NOVARTIS

Clinical Development

INC424/ruxolitinib/Jakavi®

CINC424A2401 / NCT01493414

An open-label, multicenter, expanded access study of INC424 for patients with primary myelofibrosis (PMF) or post polycythemia myelofibrosis (PPV MF) or post-essential thrombocythemia myelofibrosis (PET-MF)

Statistical Analysis Plan (SAP)
Amendment 2

Author: [redacted]

Document type: SAP Documentation

Document status: Final; Amendment 2

Release date: 26-Apr-2017

Number of pages: 28

Document History – Changes compared to previous final version of SAP

Date	Time point	Reason for update	Outcome for update	Section and title impacted (Current)
08-Mar-2012	Prior to DB lock	Creation of final version	N/A - First version	NA
04-Jul-2016		Creation of amendment 1	This SAP is developed based on the already approved study RAP M3 (approved on 08-Mar-2012). The changes in the SAP reflect changes in the study protocol amendment version 5. The language of the SAP has been updated to be aligned with the protocol amendment version 5. Key changes: 1. Included the definition of the per protocol set (PPS) and 2. Imputation rule for missing/partial dates of study drug 3. Data handling rules for death 4. Updated definitions for transfusion dependency 5. Spleen Response and Loss of Response Criteria by IWG MRT criteria	Section 1: Introduction - Section 1.1 - Section 1.2 Section 2: Patients and treatments - Section 2.3: Section 3: Statistical analysis - Section 3.1.2 - Section 3.2 - Section 3.3 - Section 3.4.1.2 - Section 3.4.2 - Section 3.4.4 - Section 3.5 - Section 3.6 - Section 3.11 Section 4
27-Sep-2016		Creation of final version of amendment 1		
13-Feb-2017		Creation of amendment 2	Key changes: 1. Modified per-protocol set definition 2. Included a listing on non-serious adverse event 3. Included text for tables required to meet leagal requirements of posting safety data on EudraCT and CT.gov sites. 4. Removed section on AEs of interest and relative details from section 16.1.9	Section 2: - Section 2.1 Section 3: - Section 3.5.1 - Section 3.5.2 Section 16.1.9

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
FAS	Full Analysis Set
eCRF	Electronic Case Report Form
INC424	Ruxolitinib
MedDRA	Medical Dictionary for Drug Regulatory Affairs
MRU	Medical Resource Utilization
o.d.	Once Daily
OS	Overall Survival
PET-MF	Post essential thrombocytopenia-myelofibrosis
PFS	Progression-Free Survival
PMF	Primary Myelofibrosis
PPS	Per-Protocol Set
PPV	Polycythemia vera
PRO	Patient-reported Outcomes
qd	Qua'que di'e / once a day
QoL	Quality of Life
RAP	Report and Analysis Process
SAP	Statistical Analysis Plan
SOC	System Organ Class
TFLs	Tables, Figures, Listings
WHO	World Health Organization

1 Introduction

This document is describing the analyses and conventions planned for reporting the INC424A2401 study.

INC424A2401 is a phase IIIB study designed to collect additional safety of INC424 (Ruxolitinib) in patients with primary myelofibrosis (PMF) or post polycythemia vera-myelofibrosis (PPV MF), or post essential thrombocytopenia-myelofibrosis (PET-MF), who have either received prior treatment with commercially available agents or never received treatment.

1.1 Study design

This is a phase IIIB, multicenter, single-arm, open-label expanded access study intended to provide additional data on the safety and efficacy of INC424 in patients with PMF or PPV MF, or PET-MF who have either received prior treatment or who have never received treatment.

Patients will be screened within four weeks prior to the first dose of INC424.

Baseline evaluations will be performed within a week of the first dose of INC424.

Visits will occur on a monthly basis for the first 3 months, then every 3 months thereafter and at study discontinuation.

In this document, study drug refers to INC424.

For this study, study drug will be administered continuously. All patients will be treated with oral INC424 at a dose of 5 - 25 mg (dose based on baseline platelet count) BID until one of the following occurs: disease progression; transplant; unacceptable toxicity; death; discontinuation from the study for any other reason; withdrawal of consent; physician decision; until the study drug becomes commercially available in each participating country or until 24- months after last patient first visit (LPFV), or the project is discontinued, whichever occurs first.

If it is determined at any study visit that the patient will discontinue study drug that day, then that study visit will be the end-of-study visit, and the end-of-study visit procedures will be followed. (Refer to protocol section 7.6.2)

An interim analysis planned for this study will take place when 50 patients with low platelets will have completed 6 months of treatment. Additional analyses may also be performed if needed to fulfill regulatory obligation or publication purpose.

The final analysis will be performed after all patients have either completed the study or have been prematurely discontinued from the study.

1.2 Study objectives and endpoints

This study is intended to collect additional safety of INC424 in patients with PMF or PPV MF, or PET-MF.

1.2.1 Primary Objective

- To collect additional safety of INC424 in patients with PMF, PPV MF, or PET-MF, who have either received prior treatment with commercially available agents or who have never received treatment.

1.2.2 Secondary Objectives:

- To document the best overall response rate to INC424 in patients with PMF, PPV MF, or PET-MF as evaluated by the Investigator.
- To collect quality of life (QoL) endpoints in patients with PMF, PPV MF, or PET-MF treated with INC424.
- To document medical resource utilization (MRU) in patients with PMF, PPV MF, or PET-MF treated with INC424.

2 Patients and Treatments:

Eligible patients for this Expanded Access study are male or female aged ≥ 18 years who are diagnosed with PMF or PPV MF or PET-MF and who have either never received treatment, or who have received prior treatment with commercially available agents or have previously received INC424, patients with high risk OR intermediate risk level 2 OR intermediate risk level 1 with an enlarged spleen.

