NOVARTIS

Clinical Development

INC424/Ruxolitinib/Jakavi®

CINC424G12201

A Phase II open-label, single-arm, multi-center study of ruxolitinib added to corticosteroids in pediatric subjects with moderate and severe chronic graft vs. host disease after allogeneic stem cell transplantation

Statistical Analysis Plan (SAP) – final CSR

Author:	Trial Statistician, [redacted] Pharmacometrician, [redacted]

Document type:	SAP Documentation
Document status:	Final
Release date:	07-Oct-2024
Number of pages:	46

Document History – Changes compared to previous final version of SAP

Date	Time point	Reason for update	Outcome for update	Section and title impacted (Current)
18-Dec-2019	Prior to DB lock	Creation of final version	N/A – First version	NA
24-Mar-2022	Prior to DB lock	Amendment 1	This SAP has been updated after protocol Amendment version 2 (release date 19-Nov-2020) and creation of TFL Shells. Main changes: • Additional analysis set "Listing only set" • Definition of analysis sets updated to consider patients with lower starting dose as a result of co-administration with strong CYP3A4/CYP2C9 inhibitors • Additional exploratory endpoint "Steroid-free remission at Cycle 7 Day 1" • Clarification of subgroup analyses by regions to add Japan and Turkey as separate region; number of subgroup analyses reduced • Analyses to investigate impact of COVID-19 pandemic added • Further changes for clarification purposes and/or to align with other studies in the project	Section: 1.2 2.1 – 2.4 2.7, 2.8, 2.13 App. 5.3
19-Nov-2022	Prior to DB lock	Amendment 2	Update to align with protocol Amendment version 2 (release date 09-Sep-2022) and to clarify some specifications after dry run. Main changes: • Exploratory endpoint "Steroid-free remission at Cycle 7 Day 1" changed into "Systemic corticosteroid-free response" and derivation clarified • Sensitivity analysis for primary endpoint added	Sections: 1.2 2.1.1, 2.2, 2.4.1, 2.5.4 2.7 2.8.4 2.13 4 App. 5.4.3

Date	Time point	Reason for update	Outcome for update	Section and title impacted (Current)
			• Analysis of BOR up to data cut-off date added and analysis of DOR up to data cut-off date added accordingly • Analysis of time to first occurrence of infections using Kaplan-Meier methodology removed • Specified how to analyze patients who change formulation during treatment • Update of protocol deviation codes in the analysis set definitions (Tab. 2-3), but content of definition keeps unchanged • More detailed specifications how to identify new or additional systemic therapy • Simplified derivation of dose reduction • Criteria for notable vital signs clarified	
07-Oct-2024	Prior to final DB lock	Final CSR	Main changes: • Analyses identified that don't need to be repeated for the final CSR • Definition of on-treatment period for safety analyses clarified, especially for subjects who tapered off ruxolitinib treatment. • Definition for duration of follow-up added. • Duration of exposure incl./excl. taper off periods added in presence of subjects who tapered off ruxolitinib and re-started after cGvHD recurrence. Definition of dose intensity adapted accordingly. • Analysis of prohibited concomitant medications removed since table on corresponding PD is sufficient.	Sections: General 2.1.1 2.4.1 2.4.1 2.4.2

Date	Time point	Reason for update	Outcome for update	Section and title impacted (Current)
			• Definition of duration of response updated according to what was used in the iCSR (specified in CSR Section 16.1.9 "Changes to the planned analysis")	2.7.1
			• Analyses on reduction of corticosteroid dose updated to consider data up to end of study	2.7.1/2.7.2
			• Analysis for safety disclosure updated to report events for on-treatment and post-treatment period, separately.	2.8.1.1
			• Criteria for notably abnormal vital signs updated according to the analysis in the iCSR (specified in CSR Section 16.1.9 "Changes to the planned analysis")	2.8.4.2
			• Specifications for analysis of time to first occurrence of infection clarified.	2.8.4.3
			• Shift table for height, weight, BMI (high/normal/low) added.	2.8.4.5
			• All specification for biomarker analyses moved into a separate analysis plan	2.12
			• Specification added how to identify subjects who tapered off ruxolitinib for the analysis of cGvHD recurrence following completion of a taper of ruxolitinib.	2.13
			• Analysis of systemic corticosteroid-free response rate clarified for subjects who didn't receive steroids at baseline.	2.13

Table of contents

List of abbreviations

AE	Adverse event
AESI	Adverse Event of Special Interest
aGvHD	Acute Graft vs. Host Disease
ATC	Anatomical Therapeutic Classification
AUC	Area Under the Curve
bid	bis in diem/twice a day
cGvHD	chronic Graft vs. Host Disease
CIBMTR	Center for International Blood and Marrow Transplant Research
CMV	Cytomegalovirus
CR	Complete Response
CSR	Clinical Study report
CTC	Common Toxicity Criteria
CTCAE	Common Terminology Criteria for Adverse Events
DAR	Dose Administration Record
eCRF	Electronic Case Report Form
FAS	Full Analysis Set
FFS	Failure-Free Survival
GvHD	Graft vs. Host Disease
HCT	Hematopoietic Cell Transplantation
MedDRA	Medical Dictionary for Regulatory Activities
NCI	National Cancer Institute
NRM	Non Relapse Mortality
ORR	Overall Response Rate
OS	Overall Survival
PAS	Pharmacokinetic analysis set
PD	Pharmacodynamics
PK	Pharmacokinetics
PPS	Per-Protocol Set
PR	Partial response
PRO	Patient-reported Outcomes
SAE	Serious Adverse Event
SAP	Statistical Analysis Plan
WBC	White Blood Cells
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 final Clinical Study Report (CSR) of study CINC424G12201, a Phase II open-label, single-arm, multi-center study of ruxolitinib added to corticosteroids in pediatric subjects with moderate and severe chronic graft vs. host disease after allogeneic stem cell transplantation. Specification of analyses done for the interim CSR but not repeated for the final CSR are kept in this document which can be seen as cumulative SAP. These are those analyses for which no new or updated data are available (for example, all efficacy at Cycle 7 Day 1, all demographic and baseline disease characteristics). Analyses that are not repeated for the final CSR are indicated in Tab. 1-1 and/or in the corresponding sections.

This SAP is based on protocol CINC424G12201 version 02, dated 09-Sep-2022.

- Changes implemented in this SAP version compared to previous versions are indicated in the document history table above. The major changes compared to the previous version are:
- Analyses specified that don’t need to be repeated in the final CSR
- Definition of on-treatment period for safety analyses clarified, especially for subjects who tapered off ruxolitinib treatment.
- Duration of exposure incl./excl. taper off periods added in presence of subjects who tapered off ruxolitinib. Definition of dose intensity adapted accordingly.
- Definition for duration of follow-up added.
- Changes specified in the interim CSR Section 16.1.9 “changes to the analysis plan” are implemented into the SAP, which are the updated
    - Definition of duration of response and
    - Criteria for notably abnormal vital signs.
- Analyses on reduction of corticosteroid dose updated to consider data up to end of study.
- Analysis for safety disclosure updated to report events for on-treatment and post-treatment period, separately.
- Shift table for height, weight, BMI (high/normal/low) added.
- All specification for biomarker analyses moved into a separate analysis plan.
- Specification added how to identify subjects who tapered off ruxolitnitb for the analysis of cGvHD recurrence following completion of a taper of ruxolitinib.
- Analysis of systemic corticosteroid-free response rate clarified for subjects who didn’t receive steroids at baseline.

