NOVARTIS

Clinical Development

INC424/Ruxolitinib/JAKAVI®

CINC424F12201 / NCT03491215

A Phase I/II open-label, single-arm, multi-center study of ruxolitinib added to corticosteroids in pediatric patients with grade II-IV acute graft vs. host disease after allogeneic hematopoietic stem cell transplantation

Statistical Analysis Plan (SAP)- Amendment 3

Author:	Trial Statistician; [redacted]

Document type:	SAP Documentation
Document status:	Final version 4.0
Release date:	10-Mar-2023
Number of pages:	49

Property of Novartis
For business use only
May not be used, divulged, published or otherwise disclosed
without the consent of Novartis

Document History – Changes compared to previous final version of SAP

Date	Time point	Reason for update	Outcome for update	Section and title impacted (Current)
Jun.27.2018		First version	Final version 1.0	NA
Feb22, 2022		Amendment 1	Final version 2.0 Changes to align with protocol Amendment 1 and 2	Section 1.1, 1.2, 1.3, 1.4, 2.1.1, 2.8.4.5, 2.9, 2.13.2, 2.14, 3
			Clarification was required	2.2, 2.4.2, 2.5.4, 2.7.1, 2.7.2,2.8.4.2, 2.10, 2.12
			Change the notable vital signs to follow the latest WHO criteria	2.8.4.2
Jun27 2022		Amendment 2	Final version 3.0 Delete Table 2-3 in Analysis sets to remove INC01 and add clarification that patients with dual CYP3A4/ CYP2C9 inhibitors will be considered as treated with full assigned dose.	Sections: 1 and 2.2
			Section 3.8.4.2 on graft failure was deleted from safety and moved in analysis of secondary efficacy	Section 3.8.4.2 deleted and moved in secondary endpoints.
06Mar2023		Amendment 3	Final version 4.0 Update to align with protocol Amendment version 3 (release date 24-Jun-2022) and to clarify some specifications after dry run. Main changes: • Clarify the last contact date • [redacted]	Sections 1.1.1 3.2

Date	Time point	Reason for update	Outcome for update	Section and title impacted (Current)
			[redacted]	3.2.1
			• Add subgroup analysis by formulation, treatment naïve vs SR-aGvHD [redacted] for efficacy and safety analysis	3.3.1 3.7.1, 3.7.2
			• Add listing on COVID-related PDs	3.4.1
			• Analysis of BOR added up to data cut-off date	3.4.2.4
			• Provide clarification on exposure, and derivation of planed dose intensity and dose changes	3.8.4.2
			• Clarify the use of CRF pages for Calcineurin Inhibitors (CNI) treatment	3.8.4.3
			• Clinical notable vital signs criteria were updated to align within the project	3.4.2.5
			• Analysis of time to first occurrence of infections using Kaplan-Meier methodoloy removed	App 6.4.3
			• More detailed specifications how to identify new or additional systemic therapy	

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

1 Introduction

This statistical analysis plan (SAP) describes all planned analyses for the Clinical Study Report (CSR) of study CINC424F12201, A Phase I/II open-label, single-arm, multi-center study of ruxolitinib added to corticosteroids in pediatric patients with grade II-IV acute graft vs. host disease after allogeneic hematopoietic stem cell transplantation.

SAP Amendment 3

The main purpose of SAP amendment 3 is to align with protocol amendment 3, provide more detailed specifications and to keep consistency across GvHD indication.

Major changes of SAP amendment 3 included:

- Clarification on last contact date
- [redacted]
- To add BOR up to data cut-off date
- To add subgroup analysis by formulation, treatment naïve vs SR-aGvHD and [redacted] for efficacy and safety analysis
- Add a listing for COVID-related PDs
- To provide clarifications on derivations for exposure, Planed Dose intensity and dose reductions
- To clarify the use of the CRF pages for derivation of the CNI treatment
- To change the clinical notable criteria of vital signs
- To remove analysis of time to first occurrence of infections
- To provide more detailed specifications on how to identify any new or additional systemic therapy

SAP amendments 1 and 2

The content of SAP amendment 1 was based on protocol CINC424F12201, versions 1 and 2. SAP Amendment 2 was also an amendment of SAP version 01, SAP amendment 1 .

The main purpose of SAP amendment 1 was to align with all the changes made in Protocol amendments 1 and 2.

Major changes of SAP amendment 1 included:

- Updates in the sample size from 39 to 45 subjects.
- The definition of analysis sets were updated to consider subjects that have been treated with the full assigned dose as a result of co-administration with strong CYP3A4 or dual (clarified at SAP amendment 2) CYP3A4/CYP2C9 inhibitors,

- [redacted]

Finally a few more changes in SAP amendment 1 were made for clarification purposes and/or for consistency purposes across GvHD indications:

- The analyses sets were clarified to include only patients who signed informed concent (standard approach).
- The presentation of prior and concomitant medications were updated to follow project requirements
- The key secondary objective was clarified as per project requirements
- Presentation of graft failure assessment was clarified as part of other safety analysis.
- Sensitivity analysis was added to primary objective in order to exclude patients with protocol deviations on aGvHD assessments due to not strictly following the standard criteria
- The notable vital sings were updated to follow the latest version of WHO growth criteria
- The presentation of prior and concomitant criteria were updated to follow project requirements
- [redacted] as well as some PK analysis were clarified that will be explored using a seperate SAP and seperate CSR report.

The main purpose of the SAP amendment 2 was to remove patients clasifications based on PD criteria since these were not applicable in the study. All patients who signed written informed consent were included in analyses sets. Therefore, Table 2-3 in Analysis sets section has been deleted from the SAP. Another change made in SAP amendment 2 was to move graft failure analysis from safety to efficacy analysis section.

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

2 Study design

This open-label, single-arm, Phase I/II multi-center study will investigate the PK, activity and safety of ruxolitinib added to the patient's immunosuppressive regimen in infants, children, and adolescents ages ≥28 days to <18 years old with either grade II-IV aGvHD or grade II-IV SR-aGvHD. This trial will utilize age groups: Group 1 includes patients ≥12y to <18y, Group 2 includes patients ≥6y to <12y, Group 3 includes patients ≥2y to <6y, and Group 4 includes patients ≥28days to <2y. Patients will remain in the age group throughout the study based on the age at the time of start of treatment. Enrollment initiation into the youngest age group, Group 4 (Phase I/II) will be subject to the review of available PK, safety, and activity data in consultation with the data monitoring committee (DMC), Pediatric Committee (PDCO), and a final decision by the Sponsor.

All patients will be enrolled and treated for 24 weeks (approximately 6 months) or until early discontinuation. All patients will also be followed for an additional 18 months (total duration = 2 years from enrollment). Should the occurrence of aGvHD flare require treatment re-initiation

or should ruxolitinib not be discontinued by the end of 24 weeks due to extended tapering, patients may continue taper ruxolitinib beyond 24 weeks up to a maximum of 48 weeks. As patients ≥12 y to < 18 y are already being included in [CINC424C2301], and treated with 10 mg BID, this dose is the recommended phase II dose (RP2D), and will be used to treat all patients in this age group. For the Phase II, all other age groups will be treated with the RP2D determined during the Phase I. Therefore, all ≥12 to <18 year old patients will automatically be enrolled in the Phase II. It is planned that the first 5 patients treated in Group 1 will undergo extensive PK sampling to inform the RP2D determination of the younger age groups in the Ph I. Additional patients may undergo extensive sampling should one or more of the first 5 patients not be evaluable or until the dose/exposure is confirmed.

2.1 Phase I

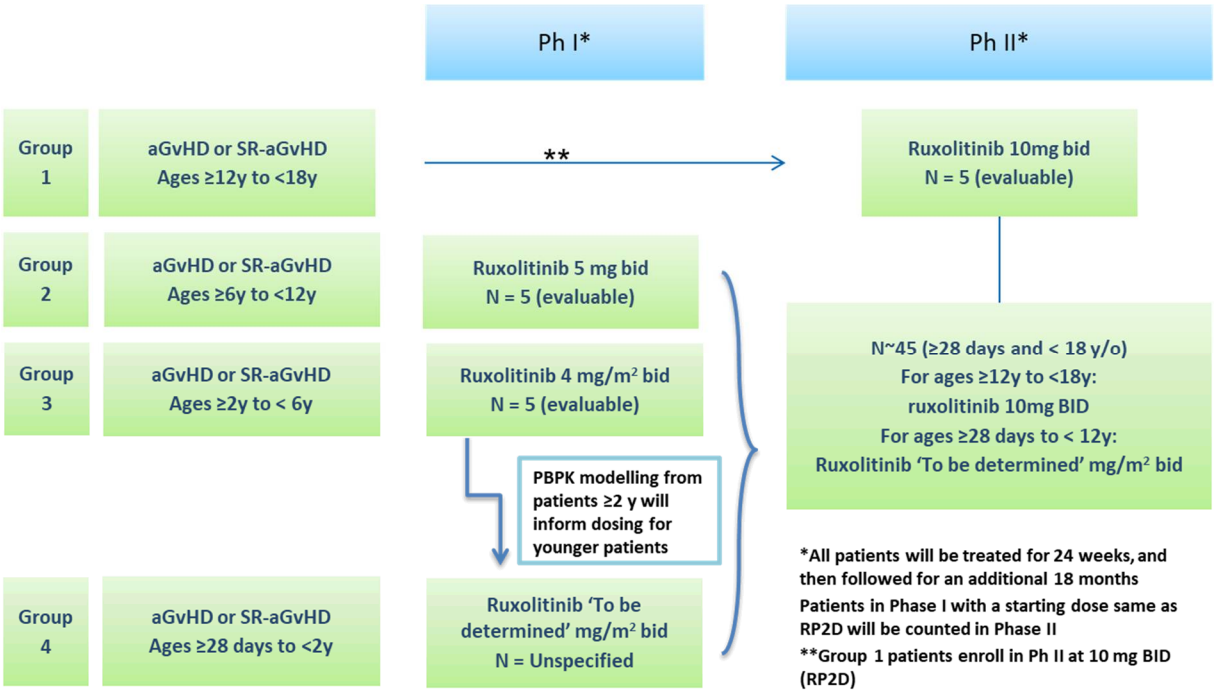
Patients will be enrolled into 4 groups, based on age, to allow appropriate dosing based on available data Table 1-1.

