NOVARTIS

Clinical Development

INC424

CINC424H12201 / NCT04097821

A randomized, open-label, phase I/II open platform study evaluating safety and efficacy of novel ruxolitinib combinations in myelofibrosis patients

Statistical Analysis Plan (SAP) – Amendment 2

Author:	Trial Statistician
Document type:	SAP Documentation
Document status:	Final v2.0
Release date:	2-Jul-2024
Number of pages:	71

Document History – Changes compared to previous final version of SAP

Date	Time point	Reason for update	Outcome for update	Section and title impacted (Current)
12-Mar-2020	In line with Protocol Amendment #2	Creation of final version	N/A - First version	NA
30-Nov-2022	In line with Protocol Amendment #8	Updated according to PA#8 and reflected the main changes made by PA #3 to 7	Provided the background of the SAP amendment in the introduction section. Updated study design and, objectives and endpoints according to PA#8. Specified the timepoints and data included in the primary and final CSRs. Introduced extension treatment phase related concepts, including the date of first administration of extension treatment and on-extension treatment period. [redacted] Modified the WOC description for third party data. For assessing COVID-19 impact, included additional analyses for protocol deviations and COVID infected AEs. Defined the exposure concepts for Rineterkib and NIS793 that were added in PA#4. Included the analyses for dose escalation and safety run-in for Rineterkib and NIS793 respectively Noted the analyses that are no longer applicable due to the permanent enrollment halt, such as analyses for Parts2&3, [redacted]. Updated the interim analyses and sample size according to PA#4 Added summary for AEs of eye disorder for rineterkib and listing display for ophthalmology assessments. Removed overall safety period summaries for AESI and death cases. Added analyses for Part 1 SVR and change in TSS MFSAF following the updated secondary objectives from PA7.	Section 1 Introduction Section 2.1 Data analysis general definitions Section 2.3 patient disposition, demographics and other baseline characteristics Section 2.4 Treatments Section 2.5 Analysis of the primary objective Section 2.7 Analysis of secondary efficacy objectives Section 2.8 Safety analyses Section 2.9 Pharmacokinetic endpoints [redacted] Section 3 Sample size calculation

Date	Time point	Reason for update	Outcome for update	Section and title impacted (Current)
2-Jul-2024	In line with Protocol Amendment #9	Updated according to PA#9	Provided the background of the SAP amendment in the introduction section. Updated study design and, objectives and endpoints according to PA#9. Add details about Hy's law.	Section 1 Introduction Section 2.7 Analysis of secondary objectives Section 2.8 Safety Analysis

Table of contents

List of abbreviations

AE	Adverse event
ATC	Anatomical Therapeutic Classification
AUC	Area Under the Curve
bid	bis in diem/twice a day
CSR	Clinical Study report
CTC	Common Toxicity Criteria
CTCAE	Common Terminology Criteria for Adverse Events
DMC	Data Monitoring Committee
FAS	Full Analysis Set
eCRF	Electronic Case Report Form
IVR	Interactive Voice Response
IWR	Interactive Web Response
MedDRA	Medical Dictionary for Drug Regulatory Affairs
NCI	National Cancer Institute
o.d.	Once Daily
OS	Overall Survival
PFS	Progression-Free Survival
PK	Pharmacokinetics
PPS	Per-Protocol Set
PRO	Patient-reported Outcomes
qd	Qua'que di'e / once a day
QoL	Quality of Life
RAP	Report and Analysis Process
RECIST	Response Evaluation Criteria in Solid Tumors
SAP	Statistical Analysis Plan
SOC	System Organ Class
TFLs	Tables, Figures, Listings
WHO	World Health Organization

1 Introduction

This statistical analysis plan (SAP) describes all planned analyses for the clinical study report (CSR) of study CINC424H12201, a randomized, open-label, phase I/II open platform study evaluating safety and efficacy of novel ruxolitinib combinations in myelofibrosis patients. DMC analyses will refer to this study SAP as appropriate.

SAP Amendment 2

This SAP is an amendment of the SAP amendment 1 that was based on study protocol CINC424H12201 Amendment version 08. The content of this SAP is based on protocol Amendment version 09. The main purpose of this amended SAP is to implement the changes made by protocol amendment 9 with details specified in the following paragraph.

The main purpose of the protocol amendment version 09 is to increase the duration of the extension treatment phase for the remaining patients to a maximum of 21 cycles. The enrollment remains halted as per protocol amendment 08.

These changes include adding an extension treatment phase and reduction of study assessments to decrease subject's burden in Part 1. The purpose of the extension phase is to allow access to the investigational drug for ongoing subjects deriving clinical benefit. In addition, definition and recording of Hy's Law has been added to ensure compliance with FDA request for expedited reporting of potential Hy's Law cases.

Notations are remaining in the contents that become not applicable due to permanent enrollment halt. All decisions regarding final analysis, as defined in the SAP document, will be made prior to the final database lock (DBL).

SAP Amendment 1

The SAP amendment 1 is an amendment of the original SAP that was based on study protocol CINC424H12201 Amendment version 01. The content of this SAP is based on protocol Amendment version 08. The main purpose of this amended SAP is to implement the changes made by protocol amendment 8 with details specified in the following paragraph. This SAP amendment also reflects the main changes made by protocol amendments 2 to 7, such as defining an 'overall safety period' for reporting safety data within the long safety follow-up period, adding two novel compounds (LTT462 and NIS793) to Part 1, increasing the number of subjects for Part 2 from 15 subjects to 25 subjects per arm, adding an interim analysis per combination treatment in Part 2, and clarifying the definition of secondary endpoints for evaluating change in MF symptoms and change in spleen size.

1.1 Study design

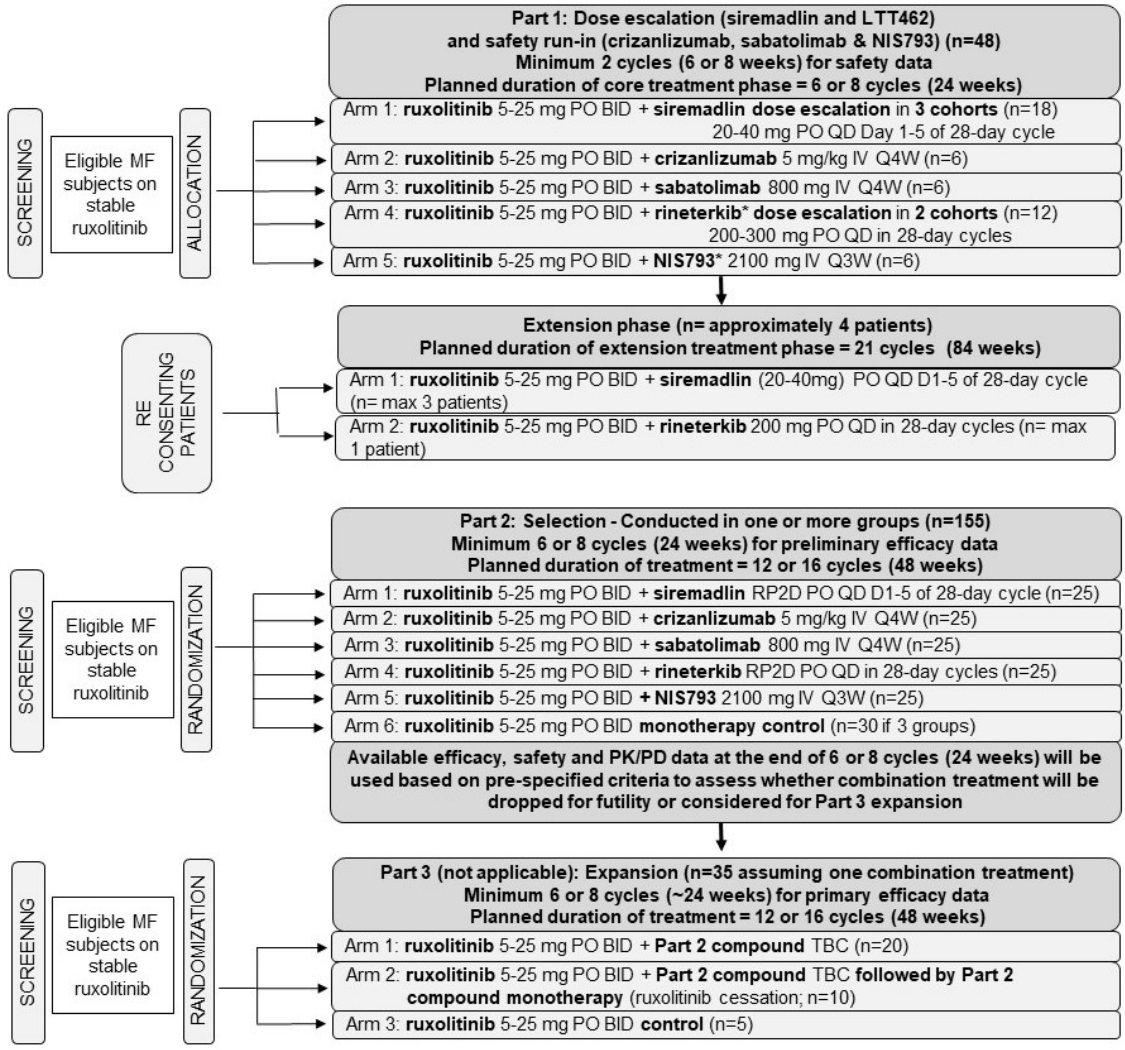
This is an open-label, multi-center, three-part, phase Ib/II open platform study to assess safety and efficacy of ruxolitinib in combination with novel compounds in myelofibrosis patients.

In the context of the permanent enrollment halt and protocol amendment 8 and 9, the study design is revised as follows:

- Part 1: Dose escalation and safety-run-in (recommended Phase II dose confirmation) including a core and an extension treatment phase
    - The end of core treatment phase will occur when amended protocol version 08 is implemented, after a minimum of 24 weeks in the core treatment phase.
    - The extension treatment phase will occur (for ongoing subjects meeting eligibility criteria who are re-consented) from the time when amended protocol version 08 is implemented until the end of extension study treatment.
- Part 2: Selection remained unchanged as the only randomized subject has discontinued prior to amended protocol version 08.
- Part 3: Expansion will not be applicable as it will never be initiated.

An overview of the study design is shown as Figure 1-1.

Figure 1-1 Overview of study design



* Rineterkib and NIS793 added in amendment 4 of the protocol

TBC: To be confirmed

Subjects treated at different dose levels of ruxolitinib will be pooled into a single treatment group for each dose level of the combination partners. A treatment group is defined by both combination partner and the dose level of the combination partner for study subjects.

The primary clinical study report (CSR) is the main primary report of the study. It will include analyses on all subject data when the last subject end the core treatment phase (Part 1) and after all subjects have completed at least, 8 cycles (arms with NIS793) or 6 cycles (all other arms) of treatment (24 weeks) or discontinued earlier (Part 1 and Part 2).

All extension treatment phase data will be reported in the final CSR after all subjects discontinue from the study or complete the extension treatment phase (maximum 21 cycles) whichever comes last. The purpose of the extension phase is to allow access to the investigational drug for ongoing subjects deriving clinical benefit. There will be a maximum of 4 subjects entering extension treatment phase and these 4 subjects will be treated with different combination treatments and/or dose levels.

Due to the permanent enrollment halt, formal analysis for Part 2 and Part 3, including interim analysis are no longer applicable. Subjects in Part 1 will be analyzed separately from the subject in Part 2.

An independent safety Data Monitoring Committee (DMC) will monitor unblinded safety, [redacted], DLT and critical efficacy variables accruing during the trial. The analysis results for the DMC reviews will be provided by Novartis Personnel (Trial Statistician and Trial Programmer).

1.2 Study objectives and endpoints

As the enrollment was permanently halted, not all planned study objectives for Parts 1, 2, and 3 could be completed. All Parts 2 and 3 objectives will not be pursued. [redacted] Table 1-1 shows the applicable objectives and related endpoints following protocol amendment version 09 Table 2-1. The objectives that will not be pursued due to the permannent enrollment halt will not require any analysis, so these are not included in Table 1-1 of this document.

Table 1 - 1 Objectives and related endpoints

Objective(s)	Endpoint(s)
Primary objective(s)	**Endpoint(s) for primary objective(s)**
• To characterize the safety, tolerability, and recommended phase 2 dose (RP2D) of each combination partner used with ruxolitinib (Part 1)	• Incidence and severity of dose limiting toxicity (DLTs) within the first 2 treatment cycles in Part 1 of the study
Secondary objective(s)	**Endpoint(s) for secondary objective(s)**
• To evaluate changes in symptoms of myelofibrosis in each treatment arm using MFSAF v4.0 (Part 1 core)	Change in MFSAF v4.0 from baseline, including proportion of subjects who achieved at least 50% reduction from baseline in MFSAF v4.0 Total Symptom Score (TSS) at the end of Cycle 4 (NIS793 arms) or Cycle 3 (all other arms), of Cycle 8 (NIS793 arms) or Cycle 6 (all other arms) and of

Objective(s)	Endpoint(s)
	Cycle 16 (NIS793 arms) or Cycle 12 (all other arms) from baseline.
• To characterize the pharmacokinetic profile of ruxolitinib administered in combination with siremadlin, crizanlizumab, sabatolimab, rineterkib and NIS793, respectively (Part 1 core)	• PK parameters (e.g., AUC, Cmax, Tmax) and concentration vs. time profiles of each investigational drug with in combination regimens
• To assess emergence of anti-crizanlizumab, anti-sabatolimab, or anti-NIS793 antibodies following one or more IV infusions (Part 1 core)	• Presence and/or concentration of anti-crizanlizumab, anti-sabatolimab or anti-NIS793 antibodies
• To evaluate the changes in spleen size in each treatment arm (Part 1 core and extension)	• Change in spleen length (by palpation) from baseline. • Change in spleen volume (by MRI/CT) from baseline, including proportions of subjects who achieved (i) at least 35% spleen volume reduction and (ii) at least 25% spleen volume reduction at the end of Cycle 8 (NIS793 arms) or Cycle 6 (all other arms, and at the end of Cycle 16 (NIS793 arms) or Cycle 12 (all other arms), or Cycle 21 (extension phase) from baseline respectively.
• To evaluate long-term safety and tolerability of ruxolitinib combination treatments in each arm (Part 1 core and extension)	• Frequency, duration and severity of adverse events, abnormalities in vital signs and laboratory test values, including ECG data
[redacted]	

2 Statistical methods

2.1 Data analysis general information

The data analysis (including analysis for DMC reviews, primary analysis, and final analysis) will be performed by Novartis personnel, Biostatistics and Statistical Programming departments. SAS version 9.4 or later and/or R version 3.0.2 or later will be used to perform all data analyses and to generate tables, figures and listings.

Data included in the analysis

At the time of protocol amendment version 08 release, all subjects have completed at least 8 cycles (arms with NIS793) or 6 cycles (all other arms) of treatment or discontinued earlier (Part 1 and Part 2). If the last core treatment phase subject entered extension treatment phase, the cut-off date for the primary analysis of study data will be established after the last subject performed the end of treatment visit (EOT) for core treatment phase (Part 1).If the last core treatment phase subject would not enter extension treatment phase, the cut-off date will be established after the last subject completed protocol defined safety follow up visit. All statistical analyses will be performed using all data collected in the database up to the data cut-off date. For clarification, Part 1 subjects might have entered extension treatment phase prior to the primary analysis cut-off date. In this case, extension treatment phase data prior to the cut-off date will also be included for the primary CSR. 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. 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 cut-off date and end date after the cut-off date will be reported as 'ongoing'. The same rule will be applied to events starting before or on the cut-off date and not having documented end date. This approach applies, in particular, to adverse event and concomitant medication reports. For these events, the end date will not be imputed and therefore will not appear in the listings.

The final analysis for the extension phase data will be performed at the end of the study. The cut-off date for the final analysis of study data will be established when all subjects have completed the study. All extension treatment phase data will be reported in the final CSR.

General analysis conventions

Pooling of center: Unless specified otherwise, data from all study centers will be pooled for the analysis. Due to expected small size of centers, no center effect will be assessed.
Qualitative data (e.g., gender, race, etc.) will be summarized by means of contingency tables by treatment group; a missing category will be included as applicable. Percentages will be calculated using the number of subjects in the relevant population or subgroup as the denominator.
Quantitative data (e.g., age, body weight, etc.) will be summarized by appropriate descriptive statistics (i.e. mean, standard deviation, median, minimum, and maximum) by treatment group.

