NOVARTIS

Clinical Development

LDE225/Sonidegib

Protocol CLDE225X2116 / NCT01787552

A Phase Ib/II, open-label, multi-center, dose-finding study to assess the safety and efficacy of the oral combination of LDE225(Sonidegib) and INC424(Ruxolitinib) in Patients with myelofibrosis

RAP Module 3 – Detailed Statistical Methodology

Author(s): [redacted]

Document type: RAP Module 3

Document status: Final V 3.0

Release date: 9 August 2016

Number of pages: 46

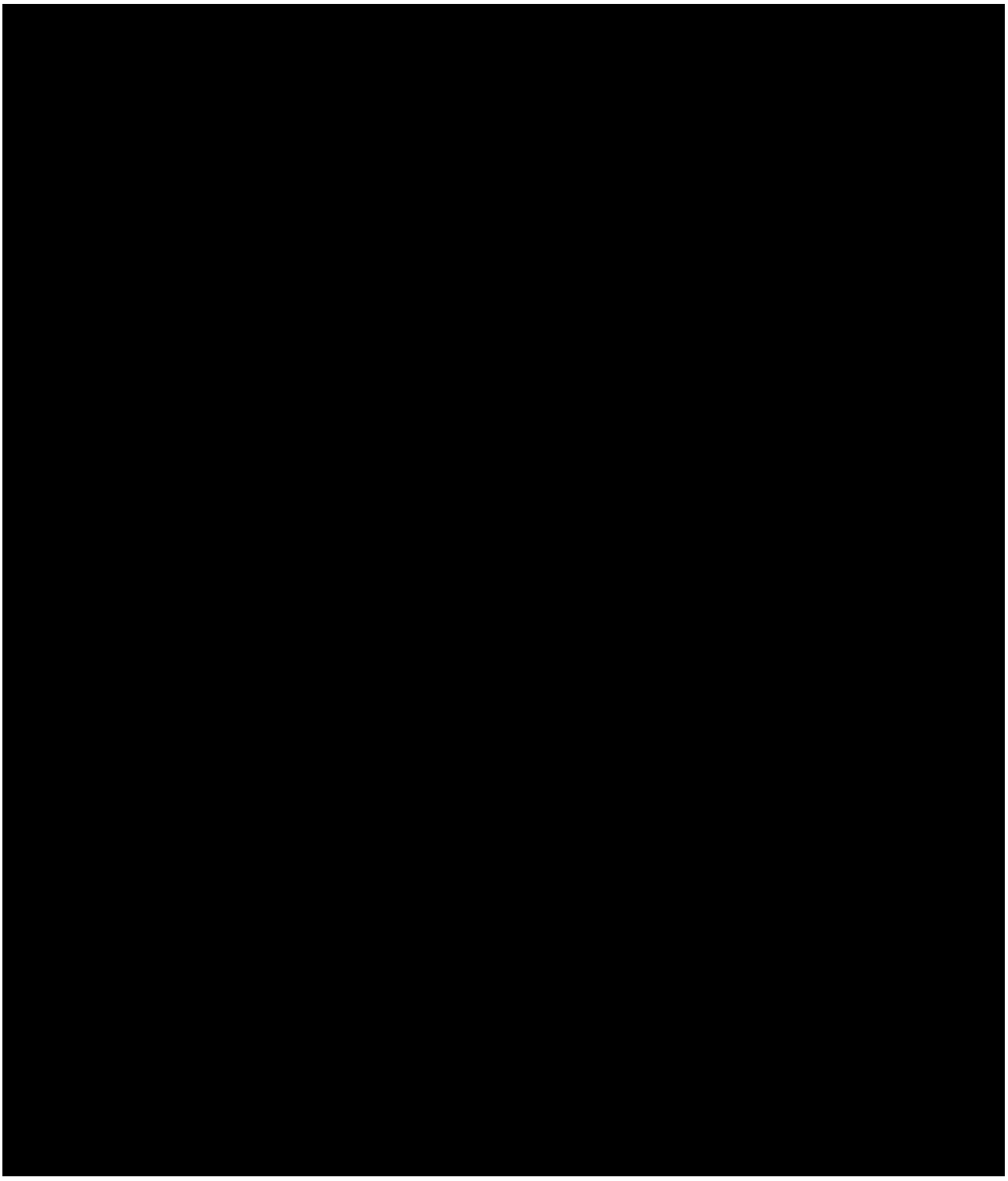


Table of contents

List of figures

List of tables

List of abbreviations

AE	Adverse Event
ATC	Anatomical Therapeutic Chemical
BAS	Biomarker Analysis Set
bpm	Beats per minute
Bw	Body Weight
CDR&R	Clinical Data Review & Reporting
CRF	Case Report Form
CSP	Clinical Study Protocol
CT	Computer Tomography
CTCAE/CTC	Common Terminology Criteria for Adverse Events
CTD	Common Technical Document
DDS	Dose-determining Set
DI	Actual Dose Intensity
DLT	Dose Limiting Toxicity
ECG	Electrocardiogram
EOS	End Of Study
FAS	Full Analysis Set
GCP	Good Clinical Practice
IA	Interim Analysis
IQR	InterQuartile Range
IRT	Interactive Response Technology
ITT	Intention To Treat
JAK2	Janus kinase 2
LLOQ	Lower Limit of Quantification
MedDRA	Medical Dictionary for Regulatory Activities
MF	Myelofibrosis
MRI	Magnetic Resonance Imaging
ms	Millisecond
NCDS	Novartis Clinical Data Standards
NCI	National Cancer Institute
PAS	Pharmacokinetic Analysis Set
PDI	Planned Dose Intensity
PK	Pharmacokinetics
PPS	Per Protocol Set
PRO	Patient Reported Outcomes
RDI	Relative Dose Intensity
RAP	Report and Analysis Plan
RT-PCR	Reverse Transcription-Polymerase Chain Reaction
SAE	Serious Adverse Event
QoL	Quality of Life

SPP	Safety Profiling Plot
SSD	Study Specification Document
UNK	Unknown
WHO	World Health Organization

1 Introduction

This document describes the detailed statistical methodology of the Report and Analysis Plan (RAP) of the study CLDE225X2116. The CLDE225X2116 is a phase Ib/II, an open-label, multi-center, dose-finding study to assess the safety and efficacy of the oral combination of Sonidegib (LDE225) and Ruxolitinib (INC424) in patients with myelofibrosis (MF) (primary myelofibrosis (PMF), post-polycythemia vera myelofibrosis (post-PV MF) or post-essential thrombocythemia myelofibrosis (post-ET MF).

This RAP describes the primary and final analysis of the data as specified in the protocol. The primary analysis of study data will be performed when all patients (Phase Ib and Phase II) have been treated for one year or discontinued treatment. The final analysis of study data for the purposes of generating CSR will be based on all patient data from phase Ib and phase II when all patients have completed at least two years of treatment or discontinued the study.

1.1 Study design

This study consists of an open-label, multi-center, dose-finding phase Ib including dose escalation and expansion followed by a phase II in patients with MF.

The study has 3 parts:

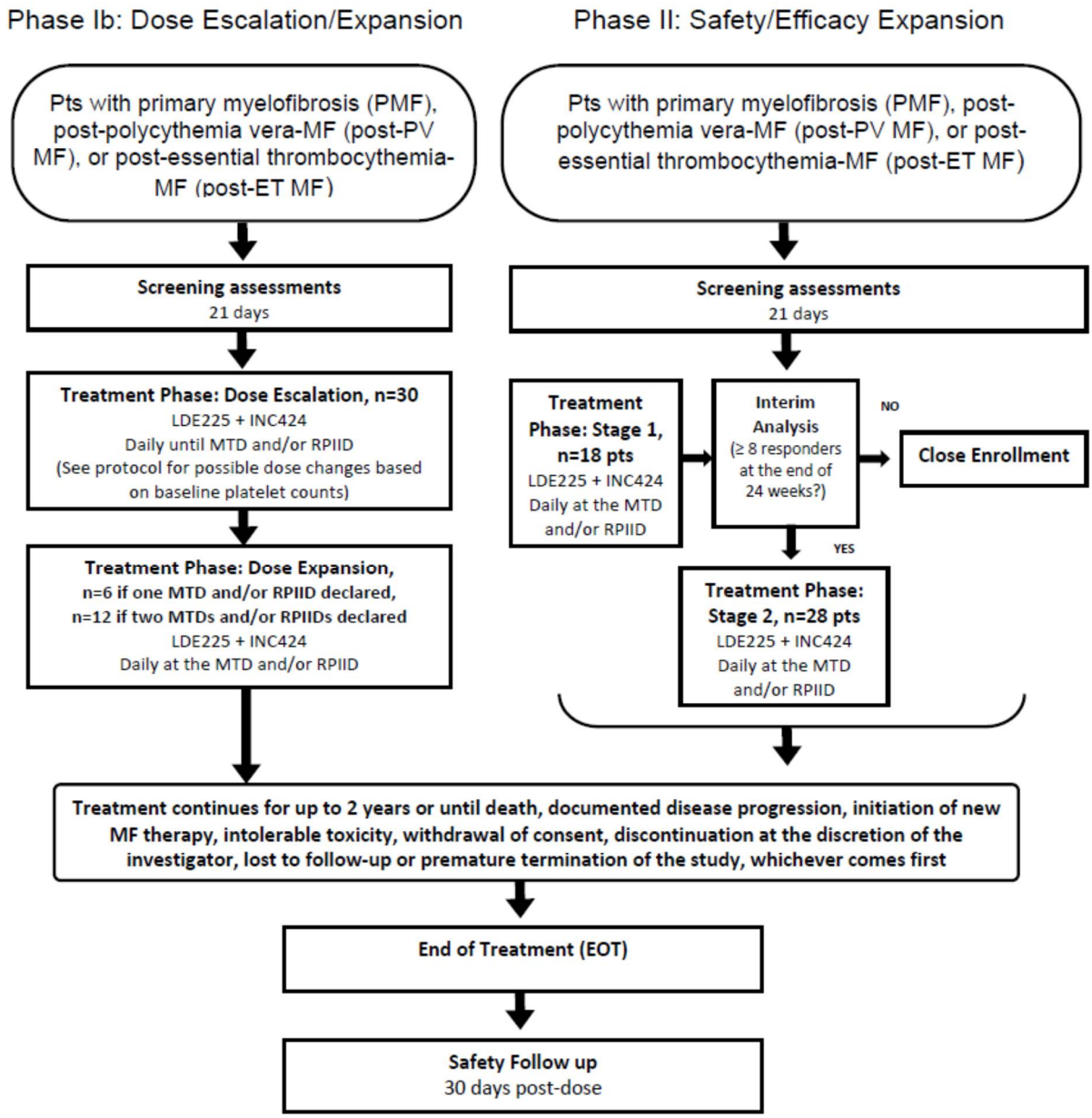
(1) Screening: 21 days screening assessment period prior to receiving the first study treatment.

