percentage (number (%)) of randomized participants included in the respective analysis sets will be presented overall and by treatment arm. The number (%) of screened and not-randomized participants and the reasons for screening failure will also be displayed. The eligibility criteria will be also summarized. The number (%) of participants in the FAS who are still on treatment, who discontinued the study phases and the reason for discontinuation will be presented overall and by treatment arm.

The following summaries will be provided (with % based on the total number of FAS, FAS_{IMA} and FAS_{2GTKI} participants, respectively):

- Number (%) of participants who were randomized (based on data from IRT system).
- Number (%) of participants who were randomized but not treated (based on DAR eCRF page).
- Primary reason for not being treated (based on “End of Treatment Phase Disposition” eCRF page).
- Number (%) of participants who were treated (based on DAR eCRF pages of each study treatment completed with non-zero dose administered).
- Number (%) of participants who are still on-treatment (based on the “End of Treatment Phase Disposition” page not completed).
- Number (%) of participants who discontinued the study treatment phase overall, before
- Week 48 and Week 96 (based on the “End of Treatment Phase Disposition” page).
- Primary reason for study treatment phase discontinuation overall, before Week 48 and Week 96 (based on the “End of Treatment Phase Disposition” page).
- Number (%) of participants who have entered the survival follow-up (based on the ‘Subject Status EOT’ page.)

If appropriate, the above disposition tables may also be presented overall, by treatment arm and by treatment arm for each pre-randomization selection of TKI stratum (i.e., by FAS_{IMA} and FAS_{2GTKI}).

2.3.1.1 Protocol deviations

The number (%) of participants in the FAS, FAS_{IMA} and FAS_{2GTKI} with any protocol deviation will be tabulated by deviation category (as specified in the Study Specification Document) overall and by treatment group. All protocol deviations will be listed. In addition, the number (%) of participants in the FAS, FAS_{IMA} and FAS_{2GTKI} respectively, with any pandemic related protocol deviation (pandemic specific protocol deviations, as well as non-specific pandemic protocol deviations with a pandemic-relationship) will be tabulated by deviation category (as specified in the Study Specification Document) overall and by treatment group.

2.3.1.2 Analysis sets

The number (%) of participants in each analysis set will be summarized by treatment arm and strata. Reasons leading to exclusion from analysis sets will be listed by treatment arm and strata, as well as tabulated overall and by treatment arm.

2.3.2 Basic demographic and background data

All demographic and baseline disease characteristics data will be summarized and listed by treatment arm. This includes categorical data such as age groups (see Section 2.2.4.1), sex, race, ethnicity, and ECOG performance status, and continuous data such as age, weight, height, body mass index (BMI), which is calculated as weight[kg] / (height[m]^2). Reasons for pre-randomization TKI selection will also be summarized.

2.3.2.1 Reporting for DSUR and PSUR

Two summary tables using the Safety Set will be produced for DSUR and PSUR reporting:

1. Per each treatment group: the number and percent of participants in the different age groups, and in each combination of sex-by-age group
2. Per each treatment group: the number and percent of participants in the different race groups

2.3.3 Baseline stratification factors

The number (%) of participants in each stratum (i.e., ELTS and PRS-TKI) based on data obtained from the IRT system will be summarized overall and by treatment arm for the FAS. Discordances between the stratum recorded in IRT at the time of randomization and the stratum recorded in the clinical database of the actual TKI a participant receives, i.e., imatinib vs. 2G TKI, through the data collected on eCRF will be cross-tabulated and listed. Such cross-table will also be produced for ELTS risk groups recorded in the IRT system and that in the eCRF.

2.3.4 Diagnosis and extent of cancer

All diagnosis and extent of cancer data will be summarized and listed by treatment arm. One summary table will include time (weeks) since initial diagnosis to the time of randomization. Another summary table with frequency (counts and percentages) of participants with extramedullary involvement: any extramedullary involvement (Yes/No) and location of extramedullary involvement (e.g.: Spleen, Liver) will be presented.

2.3.5 Medical history

Medical history and ongoing conditions, including cancer-related conditions and symptoms entered on eCRF will be summarized and listed by treatment arm. The summary will be presented by primary system organ class (SOC), preferred term (PT) and treatment arm. Medical history and current medical conditions will be coded using the Medical Dictionary for Regulatory Activities (MedDRA) terminology. The MedDRA version used for reporting will be specified in the CSR and as a footnote in the applicable tables/listings.

The cardiovascular risk factors (e.g.: heavy smoking, low physical activity, unhealthy diet, etc.), and family medical history of each participant (e.g., for ischemic heart disease, cardiac arrhythmia, sudden death, high cholesterol, diabetes mellitus, heart defects (congenital heart disease), heart failure etc.) are collected prior to randomization. A listing by treatment arm will be presented.

2.3.6 Other

All data collected at baseline, including informed consent for additional research on study data and biological samples will be listed.

2.4 Treatments (study treatment, rescue medication, concomitant therapies, compliance)

The safety analysis set will be used for the analyses related to study treatment and compliance during the *randomized treatment period*. The randomized treatment period is defined as the period between the date of FD and the date of LD, inclusively.

2.4.1 Study treatment and compliance

2.4.1.1 Duration of exposure to study treatment

Duration of exposure to a study treatment (in days) for participant *i* is defined as

$$date\ of\ LD_i - date\ of\ FD_i + 1$$

Summary of duration of exposure to study treatment include descriptive statistics. These summary statistics will be reported by the following time intervals:

- < 24 weeks
- ≥ 24 weeks
- ≥ 48 weeks
- ≥ 96 weeks

and every 48 weeks thereafter until the EoS.

The *duration of exposure in patient-years* for a treatment arm *k* is the total of the duration of exposure in years from all the participants in that treatment arm, i.e.

$$\sum_{i=1}^{n_k}(date\ of\ LD_i - date\ of\ FD_i + 1)/365.25$$

where $k = 1$ or 2 representing the asciminib or the control arm of IS-TKI, respectively; n_k is the number of participants in treatment *k*.

The duration of exposure in patient-years will be summarized for the two treatment arms using descriptive statistics.

2.4.1.2 Actual cumulative dose

The *actual cumulative dose* refers to the total actual dose administered over the duration for which a participant is on the study treatment as documented in the DAR eCRF. It is the sum of the non-zero total daily doses recorded over the dosing period. For participants who do not take any drug, the actual cumulative dose is by definition equal to zero.

Additionally, the *planned cumulative dose* for a participant is defined as the total planned dose per the protocol up to the date of LD for this participant. The calculations for the study treatments are:

- Asciminib: 80 mg/day × duration of exposure (day)
- Bosutinib: 400 mg/day × duration of exposure prior to first dose escalation (day) + 500 mg/day × duration of exposure since first dose escalation (day) + 600 mg/day × duration of exposure since second dose escalation (day), where the starting day of dose escalation is identified as the first record in the DAR eCRF with each dose increased and reason "As per protocol"
- Dasatinib: 100 mg/day × duration of exposure prior to dose escalation (day) + 140 mg/day × duration of exposure since dose escalation (day) where the starting day of dose escalation is identified as the first record in the DAR eCRF with dose increased and reason "As per protocol"
- Imatinib: 400 mg/day × duration of exposure prior to first dose escalation (day) + 600 mg/day × duration of exposure since dose escalation (day) where the starting day of dose escalation is identified as the first record in the DAR eCRF with dose increased and reason "As per protocol"
- Nilotinib: 300 mg/administration × 2 × duration of exposure (day)

The actual cumulative dose will be summarized using descriptive statistics for each of the TKI treatment separately. The planned cumulative dose is only used to define relative dose intensity (see below).

2.4.1.3 Average daily dose

Average daily dose (mg/day) for a participant is defined as:

actual cumulative dose (mg) / number of days on treatment,

where *number of days on treatment* is

(date of LD – date of FD + 1) – number of days with dose interruptions*.

* For participants treated with nilotinib, this includes the half day before or the half day after the period of at least one full day with 0 dose if the treatment is interrupted after the morning dose or resumed in the evening.

The average daily dose will be summarized by descriptive statistics for each of the study treatment separately.

2.4.1.4 Dose intensity and relative dose intensity

Dose intensity (mg/day) for a participant is defined as:

actual cumulative dose (mg) / duration of exposure (day).

For participants who do not take any drug, the dose intensity is zero.

Relative dose intensity is defined as follows:

dose intensity / planned dose intensity,

where *planned dose intensity* (mg/day) is

planned cumulative dose (mg) / duration of exposure (day).

The dose intensity and the relative dose intensity will be summarized using summary statistics separately for each of the study treatment.

2.4.1.5 Dose changes, interruptions or permanent discontinuations

2.4.1.5.1 Dose changes

A dose change occurs when total daily dose is different from the most recently planned dose. For participants in asciminib arm, there is only one planned dose, i.e., 80 mg/day. For participants in the comparator arm, the planned doses are:

- For imatinib, the initial planned dose is 400 mg/day and could be changed to 600 mg/day
- For nilotinib, the initial planned dose is 600 mg/day (300 mg b.i.d.) and no dose escalation is allowed.
- For dasatinib, the initial planned dose is 100 mg/day and could be changed to 140 mg/day
- For bosutinib, the initial planned dose is 400 mg/day and could be changed to 500 mg/day and subsequently to 600mg/day.

The field “Dose changed” from the eCRF DAR page will be used to determine whether a dose change occurs for a participant. If a dose change occurs, the following are used to determine whether it is a dose reduction or a dose increase:

Reduction: A dose change where the actual total daily dose is lower than the most recently planned dose. However, any dose change to correct a dosing error will not be considered a dose

reduction. Since only dose change is collected in the eCRF, the number of reductions will be derived programmatically based on the direction of the change.

Increase for treatments where dose escalation is allowed: dose change where the actual total daily dose is greater than the most recently planned dose. However, any dose change to correct a dosing error will not be considered a dose increase. Since only dose change is collected in the eCRF, the number of increases will be derived programmatically based on the direction of the change.

The number and percentage of participants with dose changes and the reasons for the changes will be summarized by study treatment arms and separately for each of the study treatments. Since a participant can have multiple dose changes, the frequency of reasons for changes can be higher than the number of participants experiencing them. Additionally, participant level listings will be produced.

2.4.1.5.2 Dose interruptions and duration of interruptions

The field "Dose Interrupted" in the eCRF DAR page will be used to determine whether a dose interruption occurs for a participant. When multiple entries for interruptions are entered on consecutive days with different reasons, they will be counted as separate interruptions. However, if the reason is the same in the multiple entries on consecutive days, then it will be counted as one interruption.

Duration of dose interruption (days) due to any reason will be summarized descriptively. For each participant, the duration of dose interruption will be calculated by adding days of all individual episodes of dose interruption for that participant.

The number and percentage of participants who experience dose interruptions, the reasons for the interruptions, and descriptive statistics for duration of dose interruptions will be summarized for each treatment arm and separately for each study treatment. Since a participant can have multiple dose interruptions, the frequency of reasons may be higher than the number of participants experiencing them. Additionally, participant level listings will be produced.

2.4.1.5.3 Permanent discontinuation

The field "Dose Permanently Discontinued" from the eCRF DAR page will be used to determine whether permanent discontinuation occurs for a participant.

Number and percentage of participants who discontinue and reasons for discontinuation will be summarized by the treatment arms and separately for each study treatment. Additionally, participant level listings will be produced.

2.4.2 Prior, concomitant and post therapies

Concomitant medications and significant non-drug therapies prior to and after the start of the study treatment will be listed and summarized according to the Anatomical Therapeutic Chemical (ATC) classification system, by treatment arms.

Prior anti-cancer therapy

The number and percentage of participants who received any prior anti-neoplastic medications will be summarized by treatment arm for the lowest anatomical therapeutic classification (ATC) class and preferred term. A listing will also be produced.

Anti-neoplastic medications will be coded using the WHO Drug Dictionary (WHO-DD). Details regarding WHO-DD version will be included in the footnote in the tables/listings.

The above analyses will be performed using the FAS.

The following information will be summarized for the FAS:

- Number (%) of participants that received prior treatment for CML (Imatinib or 2G TKI for ≤ 2 weeks, Hydroxyurea and/or anagrelide)
- Time on prior treatment for CML (in weeks).

Post treatment anti-cancer therapy

Anti-neoplastic therapies since discontinuation of study treatment will be listed and summarized by the lowest anatomical therapeutic classification (ATC) class, preferred term, overall and by treatment group by means of frequency counts and percentages using FAS.

Anti-neoplastic medications will be coded using the WHO-DD. Details regarding WHO-DD version will be included in the footnote in the tables/listings.

Concomitant therapies

Concomitant therapies are defined as all interventions (therapeutic treatments and procedures) other than the study treatment administered to a participant coinciding with the study treatment period. Concomitant therapies include medications (other than study drugs) and medical procedures starting on or after the start date of study treatment, or starting prior to the start date of study treatment and continuing after the start date of study treatment.

Concomitant medications will be coded using the World Health Organization (WHO) Drug Reference Listing (DRL) dictionary that employs the WHO ATC classification system and summarized by the lowest ATC class and PT using frequency counts and percentages. Surgical and medical procedures will be coded using MedDRA and summarized by SOC and PT.

The summaries for the on-treatment period using the Safety Set will include:

- Therapies starting on or after the start of randomized study treatment but no later than the end of the on-treatment period and
- Therapies starting prior to start of randomized study treatment and continuing after the start of randomized study treatment.

All concomitant therapies will be listed using the Safety Set. Any concomitant therapies starting and ending prior to the start of randomized study treatment or starting beyond end of the on-treatment period will be flagged in the listing.

The prohibited concomitant medications will be summarized by lowest ATC class and preferred term up to the end of on-treatment periods, respectively.

2.5 Analysis supporting the primary objectives

2.5.1 Primary endpoint

The primary endpoint for the two primary objectives (see Table 1-1) of the study is defined as the binary outcome (Yes/No) of whether a participant is in MMR at Week 48**.**

Only participants in MMR at Week 48 visit are considered responders. In other words, any participant who is in MMR before Week 48 visit, but is no longer in MMR at Week 48 visit, will be considered non-responder for the primary analyses.

A participant is considered to have met the primary endpoint if the result of *BCR::ABL1* analysis from the Novartis designated laboratory by RQ-PCR meets the MMR criteria (*BCR::ABL1* levels (IS) $\leq 0.1\%$) at Week 48. Participants discontinuing the randomized treatment due to any reason prior to Week 48, or participants meeting any treatment failure criteria [as per ELN 2020 criteria, (Hochhaus et al., 2020)] prior to Week 48, will be considered as not in MMR at Week 48.

2.5.2 Statistical hypothesis, model, and method of analysis

The analysis of the primary endpoint associated with the first primary objective will be performed using the FAS; and for the second primary objective using the FAS_{IMA}. Following the intent to treat principle, participants will be analyzed according to the treatment arm and stratum they are assigned to at randomization.

The null and the alternative hypotheses associated with the first primary objective are:

$$H_{10}: \pi_{a48} \leq \pi_{c48} \text{ vs.}$$

$$H_{1a}: \pi_{a48} > \pi_{c48}$$

where π_{a48} and π_{c48} are proportions of participants in MMR at Week 48 in asciminib treatment arm and in the control arm of IS-TKI, respectively.

The null hypothesis will be tested using a one-sided Cochran-Mantel-Haenszel (CMH) Chi-square test stratified by both randomization stratification factors ELTS risk score and PRS-TKI as recorded in the IRT system.