2.1.1 General definitions

Investigational drug and study treatment

Investigational drug, will refer to either ruxolitinib (INC424) or any of the novel agents. Whereas, ***study treatment*** will refer to ruxolitinib single agent or in combination with any of the novel agents.

Date of first administration of investigational drug

The date of first administration of investigational drug is defined as the first date when a nonzero dose of investigational drug is administered and recorded on the Dosage Administration Record (DAR) (e)CRF. The date of first administration of study drug will also be referred as start of investigational drug.

Date of last administration of investigational drug

The date of last administration of investigational drug is defined as is the last date when a nonzero dose of investigational drug is administered and recorded on DAR eCRF. The date of last administration of investigational drug will also be referred as end of investigational drug.

Date of first administration of study treatment

The date of first administration of study treatment is derived as the first date when a nonzero dose of study treatment was administered as per the Dosage Administration CRF. The date of first administration of study treatment will also be referred as ***start of study treatment***.

Date of first administration of extension phase study treatment

The date of first administration of extension phase study treatment is derived as the first date when a nonzero dose of extension phase study treatment was administered as per the Dosage Administration CRF for subject who entered extension treatment phase.

Date of last administration of study treatment

The date of last administration of study treatment is defined as the last date when a nonzero dose of study treatment was administered as per Dose Administration (e)CRF.

Study day

The study day describes the day of the event or assessment date, 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 date for all assessments (safety, efficacy, PK, QoL, etc.) is the start of study treatment.

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.

Time unit

A year length is defined as 365.25 days. A month length is 30.4375 days (365.25/12). 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.

Baseline

For safety and efficacy evaluations, the last available assessment on or before the date of start of study treatment is defined as "baseline" assessment.

In case time of assessment and time of treatment start is captured (e.g. pre-dose ECG), the last available assessment before the treatment start date/time is used for baseline.

For safety parameters (e.g ECGs or vital signs), where study requires multiple replicates per time point, the average of these measurements would be calculated for baseline (if not already available in the database).

If patients have no value as defined above, the baseline result will be missing.

On-treatment assessment/event and observation periods

The overall observation period will be divided into three mutually exclusive segments:

1. ***pre-treatment period***: from day of subject's informed consent to the day before first administration of study treatment
2. ***on-treatment period***: from date of first administration of study treatment to 30 days after date of last actual administration of any study treatment (including start and stop date)
3. ***post-treatment period***: starting at day 31 after last administration of study treatment.

Moreover, ***on-extension treatment period*** is defined from day of first administration of extension phase study treatment to 30 days after the date of the last administration of study treatment.

In addition, the ***overall safety period*** is defined for this study as following:
Overall safety period: from date of first administration of study treatment to 30 days after the date of the last administration of ruxolitinib, siremadlin or rineterkib, 90 days after the date of the last administration of NIS793, or 105 days after the date of the last administration of crizanlizumab, or 150 days after the date of the last administration of sabatolimab, whichever is later.

Safety summaries (tables, figures) include data from the pre-treatment period (to display the baseline status e.g. for ECG) and the on-treatment period, i.e. data from the post-treatment period with the exception of deaths should not be included unless requested from Health Authorities or external committees.

In particular, summary tables for adverse events (AEs) will summarize on-treatment events, with a start date during the on-treatment period, the so-called ***treatment-emergent*** AEs. Additionally, all (study treatment related/regardless) AEs and SAEs with a start date during the overall safety period will be summarized.

However, all safety data (including those from the post-treatment period) will be listed and those collected during the pre-treatment, post-treatment overall safety period and post-treatment period are to be flagged.

For the final analysis, all extension treatment phase safety data will be listed and those collected during post-treatment overall safety period and post-treatment period are to be flagged. The safety data in part 1 core and extension phase will be summarized in tables for the extension treatment phase clinical trial safety disclosure.

Windows for multiple assessments

In order to summarize data collected over time (including unscheduled visits), the assessments will be time slotted. The following general rule will be applied in creating the assessment windows: If more than one assessment is done within the same time window, the assessment performed closest to the target date will be used. If two assessments within a time window are equidistant from the target date, then the assessment that occurred later will be used. If multiple assessments occur on the same date then the worst case will be used. Data from all assessments (scheduled and unscheduled), including multiple assessments, will be listed.

The following time windows are defined for descriptive summary on spleen imaging, spleen length palpation, ePROs, laboratory assessments [redacted] (Tables 2-1, 2-2, and 2-3) by visit regardless of core or extension treatment phases. The end of treatment assessment will be mapped into the time points as needed.

Table 2-1 Time windows for imaging assessments (MRI/CT of abdomen)

Time Point	Planned Visit Timing	Time Window Definition
Screening/Baseline	On or before Study Day 1*	≤ Study Day 1
Cycle 7 Day 1	Study Day 168	Study Days 140 – 196
Cycle 13 Day 1	Study Day 336	Study Days 308 – 364
Every 6 cycles thereafter		
Cycle y=13+6*k Day 1 (with k = 1, 2, …)	Study Day (13+6*k-1)*28	Study Days (13+6*k-1)*28-28 to (13+6*k-1)*28+28

* Study Day 1 = Date of first administration of study treatment (start of study treatment)
EOT core and EOT extension assessments are mapped to the corresponding time points based on the defined time window.

Table 2-2 Time windows for ePROs, laboratory assessments (chemistry, coagulation, urinalysis and thyroid function) [redacted]

Time Point	Planned Visit Timing	Time Window Definition
Baseline (Cycle 1 Day 1)	On or before Study Day 1*	≤ Study Day 1
Cycle 2 Day 1	Study Day 29	Study Days 15 – 42
Cycle 3 Day 1	Study Day 57	Study Days 43 – 70
Cycle 4 Day 1	Study Day 85	Study Days 71 – 98

Cycle 5 Day 1	Study Day 113	Study Days 99 – 126
Cycle 6 Day 1	Study Day 141	Study Days 127– 154
Cycle 6+k Day 1 (with k = 1, 2, …)	Study Day (6+k-1)*28+1	Study Day (6+k-1)*28+1-14 to (6+k-1)*28+14
Post treatment		
30-day safety follow-up	Post treatment study day 30	Post treatment study days 23 – 37
90-day safety follow-up	Post treatment study day 90	Post treatment study days 83 – 97
105-day safety follow-up	Post treatment study day 105	Post treatment study days 98 – 112
150-day safety follow-up	Post treatment study day 150	Post treatment study days 136 – 164

* Study Day 1 = Date of first administration of study treatment (start of study treatment)

EOT core and EOT extension assessments are mapped to the corresponding time points based on the defined time window.

Post treatment study day 1 = end of study treatment date + 1

Post treatment visits are for laboratory chemistry assessments only

90-day safety follow-up visit for NIS7393 arm(s)

90-day safety follow-up visit, and 150-day safety follow up visit are for sabatolimab arm(s)

105-day safety follow-up visit is for crizanlizumab arm(s)

[redacted]

Table 2-3 Time windows for spleen length palpation assessments and hematology laboratory assessments

Time Point	Planned Visit Timing	Time Window Definition
Baseline (Cycle 1 Day 1)	On or before Study Day 1*	≤ Study Day 1
Cycle 1 Day 15	Study Day 15	Study Days 8 – 21
Cycle 2 Day 1	Study Day 29	Study Days 22 – 35
Cycle 2 Day 15	Study Day 43	Study Days 36 – 49
Cycle 3 Day 1	Study Day 57	Study Days 50 – 63
Cycle 3 Day 15	Study Day 71	Study Days 64 – 77
Cycle 4 Day 1	Study Day 85	Study Days 78 – 98
Cycle 5 Day 1	Study Day 113	Study Days 99 – 126
Cycle 6 Day 1	Study Day 141	Study Days 127– 154
Cycle 6+k Day 1 (with k = 1, 2, …)	Study Day (6+k-1)*28+1	Study Day (6+k-1)*28+1-14 to (6+k-1)*28+14
Post treatment		
30-day safety follow-up	Post treatment study day 30	Post treatment study days 23 – 37
90-day safety follow-up	Post treatment study day 90	Post treatment study days 83 – 97
105-day safety follow-up	Post treatment study day 105	Post treatment study days 98 – 112
150-day safety follow-up	Post treatment study day 150	Post treatment study days 136 – 164

* Study Day 1 = Date of first administration of study treatment (start of study treatment)
EOT core and EOT extension assessments are mapped to the corresponding time points based on the defined time window.
Post treatment study day 1 = end of study treatment date + 1
Post treatment visits are for laboratory hematology assessments only
90-day safety follow-up visit for NIS7393 arm(s)
90-day safety follow-up visit, and 150-day safety follow up visit are for sabatolimab arm(s)
105-day safety follow-up visit is for crizanlizumab arm(s)

Last contact date

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

Source data	Conditions
Date of Randomization	No Condition
Last date patient was known to be alive from Survival Follow-up page	Patient 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.
Efficacy assessment date if available (e.g., spleen imaging/palpation assessment, bone marrow assessment, ePRO assessment)	Evaluation is marked as 'done'.
Laboratory/PK collection dates/ [redacted]	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 or the cut-off date whichever comes first. The cut-off date will not be used for last contact date, unless the subject was seen or contacted on that date. No date post cut-off date will be used. Completely imputed dates (e.g. the analysis 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. Partial date imputation is allowed for event (death)/censoring is coming from 'Survival information' eCRF.

The last contact date will be used for censoring of patients in the analysis of survival analysis.

2.2 Analysis sets

Full Analysis Set

The Full Analysis Set (FAS) comprises all subjects that received any study drug. Subjects will be analyzed according to the treatment(s) received.

Safety Set

The Safety Set includes all subjects who received at least one dose of study treatment. Subjects will be analyzed according to the study treatment received, where treatment received is defined as the randomized/assigned treatment if the subject took at least one dose of that treatment or the first treatment received if the randomized/assigned treatment was never received.

Dose-Determining Set

The Dose-Determining Set (DDS) includes all subjects from the safety run-in and dose escalation part (Part 1) of the study who met the minimum exposure criterion and had sufficient safety evaluations, or experienced a dose-limiting toxicity (DLT) between C1D1 and C3D1. The minimum exposure required for combination treatments defined as follows:

- For ruxolitinib + rineterkib study treatment arm, a subject has met the minimum exposure criterion if the subject takes at least 75% of the planned daily rineterkib and ruxolitinib combination doses during the first two treatment cycles of dosing, i.e. at least 21 full dosing days out of the planned 28 days for the rineterkib and ruxolitinib combination, and the subject takes at least 80% of the planned daily doses of ruxolitinib i.e. at least 23 full dosing days out of the planned 28 days for ruxolitinib.
- For all other study treatment arms, a subject has met the minimum exposure criterion if the subject takes all planned doses of the combination agent during the first two treatment cycles of dosing, and at least 80% of the planned daily doses of ruxolitinib i.e. at least 23 full dosing days (17 days for NIS793) out of the planned 28 days (21 days for NIS793) for ruxolitinib.

Subjects who do not experience a DLT between C1D1 and C3D1 are considered to have sufficient safety evaluations if they have been observed until C2D28, or C2D21 for the ruxolitinib + NIS793 study treatment arm, and are considered by both the sponsor and investigators to have enough safety data to conclude that a DLT did not occur.

Pharmacokinetic analysis set (PAS)

The Pharmacokinetic Analysis Set (PAS) includes all enrolled subjects who have an evaluable PK profile. A profile is considered evaluable if all of the following conditions are satisfied:

- Subject receives the planned treatments
- Subject takes the dose of study treatments
- Subject provides at least one primary PK parameter
- Subject ingests siremadlin at least 1 h before or 2 h after a meal (only applicable for treatment arm ruxolitinib + siremadlin)

- Subject does not vomit within 4 hours after oral dosing of siremadlin or ruxolitinib (only applicable for treatment arm ruxolitinib + siremadlin).
- Subject does not vomit within 2 hours after oral dosing of ruxolitinib (all other treatment arms).

Patient Classification:

Patients 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 2-4.

Table 2-4 Subject classification based on protocol deviations and non-PD criteria

Analysis set	Protocol deviations leading to exclusion	Non protocol deviation leading to exclusion
FAS	No written inform consent	Not applicable
Safety set	No written inform consent	No dose of study medication
Dose-Determining Set	No written inform consent	Not meeting the definition of inclusion in the DDS
PK Analysis Set	No written inform consent	Not meeting the definition of inclusion in the PAS

Withdrawal of Informed Consent

Any data collected in the clinical database 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 patient withdraws full consent is recorded in the eCRF.

Third party data e.g. PK, [redacted] etc., collected in the clinical database after a subject withdraws the consent or analyzed after withdrawal of consent will not be included in the analysis.

2.2.1 Subgroup of interest

The permanent enrollment halt led to limited data that are not sufficient for subgroup analysis. Subgroup analysis is not applicable for the primary and final analysis for this early terminated study.

2.3 Patient disposition, demographics and other baseline characteristics

The Full Analysis Set (FAS) will be used for all baseline and demographic summaries and listings unless otherwise specified. Summaries will be reported by treatment group (defined in Section 1-1) and for all subjects and listings will be reported by treatment group to assess baseline comparability. Subjects in Part 1 will be analyzed separately from the subject in Part 2. Data from Part 2 will be listed but not summarized.

Basic demographic and background data

All demographic and baseline disease characteristics data will be summarized and listed by treatment group. Categorical data (e.g. gender, age groups: <65 and ≥ 65 years, rac ethnicity, ECOG performance status etc.) will be summarized by frequency counts and percentages; the number and percentage of patients with missing data will be provided. Continuous data (e.g. age, weight, height etc.) will be summarized by descriptive statistics (N, mean, median, standard deviation, minimum and maximum).

Diagnosis and other baseline characteristics

Summary statistics will be tabulated for diagnosis and other baseline disease characteristics. This analysis will include the following: initial diagnosis of disease, most recent diagnosis of disease, time since initial diagnosis, time since most recent diagnosis, hemoglobin level (g/dL), absolute neutrophil count (ANC) (/μL), platelet count (/μL), spleen palpable (Yes/No), spleen length below left costal margin (cm), spleen volume (cm3), MFSAF Total Symptom Score, prior ruxolitinib duration (weeks).

Note: Prior ruxolitinib administration is recorded on Prior/Concomitant Medications eCRF pages.

Medical history

Medical history and ongoing conditions, including cancer-related conditions and symptoms entered on the (e) CRF will be summarized and listed by treatment group. Separate summaries will be presented for ongoing and historical medical conditions. The summaries 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.

2.3.1 Patient disposition

Enrollment by country and center will be summarized for all screened patients and also by treatment group using the FAS.

The number (%) of treated patients included in the FAS will be presented overall and by treatment group. The number (%) of screened and not-treated patients and the reasons for screening failure will also be displayed. The number (%) of patients 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 group for core and extension treatment phase.

The following summaries will be provided: % based on the total number of FAS patients:

- Number (%) of patients who were treated (based on ‘DAR’ eCRF pages of each study treatment component completed with non-zero dose administered)
- Number (%) of patients who are still on-treatment (core) (based on the Disposition page and Subject Status page);

- Number (%) of patients who discontinued the core treatment phase (based on the Disposition page)
- Primary reason for study core treatment phase discontinuation (based on the Disposition page)
- Number (%) of patients who are still on-treatment (extension) (based on the Disposition page and Subject Status page);
- Number (%) of patients who discontinued the extension treatment phase (based on the Disposition page)
- Primary reason for study extension treatment phase discontinuation (based on the Disposition page)

Protocol deviations

The number (%) of patients in the FAS with any protocol deviation will be tabulated by deviation category (as specified in the study Data Quality Plan) overall and by treatment group for the FAS. All protocol deviations will be listed.

Summarize PDs that are not pandemic related and those that are, separately and in total.

For those related to the pandemic summarize "all COVID-19 pandemic related" together and by each specific relationship to COVID-19 if applicable/provided, following the COVID-19 (Coronavirus) Guidance for Clinical Trial Teams and Trial Monitoring Associates-*Protocol Deviations and Data Collection: Go Live after 01-Dec-2020 Relationship to COVID-19 via CRFs Version 2.0*. Percentage is out of total patients in the FAS.

Analysis sets

The number (%) of patients in each analysis set (defined in Section 2.2) will be summarized by treatment group.