- Phase I (Figure 1-1): Full ruxolitinib concentration-time courses, safety and activity data will be collected over 28 days for Groups 2, 3 and 4. Groups 2 and 3 will be enrolled initially, and the PK data generated from all patients (including Group 1) will be used to inform the starting dose of Group 4. Therefore, Group 4 will only open after an RP2D has been defined for Group 2, and an RP2D has been defined for Group 3.
- The PK and safety data will be used to assess the adequacy of the starting dose, which can be adapted if needed (e.g. to account for any potential difference between the expected and the observed ruxolitinib exposure).
- Should the exposure in Groups 2, 3 or 4 not be confirmed following the PK sampling in 5 evaluable patients, at least an additional 5 patients will be enrolled in that specific age group until the dose/exposure is confirmed (ie selection of the RP2D for those ages based on exposure and safety review by the DMC).
- Ruxolitinib will be administered either as tablet or as an oral pediatric formulation (capsule/liquid) twice a day".

Table 2-1 Phase I: Age groups and dosing rationale

Group	Group 1	Group 2	Group 3	Group 4
Age Range	≥12y to < 18y	≥6y to < 12y	≥2y to < 6y	≥28 days to < 2y
N (Ph I)	Not applicable	5 evaluable	5 evaluable	Undefined
Starting dose		PBPK-derived equivalent predicted to yield similar exposure to 10 mg BID adult dose =5 mg BID (Tablet /capsule formulation)	PBPK-derived equivalent predicted to yield similar exposure to 10 mg BID adult dose =4 mg/m^2 BID (Capsule/ Liquid formulation)	Will be generated based on PK data from groups 1-3

Figure 2-1 Study Design



2.1.1 Determination of recommended Phase II dose

- Group 2 (age ≥6y to <12y) and Group 3 (age ≥2y to <6y): Physiologically based pharmacokinetic (PBPK) modeling will be used to derive a starting dosing for these patients, which is predicted to yield an exposure equivalent to that of a 10 mg BID in adults. The PBPK model will incorporate PK data from adult and adolescent patients with SR-aGvHD treated on study [CINC424C2301], and these predictions will be compared to existing data in pediatric patients from other indications (Loh et al 2015). The starting doses are thereby assigned as 5 mg BID (Group 2) and 4 mg/m^2 BID (Group 3).
- Group 4 (≥ 28 days to < 2 y): the absence of existing ruxolitinib PK data in this age group warrants a conservative approach. Therefore, the starting dose for Group 4 will be determined by using PK data generated in Groups 1, 2 and 3 (Table 3-1 and Table 3-2). These data will be used to inform the PBPK model in order to enable accurate predictions for these very young patients.

Groups 2 and 3 will enroll a minimum of 5 patients each. Full ruxolitinib concentration-time courses, safety and activity data will be collected over 28 days. Should one or more of the 5 patients not be evaluable, or until the dose/exposure is confirmed, additional patients may be enrolled to ensure a minimum of 5 evaluable patients per age group. This information will be used to assess the adequacy of this starting dose, which can be adapted if needed (e.g. to account for any potential difference between the expected and the observed ruxolitinib exposure).

If the exposure from the first 5 patients does not approximate the exposure of 10mg BID in adult and adolescent patients based on AUC and Cmax, at least an additional group of 5 patients will be enrolled in that specific age group. This second group of at least 5 patients will be administered a different starting dose to again target the 10mg BID adult exposure. This process will continue until the dose-exposure relationship is confirmed (i.e selection of the RP2D for

those ages based on exposure and safety review by the DMC). If the selected RP2D is higher than the starting dose assigned to the previous group of patients, intra-patient dose escalation is allowed (Section 6.5.1.1). Inversely, if the selected RP2D is lower than the starting dose assigned to the previous group of patients, intra-patient dose de-escalation is allowed.

As mentioned above, Group 1 (patients aged ≥12y to <18y) is excluded from this process as the starting dose is already confirmed at 10mg BID. The available PK data from these patients will however be used in the determination of the RP2D for the younger patients in Groups 2, 3, and 4.

All Group 1 patients will be enrolled automatically into the Ph II, with the first 5 patients undergoing extensive PK. Once the RP2D is selected for Groups 2 and 3 any further eligible patients between the ages of ≥2y and <12y will be enrolled into the Phase II. At this point Group 4 will begin enrolling patients in the Phase I. Further, once the RP2D is established in Group 4, enrollment to Phase II for this group will be initiated.

2.2 Phase II

The Phase II aims to measure the activity of ruxolitinib in these patients assessed by overall response rate (ORR) at Day 28, as Day 28 ORR has been shown to correlate best with subsequent long-term survival (Levine et al 2010). The study's primary endpoint for the Phase II is Investgator reported ORR without requirement for addition of new systemic immunosuppressive treatment assessed 28 days after start of therapy. A key secondary endpoint is the ORR at Day 56 after start of treatment, in order to assess the durability of the primary response. Best Overall Response (BOR) up to or at Day 28 will also be calculated.

The RP2D for all groups will be assessed for both activity and safety in Phase II, over a 24 week period. At least 45 patients will be enrolled in the study regardless of age, with at least 20% of treatment naïve aGvHD and 40% with SR-aGvHD patients and be treated with the confirmed RP2D. The sample size for the Phase II objective of measuring ORR at D28 is 45 evaluable patients regardless of age. Any patient receiving the confirmed RP2D during the Phase I will be counted towards the 45 patients (Table 1-2 and Figure 1-1).

Table 2-2 Phase II: Age groups and dosing rationale

Group	Group 1	Group 2	Group 3	Group 4
Age Range	≥12y to < 18y	≥6y to < 12y	≥2y to < 6y	≥28 days to < 2y
N	N=45 evaluable patients treated with the confirmed RP2D			
RP2D dose	Same as adult dose: 10 mg BID (All patients will be enrolled in Ph II)	Same as Phase I unless correction is needed to account for exposure variations	Same as Phase I unless correction is needed to account for exposure variations	Based on PK modeling from group 1,2, and 3 data in the PK part

No formal interim analysis is planned for this study. However, summaries of safety and PK data will be produced to support the regular safety monitoring conducted by the DMC and the confirmation of RP2D. Efficacy data may also be analyzed when all patients (Phase I and Phase II) have completed 24 weeks (approximately 6 months) of treatment or discontinued earlier. The final analysis will be conducted and the clinical study report (CSR) written once all patients have completed the study (i.e. completed the long-term follow-up period of Month 24) or

discontinued earlier. Analysis cutoff date will be defined corresponding to the analysis time point and all data captured in the study up to that cutoff will be reported.

2.3 Study objectives and endpoints

Objectives and related endpoints are described in Table 2-3 below.

Table 2-3 Objectives and related endpoints

Objective(s)	Endpoint(s)
Primary objective(s)	**Endpoint(s) for primary objective(s)**
• Phase I To assess pharmacokinetic (PK) parameters of ruxolitinib for patients with aGvHD and SR-aGvHD and define an age appropriate RP2D for each of the groups 2-4 • Group 2: age ≥6y to <12y • Group 3: age ≥2y to <6y • Group 4: age ≥28days to < 2y	• Measurement of PK parameters in aGvHD and SR-aGvHD patients: AUC, Cmax, T1/2, Ctrough using extensive PK sampling in Groups 1-3 and sparse sampling in Group 4 • Age-based determination of RP2D for each of the groups 2-4, based on observed PK parameters.
• Phase II To measure the activity of ruxolitinib in patients with aGvHD or SR-aGvHD assessed by Overall Response Rate (ORR) at Day 28.	• Overall response rate (ORR) at Day 28, defined as the proportion of patients demonstrating a complete response (CR) or partial response (PR) without requirement for additional systemic therapies for an earlier progression, mixed response or non-response. Scoring of response will be relative to the organ stage at the start of the study treatment.
Secondary objective(s)	**Endpoint(s) for secondary objective(s)**
• **Key Secondary** To assess the rate of durable ORR at Day 56	• Proportion of all patients who achieve a CR or PR at Day 28 and maintain a CR or PR at Day 56
• To estimate ORR at Day 14	• Proportion of patients who achieved ORR (CR+PR) at Day 14
• To assess pharmacokinetic/pharmacodynamic relationships	• PK parameters (such as AUC, Cmax, Ctrough) versus safety, efficacy, and PD biomarkers, as appropriate
• To assess Duration of response	• Duration of response (DOR) is assessed for responders only and is defined as the time from first response until aGvHD progression or the date of additional systemic therapies for aGvHD. Onset of chronic GvHD, or death without prior observation of aGvHD progression are considered as competing risks
• To assess the cumulative steroid dose until Day 56	• Weekly cumulative steroid dose for each patient up to Day 56
• To evaluate the safety and tolerability of ruxolitinib	• Safety and tolerability including myelosuppression, infections, and bleeding will be assessed by monitoring the frequency, duration and severity of Adverse Events including occurrence of any second primary malignancies, infections, by performing physical exams, and evaluating changes in vital signs from baseline, routine serum chemistry, hematology results and coagulation profile
• To assess Overall Survival (OS)	• Overall survival, defined as the time from the start of treatment to the date of death due to any cause
• To assess Event-Free Survival (EFS)	• Event-free survival, defined as the time from start of treatment to the date of hematologic disease

Objective(s)	Endpoint(s)
	relapse/progression, graft failure, or death due to any cause
• To assess Failure-Free Survival (FFS)	• Failure-free survival, defined as the time from the start of treatment to date of hematologic disease relapse/progression, non-relapse mortality, or addition of new systemic aGvHD treatment
• To assess Non Relapse Mortality (NRM)	• Non-relapse mortality (NRM), defined as the time from start of treatment to date of death not preceded by hematologic disease relapse/progression
• To assess incidence of Malignancy Relapse/Progression (MR)	• Malignancy Relapse/Progression (MR) (refer to protocol Appendix 3), defined as the time from start of treatment to hematologic malignancy relapse/progression. Calculated for patients with underlying hematologic malignant disease
• To measure the incidence of cGvHD	• cGvHD, defined as the diagnosis of any cGvHD including mild, moderate, severe
• To estimate the rate of Best Overall Response (BOR)	• Proportion of patients who achieved OR (CR+PR) at any time point up to and including Day 28 and before the start of additional systemic therapy for aGvHD
• To assess graft failure	• Monitoring of donor cell chimerism, defined as initial whole blood or marrow donor chimerism >5% declining to <5% on subsequent measurements compared to baseline
• To describe the acceptability and palatability assessments of the ruxolitinib formulation	• Responses from the acceptability and palatability questionnaire for dose forms used after first dose, 1 month and 6 months
[redacted]	

3 Statistical methods

The final analysis will be performed by Novartis. SAS version 9.4 or later and/or R version 3.0.2 or later will be used to perform all data analyses and to generate tables, figures and listings.

3.1 Data analysis general information

The analysis cutoff date for the final analysis of study data will be established when all patients have completed the study. All statistical analyses will be performed using all data collected in the database up to the data cutoff date. All data with an assessment date or event start date (e.g. vital sign assessment date or start date of an adverse event) prior to or on the cutoff date will be included in the analysis. Any data collected beyond the cutoff date will not be included in the analysis and will not be used for any derivations.