The null and the alternative hypotheses associated with the second primary objective are:

$$H_{20}: \pi_{a,ima48} \leq \pi_{c,ima48} \text{ vs.}$$

$$H_{2a}: \pi_{a,ima48} > \pi_{c,ima48}$$

where $\pi_{a,ima48}$ and $\pi_{c,ima48}$ are proportions of participants in MMR at Week 48 in the stratum of participants who have imatinib as their PRS-TKI (as recorded in the IRT system) in asciminib treatment arm and in the control arm of IS-TKI, respectively.

The null hypothesis will be tested using a one-sided CMH Chi-square test stratified by the randomization stratification factor ELTS risk score as recorded in the IRT system.

The family-wise type I error rate will be controlled at 2.5% level via the graphical gatekeeping procedure as described in Section 2.8 of the current document. The null hypotheses will be

rejected if the one-sided p-value is less than the nominal alpha-level as specified. The study will be considered positive if the null hypothesis for either of the two primary objectives is rejected.

Additionally, the point estimates and their corresponding 95% CI for π_{a48}, π_{c48}, $\pi_{a,ima48}$, and $\pi_{c,ima48}$using Pearson-Clopper method (Clopper and Pearson, 1934) will be presented.

And lastly, the stratified Mantel-Haenszel estimate of the common risk difference between π_{a48} and π_{c48} , and between $\pi_{a,ima48}$ and $\pi_{c,ima48}$ will be provided together with their corresponding two sided 95% CIs.

All the CIs reported will be at nominal 95% confidence level and are unadjusted for multiple testing. They are presented for descriptive purposes.

2.5.3 Handling of intercurrent events

The following defines how the IE defined in Section 1.2.2 are handled:

- Taking a TKI different from their PRS-TKI (i.e., taking imatinib as IS-TKI when a 2G has been the pre-randomization selected TKI and vice versa): treatment policy strategy
- Change on study treatment per protocol (dose reduction/interruption/allowed dose escalations): treatment policy strategy
- Dosing errors (e.g., missed dose): treatment policy strategy
- Deviation in any intake of concomitant medications: treatment policy strategy
- Intake of prohibited medications: treatment policy strategy
- Meeting any treatment failure criteria prior to Week 48 or treatment discontinuation due to any reason prior to Week 48: composite
- Handling of remaining IEs: Other IEs, if occur, will be handled by the treatment policy approach.

2.5.4 Handling of missing values/censoring/discontinuations not related to intercurrent events

2.5.4.1 Handling of missing values

Participants with missing PCR evaluation at Week 48 visit will be imputed as MMR responders if they have non-missing PCR evaluations at both Week 36 and Week 60 visits, and both meet the MMR criteria (*BCR::ABL1* levels (IS) ≤ 0.1%), assuming that MMR is maintained between Week 36 and Week 60. If PCR evaluations are performed at unscheduled visits closer to the Week 48 visit (before or after), these will be taken into account for the imputation.

Otherwise, participants with missing PCR evaluations at Week 48 will be considered as non-responders.

2.5.5 Sensitivity analysis

The CMH Chi-square test of MMR rate at 48 Weeks will be repeated without the imputation rule used in the main primary analyses for participants who have missing PCR evaluations at

48 weeks. In such cases these participants are considered non-responders for the endpoint MMR at Week 48.

2.5.6 Supplementary analysis

The following supplementary analysis will be performed using FAS:

- Supplemental estimands for the endpoint of MMR at Week 48 based on the treatment policy, composite and hypothetical strategies will be produced. The supplemental estimand differs from the primary estimand on how it handles the intercurrent event of *taking a TKI different from the PRS-TKI in the comparator arm* and of misclassification of ELTS risk groups at randomization. These estimands will be presented for descriptive purposes. In other words, for this supplementary analysis, participants will be analyzed using the PRS-TKI and ELTS risk group as recorded in the baseline eCRF.
- One-sided stratified CMH test based on the ELTS and PRS-TKI strata in the IRT system is repeated (as for the primary estimand) after excluding participant(s) who (1) should have been screen failure(s) but randomized by mistake, (2) have atypical transcript at baseline.

2.5.7 Supportive analysis

If either of the primary objectives is met, subgroup analyses will be conducted for subgroups defined in Section 2.2.4.1.

2.6 Analysis supporting the key secondary objectives

2.6.1 Key secondary endpoint

The key secondary endpoint for the two key secondary objectives (see Table 1-1) of the study is defined as the binary outcome (Yes/No) of whether a participant is in MMR at Week 96. The testing of the key secondary endpoints is only conducted if the null hypothesis associated with the second primary objective, i.e., H_{20}, is rejected. Details on the sequence of testing are described in Section 2.8.

The determination of whether a participant is in MMR at Week 96 is identical to that for Week 48, only that the RQ-PCR value at week 96 is used.

2.6.2 Statistical hypothesis, model, and method of analysis

The null and alternative hypotheses for the first key secondary endpoint are

$$H_{30}: \pi_{a96} \leq \pi_{c96} \text{ vs.}$$

$$H_{3a}: \pi_{a96} > \pi_{c96}$$

where π_{a96} and π_{c96} are proportions of participants in MMR at Week 96 in asciminib treatment arm and in the control arm of IS-TKI, respectively.

The null and alternative hypotheses for the second key secondary endpoint are

$$H_{40}: \pi_{a,ima96} \leq \pi_{c,ima96} \text{ vs.}$$

$$H_{4a}: \pi_{a,ima96} > \pi_{c,ima96}$$

where $\pi_{a,ima96}$ and $\pi_{c,ima96}$ are proportions of participants in MMR at Week 96 in the stratum of participants who have imatinib as their PRS-TKI (as recorded in the IRT system) in asciminib treatment arm and in the control arm of IS-TKI, respectively.

The statistical models and methods of analysis for the key secondary endpoints are identical to that of the primary endpoints, only with MMR at Week 96 instead of Week 48.

2.6.3 Handling of intercurrent events

The handling of IEs for the key secondary endpoints are identical to that of primary endpoints, only with MMR at Week 96 instead of Week 48.

2.6.4 Handling of missing values/censoring/discontinuations not related to intercurrent events

2.6.4.1 Handling of missing values

The handling of missing MMR at Week 96 values are similar to that for MMR at Week 48, only that Week 84 and Week 108 PCR assessments are used instead of Week 36 and Week 60, respectively.

2.6.5 Sensitivity analysis

The CMH Chi-square test of MMR rate at 96 Weeks will be repeated without the imputation rule used in the main analyses for participants who have missing PCR evaluations at 96 weeks. In such cases these participants are considered non-responders for the endpoint MMR at Week 96.

2.6.6 Supplementary analysis

The following supplementary analysis will be performed using FAS:

- Supplemental estimands for the endpoint of MMR at Week 96 based on the treatment policy, composite and hypothetical strategies will be produced. The supplemental estimand differs from the primary estimand on how it handles the intercurrent event of *taking a TKI different from the PRS-TKI in the comparator arm*. These estimands will be presented for descriptive purposes. In other words, for this supplementary analysis, participants will be analyzed using the PRS-TKI as recorded in the baseline eCRF.
- One-sided stratified CMH test based on the ELTS and PRS-TKI strata in the IRT system is repeated (as for the primary estimand) after excluding participant(s) who (1) should have been screen failure(s) but randomized by mistake, (2) have atypical transcript at baseline.
- As specified in Section 2.2.4.1, subgroup analysis will be conducted if primary objectives are met.

2.7 Analysis supporting secondary safety endpoint objective

2.7.1 Secondary safety endpoint TTDAE

The study has one secondary safety objective to be tested (see Table 1-1). The associated endpoint, TTDAE, will be formally tested at the Week 96 analysis timepoint as described in the testing strategy in Section 2.8. In particular, the testing for the secondary safety endpoint is only conducted if the null hypotheses associated with the second primary objective and the key secondary objects are rejected.

The TTDAE is defined for each participant as the time from the date of FD to the date of discontinuation of study treatment due to AE, including death due to AE. For participants ongoing without study treatment discontinuation on or prior to the analysis cut-off date, the TTDAE for this participant will be censored at the analysis cut-off date. Discontinuation due to other reasons, including death due to other reasons, are considered a competing risk event.

Note that for those who discontinue from the study treatment, the reason for discontinuation will be taken from the EOT disposition page, whereas the date of discontinuation is the date of LD from the DAR records.

Other safety endpoints for which descriptive statistics and listings are provided are discussed in Section 2.11 of the current document.

2.7.2 Statistical hypothesis, model, and method of analysis

The analyses of TTDAE will be performed using the Safety Set. The comparison is between all participants receiving asciminib as their actual treatment and all participants receiving the 2G TKIs as their actual treatment. See Section 2.2.1 for definition of *actual treatment received* under Safety Set.

The hypotheses corresponding to the secondary objective for TTDAE is as follows:

- $\mathbf{H_{50}}$: the cause-specific hazard for the event of discontinuation of study treatment due to AE for participants that received asciminib is greater than or equal to that for participants that received a 2G TKI.
- $\mathbf{H_{5a}}$: the cause-specific hazard for the event of discontinuation of study treatment due to AE for participants that received asciminib is less than that for participants that received a 2G TKI.

The formal comparison of the cause-specific hazard for the event of interest will be implemented via the stratified log-rank test. The stratification factor is the ELTS category.

The family-wise type I error rate will be controlled at 2.5% level via the graphical gatekeeping procedure which is described in Section 2.8.

The cumulative incidence curve for the event of interest and the competing risk event will be plotted. The estimated cumulative incidence rates and 95% CI at specified scheduled visits will be presented for each treatment group (asciminib and the 2G TKI).

The cause-specific hazard ratio for TTDAE as well as for the competing event along with their 95% CIs will also be estimated from the cause-specific Cox regression model.

All CIs reported will be at a nominal 95% confidence level, and will be unadjusted for multiple testing. They will be presented for descriptive purposes.

2.7.3 Handling of intercurrent events

The IEs for the secondary safety objective are defined in Section 1.2.3. Below defines how they are handled:

- Discontinuation from study treatment due to other reasons: while on treatment strategy
- Change on study treatment per protocol (dose reduction, interruption, or allowed dose escalations): treatment policy strategy
- Dosing errors (e.g., missed dose): treatment policy strategy
- Deviation in any intake of concomitant medications: treatment policy strategy
- Intake of prohibited medications: treatment policy strategy
- Handling of remaining IEs: Other IEs, if occur, will be handled using the treatment policy strategy.

2.7.4 Handling of missing values/censoring/discontinuations

For participants ongoing without study treatment discontinuation on or prior to the analysis cut-off date, their TTDAE time will be censored at the at the analysis cut-off date.

2.7.5 Supplementary analysis

The competing risk analysis via the sub-distribution hazard approach will be performed. These supplementary analyses will be provided for information purpose only.

In this analysis, discontinuation due to AE (including death due to AE) is the event of interests, and discontinuation due to other reasons is the competing risks. Patients who are still ongoing at the time of the analysis are considered censored.

2.8 Testing strategy and type I error control

The testing strategy has been discussed in detail in the Section 12.7 of the CSP. For completeness it is reproduced verbatim below except for the re-numbering of Figure 2-1. Section numbers in this section refers to those in the CSP.

CSP Section 12.7: Testing Strategy and Type-I Error Control for the multiple primary and key secondary objectives; and for the secondary safety objective

The overall family wise type-I error (1-sided level of significance $\alpha = 2.5\%$) control for testing the multiple primary and key secondary hypotheses, is achieved through the graphical gate-keeping multiple testing procedure (Bretz et al 2009, Bretz et al 2011) as shown in Figure 2-1.

The multiple hypotheses are grouped into three families; those associated with the primary objectives (H1, H2), those associated with the key secondary objectives (H3, H4) and those associated with the secondary safety objective (H5).

The **primary endpoint (MMR at Week 48) family of hypotheses F_1**, are:

- **$H1_0$**: the proportion of participants that achieve MMR at Week 48 in the asciminib arm is less than or equal to the proportion of participants that achieve MMR at Week 48 in the Investigator selected TKI arm.

versus

- **$H1_a$**: the proportion of participants that achieve MMR at Week 48 in the asciminib arm is greater than the proportion of participants that achieve MMR at Week 48 in the Investigator selected TKI arm.

and

- **$H2_0$**: within the stratum of participants that have imatinib as their pre-randomization selection of TKI, the proportion of participants that achieve MMR at Week 48 in the asciminib arm is less than or equal to the proportion of participants that achieve MMR at Week 48 in the investigator selected TKI arm.

versus

- **$H2_a$**: within the stratum of participants that have imatinib as their pre-randomization selection of TKI, the proportion of participants that achieve MMR at Week 48 in the asciminib arm is greater than the proportion of participants that achieve MMR at Week 48 in the investigator selected TKI arm.

The **key secondary endpoint (MMR at Week 96) family of hypotheses F_2**, are:

- **$H3_0$**: the proportion of participants that achieve MMR at Week 96 in the asciminib arm is less than or equal to the proportion of participants that achieve MMR at Week 96 in the Investigator selected TKI arm.

versus

- **$H3_a$**: the proportion of participants that achieve MMR at Week 96 in the asciminib arm is greater than the proportion of participants that achieve MMR at Week 96 in the Investigator selected TKI arm.

and

- **$H4_0$**: within the stratum of participants that have imatinib as their pre-randomization selection of TKI, the proportion of participants that achieve MMR at Week 96 in the asciminib arm is less than or equal to the proportion of participants that achieve MMR at Week 96 in the Investigator selected TKI arm.

versus

- **$H4_a$**: within the stratum of participants that have imatinib as their pre-randomization selection of TKI, the proportion of participants that achieve MMR at Week 96 in the asciminib arm is greater than the proportion of participants that achieve MMR at Week 96 in the Investigator selected TKI arm.

The secondary safety endpoint (TTDAE) family of hypotheses F3 is:

- **$H5_0$**: the cause-specific hazard for the event of discontinuation of study treatment due to AE, for participants that received asciminib is greater than or equal to that for participants that received a 2G TKI.

versus

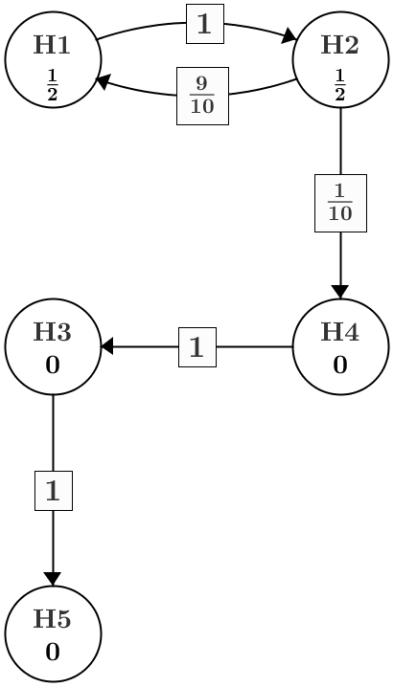
- **$H5_a$**: the cause-specific hazard for the event of discontinuation of study treatment due to AE, for participants that received asciminib is less than that for participants that received a 2G TKI.

The analyses cut-off date for testing H1 and H2 will be when all randomized participants have been treated for at least 48 weeks or discontinued from study treatment prior to Week 48.

The analyses cut-off date for testing H3 and H4 will be when all randomized participants have been treated for at least 96 weeks or discontinued from study treatment prior to Week 96.

The analyses cut-off date for testing H5 will be when all randomized participants have been treated for at least 96 weeks or discontinued from study treatment prior to Week 96.

Figure 2-1 Testing hierarchy (graphical gatekeeping procedure)



The primary and key secondary hypotheses, H1, H2, H3, H4 and H54, are represented by nodes in the graph; with the associated local weights (i.e., proportion of total alpha (1- sided α=2.5%) available to test that hypothesis). A directed edge from one hypothesis Hi, to another one Hj, means that when the null hypothesis associated with Hi is rejected, its local weight (i.e., the local alpha) can be transferred to Hj. The weight associated with that edge quantifies the proportion of the local alpha of Hi that is transferred to Hj.