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

2.4.1 Study treatment / compliance

Duration of exposure, actual cumulative dose, dose intensity (DI) and relative dose intensity (RDI) will be summarized by treatment group, separately for each component of study treatment (Ruxolitinib and appropriate novel agent). The duration of exposure to study treatment will also be presented for each treatment group. Duration of exposure will be categorized into time intervals; frequency counts and percentages will be presented for the number (%) of subjects in each interval. The number (%) of subjects who have a dose change, and the reasons, will be summarized by treatment group. Data in this section is based on the study treatment CRF pages.

Subject level listings of all doses administered on treatment along with dose change reasons will be produced.

The Safety Set will be used for all summaries and listings of study treatment.

Duration of exposure to study treatment

Duration of exposure to study treatment is calculated by taking into account the duration of exposure to ruxolitinib and/or any novel agent partner for each treatment group:

Duration of exposure to study treatment (days) for each treatment group = (last date of exposure to study treatment) – (date of first administration of study treatment) + 1.

The first date of exposure to study treatment for each treatment group is the earliest of the first dates of exposure to ruxolitinib or any novel agent partner.

The last date of exposure to study treatment for each treatment group is the latest of the last dates of exposure to ruxolitinib or any novel agent partner (see Table 2-5).

Summary of duration of exposure of study treatment in appropriate time units will include categorical summaries and continuous summaries (i.e. mean, standard deviation etc.) using appropriate units of time.

Duration of exposure to investigational drug (ruxolitinib or any novel agent partner)

Duration of exposure to investigational drug (days) = (last date of exposure to investigational drug) – (date of first administration of investigational drug) + 1.

Table 2-5 Definition of last date of exposure of study drug

Scenario	Definition of last date of exposure of study drug	Example
Ruxolitinib (INC424) Rineterkib (LTT462)	Date of last administration of a non - zero dose of the study drug.	Example 1: A patient had a permanent discontinuation of the study drug on 06Jan2013 after being put on a temporary interruption since 01Jan2013. In this case the last date of exposure is 31Dec2012.
Siremadlin (HDM201) Crizanlizumab (SEG101) Sabatolimab (MBG453) NIS793	The planned end date of the last cycle in which the last non-zero dose of the investigational drug was last administered. Note: If the patient died or was lost to follow-up before the derived last date, the last date of exposure to investigational drug is the date of death or the date of last contact, respectively. If the derived last date of exposure goes beyond the data cut-off date, it should be truncated to the date of data cut-off.	Example 2: Siremadlin is in a 28-day cycle with once a day administration for Day 1-5 of each cycle. The last date of exposure is the date of administration in the last cycle + 23 days. Example 3: Crizanlizumab is in a 28-day cycle with one iv infusion in the beginning of the cycle and with a 2nd loading dose after 14 days of initial dosing for Cycle1 only. The late dose of exposure is the date of last infusion in Cycle 1 + 13 days if Cycle 1 is the last cycle.

		The last date of exposure is the date of first infusion in the last cycle + 27 days if the last cycle is post Cycle 1. Example 4: Sabatolimab is in a 28-day cycle with one iv infusion in the beginning of the cycle. The last date of exposure is the date of first infusion in the last cycle + 27 days Example 5: NIS793 is in a 21-day cycle with one iv infusion in the beginning of the cycle. The last date of exposure is the date of first infusion in the last cycle + 20 days

Summary of duration of exposure of an investigational drug will include categorical summaries based on 12 weeks intervals and using descriptive statistics (mean, standard deviation etc). The duration includes the periods of temporary interruption.

Cumulative dose

Cumulative dose of a study treatment is defined as the total dose given during the study treatment exposure and will be summarized for each of the study treatment components. The **actual cumulative dose** refers to the total actual dose administered, over the duration for which the subject is on the study treatment as documented in the Dose Administration eCRF.

For subjects who did not take any drug the cumulative dose is by definition equal to zero.

The **planned cumulative dose** for a study treatment component is the total planned dose as per the protocol up to the last date of investigational drug administration. The planned cumulative dose will not be summarized/listed. It will be used for relative dose intensity calculations.

For each component of study treatment which is administered daily the planned cumulative dose is the planned starting dose summed over the duration of exposure for the component. For each component of study treatment that has cyclic administration, the planned cumulative dose is the sum of planned total starting dose in a cycle times the number of complete cycles the patient should have been over the duration exposure, plus the total dose expected in the last incomplete cycle.

Dose intensity and relative dose intensity

Dose intensity (DI) for subjects with non-zero duration of exposure is defined as follows:

For Ruxolitinib and Rineterkib:

DI (mg/day) = Actual Cumulative dose (mg) / Duration of exposure to study drug (in days).

For Siremadlin, sabatolimab and NIS793:

DI (mg/week) = Actual Cumulative dose (mg) / Duration of exposure to study drug (in weeks).

Example 1, the duration of exposure to siremadlin is 24 weeks (6 cycles) and the subject receives 20 mg siremadlin daily i.v. infusion on Days 1 to 5 of each cycle:

DI (mg/week) = Actual Cumulative dose (mg) / Duration of exposure to study drug (in weeks) =20*5*6 (mg) / 24 (weeks) = 25 (mg/week).

Example 2, the duration of exposure to sabatolimab is 24 weeks (6 cycles) and the subject receives 800 mg sabatolimab i.v. infusion on Day 1 of each cycle:

DI (mg/week) = Actual Cumulative dose (mg) / Duration of exposure to study drug (in weeks) =800*6 (mg) / 24 (weeks) = 200 (mg/week).

Example 3, the duration of exposure to NIS793 is 24 weeks (8 cycles) and the subject receives 2100 mg NIS793 i.v. infusion on Day 1 of each cycle:

DI (mg/week) = Actual Cumulative dose (mg) / Duration of exposure to study drug (in weeks) =2100*8 (mg) / 24 (weeks) = 700 (mg/week).

For Crizanlizumab:

DI (mg/kg/week) = Actual Cumulative dose (mg/kg) / Duration of exposure to study drug (in weeks).

For example, the duration of exposure to crizanlizumab is 24 weeks (6 cycles) and the subject receives 5 mg/kg crizanlizumab IV infusion every 4 weeks with a second loading dose after 14 days of the initial dose (7 IV infusion of crizanlizumab at 5 mg/kg):

DI (mg/kg/week) = Actual Cumulative dose (mg/kg) / Duration of exposure to study drug (in weeks)=5*7 (mg/kg) / 24 (weeks) = 1.46 mg/kg/week.

For subjects who did not take any drug the DI is by definition equal to zero.

Planned dose intensity (PDI) is defined as follows:

For Ruxolitinib and Rineterkib:

PDI (mg/day) = Planned Cumulative dose (mg) / Duration of exposure (in days).

For Siremadlin, Sabatolimab and NIS793:

DI (mg/week) = Planned Cumulative dose (mg) / Duration of exposure (in weeks).

For Crizanlizumab:

DI (mg/kg/week) = Planned Cumulative dose (mg/kg) / Duration of exposure (in weeks).

Relative dose intensity (RDI) is defined as follows:

For Ruxolitinib and Rineterkib:

RDI = DI (mg/day) / PDI (mg/day).

For Siremadlin, Sabatolimab and NIS793:

RDI = DI (mg/week) / PDI (mg/week).

For Crizanlizumab:

RDI = DI (mg/kg/week) / PDI (mg/kg/week).

DI and RDI will be summarized separately for each of the study treatment components (ruxolitinib and novel agent combination) using the duration of exposure of each of the components.

Dose changes, interruptions or permanent discontinuations

The number of subjects who have dose changes, permanent discontinuations or interruptions, and the reasons, will be summarized separately for each of the study treatment components.

'Dose changed', 'Dose interrupted', and 'Dose permanently discontinued' fields from the Dosage Administration CRF pages (DAR) will be used to determine the dose change, dose interruptions, and permanent discontinuations, respectively.

The corresponding fields 'Reason for dose change/dose interrupted' and 'Reason for permanent discontinuation' will be used to summarize the reasons.

A dose change is either 'change in prescribed dose level' or 'dosing error' where actual dose administered/total daily dose is different from the prescribed dose.

Dose interruptions are captured in the eCRF following the Oncology Data standard Reference manual and CRF Completions Guidelines. Depending on the dosing administration schedule (e.g. continuous dosing, interval dosing), data would be collected in different fashion. For instance, for continuous dosing an interruption is defined as a dose of zero in a unit of time between two non-zero dosing records. In general any rest period as part of the regimen/schedule is not considered as an interruption.

For the purpose of summarizing interruptions and reasons, in case multiple entries for interruption that are entered on consecutive days with different reasons will be counted as separate interruptions. However, if the reason is the same in this block of entries, then it will be counted as one interruption.

2.4.2 Prior, concomitant and post therapies

Prior anti-neoplastic therapy

The number and percentage of subjects who received any prior anti-neoplastic medications will be summarized by treatment group. Prior anti-neoplastic medications will be summarized by setting (e.g. adjuvant, metastatic, etc.) and also by lowest ATC class, preferred term and treatment. Summaries will include total number of regimens, indication, reason for discontinuation of therapy and best response hematological. Listings will be produced for prior anti-neoplastic medications.

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

The above analyses will be performed using the FAS.

Post treatment anti-neoplastic therapy

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

Concomitant medications

Concomitant therapy is defined as all interventions (therapeutic treatments and procedures) other than the study treatment administered to a subject coinciding with the study treatment period. Concomitant therapy includes medications (other than study drugs) starting on or after the start date of study treatment or medications 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 Anatomical Therapeutic Chemical (ATC) classification system and summarized by lowest ATC class and preferred term using frequency counts and percentages. Surgical and medical procedures will be coded using MedDRA and summarized by SOC and preferred term. These summaries will include:

1. Medications starting on or after the start of study treatment but no later than 30 days after start of last dose of study treatment and
2. Medications starting prior to start of study treatment and continuing after the start of study treatment.

All concomitant therapies will be listed. Any concomitant therapies starting and ending prior to the start of study treatment or starting more than 30 days after the last date of study treatment, or starting more than 30 days after the last date of study treatment and within the overall safety period will be flagged in the listing. The safety set will be used for all concomitant medication tables and listings.

2.5 Analysis of the primary objective

The primary objective of Part 1 of the study is to characterize the safety, tolerability, and the recommended Phase 2 dose (RP2D) of each combination partner used with ruxolitinib in subjects with myelofibrosis.

The primary objective of Parts 2 and 3 of the study is to evaluate the preliminary efficacy of the ruxolitinib combination treatments in subjects with myelofibrosis. Due to permanent enrollment halt, the Part 2 and Part 3 primary objective will not be pursued.

2.5.1 Primary endpoint

Part 1

The primary endpoint in Part 1 is the incidence of dose-limiting toxicities (DLTs) within the first two treatment cycles.

Part 2 and Part 3 (NOT applicable)

The primary endpoint in Parts 2 and 3 is the response rate (RR) at the end of Cycle 6.

The RR is the composite of anemia improvement and no spleen volume progression and no symptom worsening. For a subject to be considered a responder, all three components of the composite endpoint have to be fulfilled. The components of the composite endpoint are defined as follows:

- Anemia improvement is defined as an increase of hemoglobin from baseline of at least 1.5 g/dL. Anemia improvement requires absence of any PRBC transfusion in 12 weeks prior to achieving an increase of 1.5 g/dL. The increase in hemoglobin should be confirmed at least 2 weeks later.
- No spleen volume progression with progression defined as a spleen volume increase of 25% or more from baseline as measured by MRI/CT.
- No symptoms worsening as measured by the MFSAF version 4 patient reported outcome (PRO). An increase in total symptom score (TSS) of 10 or more from baseline is considered worsening of symptoms.

2.5.2 Statistical hypothesis, model, and method of analysis

2.5.2.1 Part 1

Identification of a recommended dose for siremadlin and rineterkib

Identification of the recommended Phase 2 dose (RP2D) of the combination treatment will be based upon the estimation of the probability of dose-limiting toxicity (DLT) within the first two cycles for patients in the Dose-Determining Set (DDS). The maximum tolerated dose (MTD) is the highest drug dosage that is not expected to cause DLT in more than 33% of the treated subjects during the DLT evaluation period. The RP2D may be chosen prior to identification of the MTD and may be lower than the MTD. The decisions on new dose levels are made by the Investigators and Novartis study personnel in a dose escalation meeting based upon the review of patient tolerability and safety information (including the BLRM summaries of DLT risk) along with PK, PD and preliminary activity information available at the time of the decision.

The dose escalation will be guided by a Bayesian analysis of dose-limiting toxicity (DLT) data from Cycle 1 and Cycle 2. The relationship between dose and the probability of DLT is modelled using logistic regression. Details of the model are as follows.

Let π(d) be the risk of DLT for the combination of ruxolitinib given at dose siremadlin dose d, the dose-DLT model is logistic:

Figure 2-1 Logistic DLT model.

$$\text{logit}\big(\pi(d)\big) = \log(\alpha) + \beta \log(d/d^*)$$

where d* is used to scale the doses. Hence, $\alpha>0$ is the odds of a DLT at d*; and $\beta>0$ is the increase in the log-odds of a DLT by a unit increase in log-dose.

Assessment of patient risk

After each subject within each cohort is followed for two cycles, the BLRM will be updated and the posterior distribution for the risk of DLTs for patients at dose levels of siremadlin will be evaluated. The posterior distributions will be summarized to provide the posterior probability that the risk of a DLT within the first two cycles lies within the following intervals (Table 2-6)::

Table 2-6 Toxicity levels

Under-dosing:	[0 , 0.16)
Targeted toxicity:	[0.16 , 0.33)
Excessive toxicity:	[0.33 , 1]

The escalation with overdose control (EWOC) principle

Dosing decisions are guided by the escalation with overdose control principle. A dose of siremadlin may only be used for newly enrolled patients if the risk of excessive toxicity at that dose is less than 25%.

Prior distributions

A meta-analytic-predictive (MAP) approach was used to derive the prior distribution for the model parameters. The MAP prior for the logistic model parameters for this study is the conditional distribution of the parameters given the historical data (see Neuenschwander et al 2014, Neuenschwander et al 2010, Spiegelhalter et al 2004). MAP priors are derived from hierarchical models, which take into account possible differences between multiple historical studies.

A full description of the application of the MAP approach to derive the prior distributions of the model parameters is given in Section 5.4.

Starting dose

The starting dose for siremadlin in combination with ruxolitinib is 20 mg/day (day 1-5 each of a 28-day cycle). For this dose the prior risk of excessive toxicity is 15.42% (see Section 5.4), which satisfies the EWOC criterion. Any dose lower than 20 mg also fulfills the EWOC criterion. A full assessment of the prior risk of toxicity is given in Section 5.4.

The starting dose for rineterkib in combination with ruxolitinib is 200 mg. For this dose the prior risk of excessive toxicity is 18.85% (see Section 5.4), which satisfies the EWOC criterion. Any dose lower than 200 mg also fulfills the EWOC criterion. A full assessment of the prior risk to patients is given in Section 5.4.

Listing of DLTs

DLTs will be listed, and their incidence summarized by primary system organ class and worst grade (CTCAE version 5.0). Listings and summaries (if data warrant) will be based on the DDS.

DLT rate for crizanlizumab, sabatolimab and NIS793

For crizanlizumab, sabatolimab, and NIS793 the proposed RP2D is known prior to the study start and based on previous single agent phase I and phase II studies in solid and hematologic malignancies. A safety run-in is performed to assess the risk of unexpected toxicities and a dose escalation procedure is not followed for these combination treatments.

For each combination treatment, the DLT rate is determined as the observed number of patients in the DDS who experienced a DLT divided by the number of patients in the DDS.

Any DLTs occurring will be listed, and their incidence summarized by primary system organ class and worst grade (CTCAE version 5.0). Listings and summaries (if data warrant) will be based on the DDS.

2.5.2.2 Part 2 (NOT applicable)

No formal hypothesis testing or between arm comparison will be conducted in Part 2. The objective of this part of the study is to explore the disease-modifying activity of the combination treatments, and to drop the arm(s) that lack promising activity.

The arms that advance into Part 3 will be selected based on the totality of the data. That includes, among other data, efficacy, safety, PK/PD, [redacted] data as appropriate, and also data from the ruxolitinib single agent arm.