All events with start date before or on the cutoff date and end date after the cutoff date will be reported as 'ongoing'. The same rule will be applied to events starting before or on the cutoff date and not having documented end date. This approach applies, in particular, to adverse event and concomitant medication reports. For these events, the end date will not be imputed and therefore will not appear in the listings.

General analysis conventions

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

3.1.1 General definitions

Investigational drug and study treatment

Investigational drug/study treatment, will refer to the ruxolitinib in this single arm study.

Date of first administration of study treatment

The date of first administration of study treatment is derived as the first date when a nonzero dose of study treatment was administered as per the DAR eCRF. The date of first administration of study treatment will also be referred as ***start of investigational drug***.

Date of last administration of study treatment

The date of last administration of study treatment is defined as the last date when a nonzero dose of study treatment was administered as per DAR eCRF. The date of last administration of study treatment will also be referred as ***end of investigational drug.***

Study day

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

The study day is defined as:

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

The reference start date for all assessments (e.g. safety, efficacy, PK, [redacted], etc.) is the start of study treatment.

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

Time unit

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

Baseline

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

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

On-treatment assessment/event and observation periods

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

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

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

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

Windows for multiple assessments

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

The following time windows are defined for descriptive summaries on aGvHD assessment, biomarkers and safety (Table 2-1) by visit. The end of treatment assessment will be mapped into the time points as approperate.

Table 2-1 Time windows for aGvHD assessment, [redacted] and safety assessment (lab, vital sign, growth, etc.)

Time Window	Planned Visit Timing	Time Window Definition

On treatment		
Baseline (Day 1)	On or before Study Day 1	≤ Study Day 1
Week 1	Study Day 7	Study Days 4 – 10
Week 2	Study Day 14	Study Days 11 – 17
Week 3	Study Day 21	Study Days 18 – 24
Week 4	Study Day 28	Study Days 25 – 31
Week 5	Study Day 35	Study Days 32 – 38
Week 6	Study Day 42	Study Days 39 – 45
Week 7	Study Day 49	Study Days 46 – 52
Week 8	Study Day 56	Study Days 53 – 59
Week 12	Study Day 84	Study Days 71 – 98
Week 16	Study Day 112	Study Days 99 – 126
Week 20	Study Day 140	Study Days 127 – 154
Week 24	Study Day 168	Study Days 155 – 182
Week 28	Study Day 196	Study Days 183 – 210
Week 32	Study Day 224	Study Days 211 – 238
Week 36	Study Day 252	Study Days 239 – 266
Week 40	Study Day 280	Study Days 267 – 294
Week 44	Study Day 308	Study Days 295 – 322
Week 48	Study Day 336	Study Days 323 – 350
Month 12*	Study Day 365	Study Days 351 – 455
Month 18*	Study Day 546	Study Days 456 - 636
Month 24*	Study Day 730	Study Days 637 - 820

Study Day 1 = start date of study treatment;
EOT assessments are mapped to the time points as appropriate;
[redacted]

Last contact date

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

Table 3-2 Last contact date data sources

Source data	Conditions
Last contact date/last date patient was known to be alive from Survival Follow-up page	Patient status is reported to be alive, lost to follow-up or unknown.
Start/End dates from new aGvHD treatment since discontinuation from study treatment	Non-missing medication/procedure term.
Start/End dates from drug administration/concomitant medication record	Non-missing dose. Doses of 0 are allowed.
Disposition event date from the dispostion log page	No condition.

Source data	Conditions
- aGvHD assessment date - any specific efficacy assessment date if available (e.g., cGvHD assessment, graft failure assessment, second primary malignacy assessement, hematologic disease relapse/progression assessment)	Non-missing assessment
Laboratory/PK collection dates/ [redacted]	Sample collection marked as 'done'.
Vital signs date	At least one non-missing parameter value
Collection date of acceptability and palatability questionnaire	Assessment performed is "Yes"
[redacted]	
Start/End dates of AE	Non-missing verbatim term

The last contact date is defined as the latest complete date from the above list on or before the data cutoff date . The cutoff date will not be used for last contact date, unless the patient was seen or contacted on that date. No date post cutoff date will be used. Completely imputed dates (e.g. the analysis cut-off date programmatically imputed to replace the missing end date of a dose administration record) will not be used to derive the last contact date. Partial date imputation is allowed for event (death)/censoring coming from 'Survival information' eCRF.

3.2 Analysis sets

If the starting dose is different from the assigned dose level due to co-administration of ruxolitinib with strong CYP3A4 or dual CYP3A4/CYP2C9 inhibitors, these patients will be included under the assigned dose level and considered that they have received the full assigned dose. This applies to all analysis sets described below.

The **Full Analysis Set** (FAS) comprises all patients to whom study treatment has been assigned and who received at least one dose of study treatment. Patients will be analyzed according to the treatment they have been assigned to.

The **Safety Set** includes all patients who received at least one dose of study treatment. Patients will be analyzed according to the study treatment received, where treatment received is defined as the assigned dose level of ruxolitinib if the patient took at least one dose of that treatment or the first dose level received if the assigned dose level was never received.

The **Pharmacokinetic Analysis Set** (PAS) includes all patients who provide at least one evaluable PK concentration. For a concentration to be evaluable, patients are required to:

- Take the dose of ruxolitinib prior to PK sample.
- For pre-dose samples, do not vomit within 2 hours after the previous dosing of ruxolitinib prior to sampling; for post-dose samples, do not vomit within 2 hours after the dosing of ruxolitinib.

The PAS will be used for Non-compartmental analysis (NCA) for patients where extensive PK sampling is obtained. This analysis set will also be used for any exposure-response analysis [redacted].

The **Efficacy Evaluable Set** (EES) comprises all patients to whom study treatment has been assigned at the Recommended Phase 2 dose (RP2D) of ruxolitinib and who received at least one dose of study treatment at that dose level. If the starting dose is different from the RP2D due to co-administration of ruxolitinib with strong CYP3A4 or dual CYP3A4/CYP2C9 inhibitors, these patients will be included in the EES.

All statisticial analysis and data presentation for all analysis sets will exclude any data collected but not required as per study protocol. Data with a Protocol Deviation Code OTH11 (Other GCP deviation- Additional samples or Data collected not required as per protocol) will not be part of the analysis data or any output presentation.

Any patients who withdrew Informed Consent or from whom written Informed Consent was not obtained (Protocol Deviation INCL01) will be excluded from all analyses sets.

Any data collected in the clinical database (other than acute GvHD biopsy, medical history, prior and concomitant medication, donor information, prior or concomitant non-drug therapies/procedures and viral load, serology from lab data, [redacted] and blood component transfusion) before signed of informed consent will not be included in the analysis. If a patient withdraws informed consent from all further participation in the trial no data will be included in the analysis after withdrawal of informed consent date. The informed consent date and the date on which a patient withdraws full consent are recorded in the eCRF.

3.2.1 Subgroup of interest

Efficacy

The primary efficacy endpoint will be summarized by the following subgroups to examine the homogeneity of treatment effect using EES:

- Age group (range defined in Table 1-1)
- Age group and formulation (tablet, capsule, liquid)
- Diagnosis (treatment naïve aGvHD and SR-aGvHD)
- [redacted] No formal statistical test of hypotheses will be performed for the subgroups, only point estimate of the treatment effect and 90% confidence intervals will be provided.

Safety

Key safety analyses will be repeated on the Safety Set in the following subgroups:

- Age group (range defined in Table 1-1)
- Age group and formulation (Tablet, capsule, liquid)
- Diagnosis (treatment naïve aGvHD and SR-aGvHD)
- [redacted]

The objective for carrying out these subgroup analyses is to identify potential safety issues that may be limited to or more commonly observed in a subgroup of patients.

The following summaries will be presented by age group, formulation and [redacted] only subgroup:

- Overview of adverse events
- Adverse events by preferred term
- Adverse events with suspected relationship to study treatment by preferred term
- Serious AEs, irrespective of causality, by preferred term
- Adverse events leading to study treatment discontinuation by preferred term
- Overview of adverse events of special interest
- On-treatment deaths
- All study deaths

The following summaries will be provided by age group and and by treatment-naive vs SR aGvHD:

- Overview of adverse events
- AEs, irrespective of causality, by primary system organ class and preferred term
- AEs with suspected relationship to study treatment, by primary system organ class and preferred term
- Serious AEs, irrespective of causality, by primary system organ class and preferred term
- Serious AEs with suspected relationship to study treatment, by primary system organ class and preferred termOn-treatment deaths, by primary system organ class and preferred term
- Overview of adverse events of special interest

3.3 Patient disposition, demographics and other baseline characteristics

The Full Analysis Set (FAS) will be used for all baseline and demographic summaries and listings unless otherwise specified. Summaries will be reported by age group (defined in Table 1-1) and for all patients and listings will be reported by age group to assess baseline comparability. If more than one dose group are assigned under one age group, summaries and listings will be reported by both age group and dose group. No inferential statistics will be provided.

Basic demographic and background data

All demographic and baseline disease characteristics data will be summarized and listed by age group and for all patients. Categorical data (e.g. gender, race, ethnicity) will be summarized by frequency counts and percentages; the number and percentage of patients with missing data will be provided. Continuous data (e.g. age, weight, height, body surface area, body mass index) will be summarized by descriptive statistics (N, mean, median, standard deviation, minimum and maximum).

Diagnosis and extent of disease

Summary statistics will be tabulated for diagnosis, overall aGvHD grade at diagnosis/baseline and stem cell transplant type.

For underlying disease, the analysis will include the following: primary diagnosis classification, details of primary diagnosis, time from diagnosis of underlying disease to start of study treatment, CIBMTR risk assessment.

For transplant related disease history, the analysis will include the following: conditioning regimen type, total HCT-specific comorbidity index score, time from transplant to the start of study treatment, time from diagnosis of underlying disease to transplant, stem cell type, cytomegalovirus status, time from engraftment to start of study treatment, time from transplant to engraftment, donor information including age, gender, HLA typing method, HLA match score, source of grafts, CMV status, T-cell depleted (Y/N), total nucleated cell dose.

For aGvHD disease history, the analysis will include the following: time from diagnosis of aGvHD to start of study treatment, Overall aGvHD grade when diagnosis of aGvHD, steroid refractory aGvHD criteria met, prior aGvHD therapy (steroid only, steroid + CNI, missing), time from diagnosis of aGvHD to steroid refractory, time from steroid refractory aGvHD to start of study treatment, Overall aGvHD grade when diagnosis of SR aGvHD, and aGvHD organ involvement.

Medical history

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

3.3.1 Patient disposition

Enrollment by country and center will be summarized for all screened patients and also by age group and for all patients. The number (%) of treated patients will be presented overall and by age group. The number (%) of screened and not-treated patients and the reasons for screening failure will also be displayed. The number (%) of patients in the FAS who are still on treatment, who discontinued the treatment and the study as well as the reason for discontinuation, will be presented overall and by age group.