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

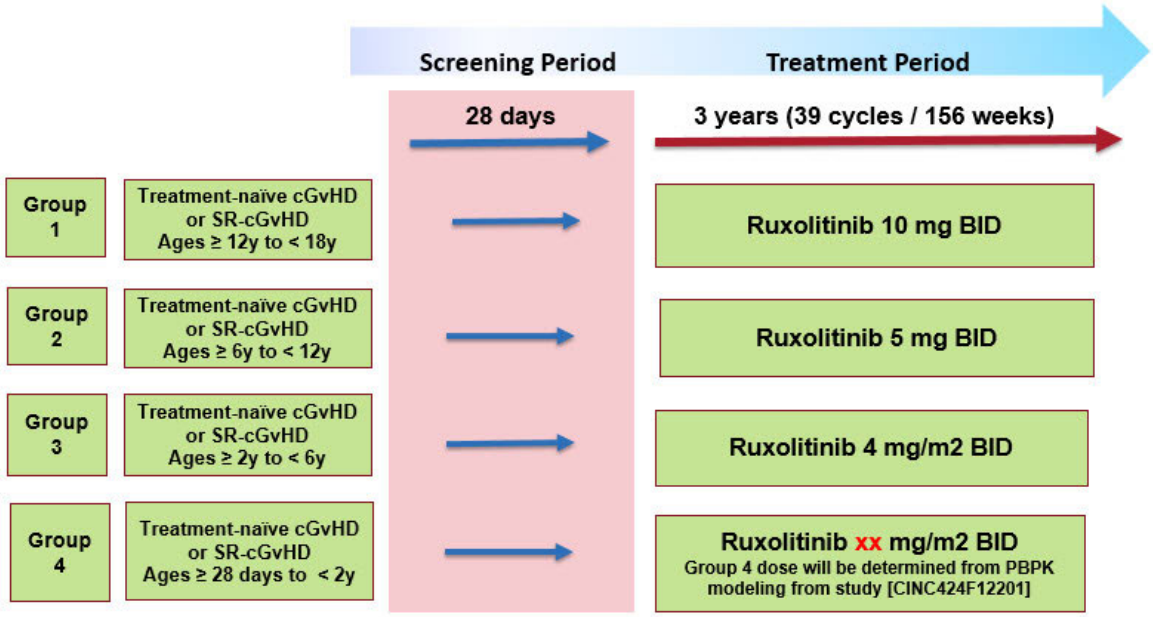
1.1 Study design

This open-label, single-arm, Phase II multi-center study will investigate the activity, pharmacokinetics and safety of ruxolitinib added to the subject's immunosuppressive regimen among infants, children, and adolescents aged ≥ 28 days to < 18 years old with either moderate

to severe treatment-naive cGvHD or SR-cGvHD. Approximately 42 subjects will be enrolled in this study. Subjects will be grouped according to their age as follows: Group 1 includes subjects ≥ 12y to < 18y, Group 2 includes subjects ≥ 6y to <12 y, Group 3 includes subjects ≥ 2y to < 6y, and Group 4 includes subjects ≥ 28days to < 2y. Subjects 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, will be subject to the availability of data in this age group from study [CINC424F12201], as well as a review of available PK, safety, and clinical activity data generated from Groups 1 to 3 in the current study, in consultation with the data monitoring committee (DMC) and a final decision by the Sponsor. At least 5 evaluable subjects per Group are needed for the primary analysis in Groups 1, 2 and 3. No minimum number of evaluable subjects are needed in Group 4.

As shown in Figure 1-1, after a screening period from Day -28 to Day -1, eligible subjects will begin the investigational treatment (ruxolitinib) on Cycle 1 Day 1 and will be treated for up to a maximum of 3 years (39 cycles/week 156) or until early discontinuation. Subjects who discontinue ruxolitinib for any reason earlier than 39 cycles will be followed every 6 months until 3 years from their first dose of ruxolitinib is reached. Subjects who tapered off ruxolitinib and all immunosuppressive therapy due to achieving a CR or PR will continue to follow the currently assigned Visit Evaluation Schedule (Table 8-1), including all safety and efficacy assessments.

Figure 1-1 Study Design



Refer to protocol Section 8.2.1 for details of the treatment period study visits. All subjects continuing to receive ruxolitinib treatment benefit following 3 years of treatment will be given the possibility to continue ruxolitinib outside the study, from another source, where permitted in accordance with local laws and regulations.

The ruxolitinib dose is based on the preliminary efficacy and safety data generated with this dose in subjects with steroid-refractory graft vs. host disease (SR-GvHD) (Zeiser et al 2015), and based on PK/safety data generated from the Phase III trials CINC424C2301 and CINC424D2301. As subjects ≥ 12 y to < 18 y are already being treated with 10 mg BID in CINC424D2301, this dose is the recommended phase II dose (RP2D), and will be used to treat all subjects in this age group.

Pediatric subjects enrolled in the ongoing pediatric aGvHD study [CINC424F12201] will provide PK data that will be used to confirm the adequacy of the doses described above and thus patients < 12 years will only be enrolled in Groups 2 to 4 in the current study once the dose is confirmed in the appropriate age group. If a different dose is confirmed in the pediatric aGvHD study [CINC424F12201], the dose will need to be adjusted accordingly in the current study. Subjects enrolled in the adult aGvHD study [CINC424C2301] and cGvHD study [CINC424D2301] may also provide PK data that will be used as additional information to confirm adequacy of this dose.

An interim CSR will be reported when all subjects have completed 1 year of treatment or discontinued earlier, at which time the primary efficacy assessment, ORR at Cycle 7 Day 1, will already be final.

1.2 Study objectives and endpoints

Objectives and related endpoints are described in Table 1-1 below.