To determine whether a combination treatment has the potential to be disease-modifying and, as such, be eligible to advance into Part 3, the combination must not cross a specified futility boundary defined for the response rate with respect to the primary composite endpoint.

A combination treatment arm that shows an early sign of outstanding efficacy by crossing a specific efficacy boundary (i.e. when at least 10 subjects of a combination treatment are evaluable) at the first interim look may expand into Part 3 seamlessly, without stopping at the second interim look.

At the second interim look, the combination treatment arm must not cross a specified futility boundary (i.e. when at least 25 subjects of a combination treatment are evaluable) defined for the RR with respect to the prima ry composite endpoint.

The threshold for the first interim look is shown in Table 2-7

Table 2-7 1st interim look (seamless expansion to Part 3) efficacy threshold

Decision rule	Action
P (RR>0.3|D) > 80%	Preliminary efficacy declared

According to the first interim threshold, an arm is considered of clinical interest and will be advanced to Part 3 expansion if the posterior probability of the RR for the primary composite endpoint being greater than 0.3 is more than 80%. This first interim decision rule was determined such that arms with outstanding efficacy that warrant further development (i.e. True RR > 50%) will be advanced with higher probability, see Section 3.1.2 for details.

Concerning the posterior probability of the response rate, let i index the study arm, let p_i be the corresponding responder rate with respect to the primary endpoint, let n_i be the number of

subjects of arm i, and let r_i be the number of responders in arm i. The number of responders r_i in arm i is modelled by a binomial distribution with parameters n_i and p_i. With the assumptions of a vague prior, *beta(1/2, 1/2)*, the posterior probability distribution of p_i is given by

$$p_i|n_i, r_i \sim beta\left(\frac{1}{2} + r_i, \frac{1}{2} + n_i - r_i\right)$$

For a given sample size n_i, the first interim threshold shown in Table 2-7 can be transformed into a number of responders with respect to the primary endpoints for which preliminary efficacy would be declared, as shown in Table 2-8.

Table 2-8 1st Interim look efficacy regions for a given sample size

n_i	Region
9	≥ 4
10	≥ 5
11	≥ 5
12	≥ 5

The futility threshold of the 2nd interim look is shown in Table 2-9.

Table 2-9 Futility threshold

Decision rule	Action
P (RR ≤ 0.2 | Data) > 50%	Declare combination arm futile

According to the futility threshold, an arm is dropped for futility if the posterior distribution of the response rate for the primary composite endpoint is less than or equal to 0.2 with a probability of more than 50%. This futility decision rule was determined such that arms with a response rate which are not of interest for further development, that are response rates of 10% or less, are likely to cross the futility threshold and would therefore be dropped at the end of Part 2, see Section 3.1.2 for details.

For a given sample size n_i, the futility threshold shown in Table 2-9 can be transformed into a number of responders with respect to the primary endpoints for which futility would be declared, see Table 2-10.

Table 2-10 Futility regions for a given sample size

n_i	Futility region
24	≤ 4
25	≤ 4
26	≤ 5
27	≤ 5

Due to the discrete nature of the binomial distribution, the futility region does not necessarily change in the sample size n_i.

The response rate (RR) as defined in Section 2.5.2.1 will be summarized using descriptive statistics (N, %) by treatment arm along with a 95% credible interval of the RR posterior

distribution. In addition, the probability that RR is less than or equal to 0.2 will be derived from the posterior distribution of RR.

2.5.2.3 Part 3 (NOT applicable)

Part 3 of the study aims to further assess the efficacy of the chosen treatment combination from Part 2 by comparing it to ruxolitinib using the FAS.

The comparison of the combination arm to ruxolitinib will be performed separately based on a Bayesian model and include data from both Part 2 and Part 3.

In detail, let n_{rux} and r_{rux} be the sample size and the number of responders with respect to the primary composite endpoint in the ruxolitinib arm, respectively. Analogously, n_{comb} and r_{comb} denote the sample size and the number of responders in the combination arm. The number of responders within each arm is modelled by a binomial distribution with response rate p_i (i = rux, comb). As in Section 2.5.2.2, a vague prior, *beta(1/2, 1/2)*, is assumed for p_i. This results in a beta-distributed posterior distribution for the response rate, see Section 2.5.2.2. The study is considered to be successful with respect to the primary endpoint when the difference in response rates between combination arm and ruxolitinib arm is greater than zero with a probability of greater than or equal to 90%:

Figure 2-2 Success criterion for primary endpoint.

$$P(p_{comb} - p_{rux} > 0 | n_{rux}, r_{rux}, n_{comb}, r_{comb}) \geq 0.9$$

The RR will be summarized using descriptive statistics (N, %) by treatment arm along with 95% creditable interval of the RR posterior distribution. The joint distribution of the posterior difference in response rates will be summarized using median and 95% credible interval. The probability that the posterior difference is greater than zero will also be presented.

2.5.3 Handling of missing values/censoring/discontinuations

Subjects with missing assessments of hemoglobin, spleen volume or MF SAF TSS at the end of Week 24 that prevent the evaluation of the primary endpoint will be considered non-responders to that treatment. The visit windows considered when defining what constitutes a missing value for hemoglobin level, spleen volume or MF SAF TSS is provided in Section 2.1.1.

2.5.4 Supportive analyses (NOT applicable)

Sensitivity analyses (NOT applicable)

As sensitivity analyses, the primary endpoint in Parts 2 and 3, i.e. the response rate (RR) at the end of Cycle 6 (Section 2.5.1), will be evaluated with different prior distributions for the response rates:

1. The prior distribution for the response rate in the combination arm is the same prior as in the primary analysis, i.e., a vague prior. The prior distribution for the response rate in the ruxolitinib arm is the robustified MAP prior based on historical data from INCB18424-351 (Comfort I) and CINC424A2352 (Comfort II). The historical data from Comfort I and

Comfort II and the statistical methodological details and the summary of robustified MAP prior are described in Section 5.5.

2. The prior distribution for the response rate in the combination arm is obtained by updating the vague prior, *beta(1/2, 1/2),* using the number of responders from Part 1. Let i index the combination study arm, let n_{iP1} be the number of subjects of Part 1 arm i, and let r_{iP1} be the number of responders in Part 1 arm i. The updated prior distribution for the RR in the combination arm i is given by

$$beta\left(\frac{1}{2} + r_{iP1}, \frac{1}{2} + n_{iP1} - r_{iP1}\right)$$

The prior distribution for the responder rate in the ruxolitinib arm is not changed compared to the primary analysis.

For the two sensitivity analyses proposed above, the RR will be summarized using descriptive statistics (N, %) by treatment arm along with 95% creditable interval of the RR posterior distribution. The joint distribution of the posterior difference in response rates will be summarized using median and 95% credible interval. The probability that the posterior difference is greater than zero will also be presented.

Supportive analyses (NOT applicable)

As supportive analyses, the definition of 'responders' will be modified compared to the primary endpoint by only considering the hemoglobin and spleen volume measurement. Thus, for supportive analyses, a subject is considered a responder if the following two criteria are fulfilled:

- Anemia improvement is defined as an increase of hemoglobin from baseline of at least 1.5 g/dL at the end of 24 weeks after randomization. Anemia improvement requires absence of any PRBC transfusion in the 12 weeks prior to achieving an increase of 1.5 g/dL and a confirmation of Hb increase at least 2 weeks later.
- No spleen volume progression is given when the spleen volume increase is smaller than 25% between baseline and the end of Week 24 measured by MRI/CT.

Subjects with missing assessment of hemoglobin or spleen volume at the end of Week 24 that prevent the evaluation of the analysis endpoint will be considered as non-responders to that treatment.

The RR will be summarized using descriptive statistics (N, %) by treatment arm along with 95% creditable interval of the RR posterior distribution. The joint distribution of the posterior difference in response rates will be summarized using median and 95% credible interval. The probability that the posterior difference is greater than zero will also be presented.

2.6 Analysis of the key secondary objective

Not applicable

2.7 Analysis of secondary efficacy objective(s)

Due to the enrollment halt, some of the secondary efficacy objectives are no longer applicable. The updated secondary efficacy objectives are listed as follows:

- To evaluate changes in symptoms of myelofibrosis in each treatment arm using MFSAF v4.0 (Part 1 core)
- To evaluate the changes in spleen size (spleen volume and spleen length) in each treatment arm (Part 1 core and extension)

In final CSR, data collected in extension phase will be listed instead of summary tables.

2.7.1 Secondary endpoints

All secondary efficacy endpoints will be analyzed using the Full Analysis Set (FAS).

Proportion of subjects achieving improvement of hemoglobin level (NOT applicable)

Hemoglobin is measured at baseline and at least once per cycle. The change from baseline will be summarized by treatment group using descriptive statistics. Individual hemoglobin level at different time points will be displayed graphically by treatment group.

The proportion of subjects achieving improvement in hemoglobin ≥ 1.5 g/dL and ≥ 2.0 g/dL, respectively, from baseline to the end of Week 24, and the end of Week 48 will be summarized using descriptive statistics (N, %) by treatment arm along with two-sided exact binomial 95% CIs [Clopper and Pearson 1934]. The proportion will be calculated using FAS subjects as the denominator; the numerator will be based on the number of subjects with a baseline measure and a post-baseline measure of the corresponding time-point.

Change in spleen volume and spleen length

Spleen volume will be measured by MRI/CT at baseline, the end of every 24 weeks (8 cycles for NIS793 arms and 6 cycles for all other arms) and at EOT (if not performed in the past 12 weeks) for Part 1 core treatment phase and extension phase. The change from baseline will be summarized at each scheduled assessment time point by treatment group using descriptive statistics. The change from baseline at the end of Week 24 will be displayed graphically for individual subjects by treatment groups.

The proportion of patients with a spleen volume progression will be summarized using descriptive statistics (N, %) at the end of Week 24 and separately at the end of Week 48 by treatment group.

The proportion of subjects achieving at least 25% and 35% reduction in spleen volume from Baseline to the end of Cycle 8 (NIS793 arms) or Cycle 6 (all other arms) and the end of Cycle 16 (NIS793 arms) or Cycle 12 (all other arms) during the core treatment phase as measured by MRI or by CT will be summarized descriptively.

Spleen length measurement will be conducted by manual palpation. Manual palpation will be performed on days 1 and 15 of cycles 1 to 3, and day 1 of subsequent cycles for Part 1 core treatment phase and extension treatment phase. The change from baseline at each planned

scheduled point will be summarized by treatment group using descriptive statistics. The change from baseline at the end of Week 24 will be displayed graphically for individual subjects by treatment groups.

The proportion of subjects who achieved at least 50% spleen length reduction from baseline for subjects with baseline splenomegaly that is palpable at greater than 10 cm, or became not palpable for subjects with baseline splenomegaly that is palpable at ≥5-≤10 cm will be calculated at the end of Cycle 4 (NIS793 arms) or Cycle 3 (all other arms), and at the end of Cycle 8 (NIS793 arms) or Cycle 6 (all other arms), and at the end of Cycle 16 (NIS793 arms) or Cycle 12 (all other arms).

Change in symptoms assessed by MFSAF and EORTC

Symptoms of myelofibrosis will be measured by MFSAF v4.0 and EORTC QLQ-C30 for the core treatment phase. Due to enrollment halt, EORTC data is not available for Parts 2 and 3 of the study. [redacted]

MFSAF v4.0

MFSAF v4.0 (Appendix 5.6) is a harmonized, consensus-based PRO questionnaire recently developed for use in MF trials by a PRO Consortium Working Group (Gwaltney et al 2017), which focuses on the 7 core symptoms of MF: fatigue, night sweats, pruritus, abdominal discomfort, pain under the ribs on the left side, early satiety and bone pain.

Subjects record symptom severity at its worst for each of the 7 symptoms on an 11-point numeric rating scale, from 0 (absent) to 10 (worst imaginable). The 7-day recall format is used in this study. The Total Symptom Score (TSS) is calculated as the average of the observed individual item responses on the 0 to 10 scale multiplied by 7, resulting in a TSS ranging from 0 to 70. At least four items must be completed to calculate the TSS from the 7-day recall format of the MFSAF v4.0. MFSAF v4.0 is collected at Screening, C1D1, Day 1 of all subsequent cycles of treatment, as well as the EOT visit.

The change from baseline in MFSAF total symptom score and individual symptoms will be summarized with descriptive statistics by treatment group at each scheduled assessment time point for Part 1. The change from baseline at the end of Week 24 will be displayed graphically for individual subjects by treatment groups. The proportion of subjects with no worsening in symptom at the end of Week 24 and Week 48 will be summarized. The proportion will be calculated using FAS subjects as the denominator; the numerator will be based on the number of subjects with a baseline measure and a post-baseline measure of the corresponding time-point. The symptom worsening is defined as an increase in TSS of 10 or more from baseline.

The proportion of subjects who achieved at least 50% reduction in MFSAF v4.0 Total Symptom Score (TSS) at the end of Cycle 4 (NIS793 arms) or Cycle 3 (all other arms), at the end of Cycle 8 (NIS793 arms) or Cycle 6 (all other arms) and at the end of Cycle 16 (NIS793 arms) or Cycle 12 (all other arms) respectively from baseline will be calculated.

Progression free survival (PFS) (NOT applicable)

PFS is defined as the time from the date of randomization to the date of the first documented progressive splenomegaly, or accelerated phase, or deteriorating cytopenia, or leukemic transformation or death due to any cause. See Protocol Section 8.3.2.4 for the details of the PFS events definition. The PFS event and the corresponding date are taken from the Progressive Disease Assessment CRF page.

PFS will be analyzed using the FAS population according to the randomized treatment group assigned at randomization. The PFS distribution will be estimated using the Kaplan-Meier method, and the Kaplan-Meier curves, medians and 95% confidence intervals of the medians will be presented for each treatment group. The hazard ratio for PFS comparing a combination treatment with ruxolitinib single agent will be calculated, along with its 95% confidence interval, using a Cox model.

PFS will be censored if no PFS event is observed before the cut-off date or the date when a new anti-neoplastic therapy or another investigational treatment is started, whichever occurs earlier. The censoring date will be the date of last adequate assessment before a new anti-neoplastic therapy or another investigational treatment is started, if it occurred, else the date of last adequate assessment before the cut-off date.

Disease Progression

Disease progression will be assessed at every clinical visit. The date of assessment, date of confirmatory progressive disease, progressive disease criteria will be listed for core and extension treatment phases.

Proportion of subjects achieving improvement in bone marrow fibrosis (NOT applicable)

The grade of bone marrow fibrosis will be measured at baseline, the end of Week 24, and every subsequent 24 weeks. The grade of bone marrow fibrosis will be summarized descriptively for each scheduled point by treatment group. The proportion of patients achieving improvement in bone marrow fibrosis of ≥ 1 grade from baseline to the end of Week 24, and the end of Week 48 will be summarized by treatment group. The proportion will be calculated using FAS subjects as the denominator; the numerator will be based on the number of subjects with a baseline measure and a post-baseline measure of the corresponding time-point.

2.8 Safety analyses

For all safety analyses, the safety set will be used, except for summaries of DLTs, for which the dose-determining set will be used in addition to the safety set. All listings and tables will be presented by treatment group. Safety summaries (tables, figures) include only data from the on-treatment period with the exception of baseline data which will also be summarized where appropriate (e.g. change from baseline summaries). In addition, a separate summary for death including on treatment and post treatment deaths will be provided. In particular, summary tables for adverse events (AEs) will summarize only on-treatment events, with a start date during the on-treatment period (treatment-emergent AEs).

The overall observation period will be divided into three mutually exclusive segments (pre-, on-, post- treatment period) in addition to overall safety period and on-extension treatment period as defined in Section 2.1.1.

In final CSR, patient safety information related to adverse events from the extension phase will be summarized together with data from phase I core.

2.8.1 Adverse events (AEs)

AE summaries will include all AEs occurring (new or worsening) during the **on treatment period**. In addition, all AEs and SAEs which started during the **overall safety period** will be summarized. AEs collected in the AE (e)CRF page will be listed along with the information collected on those AEs e.g. AE relationship to study drug, AE outcome etc. AEs with start date outside of on-treatment period (pre-treatment period, post-treatment period, and post-treatment overall safety period) will be flagged in the listings.