Patients who are **not** eligible for this Expanded Access study are patients with active hepatitis A, B, C or who are HIV-positive, and patients with coagulation parameters (PT, PTT, or INR) >=1.5 x ULN.

2.1 Analysis sets

Full Analysis Set (FAS) consists of all patients who received at least one administration of study drug.

Safety Set (SS) consists of all patients who received at least one dose of study drug and had at least one post-baseline safety assessment.

Please note: the statement that a patient had no adverse events (on the Adverse Event CRF) constitutes a safety assessment. Patients who have received at least one dose of study drug but who have no post-treatment safety data of any kind would be excluded from the safety population.

Per protocol set (PPS) consists of all FAS patients without any major protocol deviation.

The following major protocol deviations will lead to exclusion from the PPS:.

- Male or female patients aged < 18 years of age at screening (Deviation code: I03)
- Unconfirmed diagnosis of PMF, PPV-MF or PETMF at screening (Deviation code: I04).
- Secondary MF (PPV-MF or PETMF) grade 0 and 1 for Fibrosis on scale 0-3. or grade 0, 1 and 2 on scale 0-4 at screening (Deviation code: I13).

The Full analysis set will be used for all population and efficacy analyses, while Safety set will be used for treatment administration description and safety analyses.

2.2 Subgroup definitions

No sub-group analyses are planned for this study.

2.3 Assessment windows, baseline and post baseline definitions, missing data handling

Assessment windows:

On-treatment, evaluations are scheduled for study visits at Day 1, and at the end of Weeks 4, 8, 12, and every 12 weeks thereafter with time window of +/-7 days.

Baseline definition:

Baseline is the latest available assessment obtained prior to the date of first dose of study drug administration. If assessment prior to first dose of study drug administration is not available, then assessment on the day of first dose of study drug administration will be used at baseline. Otherwise baseline will be considered missing.

Post baseline definition:

All assessments obtained after the baseline assessment (see definition above) are considered as post-baseline.

Study day calculation:

In this study, we will use study day instead of actual event date.

Study day will be calculated as

- event date - start date of study drug + 1 , if the event date is on or after the start date of study drug or
- event date - start date of study drug, if the event date is prior to the start date of study drug

Month for this study will be considered as 4 weeks, so that Month 1 is at Week 4 and so on.

Similarly, quarter would be considered as 12 weeks, so that Quarter 1 is at Week 12 and so on.

Data Handling:

All patients will be followed-up for 28 days after they have discontinued INC424, data collected during this follow-up period will be included in the analysis.

Unless otherwise specified, no imputation for missing data will be performed.

3 Statistical analysis

This section describes the planned statistical analyses for Sections 10-12, 14, and 16 of the clinical study report.

The final analysis for this study will be performed after all patients have either completed or have been prematurely discontinued from the study.

Unless otherwise specified, qualitative data (e.g., gender) will be described using frequency and percentages, while quantitative data, will be described using descriptive statistics: n, mean, standard deviation, median, minimum, and maximum.

3.1 Patient disposition, background and demographic characteristics

3.1.1 Patient disposition

The number of patients included in the Full analysis set will be presented. The number and percentage of patients in the Full analysis set who completed the study, who discontinued the study and the reason for discontinuation, will be presented.

3.1.2 Background and demographic characteristics

The FAS will be used for all baseline and demographic summaries and listings.

Demographic characteristics include age, gender, race, and ethnicity. Baseline characteristics include weight at baseline, height at screening, body mass index and ECOG scores.

Disease history and disease characteristics include type of MF, time since initial diagnosis, hemoglobin, platelets, neutrophils, blast cells, prior splenectomy, hydroxurea, prior radiotherapy, transfusion history, splenic irradiation history and palpable spleen length.

Qualitative data (e.g., gender, race, ECOG performance status) will be summarized by means of contingency tables and quantitative data (e.g., age, body weight) will be summarized by appropriate descriptive statistics (n, mean, standard deviation, median, minimum, and maximum).

3.1.3 Medical history

Any condition entered on the relevant medical history / current medical conditions CRF will be coded using the latest version of MedDRA dictionary at the time of database lock. The number and percentage of patients with each medical condition will be provided for the FAS.

Medical history will be tabulated with the number and percent of patients with significant medical history for each system/organ class.

3.2 Study medication

The duration of study drug administration is defined as the time interval from first (non-zero) dose of study drug administration to last study drug administration date (i.e. last known study drug administration date).

Duration of exposure

The following algorithm will be used to calculate the duration of study drug exposure for patients who took at least 1 dose of study drug:

> *Duration of exposure (days) = [(date of last administration of study drug) – (date of first non-zero dose administration of study drug) + 1 day].*

For patients who discontinued the study drug, last dose date is the last dose date in the EOT page.
For ongoing patients, that is, when they do not have information on EOT page, last dose date will be imputed by the date of interim analysis data cutoff.

Summary statistics on exposure of study drug will be provided. Cumulative dose, average daily dose, actual dose intensity will be summarized by description statistics (n, mean, standard deviation, median, minimum, and maximum) using safety set.

Cumulative dose is defined as the total sum of all doses given during the study drug exposure.

Actual dose intensity (DI) = Cumulative dose (mg) / Duration of exposure (days)

Average daily dose = Total dose / Number of dosing days (drug free day(s) are not counted).