Protocol deviations

The number (%) of patients in the FAS with any protocol deviation will be tabulated by deviation category overall and by age group for the FAS. All protocol deviations will be listed.

To assess the impact of COVID-19 pandemic on the conduct of the study, COVID-19 specific protocol deviations will be listed separately.

Analysis sets

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

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

3.4.1 Study treatment / compliance

Duration of exposure in days to ruxolitinib will be summarized by age group and by means of descriptive statistics. It will also be categorized into time intervals; frequency counts and percentages will be presented for the number (%) of patients in each interval.

Total daily dose, actual cumulative dose, dose intensity (DI) and relative dose intensity (RDI) will be summarized for ruxolitinib by age group.

The number (%) of patients who have dose reductions or interruptions, and the reasons, will be summarized for ruxolitinib by age group.

Patient level listings of all doses administered on treatment along with dose change reasons will be produced by age group.

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

If more than one dose and/or formulation is assigned within one age group, then summaries and listings will be reported by age group as well as doses/formulations within the age group.

Duration of exposure to study treatment

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

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

Duration of exposure in patient-years

The duration of exposure in patient-years is a total of the duration of exposure in years from all the patients in the agegroup.

Duration of exposure to systemic corticosteroid

Duration of exposure to systemic corticosteroid treatment (days) = (last date of exposure to systemic corticosteroid) – (date of first administration of systemic corticosteroid) + 1.

Duration of exposure to systemic corticosteroid in patient-years

The duration of exposure to systemic corticosteroid in patient-years is a total of the duration of exposure to systemic corticosteroid in years from all the patients in the age group.

Duration of exposure to CNIs

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

Duration of exposure to CNI in patient (years)

The duration of exposure to CNI in patient-years is a total of the duration of exposure to CNI in years from all the patients in the age group.

Cumulative dose

The **planned cumulative dose** for ruxolitinib refers to the total planned dose as per protocol for each group up to the last dose date.

The **actual cumulative dose of ruxolitinib** refers to the total actual dose of ruxolitinib as documented in the Study Treatment eCRF.

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

Dose intensity and relative dose intensity of ruxolitinib

Subjects receive ruxolitinib in doses twice a day (bid). For purposes of reporting dose intensity, a summary of dose per day (total daily dose) is provided.

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

DI (mg / day or mg/m^2/day) = Actual cumulative dose (mg or mg/m^2) of ruxolitinib / Duration of exposure to ruxolitinib (days).

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

Planned dose intensity (PDI) is defined as follows:

PDI (mg / day or mg/m^2/day) = Planned Cumulative dose (mg or mg/m^2) / Duration of exposure (days).

The PDI of ruxolitinib with daily dose schedule is according to protocol for the three age groups:

- Age group 1: 20 mg/day
- Age group 2: 10 mg/day
- Age group 3: 8 mg/m2/day using the baseline BSA

Relative dose intensity (RDI) is defined as follows:

RDI = DI (mg / day or mg/m^2/day) / PDI (mg / day or mg/m^2/day).

The total daily dose, actual cumulative dose, DI and RDI up to last date of exposure to study treatment will be summarized.

Dose reductions, interruptions or permanent discontinuations

The number of patients who have dose reduction, or interruptions, and the reasons, will be summarized for ruxolitinib by age group. Number of patients with dose change due to dose tapering (change of dose due to efficacy of treatment) and patients with dose change due to

other reasons (potential safety concerns) will be sumarised. The number of patients who have dose permanent discontinuations and the reasons will be summarized by age group.

'Type of change' and 'Dose permanently discontinued' fields from the Study Treatment CRF pages will be used to determine the dose change, dose interruptions, and permanent discontinuations respectively.

The corresponding fields 'Reason for Change' and 'Reason for Permanent Discontinuation' will be used to summarize the reasons.

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

Dose Reduction: A dose change when the actual total daily dose is lower than the previous actual total daily dose. Any dose change due to dose tapering (change of dose due to efficacy of treatment) will not be considered as a dose reduction. Only dose changes that are indicated on the CRF due to reason other than dose tapering are taken into consideration. Number of reductions will be derived programmatically based on the change and the direction of the change.

3.4.2 Prior, concomitant and post therapies

3.4.2.1 Prior aGvHD therapy

The number and percentage of patients who received any prior aGvHD therapy will be summarized by lowest ATC class, preferred term and age group. Prior aGvHD procedures will be summarized by system organ class, preferred term and age group. The above analyses will be performed using the FAS. Data will also be listed.

3.4.2.2 Prior aGvhD prophylaxis

The number and percentage of patients who received any aGvHD prophylaxis prior to start of study treatment will be summarized by lowest ATC class, preferred term and group using FAS. Listings will be generated for prophylaxis.

3.4.2.3 Systemic corticosteroid prior to start of study treatment

The number and percentage of patients who received systemic corticosteroids prior to study start will be summarized by age group with number of days on steroids and peak daily dose, where doses of methylprednisolone will be converted to prednisone equivalents. This analysis is done based on Full Analysis Set and will be reported in mg/kg/day (weight adjusted).

3.4.2.4 Systemic corticosteroid during study treatment

Doses of methylprednisolone will be converted to prednisone equivalents by multiplying the methylprednisolone dose by 1.25 and will be reported in mg/kg/day

The duration of exposure will be summarized for systemic corticosteriod by age group. It will also be categorized into time intervals; frequency counts and percentages will be presented for the number (%) of patients in each interval. The actual cumulative dose, dose intensity and

relative dose intensity (relative to the starting dose of corticosteroids) will be summarized up to Day 28, Day 56 and end of on-treatment period.(Section 3.4.1) This analysis will be based on Safety Set. The data will be reported on the Study treatment eCRF through the end of treatment period per EOT Disposition eCRF, and on the Prior and Concomitant Medications eCRF for any data prior and after the end of treatment. The list of systemic corticosteroids will be documented in Data Quality Plan.

3.4.2.5 Calcineurin inhibitors (CNIs) during study treatment

The duration of exposure will be summarized for systemic CNIs (cyclosporine or tacrolimus) up to end of on-treatment period by age group. It will also be categorized into time intervals; frequency counts and percentages will be presented for the number (%) of patients in each interval. These analyses will be based on Safety Set. Data will be analysed similarly to steroids and ruxolitinib as described in section 3.4.1. These data are reported on the study treatment CRF through the end of treatment per EOT Disposition eCRF, and on the Prior and Concomitant Medications eCRF after the end of treatment. The list of CNIs will be documented in the Data Quality Plan.

3.4.2.6 Additional systemic aGvHD therapy

New additional systemic aGvHD therapy (medications and procedures) since start of study treatment will be listed and summarized by lowest ATC class, preferred term, overall and by age group by means of frequency counts and percentages using FAS. These medications will be derived as per Appendix Section 6.4.3. Procedures are not coded by ATC class and hence they will appear under uncoded on the same table.

3.4.2.7 Concomitant medications

Concomitant therapy is defined as all interventions (therapeutic treatments, blood component transfusions and procedures) other than the study treatment administered to a patient coinciding with the study treatment period. Concomitant therapy include medications (other than study drugs) starting on or after the start date of study treatment or medications starting prior to the start date of study treatment and continuing after the start date of study treatment.

Concomitant medications will be coded using the World Health Organization (WHO) Drug Reference Listing (DRL) dictionary that employs the WHO Anatomical Therapeutic Chemical (ATC) classification system and summarized by lowest ATC class and preferred term using frequency counts and percentages. Surgical and medical procedures will be coded using MedDRA and summarized by SOC and preferred term. The summaries using Safety Set will include:

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

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

All concomitant CYP3A4 inhibitors and type of blood component transfusions will be summarized by ATC class and preferred term based on Full analysis Set.

3.5 Analysis of the primary objective

The primary objective of the Phase I is to assess pharmacokinetic (PK) parameters of ruxolitinib for patients with aGvHD and SR-aGvHD and define an age appropriate RP2D for each of the groups 2-4. The primary objective of the Phase II is to measure the activity of ruxolitinib in patients with aGvHD or SR-aGvHD assessed by Overall Response Rate (ORR) at Day 28.

3.5.1 Primary endpoint

The primary endpoint of the Phase I is the list of following PK parameters (AUC, Cmax, T1/2, Ctrough, and other parameters, as appropriate) which will be derived using non-compartmental methods in subjects with extensive sampling (Groups 1, 2, and 3). These parameters will then be used to define a RP2D for Groups 2, 3, and 4. The observed PK parameters (within group) will be summarized and compared to information obtained from adult and adolescent aGvHD patients treated with ruxolitinib on study CINC424C2301. Data from patients older than 2 years old will be combined and analyzed by PBPK methods to determine the dose to be administered in patients younger than 2 years old (Group 4).

The primary endpoint for the Phase II is the investigator-reported overall response rate (ORR) at Day 28, defined as the proportion of patients with complete response (CR) or partial response (PR) without requirement for additional systemic therapies for an earlier progression, mixed response or non-response as per Table 8-3 of the protocol (Harris et al 2016). Note that response is relative to the assessment of aGvHD at the start of study treatment.

- **Complete-response** is defined as a score of 0 for the aGvHD grading in all evaluable organs that indicates complete resolution of all signs and symptoms of aGvHD in all evaluable organs without administration of additional systemic therapies for any earlier progression, mixed response or non-response of aGvHD.
- **Partial response** is defined as improvement of 1 stage in 1 or more organs involved with aGvHD signs or symptoms without progression in other organs or sites without administration of additional systemic therapies for an earlier progression, mixed response or non-response of aGvHD.
- **Lack of response** is defined as no response, mixed response, or progression.
- **No response** is defined as absence of improvement in any organ involved by aGvHD, without worsening in any involved organ.
- **Mixed response** is defined as improvement of at least 1 stage in the severity of aGvHD in at least one organ accompanied by progression in another organ or development of signs or symptoms of aGvHD in a new organ.
- **Progression** is defined as worsening in 1 or more organs by 1 or more stages without improvement in any involved organ

Patients requiring additional systemic therapy for aGvHD will be classified as non-responders, and patients with missing baseline or Day 28 aGvHD response assessment will be considered as treatment non-responders.

aGvHD Flare is defined as any increase in signs or symptoms of aGvHD that is sustained for >24h after an initial response (CR or PR) and requires re-escalation of immunosuppressive therapy (e.g. corticosteroid, CNI and/or ruxolitinib dosing). While all aGvHD flares will be captured on study whether occurring during steroid, CNI or ruxolitinib taper. Only flares in GvHD that require new additional systemic therapy, will be considered aGvHD flare failure.