AEs will be summarized by number and percentage of subjects having at least one AE, having at least one AE in each primary system organ class (SOC) and for each preferred term (PT) using MedDRA coding. A subject with multiple occurrences of an AE will be counted only once in the respective AE category. Grading of AEs is based on NCI Common Terminology Criteria for Adverse Events (CTCAE) version 5.0. A subject with multiple CTCAE grades for the same preferred term will be summarized under the maximum CTCAE grade recorded for the event . AE with missing CTCAE grade will be included in the 'All grades' column of the summary tables.

In AE summaries, the primary system organ class will be presented alphabetically and the preferred terms will be sorted within primary SOC in descending frequency. The sort order for the preferred term will be based on their frequency in the total column.

The following adverse event summaries will be produced by treatment arm; overview of adverse events and deaths (number and % of subjects who died, with any AE, any SAE, any dose reductions/interruptions, leading to treatment discontinuation, requiring additional therapy), AEs by SOC and PT, summarized by relationship (all AEs and AEs related to study treatment), seriousness (SAEs and non-SAEs), leading to treatment discontinuation, leading to dose interruption/adjustment, requiring additional therapy and leading to fatal outcome. In addition, a summary of serious adverse events with number of occurrences will be produced (an occurrence is defined as >1 day between start and prior end date of record of same preferred term).

As recommended in the COVID-19 (Coronavirus) Guidance for Clinical Trial Teams - *Developing Clinical Study Reports (CSRs) for Studies Conducted During the COVID-19 Pandemic*, all suspected or confirmed SARS-CoV-2 infections will be listed for all patients and by treatment groups. The infected cases will be selected using specified criteria as detailed in the TFL shells document.

2.8.1.1 Adverse events of special interest / grouping of AEs

Adverse events of special interest (AESI) during the on-treatment period will be tabulated. The list of AESI will also include relevant events for ruxolitinib, and the novel agents including but not limited to siremadlin, crizanlizumab, sabatolimab and NIS793, if deemed appropriate.

Data analysis of AESIs

An adverse event of special interest is a grouping of adverse events that are of scientific and medical concern specific to compound ruxolitinib and the individual novel agents, if deemed appropriate. These groupings are defined using MedDRA terms, SMQs (standardized MedDRA queries), HGLTs (high level group terms), HLT (high level terms) and PTs (preferred terms). Customized SMQs (Novartis MedDRA queries, NMQ) may also be used. A 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 and/or an existing SMQ, narrow or broad. All AESI terms for the individual compounds as well as for combination specific AESIs, if deemed appropriate, are captured in the case retrieval strategy/strategies (CRS).

For each specified AESI, number and percentage of patients with at least one event of the AESI occurring during on treatment period will be summarized. Summaries of these AESIs will be provided by treatment arm, (specifying grade, SAE, relationship, action taken, medication or therapy taken, AE outcome). A listing of all grouping levels down to the MedDRA preferred terms used to define each AESI will be generated.

2.8.1.2 Adverse events of clinical trial safety disclosure

For the legal requirements of ClinicalTrials.gov and EudraCT, two required tables on adverse events which are not serious adverse events with an incidence greater than 5% and on serious adverse events and SAE suspected to be related to study treatment will be provided by system organ class and preferred term on the safety set population.

If for a same patient, several consecutive AEs (irrespective of study treatment causality, seriousness and severity) occurred with the same SOC and PT:

- a single occurrence will be counted if there is ≤ 1 day gap between the end date of the preceding AE and the start date of the consecutive AE
- more than one occurrence will be counted if there is > 1 day gap between the end date of the preceding AE and the start date of the consecutive AE

For occurrence, the presence of at least one SAE / SAE suspected to be related to study treatment / non SAE has to be checked in a block e.g., among AE's in a ≤ 1 day gap block, if at least one SAE is occurring, then one occurrence is calculated for that SAE.

The number of deaths resulting from SAEs suspected to be related to study treatment and SAEs irrespective of study treatment relationship will be provided by SOC and PT.

2.8.1.3 Eye disorders for rineterkib

To assess the eye toxicity of rineterkib, eye disorder AEs (SOC = 'Eye disorders') will be summarized by PT and listed for rineterkib combination treatment group.

2.8.2 Deaths

Separate summaries for on-treatment and all deaths (on-treatment and post-treatment) will be produced by treatment arm, primary reason, system organ class and preferred term. All deaths will be listed, post treatment deaths and deaths occurred in the post-treatment overall safety period will be flagged. A separate listing of deaths prior to starting treatment (i.e. during screening period) will be provided for all screened subjects as applicable.

Note: "Study indication" as primary reason of death should be coded using MedDRA terms based on the diagnosis eCRF field at start of study. If not coded accordingly in the database, it still must be included in the summary table. Coded reasons for deaths will then be summarized by category 'Study indication' and 'Other' (as selected by the investigator).

2.8.3 Laboratory data

Data from all sources (central and local laboratories) will be combined. The summaries will include all laboratory assessments collected no later than 30 days after the last study treatment administration date. Laboratory data will be listed by treatment group, subject, and visit/time and if normal ranges are available abnormalities will be flagged. Summary statistics will be provided by treatment and visit/time. Shift tables using the low/normal/high/ (low and high) classification will be used to compare baseline to the worst on-treatment value.

Grading of laboratory values will be assigned programmatically as per NCI Common Terminology Criteria for Adverse Events (CTCAE) version 5.0. 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 v5.0, results will be categorized as low/normal/high based on laboratory normal ranges.

The following summaries will be generated separately for hematology, and biochemistry tests:

- Listing of all laboratory data with values flagged to show the corresponding CTCAE v5.0 grades if applicable and the classifications relative to the laboratory normal ranges. Lab data collected during the post-treatment period will be flagged.
- Listing of all CTCAE grade 3 or 4 laboratory toxicities

For laboratory tests where grades are defined by CTCAE v5.0:

- Worst post-baseline CTCAE grade (regardless of the baseline status). Each subject will be counted only once for the worst grade observed post-baseline.
- Shift tables using CTCAE v5.0 grades to compare baseline to the worst on-treatment value

For laboratory tests where grades are not defined by CTCAE v5.0:

- Shift tables using the low/normal/high/ (low and high) classification to compare baseline to the worst on-treatment value.

[redacted]

Liver function parameters

Liver function parameters of interest are total bilirubin (TBIL), ALT, AST and alkaline phosphatase (ALP). The number (%) of patients with worst post-baseline values as per Novartis Liver Toxicity guidelines will be summarized:

The following summaries will be produced:

- ALT or AST > 3xULN
- ALT or AST > 5xULN
- ALT or AST > 10xULN
- ALT or AST > 20xULN
- ALP > 1.5xULN
- TBL > 2xULN
- ALT or AST > 3xULN & TBL > 2xULN
- ALT or AST > 3xULN & ALP < 2xULN
- ALT or AST > 3xULN & TBL > 2xULN & ALP < 2xULN

2.8.4 Other safety data

2.8.4.1 ECG and cardiac imaging data

All ECG data will be listed by treatment group, subject and visit/time, abnormalities will be flagged. Summary statistics will be provided for Notable ECG values by treatment group.

Data handling

In case the study requires ECG replicates at any assessment, the average of the ECG parameters at that assessment should be used in the analyses. At baseline, a minimum of 3 sequential, individual ECGs should be recorded at least 5 minutes apart. The mean QTcF value will be calculated from the triplicate ECGs for each subject.

Data analysis

12-lead ECGs including HR, QT and QTcF will be obtained for each subject during the study. ECG data will be read and interpreted locally

The number and percentage of subjects with notable ECG values will be presented by treatment group.

- QT or 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

Cardiac imaging (rineterkib and NIS793) data will be listed for each assessment points for the corresponding treatment group (s), if deemed appropriate.

2.8.4.2 Vital signs

The following parameters are being collected: height (cm), weight (kg), body temperature (°C), heart rate (beats per minute), and 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

For analysis of vital signs the clinically notable vital sign criteria are provided in Table 2-11 below.

Table 2-11 Clinically notable changes in vital signs

Vital sign (unit)	Clinically notable criteria	
	above normal value	**below normal value**
Weight (kg)	increase > 10% from Baseline	decrease > 10% from Baseline
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%
Body temperature	>= 39.1	≤ 35.0

Vital sign (unit)	Clinically notable criteria	
	above normal value	below normal value
(Celsius)		

The number and percentage of subjects with notable vital sign values (high/low) 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. A separate listing of only the subjects with notable vital sign values may also be produced. In the listing, the assessments collected outside of on-treatment period will be flagged.

2.8.4.3 Ophthalmologic assessment

For rineterkib treatment group, ophthalmologic assessments will be listed, if deemed appropriate. The assessments include but not limited to slit lamp examination, visual acuity testing (preferably Snellen or Early Treatment Diabetic Retinopathy Study (ETDRS)), visual field testing, tonometry (Intra Ocular Pressure (IOP)), and indirect fundoscopy (with dilation) with attention to retinal abnormalities (especially signs of central serous retinopathy and RVO).

2.9 Pharmacokinetic endpoints

Pharmacokinetic parameters will be derived from the individual concentration versus time profile using a non-compartmental method as implemented in Phoenix WinNonlin® (Pharsight, Mountain View, CA) software version *8.0*. The pharmacokinetic parameters described in Table 2-12 will be determined for each investigational drug, as deemed appropriate. Additional PK parameters may be estimated as needed.

2.9.1 PK parameters

The PK parameters that will be determined are shown in Table 2-12. A list of PK parameters will be estimated for individual drug (ruxolitinib, siremadlin, crizanlizumab, sabatolimab, rineterkib or NIS793 as appropriate.

Table 2-12 Non-compartmental PK parameters

AUClast	The AUC from time zero to the last measurable concentration sampling time (Tlast)
AUCinf	The AUC from time zero extrapolated to infinity
AUC0-t	The AUC from zero to specified time points (amount x time x volume-1) for example 12, 24, 336 or 504h
Cmax	The maximum (peak) observed plasma drug concentration
Tlast	Time at which the last measurable concentration was observed (time)
Tmax	The time to reach maximum (peak) plasma, blood, serum, or other body fluid drug concentration after single dose administration (time)
T1/2	The elimination half-life associated with the terminal slope (Lambda_z) of a semi logarithmic concentration-time curve
CL (CL/F)	The total body clearance (or apparent clearance) of drug from the plasma or serum
Vz (Vz/F)	The volume (or apparent volume) of distribution during terminal phase (associated with Lambda_z)
AR	Accumulation Ratio = AUC0-t (multiple Dose)/AUC0-t (single dose)

Descriptive statistics of PK parameters for ruxolitinib, siremadlin, crizanlizumab, sabatolimab and rineterkib will be presented by treatment group for PAS. Due to enrollment halt, PK parameters for NIS793 will be listed but not summarized. Descriptive statistics will include arithmetic and geometric mean, median, SD, and coefficient of variance (CV), geometric CV, minimum and maximum. Zero concentrations will not be included in the geometric mean calculation. Since Tmax is generally evaluated by a nonparametric method, median values and ranges will be given for this parameter. The parameters that require terminal phase determination (AUCinf, T1/2, CL or CL/F and Vz or Vz/F) may not be adequately calculated by non-compartmental methods. All individual PK parameters will be listed using FAS and summarized using PAS.

2.9.2 PK concentrations

Descriptive summary statistics for ruxolitinib, siremadlin, crizanlizumab, sabatolimab, rineterkib and NIS793 concentrations will be provided by treatment group at each scheduled time point for the Pharmacokinetic analysis set. Summary statistics will include n (number of subjects with non-missing values), mean (arithmetic and geometric), SD, CV% (arithmetic and geometric), median, minimum and maximum.

Plasma concentration data for ruxolitinib, siremadlin, crizanlizumab, sabatolimab, rineterkib and NIS793 will be listed by treatment group for Full analysis set.

The mean (+/- SD) concentration-time profiles for ruxolitinib, siremadlin, crizanlizumab, sabatolimab, rineterkib and NIS793 concentrations by treatment group over time will be

displayed graphically for Pharmacokinetic analysis set on the linear view. Further graphical exploratory analyses will be carried out if deemed appropriate.

2.9.3 Handling of PK data below LLOQ or missing

Bio-fluid concentrations will be expressed in mass per volume units.

All concentration values below the lower limit of quantitation (LLOQ) (e.g.: ruxolitinib < 0.5 ng/mL; siremadlin < 1 ng/mL; sabatolimab < 1 ug/mL; crizanlizumab 150 ng/mL; rineterkib 1 ng/mL; NIS793 150 ng/mL) are set to zero by the Bioanalyst, and will be displayed in the listings as zero and flagged. LLOQ values will be treated as zero in any calculations of summary statistics, and treated as missing for the calculation of the geometric means and their CV%. The number of non-zero concentrations will also be reported in the summary statistics.

Missing values for any PK data will not be imputed and will be treated as missing.

2.9.4 Immunogenicity (only applicable to ruxolitinib combination arms with crizanlizumab, sabatolimab and NIS793)

The presence and titer of anti-crizanlizumab, anti-MBG453 and anti-NIS793 antibodies (IG) will be listed by subject for the corresponding treatment group, as appropriate. .

.

[redacted]

[redacted]

[redacted]

[redacted]

2.13 Interim analysis (NOT applicable)

No interim analysis is planned for Part 1. For the combination of ruxolitinib plus siremadlin and ruxolitinib plus LTT462, a BLRM will be used to guide dose escalation decisions.

For Part 2, an interim analysis is planned after all subjects have completed Cycle 8 (arms with NIS793) or Cycle 6 (all other arms)of study treatment or discontinued earlier. As described in Section 1.1, in this interim analysis, futile arms are dropped and the arm(s) to be considered for Part 3 are selected. In this interim analysis, no formal testing of the primary endpoint with respect to the success criterion described in Section 2.5 will be performed, i.e., no formal comparison between combination treatment arms and ruxolitinib will be performed.

No interim analysis is planned for Part 3. An efficacy analysis comparing combination treatment arms and ruxolitinib will only be performed at the end of Part 3 as described in Section 2.5.

3 Sample size calculation

3.1 Primary analysis

3.1.1 Part 1

No formal statistical power calculations were performed to determine sample size for this part of the study.

Cohorts of 3 to 6 evaluable subjects will be enrolled in the dose escalation part including at least six subjects at the RP2D, as described in Section 6.5 of the protocol. Multiple cohorts may be sequentially enrolled to the same dose level. Additional cohorts of 1 to 6 subjects may be enrolled at any dose level below the estimated MTD for further elaboration of safety and pharmacokinetic parameters as required. At least 12 subjects are expected to be treated in the

dose escalation part, for the model to have reasonable operating characteristics relating to its RP2D recommendation.

3.1.2 Part 2

The primary objective of Part 2 is to evaluate the preliminary efficacy of the combination treatments and as outlined in Section 2.5 no formal hypothesis testing will be performed. While the totality of the data will be considered when deciding whether an arm will advance to Part 3, the decision whether an arm is futile will be based on the posterior probability distribution for the responder rate within each arm.

Based on historical data from the COMFORT I and II studies, a primary endpoint response rate of less than 10% is expected for the ruxolitinib single agent arm. A treatment with a true response rate of

- less than 10% generally does not warrant further investigation,
- around 20% is potentially interesting,
- more than 30% is of interest.

Therefore, the criterion for the first interim look is defined as

$$P(p_i > 0.3 | n_i, r_i) > 0.8$$

For a sample size of $n_i = 10$ with a vague prior distribution, *beta(1/2, 1/2)*, for the RR, an arm will be considered showing preliminary efficacy and advanced into Part 3 expansion seamlessly if ≥ 5 subjects are responders. Table 3-1 lists related operating characteristics for various true RRs.

Table 3-1 Probability of any arm to meet 1st interim look criterion for n = 10 and various true RRs

True RR	Single arm	Two arms	Four arms
0.35	0.2485	0.4352	0.6811
0.4	0.3669	0.5992	0.8393
0.45	0.4956	0.7456	0.9353
0.5	0.6230	0.8579	0.9798
0.55	0.7384	0.9316	0.9953
0.6	0.8338	0.9724	0.9992

The table indicates that with the planned sample size of $n_i = 10$, an arm with outstanding efficacy, i.e. a true RR of 0.5 or higher, has more than 60% probability to meet the criterion for

the first interim look. If 2 arms each with true RR of 0.4, the probability of either arm meeting the criterion is close to 60%.

The futility decision will be guided by the criterion

$$P(p_i \leq 0.2 | n_i, r_i) > 0.5$$

For a sample size of $n_i = 25$ with a vague prior distribution, *beta(1/2, 1/2)*, for the RR, an arm fulfills the futility criterion if ≤ 4 subjects are responders. Table 3-2 lists futility related operating characteristics for various true RRs.