(2) Treatment: it includes the dose escalation and dose expansion in phase Ib and stage 1 and stage 2 in phase II. The stage 2 enrollment in phase II depends on the result for the interim analysis (IA). The treatment period is from Day 1 for up to 2 years or until death, documented disease progression, initiation of a new MF therapy, intolerable toxicity, withdrawal of content, discontinuation at the discretion of the investigator, lost to follow-up or premature termination of the study, whichever comes first. Based on the recommendation from SC and internal decision, there will be no enrollment to Phase II Stage 2.

(3) Follow up: 30 days following up after the last day of treatment.

The flow chart of the study design is in Figure 1-1.

Figure 1-1 Study design of CLDE225X2116



Note: Patients who complete 2 years of treatment and are deriving clinical benefit in the opinion of the investigator will be allowed to continue receiving treatment (LDE225 and INC424) in an extension phase of this study, until discontinuation reasons are met or an alternative setting to receive study treatment (e.g., in form of another protocol) becomes available.

1.2 Primary Objectives

The primary objective of Phase Ib is to establish the MTD and/or RPIID of the combination of LDE225 (QD) and INC424 (BID) when administered orally to patients with myelofibrosis who have not previously received therapy with a JAK inhibitor.

The primary objective of Phase II is to assess the efficacy of the co-administration of LDE225 and INC424 on spleen volume reduction as determined by centrally reviewed MRI/CT.

1.3 Secondary objectives

Phase Ib:

- To evaluate the safety of the co-administration of LDE225 and INC424 in patients with MF
- To characterize the single and multiple dose pharmacokinetics following the co-administration of LDE225 and INC424

Phase II:

- To assess the effect of the co-administration of LDE225 and INC424 on bone marrow fibrosis by central review and on disease-specific pharmacodynamic biomarkers as a function of the molecular disease characterization of MF
- To evaluate the safety of the co-administration of LDE225 and INC424 in patients with MF
- To characterize pharmacokinetics following the co-administration of LDE225 and INC424
- To assess the effect of the co-administration of LDE225 and INC424 on MF-associated symptoms burden

1.6 Patient population

The patient population will consist of adult patients, aged ≥ 18 years who have been diagnosed with PMF, post-PV MF, or post-ET MF without previous treatment with JAK inhibitor, meet intermediate or high risk prognostic criteria and exhibit palpable splenomegaly ≥ 5cm below the costal margin, and exhibit platelet counts of $\geq 75 \times 10^9/\mathrm{L}$ without the aid of transfusions.

In the CSR, patients will be reported in one of the following patient groups:

- Phase Ib group (All the Phase Ib escalation and expansion patients)
- Phase Ib expansion and Phase II stage 1 group (All the Phase Ib expansion and Phase II stage 1 patients)
- Phase Ib and Phase II stage 1 group (All the Phase Ib escalation, Phase Ib expansion and Phase II stage 1 patients)

according to the corresponding outputs and parameters analyzed.

2 Definitions and general methodology

2.1 Definitions

2.1.1 Study drug and study treatment

Study drug refers to LDE225 and INC424. Study treatment refers to LDE225 XXmg + INC424 XXmg, for example, LDE225 400mg + INC424 10mg.

2.1.2 Monotherapy vs. combination studies

This is a combination study.

The ***combination therapy*** administered is LDE225 jointly with the study treatment partner INC424.

2.1.3 Date of first administration of study treatment

The date of first administration of study drug is derived as the first date when a nonzero dose of study treatment was administered and recorded on LDE225 or INC424 DAR eCRF. The date of first administration of study treatment will also be referred as the ***start date of study treatment***.

2.1.4 Date of last administration of study treatment

The date of last administration of study drug is defined as the last date when a nonzero dose of study drug was administered and recorded on LDE225 or INC424 DAR eCRF. This date of last administration of study treatment will also be referred as the ***last date of study treatment***.

2.1.5 Study day

The term assessment will be used to represent an evaluation, a measurement or an event (e.g. AE onset, laboratory abnormality occurrence, tumor assessment, disease progression, etc)

The study day ***for all assessments*** (efficacy, safety, QoL, etc.) will be calculated as follows:

Study day = date of assessment – start date of study treatment + 1
if date of assessment is after the start date of study treatment.

Study day = date of assessment – start date of study treatment
if date of assessment is before the start date of study treatment.

For any assessment or events such as baseline disease characteristics or medical history (e.g., time since diagnosis of disease) that is supposed to occur prior to the start date of the study treatment, study day will be negative. There is no Day 0 defined.

2.1.6 Baseline

Baseline values for safety and non-safety evaluations are defined as below unless specified otherwise in the RAP.

For ***safety evaluations*** (e.g. laboratory and vital signs), the last assessment available before or on the start date of study treatment (only if such assessment is pre-dose) is taken as a "baseline" assessment, except for the definition of ECG baseline (see Section 3.8.4.1). If there are multiple assessments on the same day and it is not possible to distinguish the last one for baseline assessment, then the average of these multiple assessments will be used with exception that if those assessments are from both central lab and local lab, the assessment(s) from central lab will be used.

For ***efficacy and other non-safety evaluations*** (e.g. MRI scan or photograph), the last available assessment before or on the start of study treatment date and is taken as the "baseline" value or "baseline" assessment. Multiple assessments on the same day are not expected.

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

2.1.7 On-treatment assessment or event

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

pre-treatment period: from day of patient's informed consent to the prior start date of study treatment

on-treatment period: from day of start of study treatment to 30 days following the last date of study treatment:
[Start date of study treatment; last date of study treatment + 30 days].

post-treatment period: starting at day 31 after last dose of study medication.

On-treatment assessment is defined as any assessment reported in the on-treatment period.

2.1.8 Last contact date

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

- All assessment dates (e.g. vital signs assessment, performance status assessment, and also assessment date in third-party data such as tumor imaging, central laboratory, ECG etc.)
- Medication dates including study medication, concomitant medications, antineoplastic therapies administered after study treatment discontinuation.
- Adverse events dates including the start and stop date of AEs
- Date of treatment assignment or start/last date of study treatment
- Date of discontinuation from 'End of Treatment' eCRF

Only dates associated with patient visits or actual examinations of the patient will be used in the derivation. Dates associated with a technical operation unrelated to patient status such as the date a blood sample was processed, will not be used. An imputed partial date will not be used to derive the last contact date. The assessment dates after the cutoff date will not be applied to derive the last contact date.

2.1.9 Screening failure

The purpose of providing screening failures is to ensure that there was no bias in choosing patients who were enrolled into the study.

Screening failures are potential subjects who did not meet one or more screening criteria required for participation into the study. These patients are never treated.

Patients who met inclusion and exclusion criteria but they were never treated will be considered enrolled and will be listed separately.

2.1.10 Month

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

2.1.11 Handling of missing values

2.1.11.1 Adverse events data imputation

Date imputation is the creation of a new, complete date from a partial one according to an agreed and acceptable algorithm.
A partial date is simply an incomplete date e.g.,
ddOCT2001 the days are missing from this DDMMMYYYY date

Partial adverse event start dates, if left partial, would ultimately mean the following

- It would not be possible to place the adverse event in time.
- Therefore the treatment/dosage at the time of the event would be unknown.
- Therefore the event could not be reported/summarized appropriately – if at all.

Therefore it is important to perform date imputation to ensure that as many data events are represented as correctly as possible. Of course partial and/or missing dates should *also* be caught as edit checks and passed back to the investigator for resolution.

There will be no attempt to impute the following:

- Missing AE start dates
- AE start dates missing the year
- Missing AE end dates

Table 2-1 explains the abbreviations used.

Table 2-1 AE/Treatment date abbreviations

	Day	Month	Year
Partial adverse event start date	<not used>	AEM	AEY
Treatment start date (TRTSTD)	<not used>	TRTM	TRTY

Table 2-2 describes the possible combinations and their associated imputations. In the light grey boxes the upper text indicates the imputation (NC, A, B, C, etc.) and the lower text the relationship of the AE start date to the treatment start date (TRTSTD).

Table 2-2 Imputation algorithm

	AEM MISSING	**AEM < TRTM**	**AEM = TRTM**	**AEM > TRTM**
AEY MISSING	NC Uncertain	NC Uncertain	NC Uncertain	NC Uncertain
AEY < TRTY	(D) Before TRTSTD	(C) Before TRTSTD	(C) Before TRTSTD	(C) Before TRTSTD
AEY = TRTY	(B) Uncertain	(C) Before TRTSTD	(B) Uncertain	(A) After TRTSTD
AEY > TRTY	(E) After TRTSTD	(A) After TRTSTD	(A) After TRTSTD	(A) After TRTSTD

The legend to the above table is shown in Table 2-3.