The three families of hypotheses will be tested in the following manner. The null hypotheses in family F_1 (H1 and H2) will be examined first and tested using the weighted parametric tests (Bretz et al 2011); the hypotheses in family F_2 (H3 and H4) will be tested using the fixed-sequence testing approach; followed by the hypothesis in family F_3 (H5) will be tested using

the fixed-sequence testing approach. The null hypotheses in key secondary end-point family F_2 can be tested if the null hypothesis for H2 (or for both H1 and H2) in the primary end-point family F_1 is rejected. The null hypotheses in secondary safety end-point family F_3 can be tested if the null hypothesis for both H4 and H3 in the key secondary end-point family F_2 are rejected.

The weighted parametric test for the primary end-points family also exploits the correlation between the test statistics for H1 and H2, which arises because the population for testing H2 is a stratum from within the population for testing H1. Let T1 and T2 be the test statistics for H1 and H2 respectively, then the correlation is computed as shown in the figure below:

where n_1, n_1* are the number of participants in asciminib arm, overall and within the stratum of participants with imatinib as their pre-randomization selection of TKI; and n_0, n_0* are the number of participants in investigator selected TKI arm, overall and within the stratum of participants with imatinib as their pre-randomization selection of TKI, respectively.

Let $\alpha_1 = w_1*\alpha$ and $\alpha_2 = w_2*\alpha$ be primary end-point specific alphas for H1 and H2 respectively, let α_3 and the α_4 be key secondary end-point specific alphas for H3 and H4 respectively, and the α_5 be secondary safety end-point specific alpha for H5. To begin with, we distribute the overall global alpha level (1-sided α=0.025) such that the two primary hypotheses get equal weights (i.e., w_1=0.5 and w_2=0.5), as each of the multiple primary objectives is considered equally important. In addition, to begin with, the two key secondary hypotheses are not given any weight since achieving these objectives is not relevant if we cannot meet the primary objective H2, or meet both H1 and H2. Similarly, the secondary hypothesis H5 is also not given any weight to begin with.

Within the primary end-point family F_1:

- The testing of H1 and H2 can be performed simultaneously at the split levels of $\alpha_1 = w_1*\alpha$ and $\alpha_2 = w_2*\alpha$, respectively.
  - When using the correlation between the test statistics for H1 and H2, the test of the intersection hypothesis H1 and H2 would be done at levels that are slightly larger than w1*alpha and w2*alpha (for details please refer to Bretz et al 2011).
- If $H1_0$ is rejected at an alpha of α_1, then this w_1 is passed on to test H2 which can then be tested at the alpha of $w_1*\alpha+w_2*\alpha = \alpha$ (i.e., H2 can be tested at the full alpha).
- If $H2_0$ is rejected at an alpha of α_2, then 0.9 of w_2 is passed onto testing H1, which can then be tested at an alpha of $(w_1 + 0.9*w_2)*\alpha$

Within the key secondary end-point family F_2:

- If only $H2_0$ is rejected in the primary family, then H4 can be tested at an alpha of $\alpha_4 = 0.1*w_2*\alpha$
- If both $H1_0$ and $H2_0$ are rejected in the primary family, then H4 can be tested at the full alpha of α_4=0.025.
- If $H4_0$ is rejected at an alpha of α_4, then this α_4 is passed on to test H3
  - if only $H2_0$ from the primary family and $H4_0$ from the key secondary family are rejected, then H3 is tested at an alpha level $\alpha_3=0.1*w_2*\alpha$
  - if both $H1_0$ and $H2_0$ from the primary family, and $H4_0$ from the key secondary family are rejected, then H3 is tested at the full alpha level α_3 =0.025.

Within the secondary safety end-point family F_3:

- If both $H4_0$ and $H3_0$ are rejected in the key secondary family, then H5 can be tested at the full alpha of α_5=0.025.

The raw p-value rejection boundaries:

Let p_1 and the p_2 be the raw p-values for H1 and H2 respectively; let p_3 and the p_4 be the raw p-values for H3 and H4 respectively and let p_5 be the raw p-values for H5.

Assuming that the number of participants enrolled into the imatinib stratum based on pre-randomization selection of TKI is 50% of the total number of participants enrolled into the study, **the raw p-value rejection boundaries for the primary end-point family F1 are:**

H1 and H2 are simultaneously tested, each at local alpha= 0.01469289 (which is greater than 0.025/2 due to the positively correlated test statistics with a correlation of sqrt (0.5)).

- If $p_1 > 0.01469289$ and $p_2 > 0.01469289$ then $H1_0$ is not rejected and $H2_0$ is not rejected, and testing stops
- If $p_1 \leq 0.01469289$ and $p_2 \leq 0.01469289$ then $H1_0$ is rejected and $H2_0$ is rejected
- if $p_1 \leq 0.01469289$ and $p_2 > 0.01469289$ then $H1_0$ is rejected and H2 can be retested at the local alpha level of 0.025. If $0.01469289 < p_2 \leq 0.025$ then we can also reject $H2_0$
- If $p_1 > 0.01469289$ and $p_2 \leq 0.01469289$ then $H2_0$ is rejected and H1 can be retested at the local alpha level of 0.02375 = (1+0.9)*0.0125. If $0.01469289 < p_1 \leq 0.02375$ then we can also reject $H1_0$

The raw p-value rejection boundaries for the key secondary end-point family F2:

- If both $H1_0$ and $H2_0$ are rejected in the primary family, then H4 can be tested at the full alpha of 0.025. If $p_4 \leq 0.025$ then we can reject $H4_0$. If $H4_0$ is rejected and $p_3 \leq 0.025$, then we can reject $H3_0$.
- If only $H2_0$ is rejected in the primary family, then H4 will be tested at an alpha of 0.00125. If $p_4 \leq 0.00125$ then we can also reject $H4_0$. If $H4_0$ is rejected and $p_3 \leq 0.00125$, then we can reject $H3_0$.
- If only $H1_0$ is rejected in the primary family, or none of the null hypotheses in the primary family are rejected, then the secondary family cannot be tested.

The raw p-value rejection boundaries for the secondary safety end-point family F3:

If both $H4_0$ and $H3_0$ are rejected in the key secondary family, then H5 can be tested at the full alpha of 0.025. If $p_5 \leq 0.025$ then we can reject $H5_0$.

Equivalently, the raw p-values from the Mantel-Haenszel tests and from the log-rank test will be adjusted for the multiple testing (alpha control as described in the above graphical gatekeeping procedure), to obtain the adjusted p-values, using the "gMCP" package (Rohmeyer K, Klinglmueller F (2020)) in R. The adjusted p-values will be used to perform the 1-sided 2.5% level test for each of the hypotheses. i.e., if the *adjusted p-value* for a hypothesis is ≤0.025 we will reject the corresponding null hypothesis. The adjusted p-values for the primary end-points family will be computed as per the weighted parametric test using the correlation estimated from the observed study data.

2.9 Analysis supporting the other secondary efficacy objectives

In this section the analyses of other secondary efficacy endpoints as given in Table 1-2 are described. No confirmatory statistical testing will be performed; however, nominal p-values may be presented for descriptive purposes.

2.9.1 Other secondary endpoints for efficacy

Other secondary efficacy endpoints include response endpoints and time-to-event endpoints.

2.9.1.1 Response endpoints

Efficacy response endpoints are:

- **MMR rates at all scheduled data collection time points**, i.e., at all the protocol-planned visits (except for 48 Weeks and 96 Weeks that are already covered by primary and key secondary endpoints). Such rates are defined as the proportion of participants with MMR at the respective time points.
- **MMR rates by all scheduled data collection time points**, i.e., the protocol-planned visits. These are cumulative MMR rates by time points and are defined as the proportion of participants who achieve MMR *at or before* specified visits, i.e., if a participant achieves an MMR but then loses it before or at a specific visit, he/she will still be classed as achieving MMR by that specific time point. Baseline values are excluded from this analysis.
- **BCR::ABL1≤1%, MR4.0, MR4.5, or CHR rates at all scheduled data collection time points**, i.e., at all the protocol-planned visits. Such rates are defined as the proportion of participants with BCR::ABL≤1%, MR4.0, MR4.5, or CHR, respectively, at the respective time points.
- **BCR::ABL1≤1%, MR4.0, MR4.5, or CHR rates by all scheduled data collection time points**, i.e., the protocol-planned visits. These are cumulative BCR::ABL≤1%, MR4.0, MR4.5, or CHR rate by time points and are defined as the proportion of participants who achieve BCR::ABL≤1%, MR4.0, MR4.5, or CHR, respectively, *at or before* specified visits, i.e. if a participant achieves BCR::ABL≤1%, MR4.0, MR4.5, or CHR, but then loses it before or at a specific visit, he/she will still be classed as achieving BCR::ABL≤1%, MR4.0, MR4.5, or CHR, respectively, by that specific time point. Baseline values are excluded from this analysis.
- **Rate of each cytogenetic response (CyR) categories (including complete cytogenetic response (CCyR)) by Week 48 (Week 96)**. These are cumulative response rates by time points Week 48 (Week 96) and are defined as the proportion of participants who achieve each response categories *at or before* Week 48 (Week 96), for example, if a participant achieves an CCyR but then loses it before or at the Week 48 (Week 96) visit, he/she will still be classed as achieving CCyR by that specific time point.

Definitions of BCR::ABL≤1%, MR4.0, MR4.5, CHR, and CyR can be found in Appendix Section 5.3.3.4, Section 5.3.3.5, Section 5.3.5.1, and Section 5.3.4, respectively.

2.9.1.2 Time-to-event endpoints related to responses

Time-to-event endpoints related to responses are:

- **Time to first MMR, first MR4.0, or first MR4.5** (in weeks) is defined as: (date of first documented MMR, MR4.0, or MR4.5 - date of randomization + 1)/7. Discontinuation from treatment due to any reason (intolerance, failure, death etc.) without prior achievement of the endpoint (MMR, MR4.0, or MR4.5) will be considered as competing risk. The time to first achievement of the endpoint (MMR, MR4.0, MR4.5) will be censored at the last molecular assessment (RQ-PCR) date on treatment prior to or at the analysis cut-off date, for participants who have not experienced an event (MMR, MR4.0, or MR4.5) or a competing risk event.
- **Duration of MMR, MR4.0, or MR4.5** is defined for participants in the MMR, MR4.0, or MR4.5 Responder Set, respectively. The duration of MMR, MR4.0, or MR4.5 (in weeks) is calculated as: (end date or censoring date – date of first response + 1)/7, i.e. it is the time between the date of the first documented MMR, MR4.0, or MR4.5, and the end date of MMR, MR4.0, or MR4.5, respectively. The end date is the earliest date of loss of MMR, MR4.0, or MR4.5, treatment failure of ELN criteria (as defined in Section 5.3.6.1), progression to accelerated phase (AP)/blast crisis (BC), or CML-related death.

  The duration will be censored at the last molecular assessment (RQ-PCR) date while on treatment for participants who have not experienced any of the above events.

  In case of duration of MMR, loss of MMR must be a confirmed loss.

  Definitions of loss of MMR, MMR4.0, and MMR4.5 can be found in Appendix Section 5.3.3.3,Section 5.3.3.4, and Section 5.3.3.5, respectively.

2.9.1.3 Other time-to-event endpoints

Other time-to-event endpoints are:

- Time to Treatment Failure (TTF)
- Failure Free Survival (FFS)
- Event Free Survival (EFS)
- Progression-Free-Survival (PFS)
- Overall Survival (OS)

These endpoints are defined for each participant as the duration between the date of randomization and the earliest occurrence of a relevant event. The relevant events along with the associated endpoints are given in Table 2-7. An “x” in the columns under Endpoints indicates that the corresponding event (row-wise) is one of the relevant events for that endpoint; and a “CR” indicates that the corresponding event is a competing risk event for that endpoint.

Table 2-7 Definitions of relevant events used to define time-to-event endpoints

Event	Definition in Appendix	Endpoints				
		TTF	FFS	EFS	PFS	OS
1 Treatment failure as defined per ELN criteria	Section 5.3.6.1	x	x	x		
2 Progression to AP/BC while on treatment	Section 5.3.7	x	x	x	x	
3 Confirmed loss of MMR (in 2 consecutive tests) at any time while on study treatment	Section 5.3.3.3	x	x	x		
4 Death due to any causes while on treatment	NA	x	x	x	x	x
5 Discontinuation of study treatment due to AE	NA	x	CR	x		
6 Discontinuation from study treatment due to any other reasons	NA	x	CR	CR		
7 Progression to AP/BC during survival follow-up	Section 5.3.7		x	x	x	
8 Death due to any causes during survival follow-up	NA		x	x	x	x

The event date for "discontinuation from study treatment due to AE" or "discontinuation from study treatment due any other reasons" will be from the EoT eCRF.

Participants who do not experience any of the relevant events nor any competing risk events on or before each analysis data cut-off date or the closing of the study are considered censored for the corresponding endpoint:

- The censoring date for TTF is the date of the last study assessment while on treatment (PCR, cytogenetic, hematologic, or extramedullary) before the analysis data cut-off date.
- The censoring date for FFS, EFS, and PFS is the date of the last study assessment (PCR, cytogenetic, hematologic, or extramedullary) for those still on treatment, or last post-treatment follow-up before analysis data cut-off date, in which case the last contact date can be used.
- The censoring date for OS is the last contact date (cf. Section 2.1.3.12) before the analysis data cut-off date.

Time (in months) of each time-to-event endpoint is calculated as

(Date of earliest occurrence among the relevant events or date of censoring – date of randomization + 1) / 30.4375.

2.9.2 Statistical hypothesis, model, and method of analysis

2.9.2.1 Analyses for response endpoints

The FAS (or FAS_{IMA} or FAS_{2GTKI} respectively) will be used for the response endpoints (BCR::ABL≤1%, MMR, MR4.0, MR4.5, CHR, and CcyR) as defined in Section 2.9.1.1.

2.9.2.1.1 Molecular responses and CHR

Frequency and percentage of participants in the molecular response categories and in CHR will be presented for each scheduled visit.

The statistical models and methods for these endpoints follow that for the primary and key secondary endpoints. In particular, the response rate for each endpoint and the associated two-sided 95% CI based on the Clopper-Pearson method will be presented by the following:

- Both treatment arms
- Both treatment arms within the stratum of participants who have imatinib as their PRS-TKI.
- Both treatment arms within the stratum of participants who have a 2G TKI as their PRS-TKI.

Comparisons of proportions of responders using one-sided stratified Mantel-Haenszel test for the endpoints (at and by scheduled visits) will be conducted for:

- Asciminib vs. IS-TKI stratified by both ELTS and PRS-TKI
- Asciminib vs. IS- TKI stratified by ELTS within the stratum of participants who have imatinib as their PRS-TKI
- Asciminib vs. IS-TKI stratified by ELTS within the stratum of participants who have a 2G TKI as their PRS-TKI

The stratified Mantel-Haenszel estimates of the common risk difference will be provided, together with the corresponding two-sided 95% CI for the following:

- Asciminib and IS-TKI arms stratified by ELTS and PRS-TKI
- Asciminib and IS-TKI arm stratified by ELTS within the stratum of participants who have imatinib as their PRS-TKI
- Asciminib and IS-TKI arm stratified by ELTS within the stratum of participants who have a 2G TKI as their PRS-TKI

2.9.2.1.2 Cytogenetic responses

For each CyR categories (including CcyR) by week 48 (or week 96), frequency and percentage of responders and the associated two-sided 95% CI based on the Pearson-Clopper method will be presented by the following:

- Both treatment arms
- Both treatment arms within the stratum of participants who have imatinib as their PRS-TKI.
- Both treatment arms within the stratum of participants who have a 2G TKI as their PRS-TKI.