Table 1-1 Objectives and related endpoints

Objective(s)	Endpoint(s)	Analysis	Analysis incl. in final CSR
Primary objective(s)	**Endpoint(s) for primary objective(s)**	**Primary Analysis**	
• To evaluate the activity of ruxolitinib added to standard dose corticosteroids +/- CNI in pediatric subjects with moderate or severe treatment naive-cGvHD or SR-cGvHD measured by overall response rate (ORR) at Cycle 7 Day 1 based on all subjects in the study.	• Overall response rate (ORR) at Cycle 7 Day 1, defined as the proportion of subjects demonstrating a complete response (CR) or partial response (PR) without the requirement of additional systemic therapies for an earlier progression, mixed response or non-response. The response is assessed per NIH consensus criteria (Lee et al 2015), and scoring of response will be relative to the organ stage at the start of study treatment.	• Refer to Section 2.5.1	No
Secondary objective(s)	**Endpoint(s) for secondary objective(s)**	**Secondary analysis**	
• To assess pharmacokinetics (PK) of ruxolitinib in treatment-	• Ruxolitinib concentrations by timepoint	• Refer to Section 2.9	No

Objective(s)	Endpoint(s)	Analysis	Analysis incl. in final CSR
naïve cGvHD and SR-cGvHD pediatric subjects			
• To evaluate the safety of ruxolitinib	• Safety and tolerability will be assessed by monitoring the frequency, duration and severity of Adverse Events, including occurrence of any second primary malignancies or infections, and by performing physical exams and evaluating changes in vital signs, tanner stage, chemistry and hematology results from baseline.	• Refer to Section 2.8	Yes
• To assess duration of response (DOR)	• Duration of response (DOR) is assessed for responders only. DOR is defined as the time from first response until cGvHD progression, death, or the date of addition of systemic therapies for cGvHD.	• Refer to Section 2.7.2	Yes
• To estimate ORR at end of Cycle 3	• Proportion of subjects who achieve OR (CR+PR) at Cycle 4 Day 1		No
• To assess best overall response (BOR)	• Proportion of subjects who achieved OR (CR+PR) at any time point (until Cycle 7 Day 1 or the start of additional systemic therapy for cGvHD)		No repeat of BOR at C7D1, but for BOR up to study end
• To estimate the failure free survival (FFS)	• Composite time to event endpoint incorporating the following FFS events: i) relapse or recurrence of underlying disease or death due to underlying disease, ii) non-relapse mortality, or iii) addition or initiation of another systemic therapy for cGvHD.		Yes
• To assess cumulative incidence of malignancy relapse/recurrence (MR)	• Malignancy relapse/recurrence (MR) is defined as the time from date of treatment assignment to hematologic malignancy relapse/recurrence. Calculated for subjects with underlying hematologic malignant disease.		Yes
• To assess non-relapse mortality (NRM)	• Non-relapse mortality (NRM), defined as the time from date of treatment assignment to date of death not preceded by underlying disease relapse/recurrence.		Yes

Objective(s)	Endpoint(s)	Analysis	Analysis incl. in final CSR
• To assess overall survival (OS)	• Overall survival, defined as the time from the date of treatment assignment to the date of death due to any cause.		Yes
• To assess a reduction of at least ≥ 50% in daily corticosteroid use at Cycle 7 Day 1	• Proportion of subjects with ≥ 50% reduction from baseline in daily corticosteroid dose at Cycle 7 Day 1		No repeat of the data up to C7D1, but analysis will be done considering data up to study end
• To assess a reduction to a low dose corticosteroid dose at Cycle 7 Day 1	• Proportion of subjects with reduction from baseline in daily corticosteroid dose to ≤ 0.2 mg/kg/day methylprednisolone (or equivalent dose of ≤ 0.25 mg/kg/day prednisone or prednisolone) at Cycle 7 Day 1		No repeat of the data up to C7D1, but analysis will be done considering data up to study end
• To assess graft failure	• Assess using 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, the graft failure is declared.		Yes
Exploratory objective(s)	**Endpoint(s) for exploratory objective(s)**		
• To describe the responses from the acceptability and palatability questionnaire of ruxolitinib oral pediatric formulation	• Responses from the acceptability and palatability questionnaire for the oral pediatric dose formulation after Cycle 1 (first dose), Cycle 4 and Cycle 39, as applicable.	• Refer to Section 2.8.4.4	Yes
• To assess pharmacokinetic (PK) / pharmacodynamic (PD) relationships	• Pharmacokinetic parameters derived by population PK analysis from sparse concentration-time data, e.g., AUC, Cmax, Ctrough	• Refer to Section 2.10	No
• To evaluate effect of ruxolitinib on cytokines, cGvHD biomarkers and immune cell subsets	• Assess changes in immune cell subsets, inflammatory cytokine levels, and soluble cGvHD biomarkers.	• Refer to Section 2.12	Yes (in separate report)
• To evaluate effect of ruxolitinib on markers of	• Assess changes in soluble markers for bone resorption and	• Refer to Section 2.12	Yes (in separate report)

Objective(s)	Endpoint(s)	Analysis	Analysis incl. in final CSR
bone development in pediatric subjects	formation, including but not limited to CTX, Osteopontin, and BALP		
• To explore cGvHD recurrence following completion of a taper of systemic therapy	• Time to recurrence, assessed for patients with completion of a taper of systemic therapy only.	• Refer to Section 2.13	Yes
• To evaluate systemic corticosteroid-free response rate	• Proportion of subjects who achieved OR (CR or PR) at any time point without systemic corticosteroid therapy for at least one month prior to the disease assessment.	• Refer to Section 2.13	Yes

2 Statistical methods

2.1 Data analysis general information

The interim and 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.

The analysis cutoff date for the interim analysis of study data will be established when all patients have completed 1 year of treatment or discontinued earlier. 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).

2.1.1 General definitions

Investigational treatment and study treatment

Investigational treatment, will refer to the ruxolitinib only, while ***other study treatment*** will refer to concomitant use of corticosteroids to treat treatment-naive cGvHD or SR-cGvHD.

Study treatment, will include both ***Investigational treatment*** and ***other study treatment***.

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 ruxolitinib was administered as per the DAR eCRF. The date of first administration of study treatment will also be referred as ***start of study treatment***.

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 ruxolitinib was administered as per DAR eCRF. The date of last administration of study treatment will also be referred as ***end of study treatment***.

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, biomarker, 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 for safety analyses

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 day of first dose of ruxolitinib to 30 days after date of last actual administration of assigned treatment.
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. In analyses for the safety disclosure, all AEs collected in the post-treatment period will be included and reported separately.

Subjects who taper off ruxolitinib due to achieving a CR or PR continue efficacy and safety assessments until their EoT visit which may be only after Cycle 39 unless the subject discontinues from study earlier. Safety data from these patients collected after the last study drug intake date + 30 days and before EoT visit will not be included in any safety analyses based on the on-treatment period as defined above. Nevertheless, these data will be listed and flagged in all listings (AEs, lab, vital signs). Subject who tapered off ruxolitinib are identified as they have reason "responder" for ruxolitinib permanent discontinuation on the exposure CRF.

For subjects who tapered off study treatment and later re-started ruxolitinib treatment after a cGvHD recurrence, the on-treatment period is from date of first dose of ruxolitinib to very last dose + 30 days including the period without treatment, which are collected in the database as dose interruption. Safety data (e.g. AEs, labs) collected during this period without treatment will be included in the summary tables as treatment emergent events.

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 summary on cGvHD assessment, biomarkers, and safety (Table 2-1) by visit. The end of treatment assessment will be mapped into the time points as needed.