Table 3-2 Probability of arm(s) to fulfill futility criterion for n = 25 and various true RRs

True RR	Probability of declaring futility for a single arm	Total probability of declaring futility for both arms when there are two arms in Part 2	Total probability of declaring futility for all 4 arms
0.05	0.9928	0.9857	0.9716
0.1	0.9020	0.8136	0.6620
0.15	0.6821	0.4653	0.2165
0.2	0.4207	0.1770	0.0313
0.25	0.2137	0.0457	0.0021
0.3	0.0905	0.0082	0.0001

The table shows that with the planned sample size of $n_i = 25$, an arm with low activity, i.e. a true RR of 0.1 or smaller, will fulfill the futility criterion with a probability of larger than 90%. For potentially interesting true RRs of $p_i = 0.2$, there is an approximately 40% chance that the arm fulfills the futility criterion. For arms with a true RR of more than 30%, the probability to fulfill the futility criterion is less than 10%. If two combination arms are advanced in Part 2 and both arms have low activity, i.e. the true RR of 0.1 or smaller, the overall probability that both arms are considered futile is larger than 80%. If four combination arms are advanced in Part 2 and all four arms have low activity, i.e. the true RR of 0.1 or smaller, the overall probability that all four arms are considered futile is approximately 66% or higher.

3.1.3 Part 3

Combination treatments not stopped for futility or safety after Part 2 will be considered for expansion into Part 3. After Part 3, each combination treatment will be compared to the ruxolitinib monotherapy arm by evaluating whether the difference in the posterior distributions of the response rates for the primary endpoint is larger than zero with a probability of 90%, i.e.

$$P(p_{comb} - p_{rux} > 0 | n_{rux}, r_{rux}, n_{comb}, r_{comb}) \geq 0.9$$

Table 3-3, Table 3-4 and Table 3-5 show the probability of possible outcomes for various true RRs in the ruxolitinib monotherapy arm (rux) for n_{combo} = 45 and n_{rux} = 30, 35 and 40 respectively, at the primary analysis (PA). Calculations assume the interim analysis (IA) is conducted with 25 subjects. The four possible outcomes are based on the futility at interim analysis (Yes/No) and the treatment meet the success criterion at primary analysis (Yes/No). The first interim look boundary is in line with the second interim look boundary in respect of the require number of responders. In the first interim look, the arm can be expanded, when there are 5 or more responders out of 10 subjects. Similarly, the arm can be expanded when there are 5 or more responders out of 25 subjects because the arm is not futile. In Table 3-3, assuming a true RR of 0.1 for ruxolitinib monotherapy, and a true RR of 0.35 for combination treatment, the probability that the primary analysis is successful when the interim is not futile, is 0.8933.

Table 3-3 Probability for each possible trial outcome of a combination treatment with N combo=45 and N rux=30

True RR for ruxolitinib monotherapy	True RR for combination treatment	IA Not Futile PA Fail	IA Not Futile PA Success	IA Futile PA Fail	IA Futile PA Success
0.05	0.05	0.0037	0.0035	0.9036	0.0893
	0.15	0.0678	0.2501	0.3800	0.3021
	0.2	0.0692	0.5101	0.1659	0.2547
	0.25	0.0467	0.7396	0.0568	0.1570
	0.3	0.0233	0.8863	0.0155	0.0750
0.1	0.1	0.0686	0.0294	0.8342	0.0678
	0.2	0.2328	0.3466	0.3025	0.1182
	0.25	0.2080	0.5783	0.1278	0.0860
	0.3	0.1400	0.7695	0.0435	0.0470
	0.35	0.0746	0.8933	0.0120	0.0201
0.15	0.15	0.2533	0.0646	0.6454	0.0367
	0.25	0.4143	0.3719	0.1759	0.0379
	0.3	0.3375	0.5720	0.0670	0.0235
	0.35	0.2214	0.7465	0.0208	0.0112
	0.4	0.1196	0.8709	0.0053	0.0042
0.2	0.2	0.4935	0.0858	0.4060	0.0146
	0.3	0.5442	0.3654	0.0808	0.0097
	0.35	0.4176	0.5504	0.0269	0.0052
	0.4	0.2709	0.7196	0.0073	0.0021
	0.45	0.1505	0.8472	0.0016	0.0007
Notes: IA = Interim Analysis; PA = Primary Analysis					

Table 3-4 Probability for each possible trial outcome of a combination treatment with N combo=45 and N rux=35

True RR for ruxolitinib monotherapy	True RR for combination treatment	IA Not Futile PA Fail	IA Not Futile PA Success	IA Futile PA Fail	IA Futile PA Success

0.05	0.05	0.0031	0.0041	0.9076	0.0853
	0.15	0.0576	0.2603	0.3446	0.3375
	0.2	0.0580	0.5214	0.1455	0.2752
	0.25	0.0375	0.7487	0.0489	0.1648
	0.3	0.0173	0.8922	0.0132	0.0773
0.1	0.1	0.0672	0.0308	0.8284	0.0736
	0.2	0.2227	0.3566	0.2969	0.1238
	0.25	0.1890	0.5972	0.1251	0.0886
	0.3	0.1174	0.7921	0.0421	0.0483
	0.35	0.0567	0.9113	0.0113	0.0207
0.15	0.15	0.2548	0.0631	0.6465	0.0356
	0.25	0.3955	0.3908	0.1768	0.0369
	0.3	0.3053	0.6042	0.0669	0.0235
	0.35	0.1883	0.7796	0.0205	0.0116
	0.4	0.0955	0.8951	0.0050	0.0044
0.2	0.2	0.4935	0.0858	0.4081	0.0125
	0.3	0.5190	0.3905	0.0811	0.0093
	0.35	0.3852	0.5827	0.0268	0.0052
	0.4	0.2403	0.7502	0.0072	0.0023
	0.45	0.1266	0.8711	0.0016	0.0008

Table 3-5 Probability for each possible trial outcome of a combination treatment with N combo=45 and N rux=40

True RR for ruxolitinib monotherapy	True RR for combination treatment	IA Not Futile PA Fail	IA Not Futile PA Success	IA Futile PA Fail	IA Futile PA Success
0.05	0.05	0.0027	0.0044	0.8851	0.1077
	0.15	0.0429	0.2750	0.3260	0.3561
	0.2	0.0386	0.5407	0.1281	0.2925
	0.25	0.0225	0.7637	0.0395	0.1743
	0.3	0.0096	0.9000	0.0097	0.0808
0.1	0.1	0.0631	0.0349	0.8277	0.0743
	0.2	0.1871	0.3923	0.2812	0.1395
	0.25	0.1522	0.6341	0.1134	0.1003
	0.3	0.0914	0.8182	0.0365	0.0540
	0.35	0.0425	0.9254	0.0094	0.0226
0.15	0.15	0.2447	0.0732	0.6434	0.0387
	0.25	0.3653	0.4209	0.1709	0.0428
	0.3	0.2760	0.6335	0.0636	0.0268
	0.35	0.1648	0.8032	0.0192	0.0129
	0.4	0.0796	0.9109	0.0047	0.0048
0.2	0.2	0.4860	0.0933	0.4066	0.0141
	0.3	0.5006	0.4089	0.0802	0.0103

	0.35	0.3626	0.6054	0.0264	0.0057
	0.4	0.2177	0.7729	0.0071	0.0024
	0.45	0.1085	0.8892	0.0015	0.0008

When the true RRs in the ruxolitinib monotherapy arm and the combination arm shown in Table 3-3, Table 3-4 and Table 3-5 are identical, the probabilities in the 'IA Not Futile PA Success' column can be considered to be the Bayesian equivalent to the type I error rate, i.e., the probability to wrongfully consider an inefficacious treatment as efficacious. This probability is smaller than 10% for all proposed sample sizes (ruxolitinib monotherapy sample size = 30, 35, 40), depending on the treatment arm and group. When the true RRs for the ruxolitinib monotherapy arm is smaller than the combination arm, the probabilities in the 'IA Not Futile PA Success' column can be considered to be the Bayesian equivalent to the power. For a difference in RRs of 20%, the power is approximately 75% or higher for the considered ruxolitinib RRs for all proposed sample sizes.

Next, the probability that the difference between the posterior distributions of the RRs is calculated for several potential trial outcomes and it is illustrated for which outcomes the trial would be considered successful with respect to the primary endpoint. The outcomes are shown in Table 3-6 for $n_{combo} = 45$ and $n_{rux} = 30$, 35, and 40 respectively. The scenarios in which the trial would be considered successful are highlighted in bold.

Finally, Section 5.9 of the protocol shows the simulated probabilities for a list of operating characteristic parameters under three hypothetical Part 2 and Part 3 study scenarios.

Table 3-6 Probability that the difference in the posterior distributions of the RRs is larger than zero

		Number of responders in ruxolitinib monotherapy arm n_{rux} = 30				
		3 (10.0%)	**5 (16.7%)**	**7 (23.3%)**	**9 (30.0%)**	**11 (36.7%)**
Number of responders in combination treatment arm n_{combo} = 45	**5(11.1%)**	0.5519	0.2420	0.0797	0.0206	0.0042
	7(15.6%)	0.7519	0.4432	0.1978	0.0683	0.0186
	9(20.0%)	0.8777	0.6368	0.3617	0.1602	0.0557
	11(24.4%)	**0.9456**	0.7881	0.5399	0.2953	0.1277
	13(28.9%)	**0.9780**	0.8888	0.7004	0.4560	0.2389
	15(33.3%)	**0.9919**	**0.9473**	0.8239	0.6167	0.3816
	17(37.8%)	**0.9973**	**0.9774**	**0.9066**	0.7549	0.5371
	19(42.2%)	**0.9992**	**0.9913**	**0.9554**	0.8583	0.6837
	21(46.7%)	**0.9998**	**0.9970**	**0.9809**	**0.9263**	0.8042
	23(51.1%)	**0.9999**	**0.9990**	**0.9927**	**0.9658**	0.8911
	25(55.6%)	**1.0000**	**0.9997**	**0.9975**	**0.9859**	**0.9461**
	27(60.0%)	**1.0000**	**0.9999**	**0.9993**	**0.9949**	**0.9765**
		Number of responders in ruxolitinib monotherapy arm n_{rux} = 35				
		3 (8.6%)	**5 (14.3%)**	**7 (20.0%)**	**9 (25.7%)**	**11 (31.4%)**
	5(11.1%)	0.6408	0.3327	0.1349	0.0443	0.0120
	7(15.6%)	0.8247	0.5586	0.3004	0.1301	0.0461
	9(20.0%)	**0.9247**	0.7456	0.4971	0.2709	0.1210

Number of responders in combination treatment arm n_{combo} = 45	**11(24.4%)**	**0.9710**	0.8706	0.6794	0.4462	0.2433
	13(28.9%)	**0.9899**	**0.9415**	0.8184	0.6219	0.4015
	15(33.3%)	**0.9968**	**0.9763**	**0.9083**	0.7688	0.5701
	17(37.8%)	**0.9991**	**0.9914**	**0.9587**	0.8737	0.7217
	19(42.2%)	**0.9998**	**0.9972**	**0.9835**	**0.9386**	0.8387
	21(46.7%)	**0.9999**	**0.9992**	**0.9941**	**0.9735**	**0.9167**
	23(51.1%)	**1.0000**	**0.9998**	**0.9981**	**0.9899**	**0.9620**
	25(55.6%)	**1.0000**	**1.0000**	**0.9995**	**0.9966**	**0.9848**
	27(60.0%)	**1.0000**	**1.0000**	**0.9999**	**0.9990**	**0.9947**
		Number of responders in ruxolitinib monotherapy arm n_{rux} = 40				
		3 (7.5%)	**5 (12.5%)**	**7 (17.5%)**	**9 (22.5%)**	**11 (27.5%)**
Number of responders in combination treatment arm n_{combo} = 45	**5(11.1%)**	0.7124	0.4200	0.1994	0.0786	0.0263
	7(15.6%)	0.8756	0.6546	0.4036	0.2068	0.0892
	9(20.0%)	**0.9531**	0.8235	0.6141	0.3884	0.2079
	11(24.4%)	**0.9842**	**0.9213**	0.7822	0.5823	0.3735
	13(28.9%)	**0.9952**	**0.9691**	0.8922	0.7482	0.5555
	15(33.3%)	**0.9987**	**0.9892**	**0.9530**	0.8661	0.7192
	17(37.8%)	**0.9997**	**0.9966**	**0.9819**	**0.9372**	0.8428
	19(42.2%)	**0.9999**	**0.9991**	**0.9938**	**0.9740**	**0.9223**
	21(46.7%)	**1.0000**	**0.9998**	**0.9981**	**0.9906**	**0.9662**
	23(51.1%)	**1.0000**	**1.0000**	**0.9995**	**0.9970**	**0.9872**
	25(55.6%)	**1.0000**	**1.0000**	**0.9999**	**0.9992**	**0.9958**
	27(60.0%)	**1.0000**	**1.0000**	**1.0000**	**0.9998**	**0.9988**

4 Change to protocol specified analyses

No change from protocol specified analysis was made according to protocol amendment 8 per Protocol section 12.

5 Appendix

5.1 Imputation rules

5.1.1 Study drug

The following rule should be used for the imputation of the dose end date for a given study treatment component:

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

The patient should be treated as on-going and the cut-off date should be used as the last dosing date.

Scenario 1 should not applicable for final CSR. All patients should have EOT page complete before the Database lock for Final CSR

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

Case 1: The dose end date is completely missing, and the EOT completion date is complete, then this latter date should be used.

Case 2: Only Year (yyyy) of the dose end date is available and yyyy < the year of EOT date:

Use Dec31yyyy

Case 3: Only Year (yyyy) of the dose end date is available and yyyy = the year of EOT date:

Use EOT date

Case 4: Both Year (yyyy) and Month (mm) are available for dose end date, and yyyy = the year of EOT date and mm < the month of EOT date:

Use last day of the Month (mm)

For the primary CSR, all patients should have completed EOT page for the Part1 core treatment phase or Part 2, or have entered Part 1 extension treatment phase. The on-going status is only applicable for patients entered extension treatment phase.

All other cases should be considered as a data issue and the statistician should contact the data manager of the study.

After imputation, compare the imputed date with start date of treatment, if the imputed date is < start date of treatment:

Use the treatment start date

Patients 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, ConMeds and safety assessment date imputation

Table 5-1 Imputation of start dates (AE, CM) and assessments (LB, EG, VS)

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
	  o If stop date contains a full date and stop date is earlier than study treatment start date then set start date = 01JanYYYY
	  o 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

Missing Element	Rule
Day	• If available month and year = month and year of study treatment start date then o If stop date contains a full date and stop date is earlier than study treatment start date then set start date = 01MONYYYY. o 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 (AE, CM)

Missing Element	Rule (*if end of the on-treatment period 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 and ConMeds with partial/missing dates will be displayed as such in the data listings.

Any AEs and ConMeds which are continuing as per data cut-off will be shown as 'ongoing' rather than the end date provided.

The above imputations are only used for analyses of time to and duration of AEs and concomitant medications.

5.1.2.1 Other imputations

Incomplete date of initial diagnosis of cancer and date of most recent recurrence

Missing day is defaulted to the 15th of the month and missing month and day is defaulted to 01-Jan.

Missing death date

For cases when either day is missing or both month and day are missing for the date of death, the following imputation rules will be implemented:

- If only day is missing, then impute max [(1 mmm-yyyy), min(last contact date+1 , cutoff date)].

If both day and month are missing, then impute max [(1 Jan-yyyy, min (last contact date +1, cutoff date)].

5.2 AEs coding/grading

Adverse events are coded using the Medical dictionary for regulatory activities (MedDRA) terminology.

The latest available MedDRA version at the time of the analyses should be used. The MedDRA version used should be specified in the footnote of relevant tables.

AEs will be assessed according to the Common Terminology Criteria for Adverse Events (CTCAE) version 5.0.

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 generally 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. The latest available version of the document based on the underlying CTCAE version 5 at the time of analysis will be used. The Novartis internal CTCAE grading document should be added as appendix to the CSR.

For laboratory tests where grades are not defined by CTCAE v 5, 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.