2.9.2.2 Analyses for time-to-event endpoints related to responses

2.9.2.2.1 Duration of MMR, MR4.0, or MR4.5

The respective MMR, MR4.0, or MR4.5 Responder Set will be used for duration of MMR, MR4.0, or MR4.5. These time-to-event endpoints will be analyzed by Kaplan-Meier (KM) method and presented by KM plots. The estimated median duration along with the 95% CI (Brookmeyer and Crowley, 1982), along with the proportion of participants who are still MMR, MR4.0, or MR4.5 responders at specified scheduled visits (Week 48, Week 72, Week 96) Week and the associated 95% CI, will be presented for each treatment arm by overall and for within the stratum defined by PRS-TKI).

2.9.2.2.2 Time to First MMR, MR4.0, or MR4.5

The FAS, FAS_{IMA}, and FAS_{2GTKI} will be used for time to first MMR, MR4.0, and MR4.5. Competing risk analysis of time to first MMR, MR4.0, or MR4.5 will be performed. Discontinuation from treatment due to any reason (intolerance, failure, death etc.) without prior achievement of the endpoint (MMR, MR4.0, or MR4.5) will be considered as competing risk.

The estimated cumulative incidence rates and 95% CI at specified scheduled visits will be presented for each treatment arm for overall and for within the stratum defined by PRS-TKI. The cumulative incidence curve will be plotted. For this competing risk analysis, time to first achievement of the end-point (MMR, MR4.0, or MR4.5) will be censored at the last molecular assessment (RQ-PCR) date on treatment, or the analysis data cut-off date (whichever comes first) prior to or at the analysis cut-off date, for participants who have not experienced an event (MMR, MR4.0, or MR4.5) or a competing risk event.

2.9.2.3 Other time to event endpoints

All other time to event endpoints (TTF, FFS, EFS, PFS, and OS) are analyzed using FAS, FAS_{ima}, and FAS_{2GTKI}.

For TTF, PFS, and OS, summary statistics (including median and maximum follow-up time) and summary plot of Kaplan-Meier estimates will be presented. Stratified log-rank tests and Cox regression models for comparing the two treatment arms will also be performed.

For FFS and EFS, competing risk analysis will be performed. The estimated cumulative incidence rates and 95% Cis at specified scheduled visits will be presented for each treatment arm. The cumulative incidence curve will be plotted.

2.9.3 Handling of missing values/censoring/discontinuations

2.9.3.1 Handling of missing values for endpoints related to responses

In general, participants will be considered a non-responder at a specific time point or visit in the following situations:

- Discontinuing the randomized treatment prior to a specific time point due to any reason
- Meeting failure criteria
- Lack or missing a documented response

- Lack or missing of an available assessment for determining the response at the specified visit

Molecular response (BCR::ABL≤1%, MMR, MR4.0, MR4.5, or CcyR) at specific time points: The category "Missing" will be assigned to

- Ongoing cases, i.e., participants without assessment at the specific time point who have not discontinued study treatment and have not been treated sufficiently long for a specific time point
- Discontinued due to lack of efficacy, progressive disease, or death per EoT eCRF prior to a specific time point
- Discontinued due to other reasons prior to a specific time point

Molecular response (BCR::ABL≤1%, MMR, MR4.0, MR4.5, or CCyR) by specific time points: The category "Missing" will be assigned to participants for whom an evaluable response assessment was never provided.

Time to MMR, MR4.0, or MR4.5: For participants in the FAS, FAS_{IMA}, and FAS_{2GTKI} who have not experienced any MMR, MR4.0, or MR4.5, respectively, or a competing risk event, the time will be censored as follows in the competing risk analysis:

- If a participant does not achieve the specified response event or a competing risk event before the cut-off date for the analysis, censoring time will be the last molecular assessment (PCR) date on treatment prior to the cut-off date.
- In case no on-treatment response assessment was performed, the participant will be censored at Day 1.

Discontinuation from treatment due to any reason (intolerance, failure, progression, death etc.), without prior achievement of the endpoint (MMR, MR4.0, or MR4.5) will be considered as competing risk.

2.9.3.2 Handling of censoring for time-to-event endpoints

Censorings are defined for each time-to-event endpoint in Section 2.9.1.2 and Section 2.9.1.3. All censoring are considered right, independent censoring.

2.9.4 Supplemental analysis

Time to first MMR, MR4.0, and MR4.5 will also be analyzed by KM method. In this analysis, participants who discontinue from treatment due to any reason (intolerance, failure, death etc.) without prior achievement of MMR, MR4.0, or MR4.5 will be considered censored. The dates that are used as the dates of the competing events in CIF analysis are considered censoring dates.

The estimated 25^{th}, 50^{th} (median), and 75^{th} percentiles of time to first MMR, MR4.0, or MR4.5 along with the 95% CI, and the proportion of participants who reach MMR, MR4.0, or MR4.5 at specified timepoints and the associated 95% CI will be presented for each treatment arm and for within the stratum defined by PRS-TKI. FAS, FAS_{IMA}, and FAS_{2GTKI} are used.

2.10 Analysis supporting the secondary pharmacokinetic (PK) objectives

The PK objective is to characterize the PK of asciminib in newly diagnosed CML-CP patients.

Using PAS, summary statistics (n, mean, SD, coefficient of variation (CV) for mean, geometric mean, CV for geometric mean, median, minimum, and maximum) will be presented for plasma concentration at each scheduled time point. The mean (+/- SD) and median mean concentration-time profiles for asciminib over time will be displayed graphically for PAS on the linear and semi-log view.

Using Safety set, concentration data will be listed. Concentration values below the limit of quantification (BLQ) will be set to zero by the Bioanalyst and displayed in listings as zero with a flag. BLQ values will be handled as zero in any calculations of summary statistics but handled as missing for the calculation of the geometric means and CVs.

Pharmacokinetic parameters will be determined by non-compartmental method(s) using the PK profile of asciminib in participants with full PK sampling (i.e., full PK group, Week 2). PK parameters listed in Table 2-8 will be derived and reported, when feasible.

Population PK modeling may be performed (using all PK data) and the results may be reported in a separate population PK report. Data from this study may be combined with data from other studies for this analysis.

Table 2-8 Non compartmental pharmacokinetic parameters in full PK group

AUC_{last}	The area under the plasma concentration-time curve calculated from time zero to the last measurable concentration sampling time (tlast) ($ng*hr*mL^{-1}$)
$AUC_{0\text{-}tau}$	The area under the plasma concentration-time curve from time zero to the end of a dosing interval (tau=24h) at steady-state ($ng*hr*mL^{-1}$)
C_{max}	The maximum (peak) observed plasma drug concentration after dose administration (ng/mL)
T_{max}	The time to reach maximum (peak) plasma drug concentration after dose administration (hr.)
CL/F	The total body clearance of drug from the plasma after oral administration ($L*hr^{-1}$)

2.11 Analysis supporting the other secondary safety objectives

The study has one secondary safety objective which will be formally tested. Its associated endpoint is TTDAE, and its analysis is discussed in Section 2.7 of the current document. This section describes the analyses of all other secondary safety objectives as given in Table 1-2.

All safety analyses will be based on the safety set. All listings and tables will be presented by treatment group (overall (i.e., asciminib and investigator selected TKI); and if appropriate also by each treatment (i.e., asciminib, imatinib, nilotinib, dasatinib and bosutinib)).

With the exception of the endpoint of TTDAE which will be formally tested as described in Section 2.8, no other safety endpoint will be tested formally.

All Aes are assigned to one of the three mutually exclusive segments defined in Section 2.1.3.10. In other words, each AE is considered to have occurred either during the pre-treatment period, or the on-treatment period, or the post-treatment period. If dates are incomplete in a way that clear assignment to pre-, on-, or post-treatment period cannot be made, then the respective AE data will be assigned to the on-treatment period.

Safety summaries (tables and figures) on the Safety Set will include only data from the on-treatment period. One exception is the baseline data, which will also be summarized where appropriate (e.g., change from baseline summaries). In particular, summary tables for Aes will summarize only on-treatment events, with a start date during the on-treatment period. In addition, a separate summary for death including on-treatment and post-treatment deaths will be provided.

All safety data (including those from the pre-treatment, post-treatment) will be listed and flagged as appropriate.

2.11.1 Adverse events (AEs)

AE summaries will include all AEs occurring during the on-treatment period (i.e., that started or worsened during the on-treatment period, also known as treatment emergent AEs). All AEs collected in the AE eCRF page will be listed along with the information collected on those AEs (e.g., toxicity grade, AE relationship to study treatment, AE outcome, action taken etc.). All AEs with start date outside of the on-treatment period (i.e., with start dates during the pre-treatment or post-treatment period) will be flagged in the listings.

The number (and percentage) of subjects with treatment emergent AEs will be summarized by primary system organ class (SOC), preferred term (PT) using MedDRA coding, and maximum severity (based on Common Terminology Criteria for Adverse Events (CTCAE) grades).

AEs will be summarized by number and percentage of subjects, having at least one AE in each primary SOC and for each PT. A subject with multiple occurrences of an AE will be counted only once in the respective AE category. A subject with multiple CTCAE grades for the same preferred term will be summarized under the maximum CTCAE grade recorded for the event. AEs will be assessed according to the (CTCAE) version 5.0. AE with missing CTCAE grade will be included in the ‘All grades’ column of the summary tables.

In AE summaries, the primary SOC will be presented alphabetically, and the preferred terms will be sorted within primary SOC in descending frequency. The sorting order for the PT will be based on their frequency in the asciminib arm. The summaries will show ‘All grades’ (including AEs with missing grade) and ‘Grades ≥ 3’.

The following adverse event summaries will be produced by treatment arm for the Safety set: overview of adverse events and deaths, AEs by SOC, and PT, summarized by relationship, seriousness, leading to treatment discontinuation, leading to dose interruption, leading to adjustment, requiring additional therapy, and leading to fatal outcome.

All AEs, deaths, and serious adverse events (including those from the pre- and post-treatment periods) will be listed and those collected during the pre-treatment, post-treatment safety period will be flagged.

For posting to ClinTrials.gov and EudraCT, a summary table of on-treatment deaths and serious AEs and another summary table of non-serious AEs by treatment, both including occurrences (an occurrence is defined as >1 day between start and prior end date of record of same preferred term) and sorted by SOC and PT, will be presented as well.

In order to account for differences in exposure between the treatment arms, incidence rates of AEs and SAEs will be presented by adjusting for duration of treatment period in patient-years. (See Section 5.4.5 for calculation of exposure-adjusted incidence rate.) The first occurrence of an AE or SAE will also be reported by time intervals (i.e., period of emergence: the event is assigned to the interval when it first started): 0 to 2 months, > 2 months to 6 months, > 6 months to 12 months, and more than 12 months after the start of study treatment.

2.11.1.1 Adverse events of special interest / grouping of AEs

An adverse event of special interest (AESI) is a grouping of adverse events that are of scientific and medical concern specific to compound asciminib. These groupings are defined using MedDRA terms, SMQs (standardized MedDRA queries), HLGTs (high level group terms), HLT (high level terms) and PTs (preferred terms). Customized SMQs (Novartis MedDRA queries, NMQ) may also be used. An NMQ is a customized group of search terms which defines a medical concept for which there is no official SMQ available or the available SMQ does not completely fit the need. It may include a combination of single terms or an existing SMQ, narrow or broad. These searches will be defined in the eCRS (electronic Case Retrieval Strategy) in the DMS (Document Management System), and a listing of search terms will be provided in the CSR. The latest approved version of CRS prior to the respective database lock will be used.

For each specified AESI, number and percentage of participants with at least one event of the AESI occurring during the on-treatment period will be summarized together with the individual preferred terms in that grouping. In addition, number (%) of subjects with at least one AESIs by maximum CTC grade, related AESIs, serious AESIs as well as action taken and outcome of the respective AESI will be summarized.

Summaries of these AESIs will be provided by treatment arm (specifying grade, SAE, relationship, leading to treatment discontinuation, leading to dose adjustment/interruption, death, etc.) together with exposure-adjusted incidence rate (See Section 5.4.5 for calculation) and cumulative incidence proportions (with death and discontinuation due to any reason as competing risks) to account for difference in exposure length in the treatment groups.

A listing of all grouping levels down to the MedDRA PTs used to define each AESI will be generated.

2.11.2 Deaths

Separate summaries for on-treatment and all deaths *(including post-treatment deaths* not in the AE CRF but in the survival CRF) will be produced on the Safety set by treatment arm, system organ class and preferred term.

All deaths will be listed, where deaths occurring during the pre/post-treatment, will be flagged. A separate listing of deaths prior to starting treatment will be provided for all screened subjects.

2.11.3 Laboratory data

If multiple lab assessments are available from the same day, the average is taken first. Grading of laboratory values will be assigned programmatically as per National Cancer Institute (NCI) CTCAE version 5. The calculation of CTCAE grades will be based on the observed laboratory values only, clinical assessments will not be taken into account. CTCAE Grade 0 will be assigned for all non-missing values not graded as 1 or higher. Grade 5 will not be used. For laboratory tests where grades are not defined by CTCAE v5, results will be categorized as low/normal/high based on laboratory normal ranges. On analyzing laboratory data, all sources (central and local laboratories) will be combined. The summaries will include all assessments available for the lab parameter collected no later than 30 days after the last study treatment administration date.

The following summaries (based on the Safety Set) will be generated separately for hematology, and biochemistry tests (by laboratory parameter and treatment):

- For laboratory tests where grades are defined by CTCAE v5:
  - Worst post-baseline CTCAE grade (regardless of the baseline status). Each participant will be counted only once for the worst grade observed post-baseline in the on-treatment period.
  - Shift tables using CTCAE grades to compare respective baseline to the worst on-treatment value
- For laboratory tests where grades are not defined by CTCAE v5:
  - Shift tables using the low/normal/high/ (low and high) classification to compare respective baseline to the worst on-treatment value.

The following listings will be produced separately for hematology and biochemistry for the laboratory data:

- Listings of all laboratory data, with CTCAE v5 grades (if applicable) and classification relative to the laboratory normal range. Lab data collected during the post-treatment period will be flagged.
- Listing of all CTCAE v5 grade 3 or 4 laboratory toxicities

2.11.3.1 Liver function parameters

Liver function parameters of interest are total bilirubin (TBL), Alanine aminotransferase (ALT), Aspartate aminotransferase (AST) and alkaline phosphatase (ALP). The number (%) of participants with worst post-baseline values as per Novartis DILI Clinical safety guidelines will be summarized for the criteria defined by single lab parameter. For combination of various parameters, the worst post-baseline values from each single parameter are taken into consideration, i.e., it may not come from the concurrent measurement (i.e., same assessment):

The following summaries will be produced:

- ALT or AST > 3x upper limit of norm (ULN)
- ALT or AST > 5xULN
- ALT or AST > 8xULN

- ALT or AST > 10xULN
- ALT or AST > 20xULN
- TBL > 2xULN
- TBL > 3xULN
- ALT or AST > 3xULN & TBL > 2xULN. This defines DILI.
- ALT or AST > 3xULN & TBL > 2xULN & ALP >= 2xULN
- ALT or AST > 3xULN & TBL > 2xULN & ALP < 2xULN

2.11.4 Other safety data

2.11.4.1 ECG and cardiac imaging data

12-lead ECGs including PR, QRS, QT, QTcF and RR intervals will be obtained centrally for each subject during the study. ECG data will be read and interpreted centrally.