Acute GvHD will be assessed as per Table 8-2 of the protocol. Grade will be calculated based on the staging of the organs and recorded on the eCRF by the Investigator. The Investigator reported grade will be used for all analyses. Grade and response will be calculated by the sponsor for the purposes of data review.

3.5.2 Statistical hypothesis, model, and method of analysis

Pharmacokinetic parameters for patients in Groups 1, 2, 3 and 4 (who are part of the PAS) will be compared to adult information from study [CINC424C2301] through the use of ANOVA methods. The geometric mean ratio with 90% CI will be provided.

The response rates for ORR at Day 28 will be estimated on the Efficacy Evaluable Set (EES). Two-sided 90% confidence intervals [Clopper and Pearson 1934] will be calculated based on the exact method for binominal distribution. Summary statistics (frequencies and percentages) will be provided.

3.5.3 Handling of missing values/censoring/discontinuations

Patients with missing assessments that prevent the evaluation of the ORR will be considered non-responders. This includes aGvHD response assessments at baseline and Days 28, 56. Patients who discontinue the study treatment prior to the completion of the Day 28 visit will be considered non-responders.

The following analysis windows will be applied to the target day for assessments on overall response, where target day for Week X is X*7.

Baseline assessment is the last aGvHD assessment prior to or at the start of study treatment (Day 1).

Weeks 1, 2, 3, 4, 5, 6, 7, 8: -3 days/+3 days

Weeks 12 to 48: -13 days/+14 days

3.5.4 Supportive analyses

The subgroup analyses to assess the homogeneity of the treatment effect across demographic and baseline disease characteristics will be performed on the EES as appropriate. The subgroups are described in section 2.2.1.

A sensitivity analysis excluding subjects who had protocol deviations of organ staging or aGvHD response assessments not strictly following standard criteria defined in study protocol (PDID: OTHER03, OTHER04 per ECS) at Day 28 will be conducted.

Efficacy analyses in subgroups will be purely exploratory and are intended to explore the consistency of treatment effect. No inferential statistics (p-values) will be produced for the subgroups.

3.6 Analysis of the key secondary objective

3.6.1 Key secondary endpoint

The key secondary objective of the study is durable ORR at Day 56.

Durable ORR at Day 56 is defined as the proportions of all patients who achieve a complete response (CR) or partial response (PR) at Day 28 and maintain a CR or PR at Day 56. If a patient is a CR at Day 28 and a PR at Day 56, he/she will be considered as a durable responder. A patient will not be considered a durable responder at Day 56 if any of the following events occurs:

- Not repsonder at Day 28
- Missing aGvHD assessment at Day 56
- No CR or PR at Day 56.
- additional systemic therapy for aGvHD prior to Day 56

Durable ORR will be calculated based on the FAS using local investigators review of aGvHD assessment data.

3.6.2 Statistical hypothesis, model, and method of analysis

The durable response rates for ORR at Day 56 will be estimated with 90% confidence intervals on EES. Summary statistics (frequencies and percentages). The confidence intervals will be calculated based on the exact method for binominal distribution.

3.6.3 Handling of missing values/censoring/discontinuations

Refer to Section 2.5.3

3.7 Analysis of secondary efficacy objective(s)

3.7.1 Secondary endpoints

- **Overall Response Rate at Day 14**

Overall Response Rate at Day 14 is defined as the proportion of patients with complete response (CR) or partial response (PR) at Day 14 according to standard criteria [Harris 2016].

- **Duration of response (DOR)**

Duration of response will be calculated for patients whose overall response at Day 28 is CR or PR according to updated standard criteria (Harris et al 2016). The start date is the date of first documented response of CR or PR (i.e., the start date of response), which could be prior to or at Day 28. If it's prior to Day 28, there should not be progression or addition of systemic therapies for aGvHD between the start date of response and Day 28. The end date is defined as the date of progression or the date of addition of systemic therapies for aGvHD on or after Day28, since this constitutes a non-response. The date of progression will be the earliest date of PD recorded in the aGvHD assessment CRF page or the date of death due to aGvHD in the death CRF page.

Death without prior observation of aGvHD progression (i.e. date of death due to non aGVHD reason (Other) in the death CRF page) and onset of chronic GvHD are considered to be competing risks.

Duration of response will be censored at the last response assessment prior to or at the analysis cut-off date, if no events/competing risk occurred before or at of cut-off date.

- **Best Overall Response**

Best overall response (BOR) is defined as proportion of patients who achieved overall response (CR or PR) at any time point up to and including Day 28 and before the start of additional systemic therapy for aGvHD.

An additional analysis will include BOR at any time point up to the data cut-off date and before the start of additional systemic therapy for aGvHD. A swimmer plot for overall response by age group, SR-aGvHD vs Treatment Naive and formulation will also be provided for all treated subjects.

- **Overall survival (OS)**

Overall survival is defined as the time from date of the start of study treatment to date of death due to any cause. If a patient is not known to have died, then OS will be censored at the latest date the patient was known to be alive (last contact date on or before the cut-off date).

- **Event Free Survival (EFS)**

Event-free survival is defined as the time from the study treatment start date to the date of hematologic disease relapse/progression, graft failure or death due to any cause. If a patient is not known to have any event, then EFS will be censored at the latest date the patient was known to be alive (last contact date on or before the cut-off date).

- **Failure Free Survival (FFS)**

Failure-free survival is defined as the time from the study treatment start date to date of hematologic disease relapse/progression, non-relapse mortality (NRM) or addition of new systemic aGvHD treatment. Each type of failure as a competing risk for the other two, and onset of chronic GvHD is considered as a competing risk for all three types of failure. If a patient is not known to have any event or competing risk, then FFS will be censored at the latest date the patient was known to be alive (last contact date on or before the cut-off date).

- **Non-relapse mortality (NRM)**

Non-relapse mortality is defined as the time from the study treatment start date to date of death not preceded by hematologic disease relapse/progression. Hematologic disease relapse/progression is considered a competing risk for NRM with the date of hematologic disease relapse/progression being the earlier of documented hematologic disease relapse/progression or institution of therapy to treat potential hematologic disease relapse/progression. If a patient is not known to have died or to have relapsed/progressed, then NRM will be censored at the latest date the patient was known to be alive (last contact date on or before the cut-off date).

- **Incidence of Malignancy Relapse/Progression (MR)**

Malignancy relapse/progression is defined as the time from the study treatment start date to date of hematologic malignancy relapse/progression. Deaths not preceded by hematologic

malignancy relapse/progression are competing risks. If a patient is not known to have event or competing risks, then MR will be censored at the latest date the patient was known to be alive (last contact date on or before the cut-off date).

- **Cumulative steroid dosing until Day 56**
  The weekly cumulative steroid dose will be calculated for each patient up to Day 56 and the overall cumulative steriod dose will be calculated for each patient at Day 56. In addition, the proportion of patients with any dose reduction and any dose reduction of at least 50% in corticosteroids from baseline until EOT will be provided.

- **Incidence of cGvHD**

cGvHD is defined as the diagnosis of any cGvHD including mild, moderate, severe. Incidence of cGvHD is the time from the start of treatment to onset of cGvHD. Cumulative incidence of cGvHD will be estimated, accounting for deaths without prior onset of cGvHD and hematologic disease relapse/progression as the competing risks. If a patient is not known to have event or competing risks, then the incidence of cGvHD will be censored at the latest date the patient was known to be alive (last contact date on or before the cut-off date).

- **Graft failure**

This will be assessed by donor cell chimerism, defined as initial whole blood or marrow donor chimerism for those who had ≥5% donor cell chimerism at baseline. If donor cell chimerism declines to <5% on subsequent measurements, graft failure is declared.

3.7.2 Statistical hypothesis, model, and method of analysis

All secondary efficacy endpoint analyses will be analyzed using the EES.

- **Overall Response Rate at Day 14**

Overall response rate at Day 14 will be calculated based on the local investigators review of aGvHD assessment data. They will be summarized using descriptive statistics (N, %) with two-sided exact binomial 90% CIs [Clopper and Pearson 1934].

- **Best Overall Response (BOR)**

BOR up to and including Day 28 will be calculated based on the local investigators review of aGvHD assessment data. They will be summarized using descriptive statistics (N, %) with two-sided exact binomial 90% CIs [Clopper and Pearson 1934]. BOR up to cut-off date will also be analysed in the same way.

- **Duration of response (DOR)**

The estimated cumulative incidence rates and 95% confidence intervals at 1, 2, 6, 12, 18 and 24 months will be presented. The cumulative incidence curve will be plotted. DOR will be summarized for all patients in the EES with overall response of CR or PR at Day 28.

- **Overall survival (OS)**

The OS distribution will be estimated using the Kaplan-Meier method, and the Kaplan-Meier curves, medians, 1, 2, 6, 12, 18 and 24 month survival estimates and 95% confidence intervals [Brookmeyer and Crowley 1982] will be presented.

- **Event Free Survival (EFS)**

The EFS distribution will be estimated using the Kaplan-Meier method, and the Kaplan-Meier curves, medians, 1, 2, 6, 12, 18 and 24 month survival estimates and 95% confidence intervals [Brookmeyer and Crowley 1982] will be presented.

- **Failure Free Survival (FFS)**

Cumulative incidence curve for FFS as well as estimates at 1, 2, 6, 12, 18 and 24 months with 95% confidence intervals will be presented, accounting for onset of chronic GvHD as the competing risk.

In addition, the cumulative incidence of each of the three components considering the other two components as a competing risks will be estimated. Onset of chronic GvHD is considered as a competing risk for all three types of failure.

- **Non-relapse mortality (NRM)**

The cumulative incidence curve for NRM as well as estimates at 1, 2, 6, 12, 18 and 24 months with 95% confidence intervals will be presented.

- **Incidence of Malignancy Relapse/Progression (MR)**

The cumulative incidence curve for MR and estimates at 1, 2, 6, 12, 18 and 24 months with 95% confidence intervals will be presented for patients with underlying hematologic malignant disease, accounting for NRM as the competing risk.

- **Cumulative steroid dosing until Day 56**

Overall and weekly cumulative steroid dose for each patient up to Day 56 or discontinuation of study treatment will be tabulated. In addition, the relative dose intensity (RDI), by week, will be calculated relative to the starting dose of corticosteroids and categorized as (1) complete reduction where patients are tapered off corticosteroids by D56, (2) ≤50% RDI and (3) >50% RDI. The proportion of patients in each category and corresponding 95% confidence intervals will be presented.

- **Incidence of cGvHD**

The cumulative incidence of cGvHD and estimates at 1, 2, 6, 12, 18 and 24 months with 95% confidence intervals will be presented, accounting for deaths without prior onset of cGvHD and hematologic disease relapse/progression as the competing risks. The cumulative incidence curve will be plotted.