The duration includes the periods of temporary study drug interruption. For patients who did not take any study drug, the duration of exposure is defined as zero. The duration of exposure and duration of follow-up will be summarized by length of exposure, median, as well as by duration category (> 0 - ≤ 4 months, >4 – ≤ 8 months, >8 – ≤ 12 months….)

Duration of on-study follow-up

The duration of on-study follow-up is defined as the time interval from first (non-zero) dose of study drug administration to the last date of contact. Date of last contact will be the last available assessment date on or before the end of study visit.

> *Duration of follow-up (days) = [(last contact date – (date of first non-zero dose administration of study drug) +1 day]*

Dose reduction and dose interruption

A record with a total daily dose of 0 (zero) mg will be considered a dose interruption, if

- the immediate prior total daily dose is non-zero (>0) and
- the immediate following record has a non-zero (> 0) total daily dose.

A dose reduction occurs when a non-missing total daily dose record is followed by a record with a smaller total daily dose. A dose reduction may or may not have a dose interruption in between.

For details see protocol section 6.2.2. Some hypothetical examples are provided here:

> *If an investigator advices a patient to go from (say)*
>
> *a. 20 mg to a halt/stop and then go to 15 mg, that would be considered a dose interruption and also a dose reduction*
>
> *b. 20 mg to 15 mg without a halt/stop in between – there is a dose reduction here, but no dose interruption*
>
> *c. 20 mg to 15 mg after a halt/stop, but the patient never returns to the study, this will be considered a permanent study drug discontinuation*
>
> *In the case of a dose escalation, say, from 10 mg to 15 mg, there should not be any dose interruption in between.*

Imputation rule for missing/partial dates of study drug

The following rule should be used for the imputation of the dose end date:

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 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 (date of death is available in the EOT page), 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 or only year of EOT date is available:

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)

Case 5: Both dose end date and EOT date are completely missing:

Use the last contact data across all panels in the database.

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

Use the treatment start date

Patients with missing start dates are to be considered missing and no imputation will be made. If start date is missing then end-date should not be imputed.

3.3 Concomitant medication:

Prior medications are defined as drugs taken and stopped prior to first dose of study medication. Any medication given at least once between the day of first dose of study medication and the last day of study visit will be a **concomitant** medication, including those which were started pre-baseline and continued into the treatment period.

Prior and concomitant medications will be summarized in two separate tables.

Missing and incomplete start date for concomitant medications will be imputed according to Novartis conventions described in [RAP Module 8].

No imputation algorithm will be applied to any missing stop date.

Table 1-1 Allocation of concomitant medications

If	Concomitant medication
Start date ≤ study drug start date AND Stop > study drug start date OR Stop=missing OR ongoing	X
Start > study drug start date	X

Prior medications to treat MF

Prior medications to treat PMF, PPV MF, or PET-MF will be recorded including regimens discontinued up to 30 days prior to the study entry. This information will be summarized by ATC class and preferred term, and best response on last therapy will be summarized seperately.

3.4 Primary and secondary endpoints analyses

This section describes the planned statistical analyses for Section 11 of the clinical study report. Full analysis set will be used for efficacy data.

3.4.1 Primary endpoints

3.4.1.1 Variables

The analysis of the primary endpoint will be performed on the incidence by patients of **treatment-emergent AEs** considering or not certain AEs characteristics (e.g. relationship to study medication, severity).

Incidence of a given AE is defined as the ratio of total number of patients experiencing the given AE at least once (defined by its preferred term [PT]) divided by the total number of patients at risk (i.e. part of the safety set).

A treatment-emergent AE is an AE that first occurs or worsens during treatment (including the 28 days following last treatment administration).

3.4.1.2 Analysis

No hypotheses or models will be considered for these analyses, descriptive statistics will be provided.

In this section the term adverse event or AE refers only to treatment-emergent AEs unless specified otherwise.

The treatment-emergent AEs will be summarized, while all reported AEs will be included in listings.

The incidence of AEs will be tabulated using MedDRA (Medical Dictionary for Regulatory Activities) preferred term and system organ class (SOC).

Incidence of AEs will also be summarized by system organ class, severity (based on CTC AE grades version 3.0).

The same analysis will be repeated for SAEs regardless of drug relationship, drug related SAEs, AEs which CTCAE grade is 3 or 4 and for drug related AEs.

The following AEs will be summarized –

- AEs by primary system organ class (SOC), preferred term and maximum grade
- AEs suspected to be study drug related by primary system organ class, preferred term and maximum grade

- AEs leading to study drug discontinuation by primary system organ class, preferred term and maximum grade
- Serious adverse events will be analyzed, by primary system organ class, preferred term and maximum grade

Separate summaries for on-treatment and all on-study deaths (on-treatment and post-treatment) will be produced by system organ class and preferred term. All deaths collected in the database will be listed and post treatment and post-follow-up deaths will be flagged.

Note: The death summaries cover subjects from the Safety Set. The count of deaths in safety analyses could be different from the number of deaths in the efficacy analyses.

Data handling rules for death:

The following rule will be used for the analysis of deaths:

1. **On-treatment deaths**: If date of death is within 28 days following last treatment administration or study evaluation completion visit i.e.
   date of death ≤ min (last dose date + 28 days; study evaluation completion visit date)
2. **Post-treatment deaths:** If date of death is greater than 28 days after treatment end date but before the study evaluation completion visit i.e.
   (Last dose date +28 days) < date of death ≤ study evaluation completion visit date
3. **Post-follow up deaths:** If date of death is after the end of study evaluation completion visit i.e. *date of death > study evaluation completion visit date*

3.4.2 Secondary endpoints:

3.4.3 Analysis of secondary endpoints

In this section the term 'percent of change from baseline' at a time T is defined as per following definition:

$$\%\ \text{change at time T} = \frac{Week\ T\ value - Baseline\ value}{Baseline\ value}\ x\ 100$$

At least 50% reduction in spleen length from baseline at any time: The proportion of patients who achieved at least 50% reduction in spleen length at any time will be displayed along with the 95% CI for proportion of reduction.