Data handling

The average of the triplicate ECG parameters at each time point will be used in the analyses.

For unscheduled visits, ECGs that are reported on the same day and within 30 minutes apart from each other will be assumed to be sequential ECGs and thus will be used to compute the mean of the ECG parameters.

Unscheduled ECG measurements will not be used in computing the summary statistics for change from Baseline at each post-baseline time point. However, they will be used in the outlier analyses (e.g., QTc > 450 ms, > 480 ms, or > 500 ms at any time point, or an increase from Baseline in QTc > 30 ms or > 60 ms). End of treatment ECG measurements for discontinued participants will be considered as an unscheduled measurement in case it occurs outside a scheduled visit.

Data analysis for ECG

The number and percentage of subjects with notable ECG values will be presented by treatment arm for the Safety Set. Notable values are defined below:

- QT, QTcF
  - New value of > 450 and ≤ 480 ms
  - New value of > 480 and ≤ 500 ms
  - New value of > 500 ms
  - Increase from Baseline of > 30 ms to ≤ 60ms
  - Increase from Baseline of > 60 ms
- HR
  - Increase from baseline >25% and to a value > 100 bpm
  - Decrease from baseline >25% and to a value < 50 bpm
- PR

- Increase from baseline >25% and to a value > 200 ms
- New value of > 200 ms

- QRS
  - Increase from baseline >25% and to a value > 120 ms
  - New values of QRS > 120 ms

A listing of all ECG assessments will be produced by treatment arm and notable values will be flagged. A separate listing of only the subjects with notable ECG values will also be produced. In each listing the assessments collected during the post-treatment period will be flagged.

Change from baseline ECG parameters by timepoint will also be summarized by treatment.

2.11.4.2 Cardiovascular risk factor assessment

Prior to randomization, for each participant information on Framingham risk factors (e.g., heavy smoking, SBP etc.) are collected. Framingham CVD points and CVD 10-year risk categories are derived. Summary statistics of the points, and count and percentages of participants in each risk category will be presented. Derivations of the points and risk categories are described in Section 5.5.1.

In addition, medical history of each participant for risk factors related to Charlson comorbidity index (CCI) (e.g.: dementia, heart disease, diabetes mellitus, liver disease etc.) are also collected prior to randomization. CCI total points will be summarized with counts and percentages of participants with each index value, and the corresponding estimated 10-year survival will be summarized and reported with summary statistics as a continuous variable. Derivations of the index are described in Section 5.5.2.

2.11.4.3 Vital signs

Vital sign assessments are performed in order to characterize basic body function. The following parameters were collected: height (cm), weight (kg), body temperature (°C), heart rate (beats per minute), systolic and diastolic blood pressure (mmHg).

Data handling

Vital signs collected on treatment will be summarized. Values measured outside of on treatment period will be flagged in the listings.

Data analysis

Notable vital sign values during on-treatment period in subjects with non-notable values at baseline (e.g. systolic BP > 90 and <180 mmHg for analysis of systolic BP) will be summarized using the criteria in the Table 2-9.

The number and percentage of subjects with notable vital sign values (high/low) in systolic blood pressure, diastolic blood pressure, pulse rate, weight and temperature will be presented by treatment arm.

A listing of all vital sign assessments will be produced by treatment arm and notable values will be flagged. In the listing, the assessments collected outside of on-treatment period will be flagged.

Table 2-9 Notable vital sign values

Vital sign (unit)	Clinically notable criteria	
	above normal value	**below normal value**
Systolic blood pressure (mmHg)	>=180 with increase from baseline of >=20	<=90 with decrease from baseline of >=20
Diastolic blood pressure (mmHg)	>=105 with increase from baseline of >=15	<=50 with decrease from baseline of >=15
Pulse rate (bpm)	>=100 with increase from baseline of >25%	<=50 with decrease from baseline of > 25%
Weight (kg)	Increase >=10% from baseline	Decrease >= 10% from baseline
Body temperature (°C)	>= 39.1	-

ECOG performance status

ECOG performance status collected on treatment will be summarized. Shift tables will be provided comparing baseline with best and worst values during study for each treatment group.

2.12 Analysis supporting the secondary objectives related to patient-reported outcomes (PRO)

The respective full analyses sets (FAS, FAS_{IMA}, FAS_{2GTKI}) will be used for analyzing PRO data unless specified differently.

Two PRO measures, the EORTC QLQ-C30 (version 3.0) and the EORTC QLQ-CML24 will be used to assess patient-reported disease-related symptoms, functioning, and health-related quality of life (HRQoL) as secondary study endpoints..

The QLQ-30 consists of 30 items across (i) 5 functioning scales: physical, role, emotional, cognitive and social, (ii) 3 symptoms' scales: fatigue, nausea/vomiting and pain, (iii) 6 single-item scales: assessing, additional symptoms (dyspnea, insomnia, appetite loss, constipation, diarrhea), and financial impact and (iv) the global health status quality of life scale.

Each of the multi-item scales includes a different set of items – no item occurs in more than one scale. A high score for functional and QoL items/scales from the QLQ-30 represents better function and QOL. A high score in symptoms items from QLQ-30 represents worse symptoms.

The QLQ-CML24 consists of 24 items where (i) 22 multi-scale items: impact on daily life, symptom burden, impact on worry/mood, satisfaction with care and information, and (ii) 2 single items: body image problems, and satisfaction with social life.

A higher score on most of the item scales in QLQ-CML24 reflects a larger impairment in the corresponding domain, with the exception of the satisfaction with care and information, and problems and satisfaction with social life, where a higher score reflects a higher level of satisfaction.

Change from baseline in overall scores and individual domains for each PRO instruments will be summarized using descriptive statistics at each timepoint for each treatment arm (overall and by strata defined by PRS-TKI). Participants with an evaluable baseline score and at least one evaluable post-baseline score during the treatment period will be included in the change from baseline analyses. Baseline is as defined in Section 2.1.3.9.

All PRO measures require patient's direct completion and will be administered utilizing electronic device for data collection at scheduled time points from screening to end of treatment.

Missing data items in a scale will be handled according to the manual for each instrument. No imputation will be applied if the total or subscale scores are missing at a visit. All measures will assess differences between the treatment arms (overall and by strata defined by PRS-TKI).

Compliance to the schedule of administration of the PRO questionnaire will be summarized by treatment group, for baseline and scheduled post-baseline assessment time points. The following categories, as programmed based on the questionnaires, will be used to describe whether the questionnaire is completed at a specific time point:

1. yes, fully completed
2. yes, partly completed
3. no

2.13 Analysis supporting the exploratory objectives

Unless otherwise specified, all exploratory analysis will be performed using FAS for efficacy related endpoints, and the Safety Set for safety related endpoints.

2.13.1 Biomarker endpoints

The FAS is used to describe biomarkers and for efficacy related endpoints; Safety Set is used to assess the relationship between biomarkers and selected safety endpoints, while the PK Set is used to assess the relationship between PK parameters and biomarkers. Since no imputation is planned, the number of participants included in each analysis will reflect the number of participants in the chosen analysis set which have valid biomarker assessments.

Continuous biomarkers (e.g., gene expression) will be summarized using means, medians, standard deviations, minimums, and maximums, by visit. Both level and change from baseline levels (absolute, percent and fold changes) will be summarized for biomarkers that also have assessments at post-baseline visits.

Categorical biomarkers (e.g., % cells for immune microenvironment markers, BCR::ABL1 mutation status etc.) will be summarized using frequency counts and percentages.

Additional exploratory biomarker analyses are defined in a separate biomarker analysis plan.

BCR::ABL ratios

FAS, FAS_{ima}, FAS_{2GTKI} will be used for the following analyses.

BCR::ABL ratios (%IS) will be summarized by frequencies and proportions at each scheduled time point including baseline according to the following categories:

- 10% < BCR::ABL ratio (IS)
- 1% < BCR::ABL ratio (IS) ≤ 10%
- 0.1% < BCR::ABL ratio (IS) ≤ 1%
- 0.01% < BCR::ABL ratio (IS) ≤ 0.1%
- 0.0032% < BCR::ABL ratio (IS) ≤ 0.01%
- BCR::ABL ratio (IS) ≤ 0.0032%

For participants with no evidence of typical transcript at the time of screening, they are categorized in a separate category, and will not be counted toward any of the above categories.

Frequencies and proportions cumulative over the above categories at the scheduled time points, as well as the frequencies and proportions of the best BCR::ABL ratios of each participant by these time points will also be produced.

In addition, the following categories are defined when assessments cannot be classified as above:

- Missing assessment: an unavailable assessment at the scheduled time point
- Assessment not evaluable: an available assessment that is not evaluable as it occurs after the participant having met a treatment failure criterion.
- Discontinued due to lack of efficacy, disease progression, or death: a participant who has discontinued treatment due to lack of efficacy, disease progression, or death before the scheduled time point.
- Discontinued due to other reasons: a participant who has discontinued treatment due to other reasons before the scheduled time point.

In the “by” time point analysis, participants who have achieved any response at or before the time point will be displayed in their best response category, regardless of whether they have lost the response or have discontinued. Therefore, this response rate represents the best observed response rate up to that specific time point (including the time window). Participants for whom an evaluable response assessment is never provided will be classified as ‘Missing.’

Side-by-side boxplots summarizing the numerical BCR::ABL values at each time point for the two treatment arms from baseline to week 48 (or week 96) will be generated for visualization of changes over time.

Baseline and post-baseline mutation summary statistics

As defined in Section 2.1.3.9, for mutation related analyses, the earliest assessment taken on or before Day 84 is considered baseline. Day 84 is the mid-point between randomization and Week 24 target day (Day 168, see Table 2-1). Any assessments taken on or after Day 85 are considered post-baseline.

The following summary statistics will be generated for baseline and post-baseline mutation data:

(1) All BCR::ABL mutation data from NGS analysis will be reported using mutation counts and percentages by the mutation type in the form of contingency tables with the rows containing any mutations detected, and the treatment groups in the columns. All the

mutation categories for a gene will also be aggregated into mutant, wild type, or missing/unknown groups, and counts will be reported by these three categories as well.

(2) The functional locations of the mutation will be reported. The locations are categorized as mutations that arise in

   a. ATP BS of BCR::ABL, defined as AA 292-327 (inclusive),

   b. C lobe, defined as AA 411 (inclusive) and up, or

   c. Other, defined as anywhere outside the above two regions.

Contingency tables will be generated with three rows indicating the three locations aggregated overall post-baseline mutations, and columns for treatment groups.

All the mutation data will be listed for each participant.

Association between biomarkers and clinical outcome

The relationship between baseline BCR::ABL1 gene mutation data and the primary endpoint (MMR at week 48) and the key secondary endpoint (MMR at week 96) will be explored by contingency tables of number of participants with a given mutation, and among them the number of participants who are week 48 (or week 96) responders.

The exact same analysis will be performed for the relationship between post-baseline new BCR::ABL1 mutation up to 48 weeks and up to 96 weeks, and the primary endpoint (MMR at week 48) and the key secondary endpoint (MMR at week 96) respectively.

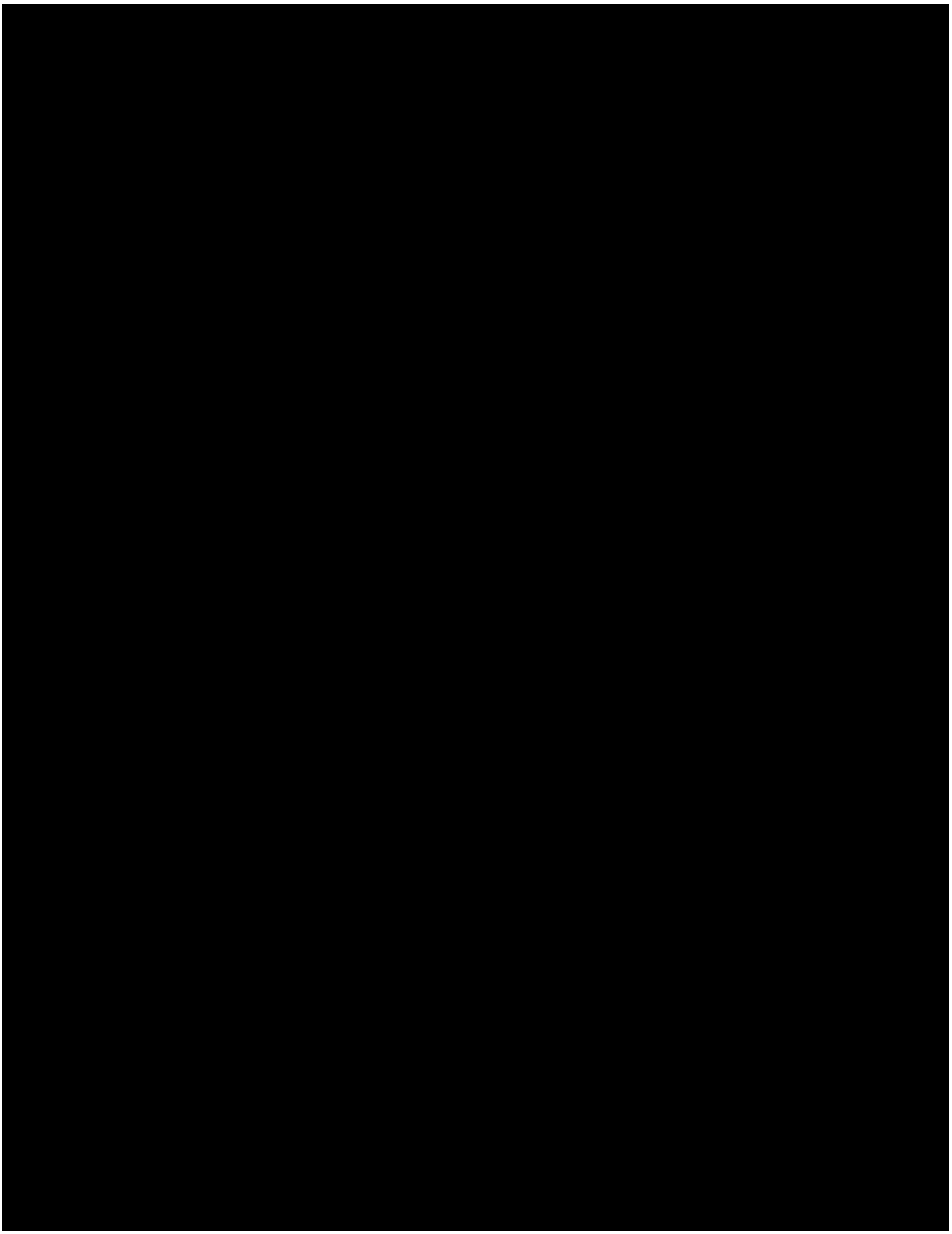
In these two analyses, the number of participants whose mutation category based on variant allele fraction (VAF) (<20% vs. ≥20%) will also be displayed.

2.13.2 Pharmacogenomics

The impact of baseline genetic variants of the [redacted] on asciminib exposure will be explored. Using PAS, summary statistics for plasma concentration at scheduled time points and/or selected PK parameters (e.g., AUC, Cmax) will be summarized by genetic variant groups. This analysis may be reported separately.

2.13.3 Other patient-reported outcomes (PRO)

[redacted]



[redacted]

2.13.4 Health care resource utilization

Health care resource utilization (HCRU) data based on FAS, FAS_{Ima}, and FAS_{2GTKI} by treatment arms will be summarized with descriptive statistics (n, mean, median, SD, min, max) for quantitative variables (e.g., number of hospitalizations or duration of hospitalizations), and count and percentage for qualitative variables (e.g., reasons for hospitalizations).