Table 2-3 Imputation algorithm legends

Relationship	
Before TRTSTD	Indicates AE start date prior to Treatment Start Date
After TRTSTD	Indicates AE start date after Treatment Start Date
Uncertain	Insufficient to determine the relationship of AE start date to Treatment Start Date
Imputation calculation	
NC / Blank	No convention/imputation
(A)	01MONYYYY
(B)	TRTSTD+1
(C)	15MONYYYY
(D)	01JULYYYY
(E)	01JANYYYY

Few examples are shown in Table 2-4.

Table 2-4 Example scenarios

Partial AE start date	Treatment start date	Relationship with TRTSTD	Imputation calculation	Imputed date
12mmyyyy	20OCT2001	Uncertain	NC	<blank>
ddmmm2000	20OCT2001	Before	(D)	01JUL2000
ddmmm2002	20OCT2001	After	(E)	01JAN2002
ddmmm2001	20OCT2001	Uncertain	(B)	21OCT2001
ddSEP2001	20OCT2001	Before	(C)	15SEP2001
ddOCT2001	20OCT2001	Uncertain	(B)	21OCT2001
ddNOV2001	20OCT2001	After	(A)	01NOV2001

2.1.11.2 Incomplete date of concomitant medication Diagnosis

The imputation of the start date of concomitant medication will follow the same convention for AE date. Partial concomitant medication end date will not be imputed.

2.1.11.3 Incomplete date of initial diagnosis of myeloproliferative disorder

For incomplete date of initial diagnosis of myeloproliferative disorder, missing day is defaulted to 15 and missing month and day is defaulted to Jan 01 respectively.

2.1.11.4 Incomplete date of anti-neoplastic therapies (myeloproliferative neoplasm directed therapy)

Prior therapies

- **Start date:** The same rule which is applied to the imputation of AE/concomitant medication start date will be used with the exception that: for scenario (B) will be replaced to be start date treatment-1.
- **End date:**
  - Imputed date = min(reference end date, DEC 31) , if month and day are missing.
  - Imputed date = min(reference end date, last day of the Month) , if day is missing.
  - Reference end date will be the start date of treatment.
- If the end date is not missing and the imputed start date is after the end date, use the end date as the imputed start date.
- If both the start date and the end date are imputed and if the imputed start date is after the imputed end date, use the imputed end date as the imputation for the start date

Post therapies

- **Start date:**
  - If Day is missing, then impute to the max (reference start date, first day of the month).
  - Day and month are missing then impute to the max(reference start date, Jan 1)
  - Reference start date will be last date of study treatment administration + 1.
- **End date:** No imputation

2.1.11.5 Incomplete dates for last contact or death

All dates must be completed with day, month and year. If the day is missing, the 1st of the month will be used for incomplete death dates or dates of last contact.

2.1.11.6 Incomplete dates for last treatment administration

Scenario 1: If the last treatment administration date is completely missing and there is no EOT page and no death date:

The patient should be treated as on-going and use the cutoff date as the last dosing date up to date.

Scenario 2: If the last treatment administration date is completely or partially missing and there is either EOT page or death date available:

If only Year is available and Year < Year of min(EOT date, death date);

Imputed date= Dec31yyyy

If both Year and Month are available, Year = Year of min(EOT date, death date) and Month

< the month of min(EOT date, death date);

Imputed date= last day of the Month

All other cases imputed date= min(EOT date, death date).

The imputed date will be compared with start date of treatment. If the imputed date is < start

date of treatment, set the last treatment administration date to be the treatment start date; otherwise use the imputed date.

2.2 Analysis sets

This section defines standard analysis sets and their use for statistical analyses.

2.2.1 Full analysis set (FAS)

Phase Ib and Phase II:

The Full Analysis Set (FAS) comprises all patients in Phase Ib and Phase II who received at least one dose of LDE225 and/or INC424.

Patients will be classified according to the assigned treatment combination. Unless otherwise specified the FAS will be the default analysis set used for all analyses.

2.2.2 Safety analysis set (SAS)

Phase Ib and Phase II:

The Safety Analysis Set (SAS) includes all patients in Phase Ib and Phase II who received at least one dose of LDE225 and/or INC424, and had at least one valid post-baseline safety assessment.

Patients will be analyzed according to the study treatment (regimen) they actually received, where treatment received for each drug is defined as (i) the treatment assigned if it was received at least once, or else (ii) the first treatment received when starting study medication. Each patient will be classified into and analyzed consistently within one (and only one) treatment group.

2.2.3 Dose Determining set (DDS)

The dose Determining set (DDS) consists of all patients from the safety set of Phase Ib who either meet the following minimum exposure criterion and have sufficient safety evaluations (as determined by the investigators and Novartis), or discontinue earlier due to DLT. This constitutes an evaluable patient for the determination of the MTD.

A patient is considered to have met the minimum exposure criterion at treatment group if the patient receives at least 75% of the planned doses of LDE225 and INC424 within 6 weeks (42 days) following the first dose (i.e., no more than 10 missed doses of LDE225 and no more than 20 missed doses of INC424 in 6 weeks are permitted). Patients who do not meet these minimum exposure and safety evaluation requirements will be regarded as ineligible for inclusion into the DDS. Patients in the DDS will be identified before database lock.

2.2.4 Pharmacokinetic analysis set (PAS)

Phase Ib and Phase II:

The pharmacokinetic analysis set (PAS) consists of all patients in Phase Ib and Phase II who receive at least one (full or partial) dose of LDE225 and/or INC424 and provide at least one evaluable PK blood sample for LDE225 and/or INC424.

2.2.5 Biomarker analysis set (BAS)

The Biomarker Analysis Set (BAS) consists of all patients in Phase Ib and Phase II who receive at least one (full or partial) dose of LDE225 and/or INC424 and provided at least one evaluable biomarker sample.

2.3 Protocol deviations

2.3.1 Anti-neoplastic therapy administered after start of study treatment

Patients who take any anti-neoplastic therapy before discontinuing study treatment (i.e. anti-neoplastic therapy other than study treatment) will be identified as protocol deviations. Their efficacy data will be censored so that response assessments made after the intake of an anti-neoplastic drug will not be included in analyses.

3 Data analysis methods

3.1 General presentation of descriptive summaries and listings

Categorical data or qualitative characteristics of a subject (e.g., gender, race, subject disposition, etc.) will be summarized by frequency count (number of subjects) and percentages. Percentages will be calculated using the number of subjects in the relevant population or subgroup as the denominator.

Continuous data (e.g., age, body weight, etc.) will be summarized by appropriate descriptive statistics (mean, standard deviation, median, minimum, and maximum).

Summary of data will be presented by patient group, treatment and overall.

3.1.1 Multiple assessments within post-baseline visits

If there are multiple measurements/samples within the same post-baseline visit, the last measurement/sample within the visit (sorted by date and as available by time or repeat measurements) will be used in the analysis by visit or overall.

For any analyses regarding outlier or abnormal assessments, all post-baseline values will be included (scheduled, unscheduled, repeat). For any other analyses, only scheduled visits will be included. This applies to quantitative and qualitative variables.

For composite assessments these rules apply for the complete sample (e.g. bone marrow differential counts, differential blood counts). It is not allowed to mix these variables across samples.

3.1.2 Center pooling

All study centers will be combined for the analysis unless otherwise specified. In general, no center effect will be assessed due to expected small size of centers.

3.2 Subject disposition, protocol deviations, demographics and other baseline characteristics

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

3.2.1 Subject disposition

FAS will be used for the patient disposition summary tables and listings. The following information will be summarized.

1. Number (%) of patients who are still on-treatment (based on the absence of the ‘End of Treatment’ page);
2. Number (%) of patients who discontinued the study treatment (based on the ‘End of Treatment’ page)
3. Reasons for study treatment discontinuation (based on ‘End of Treatment’ page).

3.2.2 Protocol deviations

The number and percentage of patients in the FAS with any protocol deviation will be tabulated by the deviation category (as specified in the data handling plan) and by treatment group and overall.

Protocol deviations will also be listed by center.

3.2.3 Protocol eligibility criteria

Protocol eligibility criteria will be provided as follows:

1. A list of screen failures along with their failed inclusion/exclusion criteria,
2. A list of informed consent signature dates,
3. A list of enrolled patients but not-treated along with the primary reason for not continuing into the study. The reason for not continuing in the study will be obtained from the end of treatment page. Additional details, if included in the comments domain or in the end of study CRF page, will be included.
4. Patients' eligibility to analysis sets defined in section 2.2 and their allocation to treatment.

3.2.4 Analysis sets

The number and percentage of patients in each analysis set (definitions are provided in section 2.2) will be summarized by treatment groups. Patients' eligibility to analysis sets will also be listed.

3.2.5 Demographics

The following demographic variables will be summarized by study phase following the general presentation specified in section 3.1:

Continuous variables:

- Age (years)
- BMI (kg/m^2)

Categorical variables:

- gender (male/ female)
- race (Caucasian/Black/ Asian/ Native American/ Pacific Islander/unknown/other)
- ECOG performance status at screening (0/ 1/ 2)

3.2.6 Other baseline characteristics

The following baseline characteristics will be summarized:

International Prognostic Scoring System (IPSS) Risk factors at the time of diagnosis, MF classification, prior PRBC transfusion status, and JAK2V617F status, bone marrow fibrosis grade (most recent result prior to study entry).