If a patient is palpable at Baseline and becomes non-palpable at post-baseline, then this post-baseline value will be imputed as 0 (zero).

The proportion of responders (achieving at least 50% reduction in spleen length at any time from baseline) will be estimated for patients whose spleen were palpable at baseline and remained palpable or became non palpable at the post-baseline visit. In addition to the FAS, this analysis will also be done by the PPS.

Time to response is defined as the time between date of start of treatment to the date of first response. Response is defined by the 50% reduction in spleen length from baseline. Summary statistics for those responders will be provided.

Duration of response is defined as the time between date of first response (≥50% reduction in spleen length from baseline) to date of loss of response or death due to MF. Loss of response is defined as going back to baseline level after first time ≥50% reduction in spleen length from baseline. In case a patient does not have loss of response or death due to myelofibrosis by the data cutoff date, DOR will be censored at the date of the last adequate assessment of spleen length OR cut-off date whichever occurs first.

Best overall response to treatment as assessed by spleen palpation is defined by the investigator.

Best overall response is defined as the best response observed during the study follow-up period. The proportion of patients in the best overall response category will be summarized.

In addition, best overall response will be evaluated and summarized by the IWG MRT criteria given in Section 3.11.2.

Change of spleen length from Baseline to end of each visit/month of therapy and at end of treatment. The change in spleen length from Baseline to end of each visit/month of therapy and at end of treatment will be described using descriptive statistics and by relevant change categories:

≤-50%, (-50%, -25%], (-25%, -5%], (-5%, 5%], (5%, 25%], (25%, 50%] and ≥50%

The descriptive summary will include the mean, standard deviation, median, minimum, maximum and 95% CI of the mean over time and at EOT visit for patients who were palpable at baseline and remained palpable or non-palpable at the post-baseline visit considered. In addition, number (n) of patients who remain non palpable, or became palpable from non-palpable at baseline will also be displayed.

WBC and platelet count changes from Baseline at end of each visit/month of therapy and at end of treatment. For each parameter, the percent of change will be computed by dividing the difference of value at assessment time and at baseline by the value at baseline, and multiplying by one hundred. WBC and platelet count changes from baseline will be summarized at end of each visit of therapy and at end of treatment.

Shift in fibrosis in the bone marrow from Baseline, is defined as the status change of fibrosis at time of worst/best response (whenever it occur) compared to fibrosis status at baseline.

Assessment of fibrosis in the bone marrow after enrollment is not mandatory and a listing will be provided with the results of the bone marrow biopsy (if performed).

Progression free survival (PFS) is the time from first study drug administration to date of documented progression (based on International Working Group for Myelofibrosis Research and Treatment Response Criteria) or deaths. Time will be censored at the last date the patient is known to be free of disease progression for patients without event. Deaths on or before the follow up/study completion evaluation visit (on-treatment and post-treatment deaths – Section 3.4.1.2) will be considered in the analyses.

Leukemia free survival (LFS) is the duration from first study drug administration to date of documented leukemia. Observation will be censored, if leukemia is not observed before the

end-of-study, death or lost to follow-up, at last date the patient is known to be free of leukemia. Deaths on or before the follow up/study completion evaluation visit (on-treatment and post-treatment deaths – Section 3.4.1.2) will be considered in the analyses.

Overall survival (OS) is the duration from first dose of study drug administration to the date of death due to any cause. If a patient is alive at the time of data cutoff, OS time will be censored at the time of last contact that is no later than the follow up/study completion evaluation visit. In case of early termination due to any reason, OS observation will be censored at the last contact date. Deaths on or before the follow up/study completion evaluation visit (on-treatment and post-treatment deaths – Section 3.4.1.2) will be considered in the analyses.

Method of analysis of secondary endpoints

The proportion of patients by best overall response category will be summarized. Proportion of responders will be estimated and its 95% confidence intervals provided.

Change in spleen length from Baseline to end of each visit/month of therapy and at end of treatment will be described using descriptive statistics and by relevant change categories (e.g. (≤50%, (-50%, -25%], (-25%, -5%], (-5%, 5%], (5%, 25%], (25%, 50%] and >50%).

WBC and platelet count changes from baseline will be summarized using descriptive statistics at the end of each visit and at the end of treatment.

Listing of patients with bone marrow biopsy will be provided with fibrosis status.

PFS, LFS, and OS will be estimated with the Kaplan Meier method and 95% confidence intervals will be provided for the estimated median.

3.4.4 Analysis of patient reported outcomes

For the analysis of quality of life assessments/patient reported outcomes (PRO), the scoring of the questionnaires will be handled as specified by the instrument developers.

To assess rate of missing data, a table of completion rates for each questionnaire at each scheduled assessment will be created. The table will include number of patients eligible for PRO assessment and number of patients with complete PRO assessment data, completion will be defined by each instruments scoring guidelines.

Change in individual scores/scales from baseline will be summarized descriptively at each visit. For FACT-Lymphoma scores, repeated measures analysis will also be used.

Change in ECOG Performance Status

The ECOG performance score will be summarized descriptively by visit. The summary will include the number and percent of patients at each reported value. A shift table including the number and percent of patients will also be produced by visit of baseline vs. post-baseline scores.