- **Graft failure**

Summary statistics of graft failure assessment will be presented over time and at any timepoint by dose level and for all patients. In case of very few patients with graft failure (e.g. less than 5 patients), descriptions using only listings will be provided instead.

3.7.3 Handling of missing values/censoring/discontinuations

Patients with missing assessments that prevent the evaluation of the ORR at Day 14 will be considered non-responders. This includes missing aGvHD assessments at baseline and/or Day 14. The time window for the Day 14 visit is defined in Section 2.5.3.

3.8 Safety analyses

For all safety analyses, the safety set will be used. All listings and tables will be presented by group. If the Phase I starting dose is different to RP2D, the safety data will be listed and summarized by age group and dose level. For safety evaluations (except for AE), the last available assessment on or before the date of start of study treatment is taken as the "baseline" assessment.

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

The on-treatment period lasts from the date of first administration of ruxolitinib to 30 days after the date of the last actual administration of ruxolitinib.

3.8.1 Adverse events (AEs)

For reporting of AEs the overall observation period will be divided into mutually exclusive categories, including pre-treatment, on-treatment, post-treatment periods as defined in Section 2.1.1.

AE summaries will include all AEs occurring (new or worsening) during on-treatment period. All AEs collected in the AE eCRF page will be listed along with the information collected on those AEs e.g. AE relationship to study drug, AE outcome etc. AEs with start date outside of on-treatment period will be flagged in the listings.

AEs will be summarized by number and percentage of patients having at least one AE, having at least one AE in each primary system organ class (SOC) and for each preferred term (PT) using MedDRA coding. A patient with multiple occurrences of an AE will be counted only once in the respective AE category. A patient with multiple CTCAE grades for the same preferred term will be summarized under the maximum CTCAE grade recorded for the event. AE with missing CTCAE grade will be included in the 'All grades' column of the summary tables.

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

The following adverse event summaries will be produced by group; overview of adverse events and deaths, AEs by SOC and PT, summarized by relationship (all AEs and AEs related to study treatment), seriousness (SAEs and non-SAEs), leading to treatment discontinuation, leading to dose interruption/adjustment, and leading to fatal outcome.

3.8.1.1 Adverse events of special interest / grouping of AEs

[redacted]

[redacted]

For each AESI, number and percentage of patients with at least one event of the AESI occurring during on-treatment period will be summarized. Summaries of AESIs will be provided by group (specifying grade, SAE, relationship, leading to treatment discontinuation, leading to dose adjustment/interruption, hospitalization, death etc.).

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

In addition to summarizing infections by CTCAE grade, they will also be summarized using infection severity (protocol Appendix 1) for the on-treatment period. A listing of infection with infection severity will be generated and the events with start date outside of on-treatment period will be flagged.

3.8.1.2 Adverse events of clinical trial safety disclosure

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

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

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

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

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

3.8.2 Deaths

Separate summaries for on-treatment and all deaths (including post-treatment death) will be produced by age group, system organ class and preferred term.

All deaths will be listed and deaths that occured outside of on-treatment period will be flagged. A separate listing of deaths prior to starting treatment will be provided for all screened patients.

3.8.3 Laboratory data

Laboratory, data from all sources (central and local laboratories) will be combined.

Grading of laboratory values will be assigned programmatically as per NCI Common Terminology Criteria for Adverse Events (CTCAE) version 4.03. The calculation of CTCAE grades will be based on the observed laboratory values only, clinical assessments will not be taken into account.

CTCAE Grade 0 will be assigned for all non-missing values not graded as 1 or higher.

For laboratory tests where grades are not defined by CTCAE v4.03, results will be categorized as low/normal/high based on laboratory normal ranges.

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

- Shift tables using CTC grades to compare baseline to the worst on-treatment value
- For laboratory tests where CTC grades are not defined, shift tables using the low/normal/high/(low and high) classification to compare baseline to the worst on-treatment value.

The following listings will be produced for the laboratory data:

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

The results from viral load testing will be listed.

3.8.4 Other safety data

3.8.4.1 ECG and cardiac imaging data

Not applicable

3.8.4.2 Vital signs

Vital sign assessments are performed in order to characterize basic body function. The following parameters were collected: height or body length (cm), weight (kg), body temperature (°C), pulse (beats per minute), systolic and diastolic blood pressure (mmHg). The body surface area (BSA in m2) and body mass index (BMI in kg/m^2) are derived using the following formulas:

BSA: (Height (cm) x Weight(kg) / 3600) ^ 0.5

BMI: Weight(kg)/ $Height^2$ (m^2)

Data handling

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

Clinically notable vital sign criteria for blood pressure or weight (see below) are based on the height/body length of the subject. Since height/body length might not be measured at each visit, the last available height/body length value before that assessment date will be used instead, i.e. using LOCF as imputation method.

Data analysis

For analysis of vital signs the clinically notable vital sign criteria are provided in Table 2-4 below.

Table 3-4 Clinically notable changes in vital signs

Vital Sign	High	Low
Systolic blood pressure [mmHg]	≥ 95th percentile of the age and height group[1]	≤ 5th percentile of the age and height group[1]
Diastolic blood pressure [mmHg]	≥ 95th percentile of the age and height group[1]	≤ 5th percentile of the age and height group[1]
Body temperature [°C]	≥ 38.4°C	≤ 35.0°C
Pulse rate [bpm] [2]	12-≤18 months > 140 18-≤24 months > 135 2-≤3 years > 128 3-≤4 years > 123 4-≤6 years > 117 6-≤8 years > 111 8-≤12 years > 103 12-≤15 years > 96 ≥ 15 years > 92	12-≤18 months < 103 18-≤24 months < 98 2-≤3 years <92 3-≤<4 years <86 4-≤6 years <81 6-≤8 years <74 8- ≤12 years <67 12-≤15 years <62 ≥ 15 years <58
Weight	increase from baseline[3] of ≥ 2 BMI- for-age percentile categories[3]	decrease from baseline[3] of ≥ 2 BMI-for-age percentile categories[3]
Height	increase from baseline of ≥ 2 height-for-age/gender percentile categories[3]	decrease from baseline of ≥ 2 height-for-age/gender percentile categories[3]

bpm=beats per minute; CDC= Centers for Disease Controls and prevention;
[1] Blood pressure percentiles are calculated for each blood BP record using the method described in Appendix B of the following reference: The Fourth Report on Diagnosis, Evaluation and Treatment of High Blood Pressure in Children and Adolescents. Pediatrics 2004; 114; 555.
[2] Fleming S, Thompson M, Stevens R, et al. Normal ranges of heart rate and respiratory rate in children from birth to 18 years of age: a systematic review of observational studies. Lancet 2011; published online March 15. DOI:10.1016/S0140-6736(10)62226-X.
[3] Gender percentiles categories (P3, P5, P10, P25, P50, P75, P85, P90, P95, P97) are obtainedfrom the CDC Growth Charts

The number and percentage of patients with notable vital sign values (high/low) will be presented.

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

3.8.4.3 Cumulative incidence of grade 3 infection

Cumulative incidence curves for time to grade >=3 infections as well as estimates at 1, 2, 6 months etc. with 95% confidence intervals will be presented overall and for each age group (if sufficient number of patients are available in the age groups) as well as by treatment-naive vs SR-cGvHD, considering death or discontinuation from study treatment (due to any reason) without prior infection as competing risks.

3.8.4.4 Acceptability and Palatability assessment

All acceptability and palatability assessment data will be sumarised and listed by age group and formulation.

3.9 Pharmacokinetic endpoints

Pharmacokinetic Analysis Set (PAS) will be used in all pharmacokinetic data analysis and PK summary statistics.

In patients with extensive PK sampling on Day 1 (samples 1 to 8 described in Section 8.6.1.1 of the protocol), PK parameters of ruxolitinib will be calculated by non-compartmental methods using Phoenix WinNonlin (Pharsight, Mountain View, CA) software. The following parameters will be calculated (Table 2-5). Additional PK parameters may be estimated as needed.

Table 3-5 Non-compartmental pharmacokinetic parameters of ruxolitinib

AUClast	The AUC from time zero to the last measurable concentration sampling time (Tlast)
AUCinf	The AUC from time zero extrapolated to infinity
AUC0-12	The AUC from time zero extrapolated to 12 hours
Cmax	The maximum (peak) observed plasma drug concentration
Ctrough	The minimum observed plasma concentration at the end of an administration interval (corresponding to the pre-dose concentration prior to the following administration)
Tmax	The time to reach maximum (peak) plasma, blood, serum, or other body fluid drug concentration after single dose administration (time)
T1/2	The elimination half-life associated with the terminal slope (Lambda_z) of a semi logarithmic concentration-time curve
CL/F	The total body clearance of drug from the plasma
Vz/F	The apparent volume of distribution during terminal phase (associated with Lambda_z)

All plasma concentrations will be analyzed using a population PK approach.

Statistical methods for pharmacokinetic analyses

All individual plasma ruxolitinib concentrations data will be listed by group. Descriptive summary statistics will be provided by group at each scheduled time point. Summary statistics will include n (number of patients with non-missing values), mean (arithmetic and geometric), SD, CV (arithmetic and geometric), median, minimum and maximum. Individual profiles with median by dose level as well as arithmetic mean with SD and geometric mean ruxolitinib plasma concentration versus time profiles by treatment will be displayed graphically.

Ruxolitinib plasma PK parameters data will be listed by group. Descriptive statistics (n, arithmetic mean, standard deviation (SD), coefficient of variation (CV)% for mean, geometric mean, geometric CV%, median, minimum and maximum) will be provided for all PK parameters by dose level except for Tmax where median, minimum and maximum will be presented.

If the Phase I starting dose is different to RP2D, PK data will be listed and summarized by age group and dose level.

Population PK approach on pooled Phase I and Phase II data

Nonlinear mixed effects modeling (population PK) or other model-based approaches, as appropriate, will be used to analyse the concentration data from all patients. Details of the analysis methods will be developed in a PK analysis plan and the population PK analysis will be documented in a separate report.

During modeling of the pharmacokinetics of study treatment, the broad principles outlined in the FDA guidance will be followed (*Guidance for Industry*: Population Pharmacokinetics; http//www.fda.gov/cder/guidance/1852fnl.pdf).

Data handling principles

All concentration values below the lower limit of quantitation (LLOQ) are set to zero by the Bioanalyst, and will be displayed in the listings as zero and flagged. LLOQ values will be treated as zero in any calculations of summary statistics, and treated as missing for the calculation of the geometric means and their CV%. The number of non-zero concentrations will also be reported in the summary statistics.