Table 2-1 Time windows for cGvHD assessment and safety assessment (lab, vital sign, etc.)

Time Window	Planned Visit Timing	Time Window Definition
On treatment		
Baseline	On or before Study Day 1	≤ Study Day 1
Cycle 1 Day 8	Study Day 8	Study Days 5 – 11
Cycle 1 Day 15	Study Day 15	Study Days 12 – 18
Cycle 1 Day 22	Study Day 22	Study Days 19 – 25
Cycle 2 Day 1	Study Day 29	Study Days 26 – 39
Cycle 3 Day 1	Study Day 57	Study Days 47 – 67
Cycle 4 Day 1	Study Day 85	Study Days 75 – 95
Cycle 5 Day 1	Study Day 113	Study Days 103 – 123
Cycle 6 Day 1	Study Day 141	Study Days 131 – 151
Cycle 7 Day 1	Study Day 169	Study Days 159 – 179
Cycle 9 Day 1	Study Day 225	Study Days 211 – 239
Cycle 12 Day 1	Study Day 309	Study Days 295 – 323
every 12 weeks	+ 84 day	+84 days to lower and upper bound
Cycle 39 Day 1	Study Day 1065	Study Days 1051 – 1179
Safety follow-up	30 days after last dose	Last dose date + 30
Study Day 1 = start date of study treatment; EOT assessments are mapped to the time points as needed.		

Last contact date

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

Table 2-2 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 cGvHD treatment since discontinuation from study treatment	Non-missing medication/procedure term.
Start/End* dates from drug administration record/concomitant medication record	Non-missing dose. Doses of 0 are allowed.
Date of discontinuation from end of treatment pages	No condition.
- cGvHD assessment date	Non-missing assessment.

Source data	Conditions
- any specific efficacy assessment date if available (e.g. graft failure assessment, hematologic disease relapse/progression assessment)	
Laboratory/PK collection dates/Biomarker collection dates	Sample collection marked as 'done'.
Vital signs date	At least one non-missing parameter value.
Assessment date of development information (e.g. tanner staging)	Evaluation is marked as 'done'.
Collection date of acceptability and palatability questionnaire	Assessment performed is "Yes".
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 cut-off date. The cut-off date will not be used for last contact date, unless the patient was seen or contacted on that date. No date post cut-off date will be used. Completely imputed dates (e.g. the analysis cut-off date programmatically imputed to replace the missing end date of a dose administration record) will not be used to derive the last contact date. Partial date imputation is allowed for event (death)/censoring if coming from 'Survival information' eCRF.

The last contact date will be used for censoring of patients in the analysis of overall survival, failure free survival, non-relapse mortality, and malignancy relapse/recurrence.

2.2 Analysis sets

If the starting dose is different from the assigned dose level due to co-administration of ruxolitinib with strong CYP3A4/CYP2C9 inhibitors, these subjects will be included under the assigned dose level and considered that they have received the full assigned dose. This applies to Full Analysis Set, Safety Set, and Pharmacokinetic analysis set as described below.

If the treatment formulation (tablet/liquid) changed during treatment, the subject will be considered under the formulation she/he started treatment. Nevertheless, a change in formulation will be flagged and may be considered for PK analyses.

The **Full Analysis Set** (FAS) comprises all subjects to whom study treatment has been assigned and who received at least one dose of study treatment.

The **Safety Set** includes all subjects who received at least one dose of study treatment. Subjects will be analyzed according to the study treatment received, where treatment received is defined as the assigned dose level of ruxolitinib if the subject 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 subjects who provide at least one evaluable PK concentration. For a concentration to be evaluable, subjects are required to:

- Take the dose of ruxolitinib prior to PK sample

- For post-dose samples on Cycle 1 Day 1: do not vomit within 2 hours after the dosing of ruxolitinib
- For pre-dose samples: have the sample collected before the next dose administration of ruxolitinib; do not vomit within 2 hours after the previous dosing of ruxolitinib prior to sampling.

The **Listing only Set** contains only those patients who are not eligible for efficacy analyses due to PD "Patient enrolled and treated beyond local regulatory requirements". These patients will also not be included in safety summary analyses, but all safety data will be provided in listings, separately from the listings of all other patients.

Patient Classification:

Patients may be excluded from the analysis populations defined above based on the protocol deviations entered in the database and/or on specific patient classification rules defined in Table 2-3.

Patients with protocol deviation OTH12 "Patient enrolled and treated beyond local regulatory requirements" will only be included in the "Listing only Set" and not in any other analysis set.

Table 2-3 Patient classification based on protocol deviations and non-PD criteria

Analysis set	Protocol deviations leading to exclusion	Non protocol deviation leading to exclusion
FAS	INCL01A, OTH12	Not applicable
Safety Set	INCL01A, OTH12	No dose of study treatment
PK Analysis Set	INCL01A, OTH12, TRT01/ TRT02	No dose of ruxolitinib, No evaluable PK concentration
Listing only Set	OTH12 (include only patients with this PD)	Not applicable

INCL01A - Written informed consent was not obtained
OTH12 - Patient enrolled and treated beyond local regulatory requirements
TRT01 - Starting dose of study drug (ruxolitinib) not per protocol-defined age-based schedule (Group 1 and 2).
TRT02 - Starting dose of study drug (ruxolitinib) not per protocol-defined age-based schedule (Group 3 and 4).

Withdrawal of Informed Consent

Any data collected in the clinical database before a patient signed informed consent or after a patient withdraws informed consent from all further participation in the trial, will not be included in the analysis. The informed consent data and the date on which a patient withdraws full consent is recorded in the eCRF.

2.2.1 Subgroup of interest

Efficacy

Subgroup analyses for efficacy endpoints will not be repreated for the final CSR.

The primary efficacy endpoint will be summarized by the following subgroups:

- Age group (please refer to Section 1.1 for the definition)
- Age group and formulation (tablet vs liquid)
- Treatment Naive cGvHD vs SR-cGvHD
- Region EU+ROW except Turkey / Turkey / Asia excl. Japan / Japan
  Note: ROW includes US, Canada, Australia and other countries (e.g. Russia).
- Chronic GvHD severity at baseline, moderate vs. severe (as reported on the staging CRF and not from the GvHD Disease History page).

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 (please refer to Section 1.1 for the definition)
- Treatment-naive cGvHD vs. SR-cGvHD
- Region EU+ROW except Turkey / Turkey / Asia excl. Japan / Japan
- Chronic GvHD severity (moderate vs. severe)

For the final CSR safety subgroup analyses will be repeated for age groups and treatment-naïve cGvHD vs. SR-cGvHD only.

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 subgroup:

- Overview of adverse events
- Serious AEs, irrespective of causality, by primary system organ class and preferred term
- On-treatment deaths, by primary system organ class and preferred term

These additional analyses will be provided only by age groups and by treatment-naive vs SR cGvHD:

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

Japanese patients

Analyses on the subgroup of Japanese patients will not be repeated for the final CSR.