3.2.7 Medical history / current medical conditions

Medical history and ongoing conditions will be summarized and listed by the redefined study treatment groups (details can be found in RAP M7.1 section 1.2.1). Separate summaries will be presented for ongoing and historical medical conditions. The summaries will be presented by primary system organ class and preferred term. Medical history/current medical conditions are coded using the medical dictionary for regulatory activities (MedDRA) terminology.

3.2.8 Prior Anti-neoplastic therapy

Prior anti-cancer therapy will be listed in three separate listings: myeloproliferative Neoplasm Directed Therapy – Medication, anti-neoplastic radiotherapy and surgery.

The number and percentage of patients recording any prior anti-neoplastic medications (myeloproliferative neoplasm directed therapy-medication), prior anti-cancer radiotherapy and prior anti-cancer surgery will be summarized.

Prior anti-neoplastic medications will be classified and summarized by therapy type (chemotherapy, hormonal therapy, immunotherapy, targeted therapy, and other), ATC class (if appropriate), preferred term and treatment.

Any indication-specific significant prior anti-neoplastic medications will be identified in the summaries mentioned above.

Anti-neoplastic therapies since discontinuation of study drug will be listed and summarized by ATC class, preferred term and treatment and overall by means of frequency counts and percentages in separate summaries.

Partial dates will be imputed according to the algorithm specified in section 2.1.11.4.

3.3 Study drug

Summaries of study drug administration characteristics will be provided for both LDE225 and INC424 separately as indicated: duration of exposure to study drug; exposure to study drug measured by cumulative dose, average daily dose, actual dose intensity (DI), and relative dose intensity (RDI); and dose reductions and interruptions, which are defined in the following sections.

Listings of study drug administration records for LDE225 and INC424, as recorded in CRFs, along with dose reductions and dose interruptions will be provided.

The safety set will be used for all summaries and listings of study treatment.

3.3.1 Duration of exposure

The number and percentage of patients treated with each study treatment will be summarized by study phase. The following formula will be used to calculate duration of study drug exposure for patients who took at least one dose of any of the components of the study treatment. Duration of exposure includes the periods of temporary interruption of study drug for any reason.

For patients *who discontinued permanently* from the study as of the data cut-off date:

Duration of exposure (days) = last dosing date of any study treatment component – first dosing date of any study treatment component + 1

For patients *who are still on-treatment* as of data cut-off date:

Duration of exposure (days) = data cut-off date – (first dosing date of any study treatment component) + 1.

Date of first administration of study treatment and date of last administration of study treatment are defined in section 2.1.3 and section 2.1.4.

In addition, duration of exposure will be categorized in intervals defined in RAP M7. The number and percentage of patients with duration of exposure per category will be summarized. Cumulative duration of exposure will also be summarized. For patients who did not take any study treatment the duration of exposure is defined as zero. The following categories for exposure to study treatment will be analyzed:

- < 2 weeks
- ≥ 2 weeks to < 4 weeks
- ≥ 4 weeks to < 8 weeks
- ≥ 8 weeks to < 12 weeks
- ≥ 12 weeks to < 24 weeks
- ≥ 24 weeks to < 48 weeks
- ≥ 48 weeks to < 106 weeks
- ≥ 106 weeks

3.3.2 Cumulative dose

Cumulative dose (mg) is defined as the total dose given during the study treatment exposure and will be summarized by study phase, treatment components separately.

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

Cumulative dose will be calculated from the dose administration pages of LDE225 and INC424.

3.3.3 Average daily dose

Average daily dose is defined as follows:

Average daily dose (mg/day) = cumulative dose / number of dosing days.

Drug free day(s) are not counted in the number of dosing days.

3.3.4 Actual dose intensity and relative dose intensity

Dose intensity (DI) [mg/day], planned dose intensity (PDI) [mg/day], and relative dose intensity (RDI) [%] will be summarized by study phase, overall (while on-treatment).

Frequencies for RDI will also be categorized in intervals for LDE225 and INC424 as follows: <70%, ≥70% - ≤90% and > 90%. The number and proportion of patients within each category will be presented.

Actual DI for patients with non-zero duration of exposure is defined as follows:

DI (mg/day) = Cumulative dose (mg)/(Duration of exposure (days)).

For patients who did not take any drug, DI will be zero.

Planned dose intensity (PDI) for patients with non-zero duration of exposure is defined as following:

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

The calculation of planned dose is based on the initial assigned dose regardless the dose interruption. Cumulative planned dose is the per-protocol planned dose cumulated over the actual duration on study treatment. For patients who did not take any drug the PDI is by definition equal to zero mg/day.

Relative dose intensity (RDI) is defined as following:

RDI (%) = DI (mg/day)/PDI (mg/day)*100.

3.3.5 Dose reduced, dose interrupted and dose discontinued permanently

Reduction: A reduction is defined as a decrease in dose from the protocol planned starting dose or a decrease from the previous non-zero dose for either INC424 or LDE225, even if this decrease has been directly preceded by an interruption. For example, in the sequence 20mg – 0mg – 10mg, the 10mg dose will be counted as a reduction.
Interruption: An interruption is defined as a zero dose on one or more days between two non-zero doses, for either INC424 or LDE225.

If, due to a dosing error, a patient receives higher than protocol planned starting dose and moves down to the planned starting dose then this is not be counted as a reduction, however if they move directly from a higher than planned starting dose down to a lower than protocol planned starting dose, then this is counted as a reduction.

Note: The last zero dose of INC424 and LDE225 (followed by permanent discontinuation) is not considered as a dose interruption. Additionally, two consecutive zero doses of INC424 (e.g. in the sequence 20 mg daily, 0 mg, 0 mg, 10 mg daily) will be counted as one interruption if the reasons for these two consecutive dose interruption are the same.

The numbers and percentage of patients who have the dose reduction, dose interruption or dose permanent discontinuation will be summarized by redefined study treatment group. For patients who had dose reductions/interruptions, counts of reductions or interruptions,

reductions only and interruptions only will be categorized and presented as specified in RAP M7.1.

3.4 Concomitant therapy

The FAS will be used for all concomitant medication tables and listings.

Concomitant therapy is defined as all interventions (therapeutic treatments and procedures) besides the study treatment that were administered to a subject preceding or coinciding with the study assessment period.

Concomitant medications entered into the database will be coded using the WHO Drug Reference List to allow for categorization by preferred term and using MedDRA for the procedures. In addition to categorizing medication data by preferred term, drugs are classified according to their ATC classification in order to present and compare how they are being utilized.

Concomitant medications and significant non-drug therapies taken concurrently with the study drug(s) will be listed and summarized by ATC class, preferred term and treatment arm by means of frequency counts and percentages. These summaries will include medications starting on or after the start of study treatment or medications starting prior to the start of study treatment and continuing after the start of study treatment.

Any prior concomitant medications or significant non-drug therapies starting and ending prior to the start of study treatment will be listed.

3.5 Primary analysis and primary variables

3.5.1 Phase Ib

3.5.1.1 Primary variable in phase Ib

The primary variable is the incidence of dose limiting toxicities (DLTs) in the first 6 weeks of treatment associated with continuous daily administration of LDE225 in combination of INC424. Estimation of the MTD of the combination treatment will be based upon the estimation of the probability of DLT in 6 weeks for patients in the dose-determining set using BLRM and guided by EWOC. This probability is estimated by the model in section 3.5.1.2.

3.5.1.2 Statistical model, and method of analysis

Statistical model:

The model is formulated in the following way: let $\pi_1(d_1)$ be the probability of a DLT if INC424 is given as a single agent as dose d_1, and let $\pi_2(d_2)$ the probability of a DLT if LDE225 is given as a single agent at dose d_2, further let $\pi_{12}(d_1,d_2)$ be the probability of a DLT when LDE225 and INC424 are given in combination at doses d_1 and d_2 respectively. The dose-DLT relationship is then modeled as:

For a single-agent INC424:

$$\text{logit}(\pi_1(d_1)) = \log(\alpha_1) + \beta_1 \log\left(\frac{d_1}{d_1^*}\right) + \gamma\, I(\text{PLT} \leq 200 \times 10^9/\text{L})$$

For a single-agent LDE225:

$$\text{logit}(\pi_2(d_2)) = \log(\alpha_2) + \beta_2 \log\left(\frac{d_2}{d_2^*}\right)$$

For the combination of LDE225 and INC424:

$$\text{Odds}(\pi_{12}(d_1,d_2)) = \frac{\pi_{12}(d_1,d_2)}{1-\pi_{12}(d_1,d_2)}$$

$$= \exp(\eta \frac{d_1 d_2}{d_1^* d_2^*}) \frac{(\pi_1(d_1) + \pi_2(d_2) - \pi_1(d_1)\pi_2(d_2))}{(1-\pi_1(d_1))(1-\pi_2(d_2))}$$

Where I is the indicator function, $\text{logit}(\pi(d))=\log[\pi(d)/\{1-\pi(d)\}]$, d_1^*=30 mg total daily dose (15 mg bid) and d_2^*=700 mg (qd) are the reference doses of INC424 and LDE225 respectively, α_1, α_2, β_1, β_2, η and $\gamma>0$ are model parameters.

The parameter γ is related to the binary baseline covariate that indicates whether a patient's platelet count at baseline is $\leq 200X10^9$/L or not, and η is an interaction term between the two drugs.

Prior Specification:

The priors used in this study are in Table 3-1. The further details on prior derivations and simulated operation characteristics can be found in the study protocol (Section 14.7, Appendix 7).