Imputation Rules

CTC grading for blood differentials is based on absolute values. However, this data may not be reported as absolute counts but rather as percentage of 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 xxx differential

xxx count = (WBC count) * (xxx %value / 100)

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:

$$\text{Corrected Calcium (mg/dL)} = \text{Calcium (mg/dL)} - 0.8\ [\text{Albumin (g/dL)} - 4]$$

In order to apply the above formula, albumin values in g/L will be converted to g/dL by multiplying by 0.1), calcium values in mmol/L will be converted to mg/dL by dividing by 0.2495. For calculation of laboratory CTC grades 0 and 1, the normal range for derived corrected calcium is set to the same limits (in mg/dL) as for calcium.

CTC grades for the derived absolute WBC differential counts (neutrophils, lymphocytes) and corrected calcium will be assigned as described above for grading

For missing normal ranges, please follow 'Normal ranges (Merck manual, July 2015) and conversion factors' (column 4 of the link: LAB - CTC grades in Novartis Oncology update for CTC Grade 5(26 Jul 2018).pdf)

5.4 Statistical considerations for siremadlin dose-escalation

This appendix provides details of the statistical model, the derivation of prior distributions from historical data, the results of the Bayesian analyses and respective dosing decisions for some hypothetical data scenarios, and a simulation study of the operating characteristics of the model.

[redacted]

5.4.2 Statistical model

Let $\pi(d)$ be the risk of DLT for the combination of ruxolitinib given at combination agent (siremadlin or rineterkib) dose d, the dose-DLT model is logistic:

Figure 5-1 Logistic DLT model.

$$\text{logit}\big(\pi(d)\big) = \log(\alpha) + \beta \log(d/d^*)$$

where d* is used to scale the doses. Hence, $\alpha>0$ is the odds of a DLT at d*; and $\beta>0$ is the increase in the log-odds of a DLT by a unit increase in log-dose.

5.4.3 Prior specifications for logistic parameters

The Bayesian approach requires the specification of prior distributions for the model parameters log(α) and log(β). A meta-analytic-predictive (MAP) approach was used to derive the prior distribution for these model parameters.

5.4.3.1 Description of the meta-analytic-predictive (MAP) approach

The aim of the MAP approach is to derive a prior distribution for the logistic parameters (log(α*), log(β*)) of the new trial using DLT data from historical studies.

Let r_{ds} and n_{ds} be the number of patients with a DLT, and the total number of patients at dose *d* in historical trial *s* ($s = 1,\ldots,S$). The corresponding probability of a DLT is π_{ds}. The model specifications for the derivation of the MAP prior are as follows:

$$r_{ds} \mid \pi_{ds} \sim \text{Bin}(\pi_{ds}, n_{ds})$$

$$\text{logit}(\pi_{ds}) = \log(\alpha_s) + \beta_s \log(d/d^*)$$

$$\big(log(\alpha_s), log(\beta_s)\big) \mid \mu, \psi \sim \text{BVN}(\mu, \psi), \qquad s = 1, \ldots, S$$

$$\big(log(\alpha^*), log(\beta^*)\big) \mid \mu, \psi \sim \text{BVN}(\mu, \psi)$$

The parameters $\mu = (\mu_1, \mu_2)$ and ψ are the mean and between-trial covariance matrix for the logistic parameters, the latter with standard deviations τ_1, τ_2, and correlation ρ. The parameters τ_1 and τ_2 quantify the degree of between trial heterogeneity. The following priors will be used for these parameters:

- normal priors for μ_1 and μ_2,
- log-normal priors for τ_1 and τ_2, and
- a uniform prior for ρ.

The MAP prior for single-agent model parameters in the new trial, $\big(log(\alpha^*), log(\beta^*)\big)$, is the predictive distribution

$$\big(log(\alpha^*), log(\beta^*)\big) \mid (r_{ds}, n_{ds} : s = 1, \ldots, S)$$

Since the predictive distribution is not available analytically, MCMC is used to simulate values from this distribution. This is implemented using JAGS version 4.2.0. The sample from this distribution is then approximated by a mixture of bivariate normal (BVN) distributions. BVN

mixtures with increasing numbers of mixture components are fitted to the sample using the expectation-maximization (EM) algorithm (Dempster et al 1977). The optimal number of components of the mixture is then identified using the Akaike information criterion (AIC) (Akaike H 1974).

5.4.3.2 Single-agent siremadlin

For the MAP model for siremadlin, reference dose $d^* = 50$ mg is used, and data from one historical study with four regimens is available, so $S = 4$ is chosen.

Weakly informative normal priors are assumed for μ_1 and μ_2, with means corresponding to a risk of DLT at the reference dose of 1/3, and a doubling in dose leading to a doubling in the odds of the risk of a DLT, respectively. Priors for τ_1 and τ_2 are assigned such that (1) their medians correspond to substantial between trial heterogeneity, and (2) their uncertainty (95% prior interval) cover plausible between-trial standard deviations (Neuenschwander et al 2014).

The prior distributions for the model used for deriving the MAP priors are specified in Table 5-3.

Table 5-3 Prior distributions for the parameters of the MAP model used to derive the prior for the single-agent model parameters

Parameter	Prior Distribution
μ_1	N (mean = logit (1/3), sd = 2)
μ_2	N (mean = 0, sd = 1)
τ_1	log-normal (mean = log (0.5), sd = log (2)/1.96)
τ_2	log-normal (mean = log (0.25), sd = log (2)/1.96)
ρ	uniform (-1,1)

Historical data from study HDM201X2101

The prior is determined based on solid tumor data from study HDM201X2101.

In this study, four different regimens and dose-schedules where studied:

- 1A: dose on day 1 of a three week cycle
- 1B: dose on day 1 and on day 8 of a four week cycle
- 2A: once dose daily for the first two weeks of a four week cycle
- 2C: once daily dosing for the first week of a four week cycle

The doses in the difference regimens are scaled to fit the schedule of siremadlin in the combination study, which is once daily for the first five days in a four-week cycle.

In study HDM201X2101, the original interest was in DLTs in cycle 1, however after observing cumulative toxicity DLTs within the first two cycles were considered for the recommended dose for expansion (document release date 28-Oct-2016). As the ruxolitinib combination study focuses on DLTs occurring in the first two cycles, the corresponding DLTs from the study HDM201X2101 were considered. When building the MAP prior distribution for siremadlin, the different regimens from study HDM201X2101 were considered as different studies (see Table 5-4).

Table 5-4 Regimens and doses from study HDM201X2101

Regimen	Dose level	Dose per cycle	Dose per day in 5 day schedule	N	DLTs (CYCLE 1)	Hematologic DLTs (CYCLE 2)	DLTs considered for MAP
1A	12.5	12.5	2.5	1	0	0	0
	25	25	5.0	1	0	0	0
	50	50	10.0	4	0	1	1
	100	100	20.0	4	0	0	0
	200	200	40.0	5	0	1	1
	250	250	50.0	6	0	1	1
	350	350	70.0	5	2	2	4
1B	120	240	48.0	9	0	2	2
	150	300	60.0	8	1	1	2
	200	400	80.0	3	0	Data n/a at cutoff	0
2A	1	14	2.8	1	0	0	0
	2	28	5.6	2	0	0	0
	4	56	11.2	4	0	0	0
	7.5	105	21.0	4	0	0	0
	15	210	42.0	4	0	1	1
	20	280	56.0	5	0	4	4
2C	15	105	21.0	8	0	1	1
	20	140	28.0	6	0	0	0
	25	175	35.0	5	2	0	2

5.4.3.3 Summary of prior distributions

To obtain the MAP prior for (log(α), log(β)), the predictive distribution obtained from data in Table 5-4 will be mixed with a vague prior assuming a weight of 0.5. This mixing takes into account the fact the that patient population from study HDM201X2101 and this combination study are not identical and that in this combination study a single-agent model is used to account for both the effect of siremadlin and potential interaction with ruxolitinib.

The MAP prior for (log(α), log(β)) is a mixture of three multivariate normal distributions and it is summarized in Table 5-5.

Table 5-5 Prior distribution for the model parameters

Prior for (log(α), log(β)) in the single agent DLT-model for the combination of siremadlin and ruxolitinib				
	Mean	**Standard Deviations**	**Correlation**	**Weight**
MAP component 1	(-0.771 ,0.492)	(0.700, 0.435)	0.221	0.3255
MAP component 2	(-0.884, 0.051)	(0.573, 0.643)	0.237	0.1745

Prior for (log(α), log(β)) in the single agent DLT-model for the combination of siremadlin and ruxolitinib				
	Mean	**Standard Deviations**	**Correlation**	**Weight**
MAP component 3	(logit(1/2), 0)	(2, 1)	0	0.5

Prior summaries for DLT rates for are summarized in Table 5-6.

Table 5-6 Summary of prior distribution of DLT rates for the combination of siremadlin and ruxolitinib

Siremadlin Dose (mg)	Prior Probabilities that P(DLT) is in the interval:			Mean	SD	Quantiles		
	(0, 0.16)	(0.16, 0.33)	[0.33,1]			2.5%	50%	97.5%
10	0.7982	0.0967	0.1051	0.1191	0.1895	0.0001	0.0443	0.7679
20	0.6526	0.1932	0.1542	0.1793	0.2103	0.0011	0.1046	0.8384
30	0.4687	0.3077	0.2235	0.2392	0.2217	0.0050	0.1713	0.8792
40	0.3035	0.3630	0.3335	0.2996	0.2297	0.0110	0.2425	0.9087

5.4.3.4 Single-agent rineterkib

For the MAP model for rineterkib, reference dose $d^* = 300$ mg is used, and data from $S = 1$ historical studies (First in Human study) is available.

Weakly informative normal priors are assumed for μ_1 and μ_2, with means corresponding to a 10% risk of DLT at the reference dose of 300 mg, and a doubling in dose leading to a doubling in the odds of the risk of a DLT, respectively. Priors for τ_1 and τ_2 are assigned such that (1) their medians correspond to substantial between trial heterogeneity, and (2) their uncertainty (95% prior interval) cover plausible between-trial standard deviations (Neuenschwander et al 2014).

The prior distributions for the model used for deriving the MAP priors are specified in Table 5-7.

Table 5-7 Prior distributions for the parameters of the MAP model used to derive the prior for the rineterkib model parameters

Parameter	Prior Distribution
μ_1	N (mean = logit (0.1), sd = 2)
μ_2	N (mean = 0, sd = 1)
τ_1	log-normal (mean = log (0.5), sd = log (2)/1.96)
τ_2	log-normal (mean = log (0.25), sd = log (2)/1.96)
ρ	uniform (-1,1)

Historical data from Study LTT462X2101

The dose-DLT data from rineterkib single agent from study LTT462X2101 are considered as the relevant information (Table 5-8) and used to derive the prior distribution for the BLRM parameters (log(α), log(β)). The clinical study LTT462X2101 is a phase I dose finding study of

oral rineterkib in adult patients with advanced solid tumors harboring MAPK pathway alterations.

The DLT observation window in the Phase I dose escalation part of this trial was 4 weeks. A review of AE records of the patients did not indicate any significant additional toxicity in the 2nd cycle of treatment, suggesting data from the 1 cycle DLT evaluation period of study LTT462X2101 is also informative for a 2 cycle DLT evaluation period.

Table 5-8 Historical DLT data from Study LTT462X2101

Dose (mg, daily/QD)	Number of patients	Number of DLTs
45	2	0
100	3	0
150	6	1
200	4	1
300	7	0
400	4	0
450	7	2
600	3	3

Summary of prior distributions for rineterkib

To obtain the MAP prior for (log(α), log(β)), the predictive distribution obtained from data in Table 5-9 will be mixed with a vague prior assuming a weight of 0.5. This mixing takes into account the fact the that patient population from study LTT462X2101 and this combination study are not identical and that in this combination study a single-agent model is used to account for both the effect of rineterkib and potential interaction with ruxolitinib.

The MAP prior for (log(α), log(β)) is a mixture of four multivariate normal distributions and it is summarized in Table 5-9.

Table 5-9 Prior distribution for the model parameters

Prior for (log(α), log(β)) in the single agent DLT-model for the combination of siremadlin and ruxolitinib				
	Mean	**Standard Deviations**	**Correlation**	**Weight**
MAP component 1	(-1.737, 0.330)	(1.135, 0.685)	-0.089	0.190
MAP component 2	(-1.518, 0.571)	(0.648, 0.478)	-0.289	0.163
MAP component 3	(-1.491, -0.211)	(0.764, 0.868)	-0.018	0.147
MAP component 4	(logit(1/4), 0)	(2, 1)	0	0.5

Prior summaries for DLT rates for are summarized in Table 5-10.

Table 5-10 Summary of prior distribution of DLT rates for the combination of rineterkib and ruxolitinib

Rineterkib Dose (mg)	**Prior Probabilities that P(DLT) is in the interval:**			**Mean**	**SD**	**Quantiles**		
	(0, 0.16)	**(0.16, 0.33)**	**[0.33,1]**			**2.5%**	**50%**	**97.5%**
100	0.7673	0.1171	0.1155	0.1273	0.1905	0.0002	0.0492	0.7552

Rineterkib Dose (mg)	Prior Probabilities that P(DLT) is in the interval:			Mean	SD	Quantiles		
	(0, 0.16)	(0.16, 0.33)	[0.33,1]			2.5%	50%	97.5%
200	0.6126	0.1989	0.1885	0.1966	0.2183	0.0040	0.1142	0.8444
300	0.4217	0.2923	0.2860	0.2705	0.2378	0.0114	0.1934	0.8999

5.4.4 Hypothetical on-study data scenarios

To illustrate the performance of the Bayesian model used to guide dose escalation, hypothetical dose escalations scenarios following the provisional dose levels specified in Section 6.5.1 in protocol are displayed. In each case, the maximum dose that can be used in the next cohort of subjects is shown. This recommended Phase 2 dose is determined using the model-based assessment of the risk of DLT in future subjects and the dose escalation rules as described in Section 6.5.1 in protocol. In practice, a dose below the maximum might be chosen based on additional safety, PK or PD information (Section 6.5 in protocol).

The first three hypothetical dose escalation scenarios in Table 5-11 are for the case that the first cohort contains six subjects being administered a combination treatment with a siremadlin dose of 20 mg. Scenarios 4, 5, 6 and 7 are for the second cohort and the case that the first cohort contains six subjects receiving 20 mg siremadlin and two subjects having DLTs. Scenarios 8 and 9 demonstrate if enrolling 2 and 3 subjects in Arm 1, the first 2 subjects experience a DLT, then further enrollment of Arm 1 will stop and the combination treatment will not open in Part 2.

Table 5-11 Hypothetical dose escalation scenarios for a cohort size of n=6 subjects receiving a dose of 20 mg

Scenario	Siremadlin Dose [mg]	Number of		Possible next dose levels		
		patients	DLTs	Dose [mg]	Median P(DLT)	P(excessive toxicity)
1	20	6	0	40	0.162	0.122
2	20	6	1	30	0.192	0.148
3	20	6	2	20	0.206	0.219
4	20	6	2			
	20	3	0	30	0.233	0.222
5	20	6	2			
	20	4	0	30	0.221	0.183
6	20	6	2			
	20	5	0	30	0.212	0.159
7	20	6	2			
	20	3	1	20	0.233	0.247
Scenario	**Siremadlin Dose [mg]**	**Number of**		**Provisional lower dose levels**		
		patients	**DLTs**	**Dose [mg]**	**Median P(DLT)**	**P(excessive toxicity)**
8	20	2	2	10*	0.554	0.696
9	20	3	2	10*	0.275	0.435

Scenario	Siremadlin Dose [mg]	Number of		Possible next dose levels		
		patients	DLTs	Dose [mg]	Median P(DLT)	P(excessive toxicity)
*The possible next dose level is not available, as the probability of excessive toxicity of the lower dose does not satisfy EWOC criteria. The combination treatment will not open in Part 2.						

5.5 Statistical considerations for MAP prior in sensitivity analysis

This appendix provides details of the statistical model, the derivation of robustified MAP prior (Schmidli et al 2014) distributions from historical data.

In Section 5.5.1, the statistical methodology for MAP priors is described and in Section 5.5.2 the robust MAP prior for the ruxolitinib arm is derived.

5.5.1 Meta-analytic-predictive (MAP) approach hierarchical model

5.5.1.1 Statistical model

Assume that there are H historical trials for ruxolitinib with the number of responders denoted as $Y_1,\ldots,Y_H$.