Subgroup analyses may also be performed for the patients treated in Japan. No selection will be done on the basis of ethnicity, the purpose being to evaluate the population of patients living in Japan, not a specific ethnic set of patients. The analysis will be done the same way as in region subgroup specified above.

2.3 Patient disposition, demographics and other baseline characteristics

Analyses on demographics, eligibility criteria, baseline disease characteristics, and other baseline data like medical history and prior treatments will not be repeated for the final CSR. The only analyses that need to be updated for the final CSR Section 14.1 are the tables on disposition and protocol deviations, as well as analysis sets.

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 and for all patients, and listings will be reported by Age group to assess baseline comparability. 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. 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 and extent of disease in underlying disease, stem cell transplant and chronic GvHD.

For underlying disease, the analysis will include the following: primary diagnosis category and subcategory, details of primary diagnosis, time since diagnosis of underlying disease, CIBMTR risk assessment.

For transplant related disease history, the analysis will include the following: conditioning regimen type, total HCT-specific comorbidity index score, time since transplant, time from diagnosis of underlying disease to transplant, stem cell type, donor information including age, gender, HLA typing method, HLA match score, source of grafts (HLA matched related donor, unrelated UCB (umbilical cord blood), related haploidentical, HLA matched/mismatched URD (unrelated donor)), T-cell depleted (Y/N), total nucleated cell dose.

For cGvHD disease history, the analysis will include the following: prior diagnosis of aGvHD, time since initial diagnosis of cGvHD, overall severity of initial cGvHD, cGvHD severity at baseline (study entry), time since steroid refractory cGvHD, cGvHD organ involvement, steroid dose at baseline (study entry).

Medical history

Medical history and ongoing conditions, including underlying disease conditions and symptoms entered on eCRF will be summarized and listed by Age group. The summaries will be presented by primary system organ class (SOC), preferred term (PT) and Age 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.

Other

All data collected at baseline including Tanner Staging will be listed.

2.3.1 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 phases as well as the reason for discontinuation, and the long term follow-up 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 summarized and listed separately.

Analysis sets

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

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

2.4.1 Study treatment / compliance

Duration of exposure 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/formulation within the Age group

Duration of follow-up

The duration of follow-up time is defined as (date of last contact) – (date of first administration of investigational treatment) + 1.

Duration of exposure to investigational treatment

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

If any subjects tapered off ruxolitinib and re-started ruxolitinib treatment due to cGvHD recurrence, two different definitions of duration of exposure will be distinguished:

- **Duration of exposure including taper off periods** is defined as the time from first to very last date of ruxolitinib treatment, which includes the interveningperiod without treatment.
- **Duration of exposure excluding taper off periods** is defined as above but deduct the interveningperiod without treatment for subjects who taper off and restart after cGvHD recurrence. These periods are captured as dose interruption with reason “dose tapering following response”.

Summary of duration of exposure of investigational 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 to investigational treatment in patient-years

The duration of exposure to investigational treatment in patient-years is a total of the duration of exposure to investigational treatment in years from all the patients in the Age group.

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.

Cumulative dose

The **planned cumulative dose** for ruxolitinib refers to the total planned dose per protocol 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 over the dose by 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).

In case the study includes subjects who taper off ruxolitinib and later re-start ruxolitinib treatment due to cGvHD recurrence, the duration of exposure excluding "taper off periods" will be used as denominator for dose intensity.

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. Separate summaries will be generated for 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). 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 changed when the actual total daily dose is lower than the previous actual total daily dose A 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 as 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.

2.4.2 Prior, concomitant and post therapies

Analyses on prior treatments or prophylaxis will not be repeated for the final CSR.

2.4.2.1 Prior prophylaxis

The number and percentage of patients who received any prophylaxis (e.g. CNI) 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.

2.4.2.2 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 (see Section 2.4.2.3).

2.4.2.3 Systemic corticosteroid during study treatment

As per protocol the steroid doses of methylprednisolone will be converted to prednisone equivalents by multiplying the methylprednisolone dose by 1.25. Prednisone doses for each subject are converted to mg/kg/day.

The duration of exposure will be summarized for systemic corticosteroid by Age group and disease history of treatment naive-cGvHD or SR-cGvHD at baseline. It will 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 the Cycle 7 Day 1 visit. Graphical display (box plots) will also be generated. These analyses will be based on Safety Set.

For the final CSR these analyses will be done considering data up to end of study/LPLV date.

2.4.2.4 Additional systemic cGvHD therapy

New additional systemic cGvHD 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.

Initiation of systemic CNI at the time of or after start of study treatment despite absence of lack of efficacy will be counted as new additional systemic treatment. Note: Although the protocol does not explicitly prohibit initiation of systemic CNI at time of randomization or initiation of study treatment potential impact of CNI on efficacy outcome in responders cannot be excluded.

See Appendix Section 5.4.3 for more detailed specifications.

2.4.2.5 Concomitant medications

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

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

In addition, a subset of concomitant medications i.e. hemopoietic cytokines and transfusions (red blood cells and platelets) will be grouped and summarized by Age group.

2.4.2.5.1 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.

2.5 Analysis of the primary objective

Analyses of the primary endpoint and related supportive analyses will not be repeated for the final CSR, if data has not changed compared to interim database lock.

The primary objective of the study is to evaluate the activity of ruxolitinib added to standard dose corticosteroids +/- CNI in pediatric subjects with moderate or severe treatment naive-cGvHD or SR-cGvHD measured by **Overall Response Rate (ORR)** at Cycle 7 Day 1 based on all subjects in the study. The analysis will be performed based on FAS.

2.5.1 Primary endpoint

The primary efficacy variable of the study is **overall response rate (ORR)** at Cycle 7 Day 1, defined as the proportion of subjects demonstrating a complete response (CR) or partial response (PR), according to the NIH Consensus Criteria (Lee et al 2015) as reported in the CRF. Note that response is relative to the assessment of cGvHD at baseline.

- **Complete response** is defined as complete resolution of all signs and symptoms of cGvHD in all evaluable organs without initiation of any new systemic therapy.
- **Partial response** is defined as an improvement in at least one organ (e.g. improvement of 1 or more points on a 4 to 7 point scale, or an improvement of 2 or more points on a 10 to 12 point scale) without progression in other organs or sites, or initiation/ addition of new systemic therapies.
- **Lack of response** is defined as unchanged, mixed response, or progression. (Please refer to Table 8-3 and Table 16-3 in the protocol for the definitions of "unchanged", "mixed response" and "progression".)

cGvHD Flare and **cGvHD Recurrence** are not considered as a treatment failure, unless they require a change or addition of another systemic treatment defined as below.