Table 3-1 Prior parameters obtained for CLDE225X2116

Parameter	Means	Standard deviations	Correlation
INC424 log(α_1),log(β_1)	(-2.141,-0.180)	0.602,0.634	0.085
LDE225 log(α_2),log(β_2)	(-2.291,0.721)	0.894,0.773	-0.117
η (normal)	0.250	2.028	n/a
γ (log-normal)	-1.426	1.61	n/a

Dose recommendation

Dose recommendation will be based on summaries of the posterior distribution of model parameters and the posterior distribution of DLT rates, including the mean, median, standard

deviation, 95%-credibility interval, and the probability that the true DLT rate for each dose combination lies in one of the following categories:

- [0%, 16%) under dosing
- [16%, 35%) targeted toxicity
- [35%, 100] excessive toxicity

Following the principle of EWOC, after each cohort of patients the recommended dose combination is the one with the highest posterior probability of DLT in the target interval [16%, 35%) among the doses fulfilling the overdose-control criterion that there is less than 25% (posterior probability) chance of excessive toxicity. In addition, the maximum inter-cohort dose escalation is limited to 100% (i.e. up to 100% and 0% increase for LDE225 and INC424 respectively or 0% and up to 100% increase for LDE225 and INC424 respectively). The proposed combination meets the EWOC criteria (i.e. less than 25% chance that the true rate of DLTs lies within the interval [35% - 100%]. A clinical synthesis of the available toxicity information (including AEs that are not DLTs), PK, PD, and efficacy information as well as the recommendations from the BLRM will be used to determine the dose combination for the next cohort at a dose escalation teleconference.

Safety Expansion

After declaration of MTD and/or RPIID, the dose will be expanded by an additional 6 patients in order to further assess tolerability, safety and early efficacy. If two MTDs/RPIIDs are established, each MTD/RPIID will enroll an additional 6 patients in the safety expansion.

Listing of DLTs

DLTs will be listed and their incidence summarized by primary system organ class, worst grade based on the CTCAE version 4.03, type of adverse event by treatment. The dose-determining set will be used for these summaries.

3.5.2 Phase II

3.5.2.1 Primary endpoints in phase II

The primary endpoint in phase II is the proportion of subjects achieving ≥ 35% reduction in spleen volume from baseline by centrally reviewed MRI/CT at the end of week 24 and week 48.

The percent change from baseline at week 24 or week 48 will be calculated by:

$$\%\ \text{change} = \frac{\text{week 24 or 48 value} - \text{baseline vlaue}}{\text{baseline value}}$$

35% Reduction at Week 24 or Week 48: A subject (responder) will be considered as having achieved 35% reduction of the spleen volume from baseline to Week 24 or Week 48 if the subject had evaluable baseline and Week 24 or Week 48 spleen volume assessments, the subject did not experience a protocol-defined progression event prior to the Week 24 or Week 48 visit, and the calculated percent change was less than or equal to –35%.

Listing of the data on response variables as assessed by the investigator will be provided. Recalculated response based on raw measurements will be listed additionally and discrepancies will be identified. Based on this approach, investigator's assessment is considered to be validated and will be used for the primary efficacy analysis.

The uniformly minimum variance unbiased estimator (UMVUE) (Jung and Kim, 2004) for patients achieving 35% spleen volume reduction from baseline as measured by MRI/CT with the exact two-sided 95% confidence interval will be provided in phase II.

Let K be the number of stages, and n_k and X_k denote the number of patients accrued and the number of responders(i.e. subjects achieve ≥35% spleen volume reduction), respectively, during stage k, $1 \le k \le K$. And let $S_k = \sum_{i=1}^{k} X_i$ denote the cumulative number of responders by stage k.

The $\hat{p}_1 = X_1/n_1$ is an unbiased estimator of the response rate p (i.e. achieving 35% spleen volume reduction), where $X_1$and n_1 denote the number of responders and the number patients accrued at stage 1, respectively.

The UMVUE of the response rate ($\tilde{p}$) is obtained by

$$\tilde{p} = E\{\hat{p}_1 | (m, s)\}$$

In a two-stage design,

$$\tilde{p} = \begin{cases} \dfrac{s}{n_1} & m = 1 \\ \dfrac{\sum_{x_1=(a_1+1)\vee(s-n_2)}^{s\wedge(b_1-1)} \binom{n_1-1}{x_1-1}\binom{n_2}{s-x_1}}{\sum_{x_1=(a_1+1)\vee(s-n_2)}^{s\wedge(b_1-1)} \binom{n_1}{x_1}\binom{n_2}{s-x_1}} & m = 2 \end{cases}$$

Where $a \wedge b$ =min(a,b) and $a \vee b$ =max(a,b). $a_1$and $b_1$is the lower and upper boundaries for stage 1. As there is only a lower boundary in the study, the upper boundary b_1 in this case is $b_1 = n_1 + 1$.

Let M denote the stopping stage and let S=S_M denote the total number of responders accumulated up to the stopping stage. The probability mass function of (M, S) in a two-stage design with lower stopping boundaries only is given as

$$f(m, s|p) = \begin{cases} p^s(1-p)^{n_1-s}\binom{n_1}{s} & m = 1, 0 \le s \le a_1 \\ p^s(1-p)^{n_1+n_2-s}\sum_{x_1=a_1+1}^{n_1 \wedge s}\binom{n_1}{x_1}\binom{n_2}{s-x_1} & m = 2, a_1+1 \le s \le n_1+n_2 \end{cases}$$

Exact confidence intervals of the UMVUE are obtained as follows:

$$\Pr(\tilde{p}(M, S) \ge \tilde{p}(m, s) | p = p_L) = \alpha/2$$

and

$$\Pr(\tilde{p}(M,S) \geq \tilde{p}(m,s)|p = p_U) = 1 - \alpha/2$$

The transition from Stage 1 to Stage 2 in phase II study will be based on the response rate in the phase II stage 1 patients in FAS. The response rate will be calculated for the final analysis in patients with phase II.

Update: since a no-go decision has been made and Phase II Stage 2 was not started, the UMVUE estimator of p would be equivalent to the simple proportion $\frac{s}{n_1}$.

3.5.2.1.1 IA design as presented in protocol section 10.7

The IA decision based on originally enrolled 18 patients is presented in Table 3-1. It is possible to terminate the trial at the interim for lack of adequate efficacy, i.e. futility. In this study, futility is equivalent to observe less than 8 responders in 18 patients (Table 3-2).

Table 3-2 Predictive probability of observing response rate of 50% (n=18) or greater at the primary analysis (46 patients) under different interim results

Observed response at IA		Predictive probability of observing response rate ≥50%
# of response	%	
4	22.2	0.002
5	27.8	0.011
6	33.3	0.047
7	38.9	0.143
8	44.4	0.318
9	50.0	0.548
10	55.6	0.762
11	61.1	0.905
12	66.7	0.972
13	72.2	0.994

3.5.2.1.2 IA design based on actual enrolled patients in phase II stage 1

Since 20 patients instead of the protocol planned 18 patients enrolled in the phase II stage 1, the calculation is updated to provide a new threshold for 'go' decision based on phase II stage 1 data (Table 3-3).

Table 3-3 Predictive probability of observing a response rate >= 50% at the final analysis with 46 patients under different interim results with 20 patients

Observed response at IA among 20 patients		Predictive probability of observing a response rate ≥50% in 46 patients
# of response	%	
5	25	0.002
6	30	0.013
7	35	0.051
8	40	0.149
9	45	0.325
10	50	0.552
11	55	0.763
12	60	0.904
13	65	0.971
14	70	0.994
15	75	0.999

If 9 or more responders are observed in IA, then 26 additional patients will be enrolled into this study.

3.5.2.1.3 Exploratory analysis based on all patients with MRI/CT assessed spleen volume

Seven patients in phase Ib expansion underwent the exact same efficacy analysis as those in phase II. This exploratory analysis includes them in the 'go' criterion calculation. As such, 27 patients will be used in IA analysis, and 53 patients (i.e. 7 patients from phase Ib expansion and 46 patients from phase II) will be used for final analysis (Table 3-3).

Table 3-3 Predictive probability of observing a response rate >= 50% at the final analysis with 53 patients under different interim results with 27 patients

Observed response at IA among 27 patients		Predictive probability of observing a response rate ≥50% in 53 patients
# of response	%	
8	29.6	0.001
9	33.3	0.007
10	37.0	0.028
11	40.7	0.086
12	44.4	0.207
13	48.1	0.393

14	51.9	0.607
15	55.6	0.793
16	59.3	0.914
17	63.0	0.972
18	66.7	0.993
19	70.4	0.999

If 12 or more responders are observed in IA, then 26 additional patients will be enrolled into this study.

3.5.3 Handling of missing values/censoring/discontinuations

The reason for discontinuation from study will be summarized and listed, along with dates of first and last study treatment, duration of exposure to study treatment and date of discontinuation for each patient. Other missing data will simply be noted as missing on the appropriate tables/listings.

Phase Ib:

Patients in the dose escalation phase are ineligible for the dose-determining set will not be replaced.

Phase II:

Patients with unknown clinical response (e.g. patients with missing assessments at the end of 24 or 48 weeks) will be considered as non-responders. No imputation of missing values will be made for the efficacy analysis.

A subject must have a baseline spleen volume in order to be included in the primary efficacy analysis. Subjects who drop out of the study due to any reason, such as lack of efficacy or treatment-related adverse events (AEs), or who have a protocol-defined event of disease progression prior to Week 24 or Week 48 visit will all be considered as having not achieved the $\geq$ 35% reduction at Week 24 or Week 48 respectively.

3.6 Secondary analysis

3.6.1 Bone marrow assessment

Bone marrow variables as assessed by aspirate and/or biopsy will be listed. Fibrosis grade assessed by bone marrow histomorphology will be summarized by the collection time points (week 1, week 25 and every 24 weeks thereafter).