2.13.5 Exploratory analysis for the primary and key secondary efficacy endpoints

FAS and FAS_{ima} will be used for these exploratory analyses.

Mantel- Haenszel estimates for the primary and the key secondary endpoints (MMR at week 48 and MMR at week 96, respectively) of the common odds ratios and the corresponding 95% CIs will be provided.

Additionally, logistic regression analyses will incorporate the key baseline variables which are significant based on subgroup analysis (see Section 2.2.4.1) into the model to further evaluate the impact of these variables on these endpoints, and to provide a treatment effect estimates which are adjusted for these variables. Adjusted odds ratios for the treatment effects with associated 95% CIs will be presented.

For the logistic regression based on FAS, both ELTS and PRS-TKI will be included as covariates. For the logistic regression based on FAS_{ima}, ELTS will be included as a covariate.

2.14 Interim analysis

No formal interim analysis is planned for this trial.

DMC safety analyses will be conducted regularly until the primary Week 48 analysis. Details of the safety data review are described in the DMC Charter.

3 Sample size calculation

Sample size calculation has been detailed in Section 12.11 of the CSP. It is reproduced below with the exception of (1) re-numbering of tables, and (2) the insertion of Table 3-2 and the paragraph before it. Otherwise, section numbers in this section refer to those in the CSP, but with reference to the current document.

Section 12.11 of the CSP:

Historical estimates for efficacy of the TKIs in the control arm were available from published literature on the pivotal trials of nilotinib, dasatinib and bosutinib respectively, in newly diagnosed patients with CML-CP.

Table 3-1 Historic MMR rates from pivotal TKI trials in newly diagnosed patients with CML-CP

		MMR rate at/by* 48 weeks (n)	MMR rate at/by* 48 weeks (n)
Study Acronym	**Investigational Drug**	**Imatinib Arm (400mg q.d.)**	**Investigational Drug Arm**
BFORE	Bosutinib 400 mg q.d.	0.37 (n=241)	0.47 (n=246)
BELA	Bosutinib 500 mg q.d.	0.27 (n=252)	0.41 (n=250)
DASISION	Dasatinib 100 mg q.d.	0.28* (n=260)	0.46* (n=259)
ENESTND	Nilotinib 300 mg b.i.d.	0.22 (n=281)	0.44 (n=283)
		MMR rate by 96 weeks (n)	**MMR rate by 96 weeks (n)**
Study Acronym	**Investigational Drug**	**Imatinib Arm (400mg q.d.)**	**Investigational Drug Arm**
BFORE	Bosutinib 400 mg q.d.	0.58 (n=241)	0.67 (n=246)
BELA	Bosutinib 500 mg q.d.	0.49 (n=252)	0.59 (n=250)
DASISION	Dasatinib 100 mg q.d.	0.47 (n=260)	0.65 (n=259)
ENESTND	Nilotinib 300 mg b.i.d.	0.44 (n=281)	0.71 (n=283)

BFORE : Cortes JE, Gambacorti-Passerini C, Deininger MW, et al. Bosutinib Versus Imatinib for Newly Diagnosed Chronic Myeloid Leukemia: Results From the Randomized BFORE Trial. *J Clin Oncol*. 2018 ;36(3) :231-237. Doi :10.1200/JCO.2017.74.7162
BELA : Cortes JE, Kim DW, Kantarjian HM, et al. Bosutinib versus imatinib in newly diagnosed chronic-phase chronic myeloid leukemia: results from the BELA trial. *J Clin Oncol*. 2012;30(28):3486-3492. Doi:10.1200/JCO.2011.38.7522
DASISION: Kantarjian H, Shah NP, Hochhaus A, et al. Dasatinib versus imatinib in newly diagnosed chronic-phase chronic myeloid leukemia. *N Engl J Med*. 2010;362(24):2260-2270. Doi:10.1056/NEJMoa1002315
ENESTND: Saglio G, Kim DW, Issaragrisil S, et al. Nilotinib versus imatinib for newly diagnosed chronic myeloid leukemia. *N Engl J Med*. 2010: 362 (24): 2251-2259. Doi: 10.1056/NEJMoa0912614

Random effects meta-analyses are performed using the historical published data and is implemented in R 3.4.3 using the package 'Meta' (Balduzzi et al. 2019).

A random effects meta-analysis of MMR at/by Week 48 from the published data gives us estimated proportions (95% CI) of participants that achieve MMR at Week 48 for imatinib of 0.28 (0.23, 0.34), and for 2G TKIs of 0.44 (0.41, 0.48).

The assumption on the MMR rate at Week 48 for asciminib is based on extrapolation of data from the FIH CABL001X2101 study and the data from the pivotal CABL001A2301 study in third line patients, to the first line setting.

On review of published data for 2G TKIs (dasatinib, nilotinib, bosutinib) it was observed that the Week 48 MMR rates for 3L to 2L to 1L treatment of CML-CP, increased by approximately 10% for each earlier line of treatment (refer Table 3-2 below). We expect a similar pattern of Week 48 MMR rate increments for asciminib as 3L of treatment to 1L of treatment for CML-CP patients.

Based on the asciminib pivotal study A2301 in 3L+ CML-CP patients, the MMR rate at 6 months was observed to be 25.5%. This rate was then extrapolated to give us an assumed MMR rate at 12 months (48 Weeks) of approximately 30-33%. Further extrapolating for response rate

increments from 3L to 1L (by ~20% as observed for other TKIs) we arrived at our assumption of MMR rate at Week 48 of 52.5% asciminib in 1L.

Table 3-2 MMR rates across lines of treatment

TKI	3L+ Efficacy (MMR @6mo)	2L Efficacy (MMR @12mo)	1L Efficacy (MMR @12mo)
Bosutinib	~15%	~30-35%	~45%
Dasatinib	n/a	~30-35%	~45%
Nilotinib	n/a	~30-35%	~45%
Imatinib	n/a	n/a	~25-30%

Khoury et al, 2012; Hochhaus, 2020; Gambacorti-Passerini, 2014; Cortes 2018; Cortes 2012; Breummendorf 2014; Shah et al 2010; Kantarjian. 2012; Kantarjian et al, 2011; Saglio, 2010

The proportions of participants that will achieve the primary Endpoint of **MMR at Week 48** are thus assumed to be:

- 0.525 for asciminib
- 0.28 for imatinib
- 0.45 for any 2G TKI

A random effects meta-analysis of MMR by Week 96 from the published data gives us estimated proportions (95% CI) of participants that achieve MMR by Week 96 for imatinib of 0.49 (0.44, 0.55), and for 2G TKIs of 0.66 (0.61, 0.71). A random effects meta-analysis of the difference in rate of MMR by Week 96 and MMR at/by Week 48, from the published data gives us estimated proportions (95% CI) for imatinib of 0.21 (0.17, 0.25), and for 2G TKIs of 0.21(0.17, 0.26). Based on internal data for ENESTnd the difference between MMR by Week 96 and MMR at week 48 is approximately 6% for imatinib and approximately 9% for nilotinib.

The proportions of participants that will achieve the primary Endpoint of **MMR at Week 96** are thus assumed to be:

- 0.635 for asciminib
- 0.43 for imatinib
- 0.56 for any 2G TKI

The study design plans for a 50% versus 50% enrollment of imatinib versus the 2G TKIs into the control arm.

Therefore, the proportion of participants that will achieve the primary Endpoint of MMR at Week 48 is assumed to be 0.365 for the Investigator selected TKI arm. This proportion is calculated as a weighted average of the 50% of participants on imatinib with the assumed proportion that are in MMR at Week 48 of 0.28, and 50% participants on a 2G TKI with the assumed proportion that are in MMR at Week 48 of 0.45.

Similarly, the proportion of participants that will achieve the primary Endpoint of MMR at Week 96 is assumed to be 0.495 for the Investigator selected TKI arm. This proportion is calculated as a weighted average of the 50% of participants on imatinib with the assumed proportion that are in MMR at Week 96 of 0.43, and 50% participants on a 2G TKI with the assumed proportion that are in MMR at Week 96 of 0.56.

Adjustment for dropouts is not made since any participant that discontinues study treatment prior to being in MMR at Week 48 will be considered a non-responder for the end-point of MMR at Week 48; and any participant that discontinues study treatment prior to being in MMR at Week 96 will be considered a non-responder for the end-point of MMR at Week 96.

The power for rejecting at least one of the two multiple primary hypotheses, tested according to the graphical gatekeeping strategy described in Section 2.8 of the current document is computed via simulations in R version 3.6.1 using the package "gMCP" (Rohmeyer K, Klinglmueller F (2020)). The underlying distribution of test statistics in power simulations is assumed to be multivariate normal with the correlation matrix that is a block diagonal matrix (with correlation between the test statistics for H1 and H2=√0.5 and correlation between the test statistic for H3 and H4=√0.5). The non-centrality parameters are the standardized test statistics computed under the alternative assumptions.

Based on a 1-sided 2.5% level of significance, **with 402 participants** and 1:1 randomization ratio between arms (i.e., 201 participants in the asciminib arm and 201 participants in the Investigator selected TKI arm) we have **94.6% power to reject at least one of the null hypotheses from the primary family (F_1={H_1, H_2})**. At this sample size, the **local power** to reject the null hypothesis for **H1 is 88.5%** and **local power** to reject the null hypothesis for **H2 is 92.7%**.

This sample size is based on 10000 simulations and from the non-centrality parameters computed from the assumptions that:

- Asciminib leads to an increase over Investigator selected TKI, in the proportion of participants that are in MMR at Week 48, from 0.365 to 0.525, which corresponds to an odds ratio of 1.92 (for asciminib versus Investigator selected TKI).
- Asciminib leads to an increase over imatinib, in the proportion of participants that are in MMR at Week 48, from 0.28 to 0.525 which corresponds to an odds ratio of 2.84 (for asciminib versus imatinib).
- Asciminib leads to an increase over Investigator selected TKI, in the proportion of participants that are in MMR at Week 96, from 0.495 to 0.635, which corresponds to an odds ratio of 1.77 (for asciminib versus Investigator selected TKI).
- Asciminib leads to an increase over imatinib, in the proportion of participants that are in MMR at Week 96, from 0.43 to 0.635 which corresponds to an odds ratio of 2.31 (for asciminib versus imatinib).

The secondary endpoint of Time to discontinuation due to AE will be assessed based on the sample size as defined for the primary endpoints. The graphical gatekeeping procedure described in Section 12.6 (in CSP or Section 2.8 of the current document), ensures preservation of the overall one-sided type I error at 0.025.

Table 3-3 Sensitivity of power to changes in the MMR assumptions (N=402)

MMR at Week 48				MMR at Week 96				Power to rej. At lea st one of H1 or H2	Local power to reject H1	Local power to reject H2	Local power to reject H3	Local power to reject H4
Asc.	Inv. Selecte d TKI	2G TKIs	Ima.	Asc.	Inv. Selecte d TKI	2G TKIs	Ima.					
0.52 5	0.365	0.45 0	0.28 0	0.63 5	0.495	0.56 0	0.43 0	94.56 %	88.46 %	92.73 %	65.70 %	74.43 %
0.52 5	0.400	0.52 0	0.28 0	0.63 5	0.530	0.63 0	0.43 0	92.13 %	69.72 %	91.69 %	41.60 %	67.52 %
0.48 0	0.365	0.45 0	0.28 0	0.59 0	0.495	0.56 0	0.43 0	80.61 %	62.13 %	78.53 %	26.59 %	41.30 %
0.50 0	0.365	0.45 0	0.28 0	0.61 0	0.495	0.56 0	0.43 0	88.26 %	75.96 %	86.11 %	43.76 %	57.20 %
Asc.=asciminib; Inv.=investigator; 2G TKIs = second generation TKIs (nilotinib/ dasatinib/ bosutinib)												

4 Change to protocol specified analyses

Not Applicable.

5 Appendix

5.1 Imputation rules

5.1.1 Study drug

The following rules should be used for the imputation of the dose end date for a given study treatment.

Scenario 1: If the dose end date is completely missing and there is no EOT page and no death date, the participant is considered as on-going:

- The participant should be treated as on-going, and the cut-off date should be used as the dose end date.

Scenario 2: If the dose end date is completely missing and the EOT page is available:

The EOT completion date should be used.

- After imputation, compare the imputed end date with start date of treatment, if the imputed date is < start date of treatment: **Use the treatment start date**

Participants with missing start dates are to be considered missing for all study treatment component related calculations and no imputation will be made. If start date is missing then end date should not be imputed.

5.1.2 AE, Concomitant medication and safety assessment date imputation

The imputations specified in this section are only used for analyses of time to and duration of AEs and ConMeds.

Table 5-1 Imputation of start dates for AEs, ConMeds and Assessments (eg: VS, LB, EG)

Missing Element	Rule
day, month, and year	No imputation will be done for completely missing dates

day, month	• If available year = year of study treatment start date then
	  • If stop date contains a full date and stop date is earlier than study treatment start date then set start date = 01JanYYYY
	  • Else set start date = study treatment start date.
	• If available year > year of study treatment start date then 01JanYYYY
	• If available year < year of study treatment start date then 01JulYYYY

day	• If available month and year = month and year of study treatment start date then
	  • If stop date contains a full date and stop date is earlier than study treatment start date then set start date= 01MONYYYY.
	  • Else set start date = study treatment start date.
	• If available month and year > month and year of study treatment start date then 01MONYYYY
	• If available month and year < month year of study treatment start date then 15MONYYYY

Table 5-2 Imputation of end dates for AEs, ConMeds and Assessments (eg: VS, LB, EG)

Missing Element	Rule (*=last treatment date plus 30 days not > (death date, cut-off date, withdrawal of consent date))
day, month, and year	Completely missing end dates (incl. ongoing events) will be imputed by the end date of the on-treatment period*
day, month	If partial end date contains year only, set end date = earliest of 31DecYYYY or end date of the on-treatment period *
day	If partial end date contains month and year, set end date = earliest of last day of the month or end date of the on-treatment period*

Any AEs, ConMeds or Assessments with partial/missing dates will be displayed as such in the data listings.

Any AEs, ConMeds or Assessments which are continuing as per data cut-off will be shown as ‘ongoing’ rather than the end date provided.

5.1.3 Imputation for missing CML disease diagnosis date

Missing CML disease diagnosis date will be imputed only for missing day, which is to be replaced as 15th of the (known) month. When the month or the year is also missing, the diagnosis date is considered missing.

5.2 AEs coding/grading

Adverse events are coded using the latest available version of Medical Dictionary for Regulatory Activities (MedDRA) terminology.

AEs will be assessed according to the Common Terminology Criteria for Adverse Events (CTCAE) version 5.

The CTCAE represents a comprehensive grading system for reporting the acute and late effects of cancer treatments. CTCAE grading is by definition a 5-point scale corresponding to mild, moderate, severe, life threatening, and death. This grading system inherently places a value on the importance of an event, although there is not necessarily proportionality among grades (a grade 2 is not necessarily twice as bad as a grade 1).

5.3 Laboratory parameters derivations

Grade categorization of lab values will be assigned programmatically as per NCI Common Terminology Criteria for Adverse Events (CTCAE) version 5. The calculation of CTCAE grades will be based on the observed laboratory values only, clinical assessments will not be taken into account. The criteria to assign CTCAE grades are given in Novartis internal criteria for CTCAE grading of laboratory parameters (embedded below). The latest available version of the document based on the underlying CTCAE version v5 at the time of analysis will be used. For laboratory tests where grades are not defined by CTCAE v5, results will be graded by the low/normal/high (or other project-specific ranges, if more suitable) classifications based on laboratory normal ranges.