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

3.10 PD and PK/PD analyses

Exposure-Response analysis

- The following are objectives for exposure-response analysisCharacterize the exposure-efficacy relationship of ruxolitinib with efficacy response defined as overall response rate at Day 28, durable response at Day 56, and Overall survival. Exposure metrics may be described in further details in a separate analysis plan.
- Characterize the exposure-safety relationship of ruxolitinib with safety response defined as 1st occurrence AEs, 1st occurrence of G3/4 AEs, 1st occurrence of AEs of interest, and appropriate liver function parameters. Exposure metrics may be described in further details in a separate analysis SAP.
- Exposure-Biomarker – [redacted]

For overall survival, if sufficient events have accrued (i.e., at analysis time points after the primary analysis), a Cox regression model with appropriate patient demographic and prognostic factors as covariates and the log-trough concentration level as a time dependent covariate will be fitted as appropriate. Goodness of fit of the model will also be examined, and other exposure measures (such as dose intensity) may be considered if appropriate. The survival time should be calculated from the time of first dose. The hazard ratio and 95% confidence interval for a

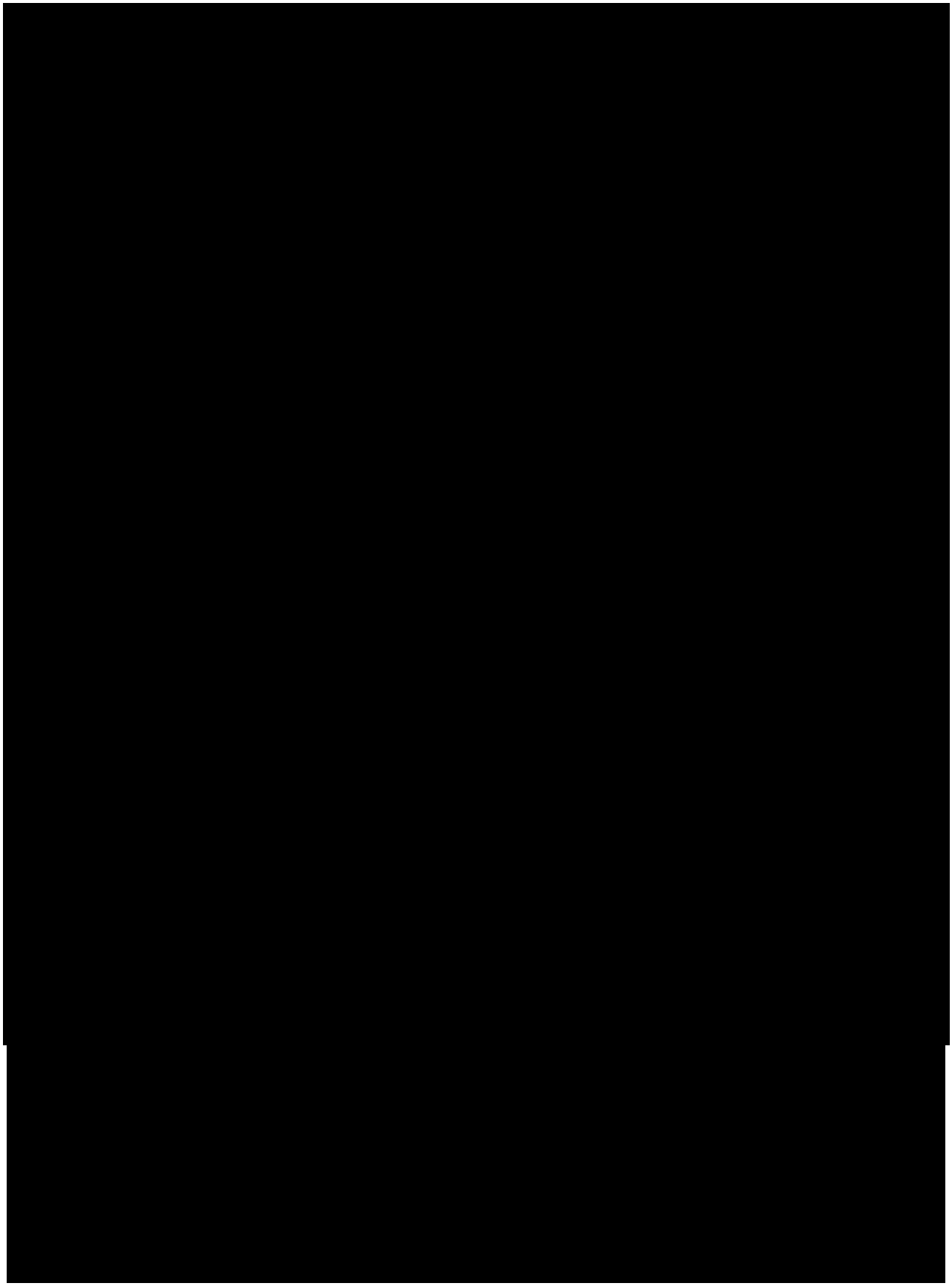
decrease by half in exposure will be displayed if appropriate. Kaplan-Meier curves may be used to summarize the data based on relevant quantiles of PK.

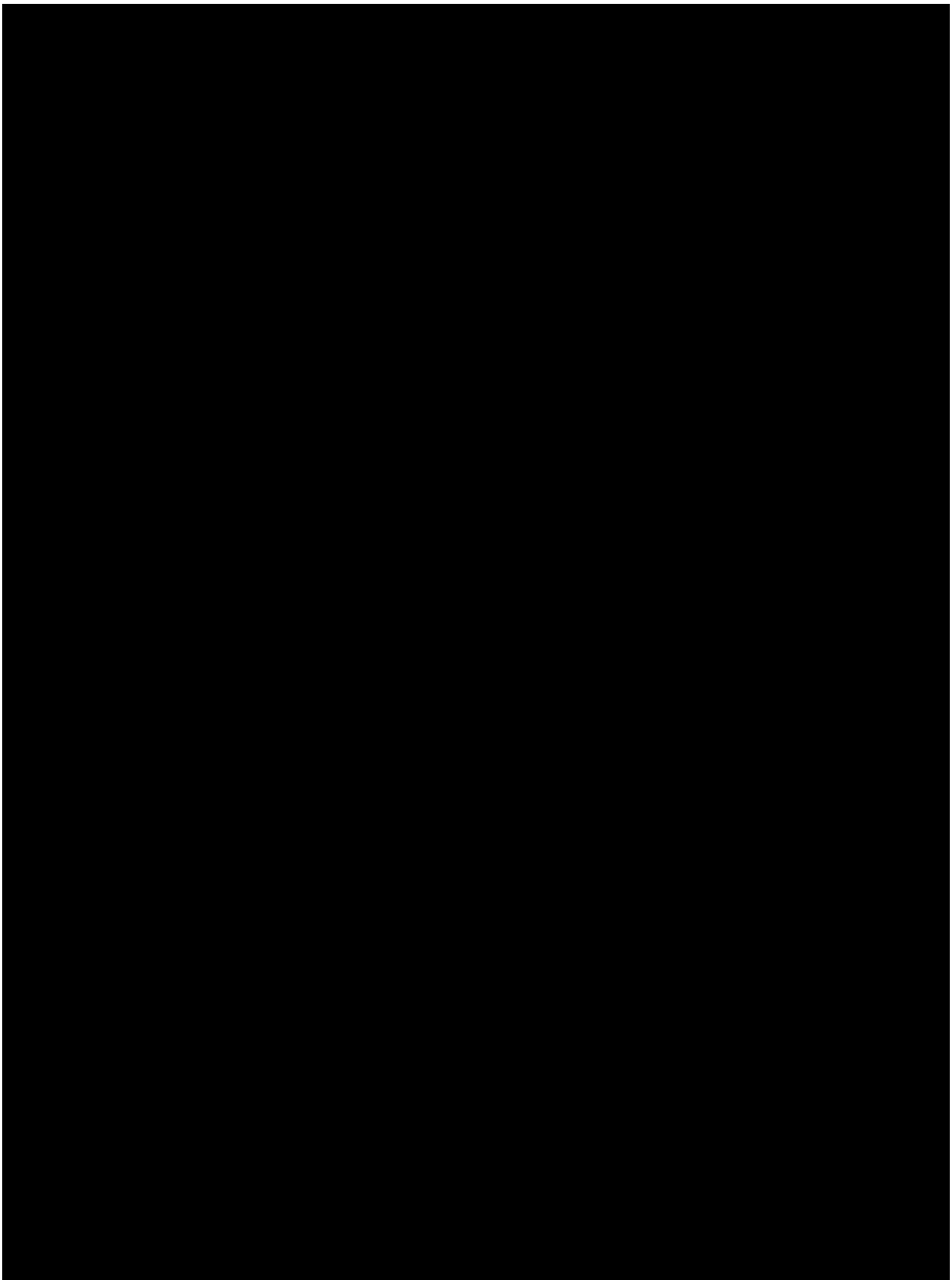
For incidence of specific AEs, Day 28 response and Day 56 durable response, logistic regression models for efficacy may be used including log- time-averaged trough concentration, and log-trough concentration (last observed prior to onset of event) for safety, as well as other demographic and prognostic covariates in the model as appropriate. The average trough concentration included in the model for Day 28 response will be the geometric mean of the Day 7, Day 14, Day 21 and Day 28 adjusted trough concentration. A similar average trough concentration will be calculated for the Day 56 response except the Day 42 and Day 56 will also be included Goodness of fit of the model will also be examined, and other exposure measures (such as dose intensity) may be considered if appropriate. The odds ratio and its 95% confidence interval for a two-fold increase in exposure will be displayed as appropriate.

Expected steady-state exposures and/or other PK parameters for the population may be computed by the population PK model accounting for dose modifications or dose interruptions up to the day prior to the day of assessments. Population PK derived parameters may be also considered for exposure-response analyses by appropriate methods.

3.11 Patient-reported outcomes

Not applicable.





[redacted]

3.14 Interim analysis

No formal interim analysis is planned for this study. However, summaries of safety and PK data will be produced to support the regular safety monitoring conducted by the DMC and the confirmation of RP2D. Efficacy data may also be analyzed when all patients (Phase I and Phase II) have completed 24 weeks (approximately 6 months) of treatment or discontinued earlier. The final analysis will be conducted when all patients have completed Month 24 (i.e., the end of the long-term follow-up period), unless the patients discontinue earlier.

In this clinical trial it is not planned to test specific efficacy hypotheses but to provide estimates of efficacy endpoints for the pediatric study population and therefore no-alpha adjustment will be made for earlier estimates of efficacy endpoints. However, the data for the primary efficacy assessment, ORR at Day 28 will already be final once all patients have completed 6 months of treatment.

4 Sample size calculation

In Phase I, 5 patients will be enrolled to each age groups 2 and 3. A minimum of 5 patients in Group 1 will also undergo extensive PK sampling during Phase II. Should one or more of the 5 patients not be evaluable for PK analysis, additional patients may be enrolled to ensure a minimum of 5 evaluable PK profiles for Groups 1, 2 and 3.