The following variables will be listed:

Bone marrow aspirate

- Was the specimen adequate for assessment (yes/no),

•Specimen quality: dry tap/hemodiluted/other

•Blast cell (%)

Bone marrow biopsy

•Was the specimen adequate for assessment? (yes/no)

•Percentage of cellularity

•Cellularity: Hypocellular/Normocellular/Hypercellular/Not assessable/Aplastic

•Blast cell (%)

•Bone marrow fibrosis grade

3.7 Exploratory analysis

3.7.2 Spleen length by manual palpation

Frequency of patients who have palpable spleen measurement will be provided. Spleen measurement by manual palpation at week 24 will be compared to the week 1 day 1 measurement. Descriptive statistics will be provided for the spleen length and percent change from baseline over time by the redefined treatment groups (details can be found in RAP M7.1 section 1.2.1) and overall. The maximum likelihood estimate and 95% exact binomial confidence interval will be also provided for patient who has ≥50% spleen length reduction by manual palpation.

Data will be summarized for the set of patients from the FAS who have a valid corresponding baseline assessment.

The percentage change in spleen length is defined as:

% Change = (Post-baseline spleen length – Baseline spleen length) ×100 / Baseline spleen length.

3.8 Safety evaluation

The assessment of safety will be based mainly on the frequency of adverse events and on the number of laboratory values that fall outside of pre-determined ranges. Other safety data (e.g. electrocardiogram, vital signs) will be considered as appropriate.

All safety outputs will be provided by study phase, treatment and overall by using the safety set. The safety summary tables will include only assessments collected no later than 30 days after study treatment discontinuation. All safety assessments will be listed and those collected later than 30 days after study treatment discontinuation will be flagged.

Adverse events of special interest will be summarized according to Section 3.8.1.5.

3.8.1 Adverse events data

3.8.1.1 Coding of AEs

AEs will be coded using the Medical dictionary for regulatory activities (MedDRA) terminology. The latest version of MedDRA prior to the database lock will be used for reporting.

3.8.1.2 Grading of AEs

AEs will be graded according to the Common Terminology Criteria for Adverse Events (CTCAE) version 4.03.

If CTCAE grading does not exist for an adverse event, grades 1 – 4 corresponding to the severity of mild, moderate, severe, and life-threatening will be used. CTCAE grade 5 (death) will not be used in this study; rather, this information will be collected on the “End of Treatment” eCRF pages. In addition, information related to the death of a patient while in the study or within the follow-up period specified in the protocol will be collected in “Death” CRF page.

3.8.1.3 General rules for AE Reporting

AE summaries will include all AEs starting on or after study day 1 (i.e. on or after the day of the first intake of study treatment) and starting no later than 30 days after the end of treatment. All AEs will be listed. AEs starting prior to study day 1 and AEs starting later than 30 days after the last treatment date will be flagged in the listings.

AEs will be summarized by dose level for the number and percentage of patients having at least one AE, in each primary system organ class and for each preferred term (using MedDRA coding). A subject with multiple occurrences of an AE will be counted only once in the corresponding AE category. Separate AE summaries will be presented by the maximum CTCAE grade. A patient with multiple CTCAE grades for an AE will be summarized under the maximum CTCAE grade recorded for the event.

In addition, a summary of serious and non-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 per EudraCT requirements.

3.8.1.4 AE summaries

The following AE summaries will be produced. Overall category will be included:

- Number (%) of patients who died, had AEs, had SAEs, discontinued, had concomitant medication or had a dose reduction or interruption, by treatment group
- AEs, regardless of study drug relationship by primary system organ class, preferred term, maximum CTC grade and treatment group
- AEs suspected to be related to the study drug by primary system organ class, preferred term, maximum CTC grade and treatment group
- Deaths by primary system organ class, preferred term and treatment group
- Serious adverse events (SAE), regardless of study drug relationship, by primary system organ class, preferred term, maximum CTC grade and treatment group
- AEs leading to study drug discontinuation, regardless of study drug relationship, by primary system organ class, preferred term, maximum CTC grade and treatment group
- AEs requiring dose adjustment or study-drug interruption, regardless of study drug relationship, by primary system organ class, preferred term, maximum CTC grade and treatment group
- AEs which are not serious adverse events, regardless of study drug relationship, by primary system organ class, preferred term and treatment group
- AEs, regardless of study drug relationship, by preferred term and treatment group
- On-treatment deaths and serious adverse events by primary system organ class, preferred term and treatment group

3.8.1.5 Adverse events of special interest

Specific groupings of AEs, referred as AEs of special interest, will be considered and the number of patients with at least one event in each grouping will be reported by the redefined study treatment. Such groups consist of AEs for which there is a specific clinical interest in connection with LDE225 or INC424 treatment (i.e. compound and class related risks linked to the drug structure, mechanism of action, formulation and mode of administration).

- Muscle-related events, including CK elevation

- Nausea and/or Vomiting
- Diarrhea
- Fatigue
- Dysgeusia
- Alopecia
- Weight decreased
- Decreased Appetite
- Thrombocytopenia

The classification and definition of AEs of special interest reflects the current version of the Safety Profiling Plan (SPP) and might be updated at the time of the data base lock for the analyses based on the latest approved version of the SPP.

The following summaries will be provided for the AEs of special interest:

- AEs of special interest, by group, preferred term and treatment group. This includes all CTC grades, suspected AEs to be related to study drug, SAEs, AEs requiring dose adjustment and AEs leading to discontinuation from study treatment.

3.8.2 Laboratory data

The summaries will include all laboratory assessments collected no later than 30 days after study treatment discontinuation. All laboratory assessments will be listed and those collected later than 30 days after study treatment discontinuation will be flagged in the listings.

The laboratory data will be classified into CTC grades according to the NCI Common Terminology Criteria for Adverse Events (CTCAE) v4.03. If value within normal limits is assigned grade 0 as well as assigned a non-zero grade by the NCI Toxicity Criteria, the severity will be assigned to grade 0. These values are not counted in the frequency distribution of the highest NCI toxicity grade even if the value met the criteria of a nonzero grade. This resolves cases where a laboratory value could simultaneously fall in the normal range and in the range of a non-zero grade under NCI toxicity Criteria.

3.8.2.1 Descriptive analyses

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

- Shift tables using CTC grades to compare baseline to the worst post-baseline value, by treatment and overall
- For laboratory parameters without CTC grade: shift tables using normal ranges to compare baseline to the worst post-baseline value, by treatment and overall

The following listings will be produced for laboratory data:

- Listing of laboratory data, with values flagged to show the corresponding CTC grades and the classifications relative to the laboratory reference ranges by treatment. This listing will be provided for hematology, biochemistry and urinalysis lab data.
- Listing of patients with laboratory abnormalities of CTC grade 3 or 4 by treatment with values flagged to show the corresponding CTC grades and the classifications relative to the laboratory reference ranges by treatment. This listing will be provided for hematology, biochemistry and urinalysis lab data combined.
- A listing of patients with grade ¾ CK elevation with values flagged to show the corresponding CTC grades and the classifications relative to the laboratory reference ranges by treatment.
- Listing of pregnancy test results by treatment

3.8.3 Vital signs

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

The criteria for clinically notable abnormalities are depicted in Table 3-2 and Table 3-3.

Table 3-2 Clinically notable elevated values

Variable	Criteria
Systolic BP	≥ 150 mmHg and an increase ≥ 20 mmHg from baseline
Diastolic BP	≥ 90 mmHg and an increase ≥ 20 mmHg from baseline
Body temperature	≥ 38°C
Weight	increase from baseline of ≥ 10%
Heart rate	≥ 100 bpm with increase from baseline of ≥20 bpm

Table 3-3 Clinically notable below normal values

Variable	Criteria
Systolic BP	≤ 90 mmHg and a decrease ≥ 20 mmHg from baseline
Diastolic BP	≤ 50 mmHg and a decrease ≥ 15 mmHg from baseline
Body temperature	≤ 35°C

Weight	decrease from baseline of ≥ 10%
Heart rate	≤ 50 bpm with decrease from baseline of ≥ 15 bpm

The following summary will be produced for heart rate, diastolic BP and systolic BP, body temperature and weight change by study phase, dose cohort and overall:

- Number and percentage of patients with at least one post-baseline vital sign abnormality (in both directions, i.e. both elevated and below normal values)

The following listing will be produced by study treatment and overall:

- Patients with clinically notable vital sign abnormalities
- All vital sign assessments will be listed by patient and vital sign parameter

In both listings, the clinically notable values will be flagged and also the assessments collected later than 30 days after the last treatment/exposure date will be flagged.

3.8.4 Cardiac assessments

3.8.4.1 Electrocardiogram (ECG) Analyses

ECG data are collected from the site. The summaries will include all ECG assessment collected no later than 30 days after the last date of study treatment. All ECG assessments will be listed, and those collected later than 30 days after the study treatment discontinuation will be flagged in the listing.

Single ECGs are collected at pre-dose for the following visits: Screening, Week 1 Day 1, then every 4 weeks and EOT visit.

All ECG analyses will be performed using the Safety set.

For all subjects, ECG baseline is defined as the average of all ECG measurements taken prior to the first dose as confirmed by date and time.

The number and percentage of patients having notable ECG interval values will be summarized by study treatment and overall. The percentages will be calculated based on the total number of patients who are at risk for a specific category.

Quantitative abnormal prolongation as determined by central review of ECG will be provided via summary of worst change from baseline in ECG intervals by dosing schedule and overall. The following categories for QT, QTcB and QTcF [ms] will be used to identify notable values:

- New >450 msec
- New >480 msec
- New >500 msec

The following categories will be used to identify patients with notable ECG changes from baseline:

QT, QTcB and QTcF [ms]

(1) >30 and <= 60 ms increase from baseline

(2) >60 ms increase from baseline

PR [ms]

(1) An increase ≥25% from baseline resulting in PR >200 at any post-baseline assessment

QRS [ms]

(1) An increase ≥25% from baseline resulting in QRS >110 at any post-baseline assessment

HR [bpm]

(1) A decrease >25% from baseline resulting in HR <50 at any post-baseline assessment

(2) An increase >25% from baseline resulting in HR >100 at any post-baseline assessment

The following summaries will be produced:

- A shift Table for QTcB and QTcF based on notable values. The patient's baseline value is compared to the worst post-baseline value,
- The number and percentage of patients with notable ECG change from baseline using the categories described above.