A severity grade of 0 will be assigned for all non-missing lab values not graded as 1 or higher. Grade 5 will not be used. For laboratory tests that are graded for both low and high values, summaries will be done separately and labelled by direction, e.g., sodium will be summarized as hyponatremia and hypernatremia.

Further derivation of laboratory parameters might be required for CTCAE grading. For instance, corrected calcium can be derived using the reported total calcium value and albumin at the same assessment using the following formula:

Corrected Calcium (mmol/L) = Calcium (mmol/L) +0.02 (40 – [Albumin (g/L)]

For calculation of laboratory CTCAE grades 0 and 1, the normal range for derived corrected calcium is set to the same limits (in mmol/L) as for calcium. CTCAE grades for the corrected calcium will be assigned as described above for grading.

5.3.1 Hematology

Immature cells (promyelocytes, myelocytes, metamyelocytes and blasts) will not be displayed in shift tables and will only be listed.

Immature cells are manually counted only if anomalies are detected during the automatic testing. Therefore, if an automatic testing has been performed but no data is transferred for immature cells, then it is assumed that no immature cells exist, and their values can be imputed to 0. Note that there should not be any imputation in case the automatic testing has not been performed or the test of immature cells is present with missing value in the database. (This would mean the test should have been performed but could not).

CTCAE grading for blood differentials is based on absolute values. However, this data may not be reported as absolute counts but rather as percentage of white blood cells (WBC).

If laboratory values are provided as '<X' (i.e., below limit of detection) or '>X', prior to conversion of laboratory values to SI unit, these numeric values are set to X.

The following rules will be applied to derive the WBC differential counts when only percentages are available for a certain differential:

Differential count = (WBC count) * (percentage value / 100).

For example, suppose WBC differential percentage for neutrophil is known to be 20%, then the neutrophil count is calculated as WBC count * 20 / 100.

The following rules will be applied to derive the WBC differential percentages when only differential counts are available for a certain differential:

Percentage value = (differential count * 100) / WBC count.

CTCAE grades for the derived absolute WBC differential counts (neutrophils, lymphocytes) will be assigned as described above for grading.

5.3.2 Biochemistry

In order to avoid double reporting of the same information, all available values for BUN and UREA will be reported under the parameter name BUN (mmol/L) in listing using the following conversion rule: UREA (mmol/L)=2.14 BUN (mmol/L) (Lamb E et al., 2012).

5.3.3 Molecular response

The BCR::ABL ratio in IS % provided by the central laboratory will be use in the analyses. However, to calculate the fold change in BCR::ABL1/ABL used to derive the loss of MMR, MR4.0, and MMR4.5 criteria, in case the BCR::ABL number of copies in the denominator is reported as 0, then the value will be replaced by 1, and the BCR::ABL ratio will be calculated as

BCR::ABL ratio (IS) (in %) = (number of BCR::ABL copies / number of control gene ABL copies) * conversion factor * 100,

where the lab conversion factor (Branford and Hughes, 2006) for this assay is 1.

5.3.3.1 Molecular response categories

Molecular response based on BCR::ABL ratio is categorized as follows:

- 10% < BCR::ABL ratio (IS)
- 1% < BCR::ABL ratio (IS) ≤ 10%
- 0.1% < BCR::ABL ratio (IS) ≤ 1%
- 0.01% < BCR::ABL ratio (IS) ≤ 0.1%
- 0.0032% < BCR::ABL ratio (IS) ≤ 0.01%
- BCR::ABL ratio (IS) ≤ 0.0032%
- No evidence of typical transcript

For participants with no evidence of typical transcript at the time of screening, they are categorized in a separate category, and will not be counted toward any of the above categories.

5.3.3.1.1 BCR::ABL ≤ 1%

As an endpoint BCR::ABL≤ 1% is considered a binary variable with participants achieving BCR::ABL≤ 1% grouped as 'responders' and participants not achieving BCR::ABL≤ 1%, or participants with missing PCR evaluations, or participants with no evidence of typical transcript grouped as 'non-responders'.

5.3.3.2 Major molecular response (MMR)

Major molecular response (MMR) is defined as a value of ≤ 0.1% of BCR::ABL ratio on the international scale (IS). This endpoint corresponds to a ≥ 3 log reduction in BCR::ABL transcripts from a standardized baseline value for untreated CML patients, which was established in the IRIS study (STI5710106). MMR will be considered as a binary variable with participants achieving MMR grouped as 'responders' and participants not achieving MMR, or participants with missing PCR evaluations, or participants with no evidence of typical transcript grouped as 'non-responders'.

5.3.3.3 Loss of MMR and confirmed loss of MMR

Loss of MMR is defined as an increase in BCR::ABL1/ABL to > 0.1% by international scale (IS) in association with a ≥ 5-fold rise in BCR::ABL1/ABL from the lowest value achieved up to that time point on study treatment and replicated by a second analysis of the same sample. Loss of MMR must be confirmed by a subsequent sample analysis within 4-6 weeks, showing loss of MMR associated with a ≥ 5-fold rise in BCR::ABL1/ABL from the lowest value achieved up to that time point on study treatment.

If there is any assessment in between indicating a BCR::ABL ratio of ≤ 0.1% or a < 5-fold increase in BCR::ABL ratio from the lowest value achieved up to that time point on study treatment, then the initial indication of loss of MMR cannot be confirmed. However, an assessment indicating (unconfirmed) loss of MMR will be considered as confirmed loss of MMR if the participant has had loss of CHR or loss of complete cytogenetic response (CcyR) after the achievement of MMR. CML-related death or progression to AP or BC will be

considered as confirmed loss of MMR in any case (if they occurred on treatment) (given that the participant achieved prior MMR).

Thus, to summarize:

- Loss of MMR is defined as *BCR::ABL1* level (IS) > 0.1% in association with a ≥ 5-fold rise in BCR::ABL1 from the lowest value achieved on study treatment and replicated by a second analysis of the same sample.
- Confirmed Loss of MMR is defined as a loss of MMR confirmed by a subsequent sample analysis taken after an interval of not less than 4 weeks and not more than 6 weeks unless associated with loss of CHR or loss of CCyR or progression to AP/BC (see Section 5.3.7) or CML related death (see Section 5.3.8).

5.3.3.4 MR4.0 and Loss of MR4.0

MR4.0 is defined as a value of ≤ 0.01% of BCR::ABL ratio on the IS (this corresponds to a ≥ 4 log reduction in BCR::ABL transcripts from a standardized baseline value for untreated CML patients). MR4.0 will be considered as a binary variable with participants achieving MR4.0 grouped as 'responders' and participants not achieving MR4.0, or participants with missing PCR evaluations, or participants with no evidence of typical transcript grouped as 'non-responders'.

Loss of MR4.0 is defined as *BCR::ABL1* IS > 0.01% confirmed by subsequent sample analysis within 12 weeks showing loss of MR4.0 associated with a ≥ 5-fold rise in BCR::ABL1 from the lowest value achieved on study treatment, unless it is associated with loss of CHR or loss of CCyR or progression to AP/BC or CML-related death.

5.3.3.5 MR4.5 and Loss of MR4.5

MR4.5 is defined as a value of ≤ 0.0032% of BCR::ABL ratio on the IS (this corresponds to a ≥ 4.5 log reduction in BCR::ABL transcripts from a standardized baseline value for untreated CML patients). MR4.5 will be considered as a binary variable with participants achieving MR4.5 grouped as 'responders' and participants not achieving MR4.5, or participants with missing PCR evaluations, or participants with no evidence of typical transcript grouped as 'non-responders'.

Loss of MR4.5 is defined as *BCR::ABL1* IS > 0.0032% confirmed by subsequent sample analysis within 12 weeks showing loss of MR4.5 associated with a ≥ 5-fold rise in BCR::ABL1 from the lowest value achieved on study treatment, unless it is associated with loss of CHR or loss of CCyR or progression to AP/BC or CML-related death.

5.3.4 Cytogenetic response

Cytogenetic response is categorized as follows using bone marrow aspiration (a review of a minimum of 20 metaphases is required):

- Complete response (CCyR): 0% Philadelphia chromosome positive (Ph+) metaphases
- Partial response (PCyR): >0 to 35% Ph+ metaphases
- Major response (MCyR = CCyR + PCyR): 0 to 35% Ph+ metaphases

- Minor response (mCyR): >35 to 65% Ph+ metaphases
- Minimal response: >65 to 95% Ph+ metaphases
- None: >95 to 100% Ph+ metaphases.

5.3.4.1 Loss of CCyR

Loss of CCyR is defined as an increase in the Ph+ bone marrow cells to > 0%. Loss of CCyR must either be confirmed by a subsequent bone marrow aspirate at least 4 weeks apart showing loss of CCyR or have led to treatment discontinuation because of lack of efficacy. In addition, CML-related death or progression to AP or BC will be considered as loss of CCyR in any case (if they occurred on treatment).

5.3.5 Hematologic response

5.3.5.1 Complete hematologic response (CHR)

CHR is defined when all of the following criteria are present at any assessment which is confirmed by another assessment at least after 4 weeks:

- White blood cells (WBC) count < 10×10^9 /L
- Platelet count < 450×10^9 /L
- Basophils < 5%
- No blasts and promyelocytes in peripheral blood
- Myelocytes + metamyelocytes < 5% in peripheral blood
- No evidence of extramedullary disease, including spleen and liver. As extramedullary disease is evaluated less frequently than hematology, the results of these evaluations are carried forward until the next assessment (unless extramedullary disease was not present at the current assessment but present at the next).

The assessment is not considered CHR, if there are any values indicative of CML in AP or BC (i.e., by blasts in bone marrow if available). The information used for hematological assessment will be obtained from the laboratory, extramedullary and bone marrow data (if available), all merged by participant and date. To accommodate for missing parameters, specific laboratory results may be carried forward up to 14 days such that assessments performed within a two-week period can be combined into one complete evaluation of hematological response. A value will be carried forward for no more than up to the subsequent valid assessment of the respective laboratory parameter. If even after applying this carry-forward algorithm, any of the above laboratory parameters is not available at a given assessment date, the response assessment will be considered missing, unless any of the available values (including those carried forward) indicates that there is no response in which case the assessment will be ‘No response’.

For confirmation of CHR, both the initial CHR as well as the confirming assessment (at least 4 weeks after the initial assessment) must satisfy all the criteria mentioned above and no assessment in between indicates ‘No response’. The terms “confirmed CHR” and “CHR” are used as synonymous given that the definition of CHR mentioned above already includes a requirement for confirmation.

Hematologic response will be assessed by CBC and physical examination at each visit.

5.3.5.2 Loss of CHR

Loss of CHR is defined by meeting any of the following:

- WBC count > 20 × 10^9/L
- Platelet count ≥ 600 × 10^9/L
- Appearance of blasts or promyelocytes in peripheral blood
- Appearance of myelocytes + metamyelocytes ≥ 5% in peripheral blood
- Progressive splenomegaly refractory to therapy (i.e., ≥ 5cm below left intercostal margin)

In addition, CML related death or progression to AP or BC will be considered as loss of CHR in any case (if they occurred on treatment).

5.3.6 Treatment Failure per ELN criteria

The following events will constitute 'treatment failure' based on ELN criteria (Hochhaus et al., 2020)

- BCR::ABL1 ratio (IS) > 10% at 3 months after initiation of therapy if confirmed within 1–3 months
- BCR::ABL1 ratio (IS) > 10% at 6 months after initiation of therapy
- BCR::ABL1 ratio (IS) > 1% at 12 months after initiation of therapy
- BCR::ABL1 ratio (IS) > 1% any time after 12 months after initiation of therapy,
- Detection of a BCR::ABL1 mutation which can potentially cause resistance to study treatment (asciminib or IS-TKI) at any time after initiation of study treatment. [Per ELN treatment guidelines for known mutations resistant to specific TKI, (Hochhaus et al. 2020)].

5.3.6.1 Treatment failure based on ELN criteria

Since the study assessment visits are scheduled with "week" as the time unit, in order to report number and percentage of subjects experiencing treatment failure per ELN criteria, timepoints in weeks associated with the milestones are defined as:

- Criterion 1: BCR::ABL1 ratio (IS) > 10% at 12 weeks after randomization if confirmed within 4–12 weeks
- Criterion 2: BCR::ABL1 ratio (IS) > 10% at 24 weeks after randomization
- Criterion 3: BCR::ABL1 ratio (IS) > 1% at 48 weeks after randomization
- Criterion 4: BCR::ABL1 ratio (IS) > 1% any time after 48 weeks after randomization

The approximation of 6 months and 12 months with 24 weeks and 48 weeks, respectively, are conservative since participants are on treatment for a shorter period.

The criteria related to the BCR::ABL1 mutation will not be used.

To determine at which timepoint the ELN treatment failure criteria is met for a participant, all BCR::ABL assessments from central lab, scheduled and unscheduled, are utilized. Per protocol

the allowed visit window for any given visit is +/- 3 days from target day of assessment. As defined in the protocol, the day of randomization is Day 1:

- If a BCR::ABL assessment is above 10% at or after day 81 (the lower end of the protocol allowed assessment window for week 12 visit) but on or before day 164 (the day before the protocol allowed assessment window for week 24 visit), and the value is confirmed within 4 to 12 weeks (or 28 to 84 days, inclusive), this assessment is considered to satisfy Criterion 1.
- If a BCR:: ABL assessment is above 10% at or after day 165 (the lower end of the protocol allowed assessment window for week 24 visit) but on or before day 332 (the day before the protocol allowed assessment window for week 48 visit), this assessment is considered to satisfy Criterion 2.
- If a BCR::ABL assessment is above 1% at or after day 333 (the lower end of the protocol allowed assessment window for week 48 visit) but on or before day 416 (the day before the protocol allowed assessment window for week 60 visit), this assessment is considered to satisfy Criterion 3.
- If a BCR::ABL assessment is above 1% after day 417 (the lower end of the protocol allowed assessment window for week 60 visit), this assessment is considered to satisfy Criterion 4.

The date when the assessment associated with a treatment failure has been collected is the date the participant is considered to have failed the treatment per ELN guideline. If a participant has multiple assessments that satisfy the treatment failure criteria, the earliest is considered the date of treatment failure.

All subsequent efficacy data after a treatment failure are no longer used for the primary and the secondary estimands, nor for any of the efficacy endpoints outlined in the previous sections.

5.3.7 CML progression to accelerated phase (AP) or blast crisis (BC)

For the evaluation of CML progression to AP or BC, the following criteria will be used. Accelerated phase (AP) is defined by any of the following:

- ≥ 15% blasts in the peripheral blood or bone marrow aspirate, but < 30% blasts in both the peripheral blood and bone marrow aspirate
- ≥ 30% blasts plus promyelocytes in peripheral blood or bone marrow aspirate
- ≥ 20% basophils in the peripheral blood
- Thrombocytopenia (<100 x 10^9/L) that is unrelated to therapy

Blast crisis (BC) is defined by any of the following:

- ≥ 30% blasts in peripheral blood or bone marrow aspirate
- Appearance of extramedullary involvement other than hepatosplenomegaly proven by biopsy (i.e., chloroma).

Any value of AP or BC within the first 4 weeks of study treatment is not considered as progression to AP/BC unless the patient discontinues study treatment due to progression or unsatisfactory therapeutic effect within the first 8 weeks.

Should a participant be classified as both AP and BC, BC takes precedence as it is a more severe state of the disease.

5.3.8 CML-related deaths

CML-related death is considered as any death during treatment or follow-up (safety or survival)

- if the principal cause of death is marked as "study indication" in the eCRF by the investigator, or
- if the death occurred subsequent to documented progression to AP/BC and the cause of death is reported as "unknown" or not reported by the investigator.