- **cGvHD Flare** is defined as any increase in symptoms or therapy for cGvHD after an initial response (CR or PR). A cGvHD flare is not considered a treatment failure unless a change or addition of another systemic treatment, including CNIs, occurs.
- **cGvHD Recurrence** is defined as the return of cGvHD disease after tapering off study treatment due to response. Following completion of a taper of systemic therapy, if worsening of cGvHD symptoms occur, the patient is allowed to resume treatment for cGvHD as per local institutional practice. For the statistical analyses, re-start of treatment for cGvHD is handled in the same way as addition or initiation of new systemic treatment.

2.5.2 Statistical hypothesis, model, and method of analysis

The response rates for ORR at Cycle 7 Day 1, will be estimated on the Full Analysis Set (FAS) for all the patients, when all subjects have completed 1 year of treatment or discontinued earlier. Confidence intervals of 90% will be calculated based on the exact method (Clopper and Pearson 1934) for binominal distribution. Summary statistics (frequencies and percentages) will be provided. It is not planned to test any specific hypotheses related to the efficacy endpoint(s), but to provide estimates of efficacy endpoints in the pediatric population.

2.5.3 Handling of missing values/censoring/discontinuations

Patients with missing assessments that prevent the evaluation of the primary endpoint will be considered non-responders. This includes missing overall cGvHD response assessments at baseline and/or Cycle 7 Day 1. The time window for the Cycle 7 Day 1 visit is defined in Table 2-1 above.

No data imputation will be applied. Patients who discontinue study treatment should return for the regular assessments indicated in Protocol Section 9.1. Addition or initiation of a new systemic therapy before Cycle 7 Day 1 will be considered a treatment failure, and patients will be counted as non-responder in the primary analysis.

2.5.4 Supportive analyses

Supportive analysis will include:

- A detailed description of response rates (CR, PR. Unchanged, mixed response and progression) at Cycle 7 Day 1
- A detailed description of the organ-specific response for all organs at Cycle 7 Day 1
- If at least 18 subjects are enrolled with either first line vs. SR-cGvHD respectively, the ORR will be explored for treatment naive and SR subgroup subjects.
  This and further subgroups are described in Section 2.2.1.

The supportive analyses will be performed on the FAS.

In addition, a sensitivity analysis will be done, where subjects with protocol deviations that may affect the primary endpoint are excluded from the analysis. These are following:

- Prohibited concomitant medication administered during treatment period (COMD01)
- Treatment naive or steroid refractory cGvHD corticosteroid treatment criteria not met (INCL04B)
- Missing individual organ assessment at C1D1 (unless not possible) (OTH02)
- Missing individual organ assessment at C4D1 or C7D1 (unless not possible) (OTH03)
- Starting dose of study drug (ruxolitinib) not per protocol-defined age-based schedule (Group 1 and 2) (TRT01)
- Starting dose of study drug (ruxolitinib) not per protocol-defined age-based schedule (Group 3 and 4) (TRT02)
- In treatment naive cGvHD subjects, in addition to ruxolitinib, methylprednisolone (or equivalent prednisone) treatment was initially not given as part of study treatment (TRT03)
- Patient did not take the ruxolitinib assigned dose (TRT08)
- The ruxolitinib dose assigned on C1D1 (based on age) was changed before the subject completed C7D1 assessments (excluding dose adjustments for safety) (TRT09)

- Corticosteroid taper was initiated earlier than allowed as per protocol (excluding dose adjustments for safety) (TRT13)
- Subject had a dose interruption of >21 days from ruxolitinib but was not permanently discontinued from ruxolitinib (WITH01)
- Individual organ response assessment / overall response assessment not as per protocol response criteria (OTH14)
- Patient was not compliant with study drug (Ruxolitinib); More than 7 days of consecutive insufficient treatment or overdosing or cumulated insufficient treatment over a period of 24 weeks is observed (TRT16)

Supportive analysis 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.

2.6 Analysis of the key secondary objective

No key secondary objective has been planned.

2.7 Analysis of secondary efficacy objective(s)

The secondary objectives in this study include the assessment of: failure free survival (FFS), ORR at Cycle 4 Day 1 (end of Cycle 3), duration of response (DOR), overall survival (OS), non-relapse mortality (NRM), best overall response (BOR), cumulative incidence of malignancy relapse/recurrence (MR), reduction and successful tapering of corticosteroid treatment (Cycle 7 Day 1) and graft failure.

2.7.1 Secondary endpoints

Failure Free Survival

FFS is a composite time to event endpoint incorporating the following FFS events: i) relapse or recurrence of underlying disease or death due of underlying disease, ii.) non-relapse mortality, or iii.) addition or initiation of another systemic therapy for cGvHD. If a patient did not experience any of these events, FFS will be censored at the latest contact data (on or before the cut-off date).

Best overall response (BOR)

Analyses of BOR at Cycle 7 Day 1 will not be repeated, whereas BOR up to end of study will be analyzed for the final CSR.

Best overall response (BOR) is defined as proportion of subjects who achieved overall response (CR or PR) at any time point up to and including Cycle 7 Day 1 and before the start of additional systemic therapy for cGvHD.

An additional analysis will include all response assessments up to End of Treatment of the subjects or data cut-off date and before the start of additional systemic therapy for cGvHD. Note,

for subjects who tapered off, all response assessments, including those collected after ruxolitinib discontinuation, will be considered for the analysis.

BOR will be calculated based on the FAS using local investigators' overall response assessments.

ORR at Cycle 4 Day 1 (end of Cycle 3)

Analysis of ORR at Cycle 4 Day 1 will not be repeated for the final CSR.

ORR at Cycle 4 Day 1, defined as the proportion of patients with complete response (CR) or partial response (PR), according to the NIH Consensus Criteria (Lee 2015). ORR will be calculated based on the FAS using local investigators' overall response assessed at the Cycle 4 Day 1 visit and taking into account initiation or addition of new systemic therapy before this time point.

Duration of Response

Duration of response (DOR) is assessed only for responders up to Cycle 7 Day 1 (i.e. all subjects with BOR = CR or PR considering response assessments up to Cycle 7 Day 1). DOR is defined as the time from first response until cGvHD progression, death, or the date of additional systemic therapies for cGvHD. Subjects without event will be censored at the date of their last response assessment prior to or at the analysis cut-off date if no events occurred on or before 12 weeks (84 days) after the last GvHD assessment.

An additional analysis of duration of response for responders up to data cut-off date will be done on all subjects with BOR considering all response assessments up to End of Treatment of the subjects or data cut-off date. Note, for subjects who tapered off, all response assessments up to End of Treatment even those collected after ruxolitinib discontinuation will be considered for the analysis.

Overall survival (OS)

Overall survival is defined as the time from date of treatment assignment 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).

Non-Relapse Mortality (NRM)

Non-relapse mortality (NRM), defined as the time from date of treatment assignment to date of death not preceded by underlying malignant or non-malignant disease relapse/recurrence. Underlying disease relapse/recurrence is considered as competing events. If a patient is not known to have died or experienced the competing event, then NRM will be censored at the latest date the patient was known to be event-free (last contact date on or before the cut-off date).