Change in FACT-Lym scores

The FACT-Lym questionnaire consists of a total of 42 questions contributing to 5 subscales (i.e., physical well-being, social/family well-being, emotional well-being, functional well-being and lymphoma subscale) which in turn contribute to three total sum scores, namely

FACT-Lym Trial Outcome Index (TOI), FACT-General (FACT-G) total score and FACT-Lym total score. More details on the questionnaire and derivation of the score can been found in Section 16.1.9.

Data from FACT-Lymphoma scores will be analyzed in the three sum scores and one subscale LymS. Change and percent change from baseline to each scheduled visit when the FACT-Lymphoma scores were collected will be calculated. The scores will be summarized descriptively by visit. The change from baseline score at each time point will be estimated by using repeated measurement analysis for patients with at least one post-baseline scale score. Patients without baseline value will not be considered in the analysis and baseline is the data collected from the last visit prior to or at Day1.

Time on treatment time will be treated as a discrete variable to allow for a nonlinear relationship between PRO scores and time. Time, DIPSS status and baseline score (as a covariate) will be included in the model as fixed effects, while the patient-specific intercept will be included as a random effect.

DIPSS status will be calculated as per the criteria given in section 3.11.3

Change in FACIT Fatigue scores:

Data from FACIT-Fatigue scores will be analyzed in the sum scores. Change and percent change from baseline to each scheduled visit when the FACIT Fatigue scores were collected will be calculated with the data collect from the last visit prior to and at Day 1 as the baseline. The scores will be summarized descriptively by visit.

Responder Analysis and Time to first improvement:

The proportion of patients showing improvement will be estimated for ECOG performance score, the three FACT Lym sum scores and one subscale LymS and FACIT Fatigue scale.

The improvement (i.e. the responder) will be defined by the upper limit of the MID. Patients with the best possible score at baseline will be excluded from this analysis because their HRQoL cannot be further improved.

Additionally, the proportion of patients achieving response (HRQoL) by the lower limit of MID will be presented by visit.

Responders and non-responders for a PRO endpoint will be defined based on change from baseline scores using pre specified cutoff points. Patients with an improved PRO score compared to baseline, for which the magnitude of the change was at least the cutoff value, will be classified as responder; otherwise, non-responder.

To define PRO responder, the cutoff points are chosen based on minimum important difference from literature as listed in the following table.

PRO Score	Range	MID	Responder Cutoff
FACT –G total	0-108	3-7[c]	7
ECOG performance status	0-5	NA	1
FACT-Lym subscale	0-60	2.9-5.4[d]	5.4
FACT-Lym TOI	0-116	5.5-11[d]	11

PRO Score	Range	MID	Responder Cutoff
FACT-Lym Total	0-168	6.5-11.2[d]	11.2
FACIT - Fatigue	0-52	3	3

[a] Kvam et al., 2010; [b] Osaba et al., 1998; [c] Webster et al., 2003; [d] Carter et al., 2008.

The median time to first improvement will be estimated using Kaplan Meier method for ECOG performance score, the three FACT Lym sum scores and one subscale LymS and FACIT Fatigue scale. The time to improvement event will be determined based on upper bound of the MID. The time to improvement is calculated from the date of first study drug administration.

Patients, who prematurely discontinued the study drug, will be censored at the date of their last assessment before study discontinuation.

Patients with no baseline score will be censored at study day 1.

3.4.4.1 Handling of missing values/discontinuations in efficacy analyses

No imputation for missing data will be performed.

Patients with missing baseline assessments will be excluded from analysis where baseline value is required to compute the corresponding endpoint.

3.4.4.2 Center pooling

All data from all centers participating in this trial will be combined and analyzed together.

Unless specified otherwise, all assessments (scheduled and unscheduled) will be used for every patient in the analysis.

3.5 Safety evaluation

This section describes the planned statistical analyses for Sections 10, 12, 14, and 16 of the clinical study report.

Safety analysis will be based on safety population and will use the data collected during the study. Cut off for the safety reporting is end-of-treatment date + 28 days.

The following safety endpoints will be analyzed:

- Treatment-emergent AEs and serious adverse events (SAEs).
- AEs leading to study drug discontinuation.
- Adverse events of interest.
- Clinical laboratory values.
- Vital signs.
- Targeted physical examinations.
- 12-lead electrocardiograms (ECGs).

3.5.1 Adverse events

In this section the term adverse event or AE refers only to treatment-emergent AEs unless specified otherwise.

Only treatment emergent AEs will be summarized, i.e. events started after the first dose of study medication or events present prior to the first dose of study medication but increased in severity during study treatment based on PT, while all reported AEs will be included in listing.

The incidence of AEs will be tabulated by MedDRA (Medical Dictionary for Regulatory Activities) PT and SOC. Incidence of AEs will also be summarized by system organ class, severity (based on CTCAE grades version 3.0).

The same analysis will be repeated for SAEs regardless of drug relationship, drug related SAEs, AEs which CTCAE grade is 3 or 4 and for drug related AEs.

Deaths reportable as SAEs and non-fatal SAEs will be listed by patient and tabulated by type of AE.

Separate summaries for on-treatment and all on-study deaths (on-treatment and post-treatment) will be produced by system organ class and preferred term. All deaths will be listed and post treatment deaths will be flagged.

The most up to date MedDRA will be used for reporting AEs in the study. If a patient reported more than one AE with the same PT, the AE with the greatest severity will be presented. If a patient reported more than one AE within the same primary SOC, the patient will be counted only once with the greatest severity at the SOC level, where applicable.

Separate summaries will be provided for study medication related AEs, death, SAEs, other significant adverse events leading to discontinuation and AEs leading to dose adjustment. Non-serious adverse events will be listed separately.