With respect to the second bullet, as "unknown" cause of death will be coded to the Medical Dictionary for Regulatory Activities (MedDRA) preferred term 'Death', this MedDRA coding will be used in the derivation of CML-related death.

5.4 Statistical models and analyses

5.4.1 Analysis of the primary endpoint

5.4.1.1 Testing of the null hypotheses

The null hypotheses for the primary endpoint (cf. Section 2.5.2) will be tested using the one-sided Cochran-Mantel-Haenszel chi-square tests. The test will be stratified by the randomization stratification factors, i.e., the ELTS risk group (low, intermediate, high) and the PRS-TKI, for testing H_{10}; and only ELTS risk group for testing H_{20}.

The test result can be obtained by SAS procedure FREQ with CMH option in the TABLES statement. The raw (un-adjusted) two-sided p-values (denoted as p) corresponding to the CMH test for "general association" will be used to derive the one-sided p-value as the following:

- If the MMR response rate in the asciminib arm is higher than or equal to that in the IS-TKI arm, then the un-adjusted one-sided p-value for the test is p/2.
- If the MMR response rate is higher in the IS-TKI arm than in the asciminib arm, then the un-adjusted one-sided p-value is 1-p/2.

The family-wise type I error rate will be controlled at 2.5% alpha level via the gate-keeping procedure as described in Section 2.8, where the rejection boundaries for the raw (unadjusted) p-values are also given. In addition, adjusted p-values are produced using the R package gMCP in DaVinci RStudio.

The adjusted one-sided p-values will be displayed with other information for MMR at week 48.

5.4.1.2 Estimating the response rates and the difference in response rates

Unstratified proportions (or rates) of responders in each treatment arm, and in each the two strata within the IS-PRS arm along with their two-sided Clopper-Pearson 95% CI's will be reported. SAS procedure FREQ with the EXACT statement for a one-way table is implemented for this purpose.

Two estimates of the difference in the response rates will be generated. The first is the unstratified difference in the response rates and its Wald's 95% CI. These can be obtained by SAS procedure FREQ with RISKDIFF option in the TABLES statement with the default METHOD=WALD and VAR=SAMPLE.

The second is the Mantel-Haenszel estimate of common risk difference and its 95% CI, which can be obtained with the SAS procedure FREQ with RISKDIFF(COMMON) option in the TABLES statement. Since this is a stratified estimate, the stratification factor(s) needs to appear as in the TABLE statement. The estimated difference and its CI are taken from the "Confidence Limits for the Common Risk Difference" output table under Mantel-Haenszel method.

5.4.2 Analysis of the key secondary endpoints

Identical analyses as described in Section 5.4.1 will be conducted for the key secondary endpoints and their corresponding null hypotheses H_{30} and H_{40}.

The adjusted one-sided p-values will be displayed with other information for MMR at week 96.

5.4.3 Analysis of the secondary safety endpoint and other time-to-event endpoints

Cumulative incidence

The cumulative incidence proportion (CIP) will be estimated using SAS procedure LIFETEST or PHREG with EVENTCODE=Code for event of interest (e.g., 1) as option in the MODEL statement, whereas code=0 for censored subjects and any other code (e.g., code=2) for subjects who dropped out due to a competing risk. The estimated CIP at the defined time points will be presented with 95% CI together with number of subjects with events, number of subjects with competing risks, and number of subjects censored.

Sub-distributional hazard

The sub-distributional hazard model can also be fitted with PHREG in SAS with the EVENTCODE option.

5.4.4 Other analyses

Mantel-Haenszel common odds ratio

To obtain Mantel-Haenszel estimates of the common odds ratio and the corresponding 95% confidence interval in exploratory analyses, it requires SAS procedure FREQ with CMH and RELRISK options in the TABLES statement.

Logistic Regression

Odds ratio will be used as a measure of association between treatment and response in exploratory analyses (Section 2.14). The odds ratio will be derived from the logistic regression model (implemented using SAS procedure LOGISTIC, with treatment specified as an explanatory variable in the CLASS statement) which allows for including not only the

stratification factor but also for adjustments for other covariates (both categorical and continuous). The odds ratio will be presented with 95% Wald confidence limits.

In cases where an exact test has been used to compare response rates, the odds ratio should be determined using exact logistic regression, and the odds ratio presented with exact 95% confidence limits. In these cases, SAS PROC LOGISTIC with EXACTONLY option will be used.

Kaplan-Meier estimates

An estimate of the survival function in each treatment group will be constructed using Kaplan-Meier (product-limit) method as implemented in PROC LIFETEST with METHOD=KM option. The PROC LIFETEST statement will use the option CONFTYPE=LOGLOG. Median survival for each treatment group will be obtained along with 95% confidence intervals calculated from PROC LIFETEST output using the method of [Brookmeyer and Crowley 1982]. Kaplan-Meier estimates of the survival function with 95% confidence intervals at specific time points will be summarized. The standard error of the Kaplan-Meier estimate will be calculated using Greenwood's formula [Collett 1994].

5.4.5 Calculation of exposure-adjusted incidence rate

To adjust for different durations of exposure across treatment arms, the incidence rate (IR) per 100 patient-years of exposure (exposure-adjusted incidence rates of adverse events) will be calculated.

The IR per 100-patient-years is defined as Numerator/Denominator, where

- Numerator = 100 * number of participants with the adverse events of interest (not the number of events; one participant may have more than one event).
- Denominator = patient-years = (among all participants in the population, sum of the duration of exposure (in days) until the first onset of the event of interest, if the participant experienced the event, or until the date of last dose if the participant did not experience the event) / 365.25.

5.5 Calculation of cardiovascular risk scores

5.5.1 Framingham cardiovascular disease risk score

The score for each participant is derived using algorithm given in (D'Agostino et al. 2008). In particular, Table 5-3 and Table 5-4 [cf. Tables 5 and 6 of D'Agostino et al. 2008] are used to determine the risk scores for female participants, and Table 5-5 and Table 5-6 [cf. Tables 7 and Table 8 of D'Agostino et al. 2008] are used for male participants.

Note that high-density lipoprotein cholesterol (HDL-C) and total cholesterol in Table 5-3 and Table 5-5 are in the unit of mg/dL. To convert them to mmol/L, use the following relations:

- 1 mmol/L = 38.67 mg/dL, or
- 1 mg/dL = 0.02586 mmol/L

5.5.1.1 CD points and risk (%) for female participants

For a female participant, a total point of cardiovascular disease (CVD) is calculated by adding up the points associated with each risk factor given in Table 5-3.

Table 5-3 CVD points for women

Points	Age (years)	HDL-C (mg/dL)	Total Cholesterol (mg/dL)	SBP (mm Hg) Not Treated	SBP (mm Hg) Treated	Smoker	Diabetic
-3				< 120			
-2		60+					
-1		50-59			< 120		
0	30-34	45-49	< 160	120-129		No	No
1		35-44	160-199	130-139			
2	35-39	< 35		140-149	120-129		
3			200-239		130-139	Yes	
4	40-44		240-279	150-159			Yes
5	45-49		280+	160+	140-149		
6					150-159		
7	50-54				160+		
8	55-59						
9	60-64						
10	65-69						
11	70-74						
12	75+						

As an example, for a female participant who is 43 years of age (+4) with HDL-C 65 mg/dL (-2), total cholesterol 150 mg/dL (+0), treated SBP 125 mm Hg (+2), smoker (+3), and non-diabetic (+0), her CVD total point is 7.

The total point for a female participant is converted to CDV risks using the conversion in Table 5-4.

Table 5-4 Estimated 10-year CVD risk for women

Points	Risk, %	Points	Risk, %
-2	< 1	10	6.3
-1	1.0	11	7.3
0	1.2	12	8.6
1	1.5	13	10.0
2	1.7	14	11.7
3	2.0	15	13.7
4	2.4	16	15.9
5	2.8	17	18.5
6	3.3	18	21.5
7	3.9	19	24.8
8	4.5	20	28.5
9	5.3	21+	>30

5.5.1.2 CVD points and risk (%) for male participants

For a male participant, a total point of CVD is calculated by adding up the points associated with each risk factor given in Table 5-5.

Table 5-5 CVD points for men

Points	Age (years)	HDL-C (mg/dL)	Total Cholesterol (mg/dL)	SBP (mm Hg) Not Treated	SBP (mm Hg) Treated	Smoker	Diabetic
-2		60+		< 120			
-1		50-59					
0	30-34	45-49	< 160	120-129	< 120	No	No
1		35-44	160-199	130-139			
2	35-39	< 35	200-239	140-159	120-129		
3			240-279	160+	130-139		Yes
4			280+		140-159	Yes	
5	40-44				160+		
6	45-49						
7							
8	50-54						
9							
10	55-59						
11	60-64						
12	65-69						
13							
14	70-74						
15	75+						

The total point for a male participant is converted to CDV risks using the conversion in Table 5-6.

Table 5-6 Estimated 10-year CVD risk for men

Points	Risk, %	Points	Risk, %
-3 or less	< 1	8	6.7
-2	1.1	9	7.9
-1	1.4	10	9.4
0	1.6	11	11.2
1	1.9	12	13.2
2	2.3	13	15.6
3	2.8	14	18.4
4	3.3	15	21.6
5	3.9	16	25.3
6	4.7	17	29.4
7	5.6	18+	>30

5.5.1.3 CVD risk categories

Framingham risk categories are assigned for each participant (male or female) as the following (Jahangiry 2017):

- Low: risk < 10%
- Intermediate: 10% ≤ risk < 20%
- High: risk ≥ 20%

5.5.2 Charlson Comorbidity Index (CCI)

5.5.2.1 Derivation of the index

Each participant is assigned a score (or index) based on the presence or absence of certain comorbidities, or the severities of them. The total of the individual item scores is the CCI.

The scoring algorithm can be found at the website (last accessed on 13th September 2023):

http:www.mdcalc.com/calc/3917/charlson-comorbidity-index-cci

The algorithm is reproduced in Table 5-7.

Table 5-7 Charlson Comorbidity Index

Factor or comorbidity	Category	Score
Age	< 50 years	0
	50-59 years	+1
	60-69 years	+2
	70-79 years	+3
	≥ 80 years	+4
Myocardial infarction (History of definite or probable MI (EKG changes and/or enzyme changes))	No	0
	Yes	+1
CHF (Exertional or paroxysmal nocturnal dyspnea and has responded to digitalis, diuretics, or afterload reducing agents)	No	0
	Yes	+1
Peripheral vascular disease (Intermittent claudication or past bypass for chronic arterial insufficiency, history of gangrene or acute arterial insufficiency, or untreated thoracic or abdominal aneurysm (≥6 cm))	No	0
	Yes	+1
CVA or TIA (History of a cerebrovascular accident with minor or no residua and transient ischemic attacks)	No	0
	Yes	+1
Dementia (Chronic cognitive deficit)	No	0
	Yes	+1
COPD	No	0
	Yes	+1
Connective tissue disease	No	0
	Yes	+1
Peptic ulcer disease	No	0
	Yes	+1

Factor or comorbidity	Category	Score
(Any history of treatment for ulcer disease or history of ulcer bleeding)		
Liver disease (Severe = cirrhosis and portal hypertension with variceal bleeding history, moderate = cirrhosis and portal hypertension but no variceal bleeding history, mild = chronic hepatitis (or cirrhosis without portal hypertension))	None Mild Moderate to severe	0 +1 +3
Diabetes mellitus	None or diet-controlled Uncomplicated End-organ damage	0 +1 +2
Hemiplegia	No Yes	0 +2
Moderate to severe CKD (Severe = on dialysis, status post kidney transplant, uremia, moderate = creatinine >3 mg/dL (0.27 mmol/L))	No Yes	0 +2
Solid tumor	None Localized Metastatic	0 +2 +6
Leukemia	No Yes	0 +2
Lymphoma	No Yes	0 +2
AIDS	No Yes	0 +6

For example, for a participant who is 62 years of age (+2), with a history of myocardial infarction (+1), and severe living disease (+3), this participant's CCI is 6.

5.5.2.2 Estimated 10-year survival

The estimated 10-year survival is calculated as the following two steps [Charlson 1987]:

1. $x = \exp(0.9 \times CCI)$
2. estimated 10-year survival = 0.983^x, i.e., 0.983 raised to the power of x.

For example, for the participant whose CCI is 6:

- $x = \exp(0.9 \times 6) = 221.4$
- 0.983^221.4 = 0.022 = 2.2%

6 Reference

Alves R, Gonçalves AC, Rutella S, et al. (2021). "Resistance to Tyrosine Kinase Inhibitors in Chronic Myeloid Leukemia—From Molecular Mechanisms to Clinical Relevance", Cancers, 13:4820. https://doi.org/10.3390/cancers13194820.

Balduzzi S, Rücker G, Schwarzer G (2019) "How to perform a meta-analysis with R: a practical tutorial". Evid Based Ment Health, 153-160.

Branford S and Hughes T. (2006). "Diagnosis and monitoring of chronic myeloid leukemia by qualitative and quantitative RT-PCR". Methods Mol Med; 125:69-92.

Bretz F, Maurer W, Brannath W, et al. (2009). "A graphical approach to sequentially rejective multiple test procedures". Stat Med, 586-604.

Bretz F, Posch M, Glimm E, et al. (2011). "Graphical approaches for multiple comparison procedures using weighted Bonferroni, Simes, or parametric tests". Biom J, 894-913.

Brookmeyer, R and Crowley J. (1982). "A confidence interval for the median survival time". Biometrics 38:2941.

Charlson ME, Pompei P, Ales KL, et al. (1987). "A new method of classifying prognostic comorbidity in longitudinal studies: Development and validation." Journal of Chronic Diseases, 40(5):373-383. https://doi.org/10.1016/0021-9681(87)90171-8.

Clopper CJ and Pearson ES. (1934). "The Use of Confidence or Fiducial Limits Illustrated in the Case of the Binomial", Biometrika 26:404-413.

Collett (1994). Modelling survival data in medical research. London, Chapman & Hall.

D'Agostino RB Sr, Vasan RS, Pencina MJ, et al. (2008). "General cardiovascular risk profile for use in primary care: the Framingham Heart Study". Circulation, 117(6):743-753. doi: 10.1161/CIRCULATIONAHA.107.699579.

Hochhaus A, Baccarani M, Silver RT, et al. (2020), "European LeukemiaNet 2020, recommendations for treating chronic myeloid leukemia". Leukemia, 966-984.

Jahangiry L, Farhangi MA, and Rezaei F (2017). "Framingham risk score for estimation of 10-years of cardiovascular diseases risk in patients with metabolic syndrome". Journal of Health, Population and Nutrition; 36:36. DOI 10.1186/s41043-017-0114-0.

Lauseker M, Hanfstein B, Haferlach C, et al. (2014). "Equivalence of BCR-ABL transcript levels with complete cytogenetic remission in patients with chronic myeloid leukemia in chronic phase". J Cancer Res Clin Oncol; 140:1965-9.

Lamb E, et al. (2012). "Kidney Function Tests In Elsevier Health Sciences" (Ed), Tietz Textbook of Clinical Chemistry and Molecular Diagnostics (Chapter 25). Also retrieved from https://books.google.es/books?id=BBLRUI4aHhkC&lpg=PA685&hl=es&pg=PA685#v=onepage&q&f=false on 13th October 2021.

NCCN (2020). National Comprehensive Cancer Network CML treatment guidelines ver 3.

Rohmeyer K and Klinglmueller F. (2020). "*gMCP: Graph Based Multiple Test Procedures"*. R package version 0.8-15.