The percentage of patients having notable ECG interval values is based on the total number of patients who are at risk for a specific category.

Assessments collected later than 30 days after the study treatment discontinuation will be flagged in the listings. The following listings will be produced:

- A listing of patients with abnormal ECG intervals,
- All ECG intervals with corresponding baseline values and notable values flagged.

Patients with abnormal ECG intervals will be listed and the corresponding notable values will be flagged in the listings. An overall listing of ECG data with notable values flagged will also be provided.

Patients will be considered evaluable (included in "Total") for absolute QT/QTc outliner analysis if they have at least 1 post-drug ECG measurement. For change from baseline analyses they have to have at least 1 baseline ECG measurement and 1 post-baseline measurement.

3.9 Patient reported outcomes (PRO)

The FAS will be used for all PRO summaries and listings.

3.9.1 EORTC Quality of Life Questionnaire (QLQ-C30)

The EORTC QLQ-C30 includes 5 functional scales (physical, role, emotional, cognitive and social), global health status/quality-of-life scale and 9 symptom scale/items (fatigue, nausea and vomiting, pain, dyspnea, insomnia, appetite loss, constipation, diarrhea, and financial difficulties). For functional and global health status/Qol scales, higher scores indicate better QoL and level of functioning; for symptom scales, higher scores indicate greater level of symptoms or difficulties.

The EORTC QLQ-C30 will be collected at Week 1 Day 1, then every 4 weeks and at the EOT visit.

3.9.1.1 Score generation

EORTC QLQ-C30 scale scores will be generated according to the scoring procedures and SAS code guidelines for version 3.0 provided in *The EORTC QLQ-C30 Scoring Manual (3rd Edition)*. Scores for each functional scale range from 0 to 100. As a general rule, scores values are set to missing when more than half of the items for that scale are missing. If the functional scale consists of one item and that item is missing, the score is automatically missing.

Change from baseline scores is defined as:

score value of given visit or time-point − baseline score,

where baseline scores are those obtained from Week 1visit assessments.

3.9.2 Seven-day modified MFSAF v2.0

The Seven-day Total Symptom Score (TSS) and individual item scores will be summarized descriptively along with changes from baseline to each visit where the variable is measured.

At week 24 and week 48 the proportion of patients who have a ≥50% reduction from baseline in TSS based on the Seven-day modified MFSAF v2.0 will be determined as follows.

The Seven-day TSS will be defined as the sum of 6 individual symptom scores (each with a 0-10 point scale) collected on the same day; the score will be missing if there are any missing individual scores and inactivity will not be included.

The percent change will be calculated by:

% change=100×(Week 24 or Week 48 total score- baseline total score)/ baseline total score

The percent of patients who have ≥50% reduction in total symptom score at week 24 and week 48 will be estimated with a two-sided 95% confidence intervals.

3.9.3 Compliance

Compliance rate calculated as the number of patients who filled out the questionnaire will be summarized for each visit for EORTC QLQ-C30 by study treatment.

3.9.4 Descriptive statistics

PRO scaled scores and change from baseline scores will be summarized by study treatment and by visit for patients in the FAS. All scores will be listed, and those assessments performed later than 30 days after the last date of study treatment will be flagged in the listings but not included in the summaries. Unscheduled assessments will not be included in the summaries.

The following listings will be produced:

- A listing of individual responses to questionnaires.

3.9.5 Waterfall graphs

Waterfall graphs to display the change from baseline at Week 24 and 48 in Phase Ib dose expansion and Phase II Stage 1 will be constructed for the following scores:

EORTC QLQ-C30

- Physical functioning
- Social functioning
- Fatigue
- Pain

MFSAF Total Symptom Score

3.10 Chest X-ray

All chest X-ray variables will be listed.

3.11 Biomarkers

Since the study is not adequately powered to assess specific biomarker-related hypotheses, the statistical analyses of these data should be considered exploratory in nature. Analytical results from such analyses may be used to generate additional hypotheses that must then be verified with data derived from subsequent clinical trials. Furthermore, additional post hoc exploratory assessments are expected and may be performed. Such analyses may include, but are not limited to i) the meta – analysis of data from this study combined with data from other studies, ii) the analysis of pharmacogenetic biomarkers etc. Details of additional exploratory analysis will be described in separate Biomarker analysis plans and the results of such assessments may also be documented in separate biomarker reports.

The standard analysis sets will be used instead according to the purpose of a given analysis (e.g. FAS to describe biomarkers, safety/per-protocol set to assess the relationship between biomarkers and selected safety/efficacy endpoints). The purpose of the proposed data analysis is:

- To describe the distribution of biomarkers at baseline and changes from baseline

- To assess the relationship between biomarkers and selected efficacy endpoints, if applicable:
  - Response: achieving 35% spleen volume reduction (Yes/No)
  - Bone marrow: blast count (%), blast count change from baseline (%)
  - Gene expression signature at baseline correlated with efficacy endpoints

The following categories of biomarkers will be analyzed:

- Gene expression markers
- Gene mutation markers

3.11.1 Data handling and pre-processing

Gene expression data will be normalized to remove systematic bias. Values below the lower limit of quantification (LLOQ) will be handled as follows. If the proportion of such measures is below 15%, different imputation rule will be used depending on assay as specified below. Standard summary statistics will be used. Summary measures of changes in biomarkers collected longitudinally (if available) during the study will be derived as the difference from baseline at the time of response assessments.

Unscheduled and missing measurements will be handled in a similar way as laboratory data.

Table 3-4 Data pre-processing and handling of biomarkers in LDE225 and INC424

Biomarker type / assay	Assay and measurement	Data pre-processing and handling
RT-PCR	Measures the relative RNA expression of a particular gene Assay output is the number of cycles (CT) to reach detection threshold (where the amplification becomes linear) Raw or normalized CT values will be reported. The greater the starting amount of RNA, the fewer number of cycles to reach threshold The higher the raw CT for a target gene, the weaker the level of expression for this gene.	Raw CT results will be normalized[1]. ΔCT is in general the average CT of the control genes minus the raw CT of the target gene. ΔCT is, by definition, on the log-2 scale. No transformation will be applied to this data. Values below LLOQ will be imputed as the maximum reliable CT, in general 36 (contact BTH or GCSL to confirm this information).
Somatic mutation (SMU)	Qualitative data - Usually 2 categories: wild type and mutant	Mutant may refers to several forms of mutations of the same gene (e.g. exon 20 and exon 9 for PIK3CA gene) and may be reported together and separately.

1. Algorithms and formulas for normalization depend on the sample type will be defined in consultation with the RT-PCR Lab Scientist or BTH.

The mean of all pre dose assessments will be used as the baseline value.

When more than one biomarker data value are available for a subject at any time point, the mean of the replicate values will be used for all statistical analysis.

Except for RT-PCR biomarkers, measures below the LLOQ will be handled according to the following algorithm:

1. Numerical values that are below the LLOQ for the assay will need to be imputed as LLOQ/2, except for RT-PCR biomarkers values which will be imputed following rules provided in Table 3-5. Sometimes numerical values for the biomarker are reported, even if the values are below LLOQ. In this case, the imputation is still needed.
2. If both baseline and post baseline values are below LLOQ, then change from baseline is not imputed and considered missing.
3. If <15% of the values (at a given time point/ change from baseline at a given time point) are below LLOQ then these will be summarized.
4. If the proportion of such measures below LLOQ is >15% but <50% then only percentiles are reported as summary statistics
5. If >50% of such values are below LLOQ then data is only listed for this measure.

3.11.1.1 RT-PCR biomarkers

RT-PCR data will be reported as cycle time (CT) values or delta CT (ΔCT). ΔCT is, by definition, on the $\log_2$ scale and no transformation will be applied to this data. Values below the LLOQ will be imputed as the maximum reliable CT, in general 36. Maximum reliable CT values must be confirmed with the BTH or GCSL.

Change from baseline is defined as:

$$\text{Diff} = \log_2 (\Delta CT)_{post} - \log_2 (\Delta CT)_{pre} = \log_2 \{ (\Delta CT)_{post} / (\Delta CT)_{pre} \}$$

Where $(\Delta CT)_{pre}$ is the value at baseline and $(\Delta CT)_{post}$ is the value post-baseline.

Fold-change from baseline is defined as:

$$\text{Fold-change} = 2^{\wedge(\text{Diff})} \text{ ; if Diff} \geq 0 \text{ i.e. Fold-change is } \{ (\Delta CT)_{post} / (\Delta CT)_{pre} \}$$

$$\text{Fold-change} = -1 \times \{2^{\wedge(-1 \times \text{Diff})}\} \text{; if Diff} < 0 \text{ i.e. Fold-change is } -1 \times \{ (\Delta CT)_{pre} / (\Delta CT)_{post} \}$$

Percent inhibition is defined as:

$$\% \text{ inhibition} = 100 \times (1 - 2^{\wedge(\text{Diff})})$$

Example: If D_28 is the $\log_2$ ΔCT of day 28, and D_0 is the $\log_2$ ΔCT at baseline, then Diff = D_28 − D_0 (in this case it is the $\log_2$ fold change):

```
data lde_diff;
set lde_diff;
diff = D_28 - D_0;
if diff < 0 then fold = -1*(2**(-1*diff));
else fold = (2**(diff));
run;
```

New categorical variables will be derived for RT-PCR biomarkers using the median levels as the threshold or based on the post hoc analysis of the distribution of the data. For percent inhibition of Gli1, the following thresholds will be used: <80%, [80% - 90%], > 90%.