A comparison of PK parameters (Cmax) from 5 pediatric patients (in Groups 2 and 3 separately) to ~25 adult patients (from study CINC424C2301) will be performed. With expected similarity in the point estimates across groups (geo-mean ratio, GMR = 1), and accounting for expected higher variability in the pediatric patients (CV% ~55.7%, (Loh et al 2015, detailed data on file) compared to ~40% in adults), the confidence interval for a GMR of 1 would be [0.609;1.641] which demonstrates clinically relevant comparability of exposure to adult exposure (within 2-fold).

Therefore, with a minimum of 5 evaluable profiles in each of Groups 2 and 3, combined with 5 evaluable profiles from Group 1 and further sparse PK samples in the Phase II of the study, there is sufficient precision to support the PK objectives of the study.

Should the exposure in Group 2, 3 or 4 not be confirmed following the PK sampling in at least 5 patients, additional patients will be enrolled in that specific age group until the dose/exposure is confirmed (i.e., selection of the RP2D for those ages based on exposure and safety review by the DMC). Once the RP2D is selected for Groups 2 and 3 any further eligible patients between the ages of 2 years and 12 years will be enrolled into the Phase II, and Group 4 will begin enrolling patients in the Phase I.

Although several studies of treatments for aGvHD and SR-aGvHD have included pediatric patients, it is difficult to find recent estimates for steroid response in de novo aGvHD pediatric

patients or BAT response in SR-aGvHD pediatric patients. A review of 443 patients (including 175 patients < 20 years, reported a CR/PR rate of 55% at Day 28, no association with ORR rate and patient age was found (MacMillan et al 2010). Other papers report response rates in the ranges 50-60% for de novo and SR-aGvHD (Martin et al 2012).

The sample size for the Phase II objective of measuring ORR at D28 is 45 patients regardless of age. Of these, there must be at least 20% of the patients with treatment naïve aGvHD and 40% of the patients with SR-aGvHD to ensure the sample is representative of the study population. The remaining can be filled with either diagnosis. Any patient receiving the confirmed RP2D during the Phase I will be counted towards the 45 patients.

The sample size calculation for the Ph II activity objective is based on the ORR at Day 28. Assuming the true ORR at Day 28 of the study population is 80%, an overall sample size of 45 patients would have 90% probability to have a 90% CI for ORR with lower limit ≥ 60%. In addition, considering the Saw-Toothed behavior of power waving for single binomial proportion using an exact method (Chernick and Liu 2002). A minimum sample size of 45 subjects would provide >85% of the probability to have a 90% CI with lower limit ≥ 60%. Table 12-2 provides estimates of power, minimum number of responders required to have a 90% CI with lower limit of ≥ 60% and two-sided 90% Clopper-Pearson CIs for various sample size. Patients treated at the RP2D from the Phase I contribute to this analysis.

Table 4-1 Probability to have a 90% CI with lower limit ≥ 60% and 90% confidence intervals for different number of patients

Sample size	Minimum No. of responders	Response Rate	Probability to have a 90% CI with lower limit ≥ 60%	90% CI	
37	28	0.76	0.809	0.614	0.867
38	29	0.76	0.784	0.623	0.871
39	29	0.74	0.859	0.604	0.854
40	30	0.75	0.839	0.613	0.858
41	31	0.76	0.818	0.621	0.861
42	31	0.74	0.882	0.604	0.846
43	32	0.74	0.864	0.612	0.849
44	33	0.75	0.846	0.620	0.853
45	33	0.73	0.901	0.604	0.838
46	34	0.74	0.886	0.612	0.842
47	35	0.74	0.869	0.619	0.846
48	35	0.73	0.916	0.604	0.832
49	36	0.73	0.903	0.612	0.835
50	37	0.74	0.889	0.619	0.838

5 Change to protocol specified analyses

No change from protocol

6 Appendix

6.1 Imputation rules

6.1.1 Study drug

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

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

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

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

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

Please note that date of assessment on EOT eCRF might be very different from last date of dose.

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

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

Use Dec31yyyy

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

Use EOT date

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

Use last day of the Month (mm)

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

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

Use the treatment start date

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

6.1.2 AE, ConMeds and safety assessment date imputation

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

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

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

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

Table 5-2 Imputation of end dates (AE, CM)

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

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

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

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

6.1.2.1 Other imputations

Incomplete date of initial diagnosis of aGvHD

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

6.2 AEs coding/grading

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

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

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

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

6.3 Laboratory parameters derivations

Grade categorization of lab values will be assigned programmatically as per NCI Common Terminology Criteria for Adverse Events (CTCAE) version 4.03. The calculation of CTCAE grades will be based on the observed laboratory values only, clinical assessments will not be taken into account. The criteria to assign CTCAE grades are given in Novartis internal criteria for CTCAE grading of laboratory parameters. The latest available version of the document based on the underlying CTCAE version 4.03 at the time of analysis will be used. For laboratory tests where grades are not defined by CTCAE v4.03, results will be graded by the low/normal/high (or other project-specific ranges, if more suitable) classifications based on laboratory normal ranges.

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

Imputation Rules

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

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

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

Further derivation of laboratory parameters might be required for CTCAE grading.

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

6.4 Statistical models

6.4.1 Primary analysis

Confidence interval for ORR

Responses will be summarized in terms of percentage rates with two-sided 90% confidence interval using exact binomial confidence interval (implemented using SAS procedure FREQ with EXACT statement [Clopper and Pearson 1934].

6.4.2 Key secondary analysis

Same instruction for the confidence interval for ORR as in section 5.4.1

6.4.3 Definition of new or additional systemic treatment for aGvHD

The data source to search for the new or additional systemic therapy is the CRF of "Prior and Concomitant Medication" (CONMED) or CRF "Prior or Concomitant non-drug therapies/procedures" and "Study treatment CNI" .

Systemic therapies are identified using the Route (if CMROUTE is Intramuscular, Intravenous, Oral, Subcutaneous, or Parenteral).

Any of the following therapies represent new or additional systemic therapy for aGvHD:

- Any CNI therapy being initiated newly as 'treatment for aGvHD' after the baseline (as recorded on Study treatment CNI and with indication = "treatment for aGvHD" on CONMED page). That means when a subject starts a new systemic therapy during study treatment for the indication 'treatment for aGvHD', or a subject received the same therapy before as prophylaxis and then changed to treatment of aGvHD as indication, this will be considered as new or additional systemic treatment.

  CNI therapy given as 'treatment for aGvHD' that started after Day 1 will be considered as new or additional systemic treatment, even if the same therapy was given as treatment for aGvHD already before start of study treatment but the patient was not receiving such therapy on Day 1. Whereas a stop of such therapy after Day 1 and later re-start of the same therapy will not be considered as new or additional therapy.

  Change of the drug within the defined group of CNIs or only change of the dose is not considered as new CNI treatment.

- Any other systemic treatments (excluding CNI and steroids) being initiated newly as 'treatment for aGvHD' or after the baseline and documented in CONMED (according to the same definitions as for CNI treatments above).

- Any other non-drug therapy/procedure being initiated newly as 'treatment for aGvHD' after the baseline (according to the same definitions as for CNI above) and documented on concomitant non-drug therapies/procedures (e.g. ECP).

Changes in the systemic corticosteroid therapy (captured on study treatment eCRF), e.g. change of the drug, the dose or any interruption and re-start, will not be taken into account for identification of new or additional systemic therapy, nor will any other corticosteroid therapy (in CONMED) be considered.

7 Reference

1. Anderson JR, Cain KC and Gelber RD (1983). Analysis of survival by tumor response. Journal of Clinical Oncology, 1(11), 710-719.
2. Armitage P, McPherson CK and Rowe BC (1969). Repeated significance tests on accumulating data. J .R. Statist. Soc. A, 132, 232-244.
3. Brookmeyer R and Crowley J (1982). A Confidence Interval for the Median Survival Time. Biometrics, 38, 29 - 41.
4. Chernick, M. R. and C. Y. Liu (2002) The saw-toothed behavior of power versus sample size and software solutions: single binomial proportion using exact methods p. 56(2): 149-155.
5. Clopper CJ and Pearson ES (1934). The use of confidence or fiducial limits illustrated in the case of the binomial. Biometrical, 26, 404-413.
6. FDA briefing Document Oncologic Drugs Advisory Committee Meeting, July 24, 2012, Evaluation of radiologic review of progression-free survival in Non-hematologic malignancies. http://www.fda.gov/downloads/advisorycommittees/committeesmeetingmaterials/drugs/oncologicdrugsadvisorycommittee/ucm311141.pdf
7. Harris AC, Young R, Devine S, Hogan WJ, Ayuk F, Bunworasate U, Chanswangphuwana C, Efebera YA, Holler E, Litzow M, Ordemann R, Qayed M, Renteria AS, Reshef R, Wölfl M, Chen YB, Goldstein S, Jagasia M, Locatelli F, Mielke S, Porter D, Schechter T, Shekhovtsova Z, Ferrara JL, Levine JE. (2016) International, Multicenter Standardization of Acute Graft-versus-Host Disease Clinical Data Collection: A Report from the Mount Sinai Acute GVHD International Consortium. Biol Blood Marrow Transplant; 22(1): 4-10.
8. Korhonen P, Laird NM and Palmgren J. Correcting for non-compliance in randomized trials: an application to the ATBC study (1999). Statistics in Medicine, 18, 2879-2897.
9. Loh ML, Tasian SK, Rabin KR, et al (2015) A phase 1 dosing study of ruxolitinib in children with relapsed or refractory solid tumors, leukemias, or myeloproliferative neoplasms: A Children's Oncology Group phase 1 consortium study (ADVL1011). Pediatr Blood Cancer p. 1717-24.
10. Mantel, N. and Joseph L Fleisś (1980). Minimum expected cell size requirements for the Mantel-Haenszel one degree-of-freedom Chi-Square Test and a related rapid procedure. American Journal of Epidemiology, 112, 129-134.
11. Rauchenzauner M, Schmid A, Heinz-Erian P, et al (2007) Sex- and age-specific reference curves for serum markers of bone turnover in healthy children from 2 months to 18 years. J. Clin. Endocrinol. Metab. p. 443-9.
12. Robins JM and Tsiatis AA (1991). Correcting for non-compliance in randomized trials using rank preserving structural failure time models (RPSFTM). Commun. Statist.-Theory Meth., 20(8), 2609-31.
13. Van Groningen, LF, Liefferink AM, de Haan AF, Schaap NP, Donnelly JP, Blijlevens NM, van der Velden WJ. (2016) Combination Therapy with Inolimomab and Etanercept

for Severe Steroid-Refractory Acute Graft-versus-Host Disease Biol Blood Marrow Transplant. 22(1):179-82.