For the legal requirements of ClinicalTrials.gov and EudraCT, two required tables on on-treatment adverse events which are not serious adverse events with an incidence greater than 5% and on on-treatment 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.

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

The relationship or association of the AE to a study drug will be characterized as Suspected or Not suspected.

Any missing onset date, causality, or severity must be queried for resolution. Unresolved missing values will be imputed according to the following:

- An unresolved missing causality will be imputed by being suspected.
- An unresolved missing severity will be left as missing in the AE tables.
- An unresolved missing onset date will be imputed by the Study Day 1, which will force the AE to be treatment emergent, with the following exceptions:
  - If the stop/resolution date is prior to Study Day 1, the AE will then be considered as not being treatment-emergent.
  - If both the month and day are missing, and the last day of the year is prior to Study Day 1, the AE will then be considered as not being treatment-emergent.
  - If only the day is missing, and the last day of the month is prior to Study Day 1, then the AE will be considered as not being treatment-emergent.

Date imputation for AEs will be imputed according to Novartis conventions described in [RAP Module 8].

3.5.2 Clinical laboratory values

The summary of laboratory evaluations will be presented for three groups of laboratory tests (Hematology, Serum chemistry and coagulation).

All laboratory values will be converted into SI units and the severity grade calculated using the appropriate common toxicity criteria (CTC).

Descriptive summary statistics (mean, median, standard deviation, minimum and maximum) for the change from baseline to each study visit will be presented. These descriptive summaries will be presented by laboratory test. Change from baseline will only be summarized for patients with both baseline and post baseline values and will be calculated as:

change from baseline = post baseline value – baseline value

The following summaries will be produced for the laboratory data (by laboratory parameter):

- Number and percentage of patients with worst post-baseline CTC grade (regardless of the baseline status). Each patient will be counted only for the worst grade observed post-baseline.
- Shift tables using CTC grades to compare baseline to the worst post-baseline value will be produced for hematology biochemistry laboratory parameters and coagulation with CTC grades.

- For laboratory parameters where CTC grades are not defined, shift tables to the worst post-baseline value will be produced using the low/normal/high classifications based on laboratory reference ranges.
- A listing of laboratory values will be provided by laboratory parameter. A separate listing will display notable laboratory abnormalities (i.e. newly occurring CTCAE grade 3 or 4 laboratory toxicities)

3.5.3 Vital signs

Analysis of vital sign measurement using descriptive summary statistics (mean, median, standard deviation, min, max) for the change from baseline to worst post-baseline value will be performed. These descriptive summaries will be presented by vital sign. Change from baseline will only be summarized for patients with both baseline and post-baseline values and will be calculated as:

% of change from baseline = [(post-baseline value – baseline value)/ baseline value] x 100;

The change and percent change in body weight from baseline to each assessment time when body weight was measured will be summarized descriptively. Summary will include the mean, standard deviation, median, minimum, and maximum.

Frequency of weight relative change by at least 20% will be summarized.

The number and percentage of patients with clinically notable vital signs changes from baseline will be presented. Clinically notable vital sign results are provided in Table 1-2 below.

Table 1-2 Criteria for clinically notable vital sign abnormalities

Vital sign (unit)	Clinically notable criteria
Weight (kg)	decrease >= 20% from Baseline
	increase >= 20% from Baseline
Systolic blood pressure (mmHg)	<=90 and decrease from baseline of >=20
	>=180 and increase from baseline of >=20
Diastolic blood pressure (mmHg)	<=50 and decrease from baseline of >=15
	>=105 and increase from baseline of >=15
Pulse (bpm)	<=50 and decrease from baseline of >= 15
	>=120 and increase from baseline of >=15
Respiratory rate (bpm)	>=30 bpm
	<=10 bpm
Body temperature	>=37.5°C
	<=35°C

bpm = beats per minute.

3.5.4 Electrocardiogram (ECG)

A listing of all newly occurring or worsening abnormalities will be provided, as well as a by-patient listing of all ECG parameters.

3.6 Resource Utilization

Frequency of hospitalization, of emergency room visits, general practitioner, Specialist, urgent care, splenectomy and splenic irradiation will be summarized at end of each quarter (i.e. every three months after start of study medications).

Overall duration and number of hospitalization at the end-of-study will be summarized.

Change in blood transfusion dependency

The proportion of patients who are transfusion dependent as well as the proportion of patients, whose transfusion status (dependent or independent) changed (from dependent to independent or vice versa) , will be tabulated with summary statistics. For the purposes of this assessment

- **Transfusion dependence at baseline** will be defined as patient who received 6 or more units of red blood cell product(s) during the 12 weeks prior to first study drug administration.
- **New onset of transfusion dependence** will be defined as the use of 6 or more units of red blood cell product(s) during any consecutive 12 weeks prior to end of study for patients who were not transfusion dependent at baseline.
- **New transfusion independency** will be defined as no (0) use of red blood cell product(s) during any consecutive 12 weeks prior to end of study for patients who met the definition of transfusion dependence at baseline.
- **Transfusion dependence response:** will be defined as the use of 5 or less units of red blood cell product(s) during any consecutive 12 weeks prior to end of study for patients who were transfusion dependent at baseline.

Patients who were not transfusion-dependent at baseline and who required transfusions during the study, but not during the last 12 weeks of observation, will be considered transfusion independent.

The number and percent of patients who received any blood component transfusions will be summarized by blood component (PRBC (Packed Red Blood Cells), platelets), and visit when the transfusion data were collected.

Prior and concomitant medications, as listed in protocol Appendix IX will be summarized.