3.11.2 Statistical methods for Biomarker

Mutations involved in the molecular disease characterization of MF will be listed by patient and summarized using frequency tables by study treatment.

The expression levels of Hh pathway related biomarkers (such as Gli1) and JAK signaling pathway related biomarkers (such as JAK2V617F allele burden; reported as percent mutated) and cytokine levels will be reported and summarized as follows:

All data will be listed by patient in each cohort and time point. Summary tables of the expression levels will be created at each measured time point by study treatment in Phase Ib and Phase II Stage 1. Summary tables will include number of samples (n), % of those samples below LLOQ, mean, standard deviation, median, %CV, minimum and maximum. For all post-baseline time-points; fold change in biomarker expression levels will also be summarized. The summary statistics will include, number of samples, mean, geometric mean, %CV of geometric mean, minimum, maximum etc. Additionally frequency tables (number and percentage of subjects) for JAK2 mutation status (positive/negative) at baseline will be reported.

Longitudinal plots of the mean fold-change from baseline along with 95% CI's or standard error bars will be created to explore the trend over time.

The associations between biomarkers and efficacy outcomes may also be explored via more sophisticated statistical models if sufficient data are available. Advanced statistical analysis may be carried out to explore the data in a post-hoc fashion. Details of such methodology if applied to the data will be described in a separate Biomarker Analysis Plan document.

3.12 Pharmacokinetic analyses

The pharmacokinetic analyses will be based on the PAS described in section 2.2.4, unless otherwise specified.

Listings will be presented by study treatment using FAS and ordered according to instructions detailed in the general guidance section of RAP M7.

As convention, a PK sample should be excluded from analyses if the patient vomited within the first 4 hours following the last oral dose of study drug. Other PK samples will be excluded as deemed appropriate by the Pharmacokineticist. The excluded concentrations will be flagged in the listings.

All concentrations below the limit of quantification or missing data will be labeled as such in the concentration data listings. Concentrations below the limit of quantification will be treated as zero in summary statistics and for the calculation of pharmacokinetic parameters.

3.12.1 Trough and profile concentration data

Plasma concentrations of LDE225 and INC424 will be summarized at each scheduled time point by study phase and dose cohort. Profile concentration data for LDE225 refers to the data collected on week 1 day 1 (also include pre-dose and week 1 day2) and week 9 day 1 (also include pre-dose and week 9 day 2). Profile concentration data for INC424 refers to the data collected on week 1 day 1 within the 12 hours after the first dose and before the second dose, and week 9 day 1 within the 12 hours after the first dose and before the second dose. Trough concentration data refers to pre-dose samples eg week 3 day1, week 5 day 1, week 7 day1, week 13 day1, etc. Descriptive graphical plots of individual plasma concentration time profiles with median will be generated, along with geometric mean and mean (SD) concentrations time profiles trough concentrations for LDE225 and INC424. Individual plots of trough concentrations for LDE225 and INC424 will also be produced. A list of individual concentrations will be provided using FAS. Concentrations excluded from analyses will be flagged. Descriptive statistics will include n (number of patients), m (number of non-zero concentrations), arithmetic mean, geometric mean, median, SD, coefficient of variation CV (%), geometric CV (%), minimum and maximum.

Coefficient of variation CV (%) is calculated as follows: 100×(SD/arithmetic mean).

Geometric CV (%) is calculated as follows: 100×sqrt(exp[Var(log(x))]-1), where Var(X) denotes the variance of X.

3.12.2 Pharmacokinetic parameters

For Phase Ib and Phase II Stage 1 only, PK parameters listed in Table 3-6 and Table 3-7 will be determined for all PK-evaluable patients using non-compartmental method(s) using WinNonlin or Phoenix WinNonlin (Pharsight, Mountain View, CA). PK parameters will be estimated for Week 1 and Week 9 and reported, when feasible. PK parameters in Table 3-6 and Table 3-7 for LDE225 and INC424 will be descriptively summarized for Week 1 and Week 9 by dose level.

Descriptive statistics (n, mean, SD, CV% or median [range], and geometric mean and CV%) will be presented for all parameters with the exception of Tmax for which median values and ranges will be presented. When a geometric mean is presented, it will be stated as such.

The following pharmacokinetic parameters are determined using non-compartmental method(s) (Table 3-6 and Table 3-7).

Table 3-5 Non-compartmental pharmacokinetic parameters for LDE225 (Phase Ib and Phase II Stage 1 only)

Term	Definition
Cmax	Maximum observed plasma concentration after drug administration [mass x volume^{-1}]
Tmax	Time to reach C_{max} [time]
AUC0-24h	Area under the concentration-time curve from time zero to 24 hours [mass x time x volume^{-1}]
CL/F	Apparent total plasma clearance of drug after oral administration [volume x time^{-1}]

Term	Definition
Racc	Accumulation ratio calculated as AUC0-24h on Week 9 Day 1 divided by AUC0-24h on Week 1 Day1 [fold]

Table 3-6 Non-compartmental pharmacokinetic parameters for INC424 (Phase Ib and Phase II Stage 1 only)

Term	Definition
Cmax	Maximum observed plasma concentration after drug administration [mass x volume^{-1}]
Tmax	Time to reach Cmax [time]
AUClast	Area under the concentration-time curve from time zero to the time of last measurable concentration [mass x time x volume^{-1}]
AUC0-12h	Area under the concentration-time curve from time zero to 12 hours extrapolate from AUClast[mass x time x volume^{-1}]
AUCinf	Area under the concentration-time curve from time zero to infinity with extrapolation of the terminal phase [mass x time x volume^{-1}]
T1/2	Elimination half-life associated with the terminal slope (lambda_z) of a semi logarithmic concentration-time curve [time]
CL/F	Apparent total plasma clearance of drug after oral administration [volume x time^{-1}]
Racc	Accumulation ratio calculated as AUC0-12h on Week 9 Day 1 divided by AUC0-12h on Week 1 Day 1 [fold]
Vss/F	Apparent volume of distribution at steady state after oral administration [volume]

3.12.3 Analysis of relationship between LDE225/INC424 and efficacy/safety endpoints

3.12.3.1 Exposure-response relationship

Box plots for LDE225 and INC424 AUC and Cmax will be presented by responders and non-responders.

3.12.3.2 Exposure-safety relationship

Box plots for on-treatment LDE225 AUC and Cmax will be presented separately by worst CTC grade for CK.

3.13 Sample size and power considerations

The total sample-size for the study will be approximately 82.

Phase Ib: Approximately 36 patients are expected to be enrolled into dose escalation cohorts. Cohorts of 3 to 6 evaluable patients per dose level will typically be enrolled in the dose escalation part including at least 6 evaluable patients at the MTD level. Assuming a total of 5 cohorts, the number of patients during dose escalation is expected to be approximately 30. An additional 6 patients will be treated at the MTD/RPIID during the safety expansion. If two

MTDs/RPIIDs are established, each MTD/RPIID will enroll additional 6 patients in the safety expansion.

Phase II: Following the determination of an MTD or RPIID in Phase Ib, 18 patients will be enrolled in the first stage and if the criterion is met to continue to stage 2, an additional 28 patients will be enrolled.

4 References

Babb J, Rogatko A, Shelemyahu Z (1998) Cancer phase I clinical trials: efficient dose escalation with overdose control. Statistics in Medicine, 17, 1103-1120

Bishop YMM, Fienberg SE, Holland PW (1991) Discrete multivariate analysis. Cambridge MA, 11th ed. The MIT Press

Clopper CJ, Pearson ES (1934). The use of confidence or fiducial limits illustrated in the case of the binomial. Biometrika, 26; 404–413.

CPMP (2002). Points to Consider on Multiplicity Issues in Clinical Trials.

FDA (2007). Guidance for Industry Clinical Trial Endpoints for the Approval of Cancer Drugs and Biologics, US Department of Health and Human Services.

FDA (2009). Guidance for Industry Patient-Reported Outcome Measures: Use in Medical Product Development to Support Labeling Claims, US Department of Health and Human Services.

Gale RP, Barosi G, Barbui T, et al (2011) What are RBC-transfusion-dependence and independence? Leuk Res; 35(1): 8-11.

Garwood, F (1936). Fiducial limits for the Poisson distribution. Biometrika, 46; 441–453.

Grouin JM, Maylis C, Lewis J (2005). Subgroup analyses in randomized clinical trials: statistical and regulatory issues. Journal of Biopharmaceutical Statistics, 15, 869-882.

Jung S and Kyung MK (2004) On the estimation of the binomial probability in multistage clinical trials. Statist. Med; 23(6):881–896.

Neuenschwander B, Branson M, Gsponer, T (2008) Critical aspects of the Bayesian approach to phase I cancer trials. Statistics in Medicine, 27, 2420-2439

Sahai H, Khurshid Anwer (1993). Confidence intervals for the mean of a poisson distribution: a review. Biom J, 35 (7); 857-867

Storer BE (1989) Design and analysis of phase I clinical trials. Biometrics, 45, 925-937.

Thall PF, Millikan RE, Mueller P, Lee SJ (2003) Dose-finding with two agents in phase I oncology trials. Biometrics 59, 487-496

Thiele J, Kvasnicka HM, Facchetti F, et al (2005) European consensus on grading bone marrow fibrosis and assessment of cellularity. Haematologica; 90(8): 1128-32.

Ulm K (1990). A simple method to calculate the confidence interval of a standard mortality ratio. American Journal of Epidemiology, 131(2); 373-375