Incidence of Malignancy Relapse/Recurrence (MR)

Incidence of Malignancy Relapse/Recurrence (MR), defined as the time from date of treatment assignment to date of relapse/recurrence of the underlying hematologic malignancy disease. NRM is considered as competing event. If a patient is not known to have a hematologic malignancy relapse/recurrence or died without relapse/recurrence, then incidence of MR will

be censored at the latest date the patient was known to be event-free (last contact date on or before the cut-off date).

Reduction of daily corticosteroid dose

Reduction of daily corticosteroid dose, is defined as the proportion of patients with ≥50% reduction from baseline in daily corticosteroid dose due to disease improvement at Cycle 7 Day 1.

For the final CSR the analyses will be done considering data up to end of study.

Reduction to low dose corticosteroids

Reduction to low dose corticosteroids, is defined as the proportion of subjects with reduction from baseline in daily corticosteroid dose to methylprednisolone-equivalent steroid dose of ≤0.2mg/kg/day (or equivalent dose of ≤0.25mg/kg/day prednisone or prednisolone) due to disease improvement at Cycle 7 Day 1.

For the final CSR the analyses will be done considering data up to end of study.

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.

2.7.2 Statistical hypothesis, model, and method of analysis

Failure Free Survival

The FFS distribution will be estimated using the Kaplan-Meier method, and the Kaplan-Meier curves, medians, 3, 6, 12, 18, 24 and 36 months FFS estimates and 95% confidence intervals (Brookmeyer and Crowley J. 1982) will be presented based on all subjects in the FAS.

The cumulative incidence curve of each of the three FFS components (considering the other two components as a competing risk) as well as estimates at 3, 6, 12, 18, 24 and 36 months will also be presented with 95% confidence intervals.

Best overall response (BOR)

BOR up to Cycle 7 Day 1 as well as BOR up to data cut-off date and its 90% confidence interval will be presented based on all subjects in the FAS.

ORR at Cycle 4 Day 1 (end of Cycle 3)

ORR (CR+PR) at Cycle 4 Day 1 and its 90% confidence interval will be presented based on all subjects in the FAS.

Duration of Response

Duration of response will be calculated for all patients with BOR = CR or PR up to Cycle 7 Day 1 as well as for all patients with BOR up to data cut-off date = CR or PR. Kaplan-Meier method and the Kaplan-Meier curves, medians, 3, 6, 12, 18, 24 and 36 months survival

probabilities with 95% confidence intervals will be presented based on these responder subjects in the FAS.

Overall survival (OS)

The OS distribution will be estimated using the Kaplan-Meier method, and the Kaplan-Meier curves, medians, 3, 6, 12, 24 and 36 month survival probabilities and 95% confidence intervals (Brookmeyer and Crowley J. 1982) will be presented based on all subjects in the FAS.

Non-Relapse Mortality (NRM)

Cumulative incidence of NRM and derived probabilities at Months 1, 2, 6, 12, 18, 24 and 36 with 95% confidence intervals will be estimated based on all subjects in the FAS, considering underlying disease relapse/recurrence as competing events (Gooley et al. 1999).

Incidence of Malignancy Relapse/Recurrence (MR)

The cumulative incidence curve for MR and estimates at 3, 6, 12, 18, 24 and 36 months with 95% confidence intervals will be presented for subjects with underlying hematologic malignant disease, accounting for NRM as the competing risk. In addition, the proportion of subjects who had hematologic malignancy relapse/recurrence and its 95% confidence interval at 3, 6, 12, 18, 24 and 36 months will be presented for subjects with underlying hematologic malignant disease.

Reduction of daily corticosteroid dose

The proportion of subjects with ≥ 50% reduction from baseline in daily corticosteroid dose due to disease improvement at Cycle 7 Day 1, and its 95% confidence interval will be presented based on all subjects in the FAS. For the final CSR the analyses will be done considering data up to end of study.

Reduction to low dose corticosteroids

The proportion of subjects with reduction from baseline in daily corticosteroid dose to methylprednisolone-equivalent steroid dose of ≤0.2mg/kg/day (or equivalent dose of ≤0.25mg/kg/day prednisone or prednisolone) due to disease improvement at Cycle 7 Day 1, and its 95% confidence interval will be presented based on all subjects in the FAS. For the final CSR the analyses will be done considering data up to end of study.

Graft Failure

The percentage of graft failure with 95% confidence intervals at 3, 6, 12, 18, 24 and 36 months will be presented based on all subjects in the FAS. In case of very few patients with graft failure (e.g. less than 5 patients), case descriptions using listings will be provided instead.

2.7.3 Handling of missing values/censoring/discontinuations

Patients with missing assessments that prevent the evaluation of the ORR at Cycle 4 Day 1 and BOR will be considered non-responders. This includes missing overall cGvHD response assessments at baseline and/or Cycle 4 Day 1. The time window for the Cycle 4 Day 1 visit is defined in Table 2-1 above.

2.8 Safety analyses

For all safety analyses, the safety set will be used. All listings and tables will be presented by Age group. 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, see definitions in Section 2.1.1.

2.8.1 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 the ruxolitinib arm.

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.

In addition, summary table and listing will be provided for COVID-19 related Aes.

2.8.1.1 Adverse events of clinical trial safety disclosure

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

greater than 5% and on-treatment/post-treatment deaths, serious adverse events and SAEs suspected to be related to study treatment will be provided by system organ class and preferred term on the safety set population. Events that occurred during on-treatment period and those in post-treatment period (see definitions in Section 2.1.1) will be reported separately.

If for the 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.

2.8.1.2 Adverse events of special interest / grouping of AEs

An adverse event of special interest is a grouping of adverse events that are of scientific and medical concern specific to compound ruxolitinib. These groupings are defined using MedDRA terms, SMQs (standardized MedDRA queries), HGLTs (high level group terms), HLT (high level terms) and/or PTs (preferred terms). Customized SMQs (CMQs) may also be used. A CMQ is a customized group of search terms which defines a medical concept for which there is no official SMQ available or the available SMQ does not completely fit the need. It may include a combination of single terms and/or an existing SMQ, narrow or broad.

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.

2.8.2 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 occurred outside of on-treatment period will be flagged. A separate listing of deaths prior to starting treatment will be provided for all screened patients.

2.8.3 Laboratory data

On analyzing laboratory, data from all sources (central and local laboratories) will be combined. The summaries will include all assessments available for the lab parameter collected no later than end of the on-treatment periods.

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 Age 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

In addition to the above mentioned tables and listings, other exploratory analyses, for example figures plotting time course of raw or change in laboratory tests over time or box plots may be produced.

2.8.4 Other safety data

2.8.4.1 ECG and cardiac imaging data

Not applicable.