Concomitant medication at week 24, as listed in protocol Appendix IX, will be summarized.

3.7 Interim analyses

An interim analysis will be performed to closely monitor safety, and ensure an appropriate risk/benefit ratio in the patient population with low platelet counts, between 50,000 and 100,000/μL. In order to adequately assess the risk/benefit ratio, this analysis will also be performed on non-low platelet patients (≥100,000/μL), which will be further detailed in analysis plan, who will serve as an internal reference population. This interim analysis will take place when 50 patients with low platelets have completed 6 months of treatment.

Analyses may be performed if needed to fulfill regulatory obligations, to comply with post-approval commitments or for publication purposes.

3.8 Data monitoring committee (DMC)

No DMC are planned for this study.

3.9 Sample size and power considerations

This is an expanded access study and thus no sample size and no power computation was performed. Patient will be enrolled until INC424 is commercially available at their sites of 24 months after LPFV. The maximum number of patients that will be treated in this study is up to 2500 and is based on country assessments of the number of potential patients. Actual number of enrolled patients may vary and reflect individual site conditions, which may lead to imbalance between centers, or countries.

3.10 Pharmacokinetic analyses

Not applicable.

3.11 Other topics

3.11.1 Handling of missing data

Patients with missing baseline values where a change from baseline is calculated or when the baseline is used as a covariate will be excluded from the analysis of that variable. Missing values in the QoL will be handled following the instruction of the instruments.

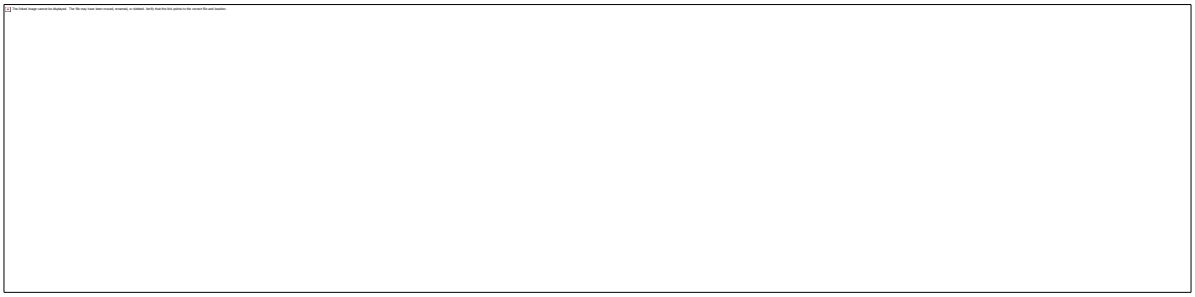
3.11.2 Table of Spleen Response and Loss of Response Criteria by IWG MRT criteria

Patients with spleen length at baseline	**Response criteria**	**Loss of response criteria 1**
Less than 5 cm	Non evaluable for response	Non evaluable for loss of response
Between 5 and 10 cm	Responder: achieve non palpable spleen	Splenomegaly 5 cm or more
	Stable disease: does not meet criteria for response NOR for disease progression	Meet criteria of progressive disease
	Progressive disease: increase of 100% from baseline	-
More than 10 cm	Responder: reduction of 50%	Loss of the 50% reduction
	Stable disease: does not meet criteria for response NOR for	Meet criteria of progressive disease

	disease progression	
	Progressive disease: increase of 50% from baseline	-

3.11.3 Dynamic International Prognostic Scoring System (DIPSS)

The DIPSS can be used to assess a patient's progrnosis as conditions change over time.



The risk level is assigned by the total score:

4 Change to protocol specified analyses

The definitions for the transfusion dependence at baseline, new onset of transfusion dependence and new transfusion independency (protocol v05 section 11.6.4.1) are not current. For the final analysis these definition have been updated as per IWG MRT 2013 (section 3.6 of this SAP).

Clinical Study Report

Appendix 16.1.9 Documentation of statistical methods

FACIT-Fatigue Scale questionnaire:

FACIT-Fatigue scale consists of a total of 13 items (an item is equivalent to a question); each item is scored on a scale of 0 to 4. For some questions, a reversal of the raw item score is required to make all item scores in the same direction: higher the score is, better the QOL is. For each raw item score where reversal is required, the reversal will always be calculated by:

$$New\ Item\ Score = 4 - Raw\ Item\ Score$$

Based on the 13 individual item scores, the total score will be derived:

FACIT- Fatigue total score = sum of all the individual scores

FACIT-Fatigue scoring guidelines

Instructions:

1. Record answers in "item response" column. If missing, mark with an X.
2. Perform reversals as indicated, and sum individual items to obtain a score.
3. Derive total score.
4. The higher the score, the better the QOL.

The FACIT-Fatigue Scale questionnaire consists of the following 13 questions.