For historical studies:

$$Y_h \sim Binomial\ (\pi_h, n_h), \text{ and } \theta_h = logit(\pi_h), \text{ where } h = 1, \ldots, H.$$

For new study:

$$Y_* \sim Binomial\ (\pi_*, n_*), \text{ and } \theta_* = logit(\pi_*).$$

The parameters $\theta_1, \ldots, \theta_H, \theta_*$ are assumed to be exchangeable and normally distributed

$$\theta_1, \ldots, \theta_H,\ \theta_*|\beta, \tau^2 \sim N(\beta\ , \tau^2),$$

where the prior distributions for hyper-parameters population mean μ and between-trial standard deviation τ are as follows:

$$\beta \sim N(0, \sigma_\beta^2)$$
$$\tau \sim HalfNormal\ (0, \sigma_\tau^2)$$

5.5.1.2 MAP prior and robustification

The MAP prior for the parameter in the new study π_* is the predictive distribution

$$\pi_*|Y_h, n_h, \text{ where } h = 1, \ldots, H$$

The predictive distribution is not available analytically. A very large sample from this distribution using Markov chain Monte Carlo (MCMC) approach is simulated. This is implemented using JAGS version 4.2.0. The sample from this distribution is then approximated by a mixture of beta distributions. The beta mixtures with increasing numbers of mixture components are fitted to the sample using the expectation-maximization (EM) algorithm (Dempster et al 1977). The optimal number of components of the mixture is then identified using the Akaike information criterion (AIC) (Akaike H 1974). The mixture components k are defined by a triplet (w_k, a_k, b_k), where w_k is the prior weight of the kth component, $a_k\ and\ b_k$ are the two prior beta distribution parameters of the kth component. The approximated MAP prior is

$$p(\pi_*) = \sum_{k=1}^{K} w_k Beta(\pi_* | a_k, b_k)$$

with positive weights w_k summing up to one.

A robust version of the MAP prior, is introduced by (Schmidli et al 2014) in consideration of potential prior-data conflict as follows.

$$p_R(\pi_*) = (1 - w_R)p(\pi_*) + w_R p_R(\pi_*),$$

where $p(\pi_*)$ is the approximated MAP prior, $p_R(\pi_*)$ is a vague conjugate prior, and w_R is the prior probability that the new trial differs systematically from the historical trials. The prior information is discarded with increasing prior-data conflict due to the heavy tail of the robustified MAP prior.

5.5.2 Robustified MAP prior calculation for ruxolitinib monotherapy

The robustified MAP prior will be used for the response rate in the ruxolitinib monotherapy arm in the sensitivity analysis 1 as specified in Section 2.5.4. The MAP approach as described in Section 5.5.1 will be used for the prior derivation based on historical data from INCB18424-351 (Comfort I) and CINC424A2352 (Comfort II). The vague prior for the response rate in the combination arm is *beta(1/2, 1/2),* the same as the primary analysis.

The prior distributions for the parameters of the MAP model, outlined in Section 5.5.1.1, for the ruxolitinib monotherapy arm are as follows:

$$\beta \sim N(0, 10^2)$$
$$\tau \sim HalfNormal\ (0, \sigma_\tau^2)$$

where a vague prior is set for β, and the Half-normal distribution for τ assumes a substantial between trial heterogeneity with standard deviation 0.91.

Historical data from Comfort I and Comfort II studies were used for deriving the MAP prior of the ruxolitinib monotherapy arm. The response rate was derived following the definition described in Section 1.2, as applicable, for Comfort I and Comfort II. The MF symptoms were measured using MFSAF v2.0 questionnaire with total symptom score (TSS) range from 0-60. Follow the same definition, a TSS in increase of 10 or more from baseline is considered as symptoms worsening for Comfort I. MF symptoms were not assessed for Comfort II and the response rate is calculated based on two criteria: anemia improvement and no spleen volume progression. The Comfort I RR at Week 24 is 5.81%, 9 responders out of 155 subjects in the ruxolitinib arm. The Comfort II RR at Week 24 is 6.16%, 9 responders out of 146 subjects in the ruxolitinib arm.

To obtain the robustified MAP prior for π_* the predictive distribution obtained from the Comfort I and Comfort II RR data will be mixed with a vague prior assuming a weight of 0.5. This mixing takes into account the differences in patient population and symptom assessments between the COMFORT studies and CINC424H12201.

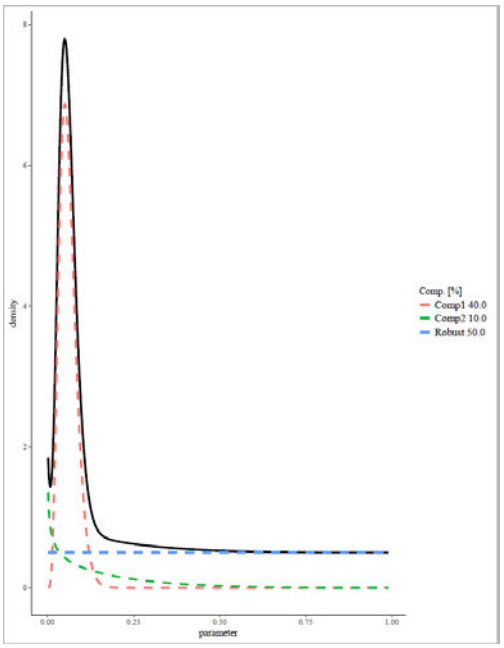
The MAP prior for π_*, is a mixture of three beta distributions and it is summarized in Table 5-12. The prior has mean 0.29, standard deviation 0.30, median 0.11, 95% credible interval (0.02, 0.95).

Table 5-12 Prior distribution for the model parameters

MAP Prior Components	a	b	Weight
MAP component 1	5.56	86.70	0.40
MAP component 2	0.69	4.19	0.10
Robust	1.000	1.000	0.500

The density plot of the MAP prior distribution is displayed in Figure 5-1.

Figure 5-2 Density plot of the MAP prior distribution



5.6 MFSAF v4.0 PRO

Myelofibrosis Symptom Assessment Form version 4.0 7-Day Recall
(MFSAF v4.0 7-Day Recall)

Instructions: The following questions refer to symptoms that you may experience as a result of your myelofibrosis. Please read through and complete the questions on the following screens. There are no right or wrong answers. Please select the answer that best applies to you.

Items:

1. During the past 7 days, how severe was your worst fatigue (weariness, tiredness)?

0	1	2	3	4	5	6	7	8	9	10
Absent										Worst Imaginable

2. During the past 7 days, how severe were your worst night sweats (or feeling hot or flushed)?

0	1	2	3	4	5	6	7	8	9	10
Absent										Worst Imaginable

3. During the past 7 days, how severe was your worst itching?

0	1	2	3	4	5	6	7	8	9	10
Absent										Worst Imaginable

4. During the past 7 days, how severe was your worst abdominal discomfort (feeling pressure or bloating)?

0	1	2	3	4	5	6	7	8	9	10
Absent										Worst Imaginable

5. During the past 7 days, how severe was the worst pain under your ribs on your left side?

0	1	2	3	4	5	6	7	8	9	10
Absent										Worst Imaginable

6. During the past 7 days, what was the worst feeling of fullness you had after beginning to eat?

0	1	2	3	4	5	6	7	8	9	10
Absent										Worst Imaginable

7. During the past 7 days, how severe was your worst bone pain (not joint or arthritis pain)?

0	1	2	3	4	5	6	7	8	9	10
Absent										Worst Imaginable

Myelofibrosis Symptom Assessment Form version 4.0 7-Day Recall (MFSAF v4.0 7-Day Recall)

5.7 EORTC QLQ-C30 PRO

ENGLISH

EORTC QLQ-C30 (version 3)

We are interested in some things about you and your health. Please answer all of the questions yourself by circling the number that best applies to you. There are no "right" or "wrong" answers. The information that you provide will remain strictly confidential.

Please fill in your initials: ⌊ ⌋⌊ ⌋⌊ ⌋⌊ ⌋

Your birthdate (Day, Month, Year): ⌊ ⌋⌊ ⌋ ⌊ ⌋⌊ ⌋⌊ ⌋⌊ ⌋⌊ ⌋⌊ ⌋

Today's date (Day, Month, Year): 31 ⌊ ⌋⌊ ⌋ ⌊ ⌋⌊ ⌋⌊ ⌋⌊ ⌋⌊ ⌋⌊ ⌋

		Not at All	A Little	Quite a Bit	Very Much
1.	Do you have any trouble doing strenuous activities, like carrying a heavy shopping bag or a suitcase?	1	2	3	4
2.	Do you have any trouble taking a long walk?	1	2	3	4
3.	Do you have any trouble taking a short walk outside of the house?	1	2	3	4
4.	Do you need to stay in bed or a chair during the day?	1	2	3	4
5.	Do you need help with eating, dressing, washing yourself or using the toilet?	1	2	3	4

During the past week:		Not at All	A Little	Quite a Bit	Very Much
6.	Were you limited in doing either your work or other daily activities?	1	2	3	4
7.	Were you limited in pursuing your hobbies or other leisure time activities?	1	2	3	4
8.	Were you short of breath?	1	2	3	4
9.	Have you had pain?	1	2	3	4
10.	Did you need to rest?	1	2	3	4
11.	Have you had trouble sleeping?	1	2	3	4
12.	Have you felt weak?	1	2	3	4
13.	Have you lacked appetite?	1	2	3	4
14.	Have you felt nauseated?	1	2	3	4
15.	Have you vomited?	1	2	3	4
16.	Have you been constipated?	1	2	3	4

Please go on to the next page

ENGLISH

During the past week:	**Not at All**	**A Little**	**Quite a Bit**	**Very Much**
17. Have you had diarrhea?	1	2	3	4
18. Were you tired?	1	2	3	4
19. Did pain interfere with your daily activities?	1	2	3	4
20. Have you had difficulty in concentrating on things, like reading a newspaper or watching television?	1	2	3	4
21. Did you feel tense?	1	2	3	4
22. Did you worry?	1	2	3	4
23. Did you feel irritable?	1	2	3	4
24. Did you feel depressed?	1	2	3	4
25. Have you had difficulty remembering things?	1	2	3	4
26. Has your physical condition or medical treatment interfered with your family life?	1	2	3	4
27. Has your physical condition or medical treatment interfered with your social activities?	1	2	3	4
28. Has your physical condition or medical treatment caused you financial difficulties?	1	2	3	4

For the following questions please circle the number between 1 and 7 that best applies to you

29. How would you rate your overall health during the past week?

1	2	3	4	5	6	7
Very poor						Excellent

30. How would you rate your overall quality of life during the past week?

1	2	3	4	5	6	7
Very poor						Excellent

5.8 Scoring the QLQ-C30

Table 1: Scoring the QLQ-C30 version 3.0

	Scale	Number of items	Item range*	Version 3.0 Item numbers	Function scales
Global health status / QoL					
Global health status/QoL (revised)†	QL2	2	6	29, 30	
Functional scales					
Physical functioning (revised)†	PF2	5	3	1 to 5	F
Role functioning (revised)†	RF2	2	3	6, 7	F
Emotional functioning	EF	4	3	21 to 24	F
Cognitive functioning	CF	2	3	20, 25	F
Social functioning	SF	2	3	26, 27	F
Symptom scales / items					
Fatigue	FA	3	3	10, 12, 18	
Nausea and vomiting	NV	2	3	14, 15	
Pain	PA	2	3	9, 19	
Dyspnoea	DY	1	3	8	
Insomnia	SL	1	3	11	
Appetite loss	AP	1	3	13	
Constipation	CO	1	3	16	
Diarrhoea	DI	1	3	17	
Financial difficulties	FI	1	3	28	

* *Item range* is the difference between the possible maximum and the minimum response to individual items; most items take values from 1 to 4, giving *range* = 3.
† (revised) scales are those that have been changed since version 1.0, and their short names are indicated in this manual by a suffix "2" – for example, PF2.

For all scales, the *RawScore*, *RS*, is the mean of the component items:

$$RawScore = RS = (I_1 + I_2 + ... + I_n)/n$$

Then for **Functional scales**:

$$Score = \left\{1 - \frac{(RS-1)}{range}\right\} \times 100$$

and for **Symptom scales / items** and **Global health status / QoL:**

$$Score = \{(RS-1)/range\} \times 100$$

Examples:

Emotional functioning

$$RawScore = (Q_{21} + Q_{22} + Q_{23} + Q_{24})/4$$
$$\text{EF } Score = \{1 - (RawScore - 1)/3\} \times 100$$

Fatigue

$$RawScore = (Q_{10} + Q_{12} + Q_{18})/3$$
$$\text{FA } Score = \{(RawScore - 1)/3\} \times 100$$

5.9 Simulation for hypothetical Part 2 and Part 3 study scenarios

The simulation assumed three groups for Part 2 and Part 3 of the study. Group 1 included Combination 1, Combination 2 and Ruxolitinib monotherapy; Group 2 included Combination 3, Combination 4 and Ruxolitinib monotherapy; Group 3 included Combination 5 and Ruxolitinib monotherapy. The three groups are assumed to be enrolled in a sequential manner within Part 2 and Part 3. Three scenarios are simulated with the true RR for each treatment arm shown in Table 5-13. The probabilities for a list of operating characteristic parameters for the three hypothetical scenarios are presented in Table 5-14.

Table 5-13 Hypothetical true RR for each treatment arm

Scenarios	True RR					
	RUX	COMB1	COMB2	COMB3	COMB4	COMB5
Scenario1	0.05	0.05	0.05	0.05	0.05	0.05
Scenario2	0.05	0.3	0.05	0.05	0.05	0.05
Scenario3	0.05	0.05	0.2	0.3	0.05	0.25

Table 5-14 Probabilities for a list of operating characteristic parameters for hypothetical scenarios of the study

Scenarios	Probabilities			
	All combo arms are futile	All combo arms are futile at IA or not success at PA	Any combo arm(s) is not futile at IA and success at PA	Any combo arm(s) with true RR>0.2 is not futile at IA and success at PA
Scenario1	0.9651	0.9800	0.0200	0.0000
Scenario2	0.0894	0.1145	0.8855	0.8841
Scenario3	0.0100	0.0177	0.9823	0.9713

Scenario 1 assumed all treatment arms have equal true RR and do not warrant further development (RR=0.05). The column 'All combo arms are futile at IA or not success at PA' in Table 5-14 shows the probability to reject all inefficacious treatments at interim or primary analysis. Scenario 2 assumed one combination treatment in Group 1 has a true RR of interest (i.e. RR>0.2). Scenario 3 assumed one combination treatment in each Group has true RR of interest (i.e. RR>0.2). The column 'Any combo arm(s) with true RR>0.2 is not futile at IA and success at PA' shows the probability to select any combination arm or arms that are of interest for the study.

6 Reference

1. Aaronson N. et Al (1993) The European Organization for Research and Treatment of Cancer QLQ-C30: A Quality-of-Life Instrument for Use in International Clinical Trials in Oncology journal of National Cancer Institute, Vol. 85, No. 5 Marche 3. 1993 p. 365-76.
2. Akaike H (1974) A new look at the statistical model identification. Selected Papers of Hirotugu Akaike. Springer, New York, NY, p. 716-23.
3. Dempster AP, Laird NM, Rubin DB (1977) Maximum likelihood from incomplete data via the EM algorithm. Journal of the Royal Statistical Society, Series B p. 39(1):1-38.
4. Clopper CJ and Pearson ES (1934). The use of confidence or fiducial limits illustrated in the case of the binomial. Biometrical, 26, 404-413.Gwaltney C, Paty J, Kwitkowski VE, et al (2017) Development of a harmonized patient-reported outcome questionnaire to assess myelofibrosis symptoms in clinical trials. Leuk. Res. p. 26-31.
5. Neuenschwander, B., et al. (2010). "Summarizing historical information on controls in clinical trials." Clinical Trials 7(1): 5-18.
6. Neuenschwander B, Matano A, Tang Z, et al (2014) A Bayesian Industry Approach to Phase I Combination Trials in Oncology. Statistical Methods in Drug Combination Studies p. 96 - 135.
7. Schmidli, H., et al. (2014). "Robust meta‐analytic‐predictive priors in clinical trials with historical control information." Biometrics 70(4): 1023-1032.
8. Spiegelhalter DJ, Abrams KR and Myles JP (2004). Bayesian Approaches to Clinical Trials and Health-care Evaluation. Chichester, Wiley.