1. I feel fatigued
2. I feel weak all over
3. I feel listless (“washed out”)
4. I feel tired
5. I have trouble starting things because I am tired
6. I have trouble finishing things because I am tired
7. I have energy
8. I am able to do my usual activities
9. I need to sleep during the day
10. I am too tired to eat
11. I need help doing my usual activities
12. I am frustrated by being too tired to do the things I want to do
13. I have to limit my social activity because I am tired

FACT-Lymphoma questionnaire

FACT-Lymphoma questionnaire consists of a total of 42 questions; each question takes an item score of 0 to 4. For some questions, a reversal of the raw item score is required to make all item scores in the same direction: higher the score is, better the QOL is. For each raw item score that the reversal is required, the reversal will always be calculated by

$$New\ Item\ Score = 4 - Raw\ Item\ Score$$

Based on the 42 individual item scores, the following 5 subscales will be derived:

1. **Physical Well-Being (PWB)**: it is the mean of non-missing individual item scores from Questions GP1, GP2, GP3, GP4, GP5, GP6, and GP7 multiplied by 7, where each item score is the reverse of the raw item score.
2. **Social/Family Well-Being (SWB)**: it is the mean of non-missing individual item scores from Questions GS1, GS2, GS3, GS4, GS5, GS6, and GS7 multiplied by 7.
3. **Emotional Well-Being (EWB)**: it is the mean of non-missing individual item scores from Questions GE1, GE2, GE3, GE4, GE5, and GE6 multiplied by 6, where each item score is the reverse of the raw item score except for the 2nd item score from Question GE2.
4. **Functional Well-Being (FWB)**: it is the mean of non-missing individual item scores from Questions GF1, GF2, GF3, GF4, GF5, GF6 and GF7 multiplied by 7.
5. **Lymphoma Subscale (LymS)**: it is the mean of non-missing individual item scores from Questions P2, LEU1, BRM3, ES3, LYM1, LYM2, BMT6, C2, GA1, H18, N3, LEU6, LEU7, BRM9, and LEU4 multiplied by 15, where each item score is the reverse of the raw item score.

A subscale will be missing if the corresponding item scores are all missing. Based on the 5 subscale scores, the following 3 total scores will be calculated:

- FACT-Lymphoma Trial Outcome Index (TOI) = PWB + FWB + LymS
- FACT-G total score = PWB + SWB + EWB + FWB
- FACT-Lymphoma total score = PWB + SWB + EWB + FWB + LymS

A total score will be missing if any of the corresponding subscales is missing.

FACT-Lymphoma scoring guidelines

Instructions:

5. Record answers in "item response" column. If missing, mark with an X.
6. Perform reversals as indicated, and sum individual items to obtain a score.
7. Multiply the sum of the item scores by the number of items in the subscale, then divide by the number of items answered. This produces the subscale score.
8. Add subscale scores to derive total scores (TOI, FACT-G & FACT-Lymphoma).
9. The higher the score, the better the QOL.

Subscale	Item code	Reverse item?		Item response	Item score
PHYSICAL WELL-BEING (PWB)	GP1	4	–	_____	=_____
	GP2	4	–	_____	=_____
	GP3	4	–	_____	=_____
Score range: 0-28	GP4	4	–	_____	=_____
	GP5	4	–	_____	=_____
	GP6	4	–	_____	=_____
	GP7	4	–	_____	=_____

Sum individual item scores: _____

Multiply by 7: _____

Divide by number of items answered: _____=PWB subscale score

SOCIAL/FAMILY WELL-BEING (SWB)	GS1	0	+	_____	=_____
	GS2	0	+	_____	=_____
	GS3	0	+	_____	=_____
Score range: 0-28	GS4	0	+	_____	=_____
	GS5	0	+	_____	=_____
	GS6	0	+	_____	=_____
	GS7	0	+	_____	=_____

Sum individual item scores: _____

Multiply by 7: _____

Divide by number of items answered: _____=SWB subscale score

EMOTIONAL WELL-BEING (EWB)	GE1	4	–	_____	=_____
	GE2	0	+	_____	=_____
	GE3	4	–	_____	=_____
Score range: 0-24	GE4	4	–	_____	=_____
	GE5	4	–	_____	=_____
	GE6	4	–	_____	=_____

Sum individual item scores: _____

Multiply by 6: _____

Divide by number of items answered: _____=EWB subscale score

FUNCTIONAL WELL-BEING (FWB)	GF1	0	+	_____	=_____
	GF2	0	+	_____	=_____
	GF3	0	+	_____	=_____
Score range: 0-28	GF4	0	+	_____	=_____
	GF5	0	+	_____	=_____
	GF6	0	+	_____	=_____
	GF7	0	+	_____	=_____

Sum individual item scores: _____

Multiply by 7: _____

Divide by number of items answered: _____=FWB subscale score

FACT-Lymphoma scoring guidelines (Version 4) – Page 2

Subscale	**Item code**	**Reverse item?**		**Item response**	**Item score**
LYMPHOMA SUBSCALE (LYMS)	P2	4	–	_____	=_____
	LEU1	4	–	_____	=_____
	BRM3	4	–	_____	=_____
Score range: 0-60	ES3	4	–	_____	=_____
	LYM1	4	–	_____	=_____
	LYM2	4	–	_____	=_____
	BMT6	4	–	_____	=_____
	C2	4	–	_____	=_____
	GA1	4	–	_____	=_____
	HI8	4	–	_____	=_____
	N3	4	–	_____	=_____
	LEU6	4	–	_____	=_____
	LEU7	4	–	_____	=_____
	BRM9	4	–	_____	=_____
	LEU4	4	–	_____	=_____

Sum individual item scores: _____

Multiply by 15: _____

Divide by number of items answered: _____=LYM subscale score

To derive a FACT-Lymphoma Trial Outcome Index (TOI):

Score range: 0-116

_____ + _____ + _____ =_____=**FACT-Lymphoma TOI**

(**PWB** score) (**FWB** score) (**LymS** score)

To derive a FACT-G total score:

Score range: 0-108

_____ + _____ + _____ + _____=_____=**FACT-G Total score**

(**PWB** score) (**SWB** score) (**EWB** score) (**FWB** score)

To derive a FACT-Lymphoma total score:

Score range: 0-168

(**PWB** score) + (**SWB** score) + (**EWB** score) + (**FWB** score) + (**LymS** score) = **FACT-Lymphoma Total score**