2.8.4.2 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 m^2) and body mass index (BMI in kg/m^2) are derived using the following formulas:

BSA: $(\text{Height (cm)} \times \text{Weight (kg)} / 3600)^{0.5}$

BMI: $\text{Weight (kg)} / \text{Height}^2 (m^2)$

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 of the subject. Since height might not be measured at each visit, the last available height 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 depending on the patient's age are provided in Table 2-4 below. Criteria should be applied according to the age at vital sign assessment rather than age at baseline.

Table 2-4 Criteria for notably abnormal vital signs

	Clinically notable criteria			
Vital sign (unit)	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	12-<18 months	< 103
	18-<24 months	> 135	18-<24 months	< 98
	2-<3 years	> 128	2-<3 years	< 92
	3-<4 years	> 123	3-<4 years	< 86
	4-<6 years	> 117	4-<6 years	< 81
	6-<8 years	> 111	6-<8 years	< 74
	8-<12 years	> 103	8-<12 years	< 67
	12-<15 years	> 96	12-<15 years	< 62
	≥ 15 years	> 92	≥ 15 years	< 58
Weight[3]	increase from baseline of ≥ 2 BMI-for-age/gender percentile categories		decrease from baseline of ≥ 2 BMI-for-age/gender percentile categories	
Height[4]	increase from baseline of ≥ 2 height-for-age/gender percentile categories		decrease from baseline of ≥ 2 height-for-age/gender percentile categories	

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; 377: 1011-18.

[3] BMI-for-age/gender percentiles categories (P3, P5, P10, P25, P50, P75, P85, P90, P95, P97) are obtained from the CDC Growth Charts.

[4] Height-for-age/gender percentiles categories (P3, P5, P10, P25, P50, P75, P90, P95, P97) are obtained from the CDC Growth Charts.

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

Graphical presentations of vital signs (e.g., body weight, body mass index, systolic blood pressure, diastolic blood pressure) over time for change from baseline may be produced via boxplots based on time windows and corresponding tables displaying the statistics used for the box plots by the selected time points.

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.

2.8.4.3 Cumulative incidence of severity grade ≥ 3 infection

Time to first occurrence of infection is defined as time from start of study treatment to the date of first occurrence of an infection of severity grade ≥ 3, i.e. time in days is calculated as (start date of first occurrence of infection with CTC grade ≥ 3) – (start of study treatment) +1.

The cumulative incidence analysis will consider death or discontinuation from study treatment (due to any reason) without prior infection as competing risks.

In the absence of an event during the on-treatment period, the censoring date applied will be **the earliest** of the following dates:

- data cut-off date
- withdrawal of informed consent date

whereas dates for competing risks are:

- death date
- end date of on-treatment period

Cumulative incidence curves for time to grade ≥ 3 infections as well as summary tables for the number of subjects with event/competing risk/censored and cumulative incidence 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.

2.8.4.4 Acceptability and Palatability assessment

All acceptability and palatability assessment data of all subjects using liquid formulation will be summarized and listed by Age group in the final study report.

2.8.4.5 Development (Growth and sexual maturation)

Growth

Graphical displays of percentile of height and weight over time will be provided for signal detection of impact on growth development. In case of a signal, further analyses may be provided.

In addition, a shift tables for height, weight, and BMI will be provided using following low/normal/high/(low and high) classification to compare baseline to the worst on-treatment value. Low: below the 5th percentile, Normal: between 5th and 95th percentiles, High: above the 95th percentile. Percentiles are based on CDC Growth Charts, derivation is described below.

Height and BMI will be summarized at 6-month intervals, using the standard deviation scores (SDS, also called z-score), velocity and velocity SDS as appropriate. The relevant height and weight values for each 6-month period are defined using time windows, according to Table 2-1. The z-scores will allow identification of potential outliers.

SDS will be calculated using the current formulae provided by the CDC as follows:

$$\text{Calculate } z_{ind} = \frac{\left(\frac{X}{M}\right)^{L} - 1}{LS}$$

where:

- X is height in centimeters or BMI in kilograms/m^2,
- L, M and S are height or BMI-, sex- and age-specific reference values from the CDC Growth Charts.

The CDC Height-for-age and BMI-for-age L, M and S reference values for males and females are available under https://www.cdc.gov/growthcharts. The age category immediately above the patient's exact age should be used. SDS is actually a Z score that measures the distance from the population mean in units of standard deviations. That is, SDS < 0 refers to values lower than the population mean, and for example SDS ≤ -1.645 refers to values in the lowest 5%. The percentile can be derived from the z-score by a normal distribution.

Height velocity is defined as follows:

Height velocity (cm/6-months) = (height in time window *k* – height in time window *k-1*) ÷ ([assessment date in time window *k* – assessment date in time window *k-1*] ÷ [365.25/2]),

and similarly for weight velocity.

Velocity SDS is calculated as (velocity – mean) / SD, where mean and SD are obtained as the height-, weight-, sex- and age-specific values (Baumgartner et al 1986), where the age category immediately above the patient's exact age (at the assessment date in time window *k*) should be used. Velocity SDS will only be calculated for time window *k* if data also exists for time window *k-1*, since calculating across multiple units of 6 months requires more than one reference value to be taken into account.

All height/length, weight, BMI, standardized scores (percentile and SDS) will be listed.

Sexual maturation

Sexual maturation will be monitored by the Tanner stages. Tanner stage includes two components for boys, namely genitalia stage and pubic hair stage, and two components for girls: breast development stage and pubic hair stage. The overall Tanner stage is derived by the maximum of the two stages (Tanner stage of pubic hair and Tanner stage of genitalia/breast development) from the same assessment.

Tanner stage will only be analyzed among pre-pubescent patients. Therefore, analysis will be based on patients from the Safety Set who were classified as Tanner Stage 4 or lower at the latest assessment prior to the start of study treatment.

It is expected that data will become available during the trial on a proportion of patients as they go through puberty attaining higher levels of the Tanner Stage. The age at which Tanner Stages 2-5 are achieved, and the age at menarche (females), will be displayed graphically and summarized descriptively (if sufficient number of patients are available).

All available sexual maturation data will be included in the figures and listed.

2.9 Pharmacokinetic endpoints

Analyses of PK data will not be repeated for the final CSR, since no PK samples were collected since interim CSR analyses..

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

Pharmacokinetic variables:

Ruxolitinib concentration sparse profiles will be summarized by time point. Pharmacokinetic parameters, e.g. AUC0-12h and Cmax and Ctrough, may be determined using Population PK method(s) for ruxolitinib, and methods detailed in an independent analysis plan.

Statistical methods for pharmacokinetic analyses

Ruxolitinib concentrations data will be listed by Age group and formulation. Descriptive summary statistics will be provided by Age group/formulation at each scheduled time point. Summary statistics will include n (number of subjects with non-missing values), mean (arithmetic and geometric), standard deviation (SD), coefficient of variation (CV%) (arithmetic and geometric), median, minimum and maximum. Individual profiles with median by Age group/formulation 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 Age group/formulation. Descriptive statistics (n, arithmetic mean, SD, CV% for mean, geometric mean, geometric CV%, median, minimum and maximum) will be provided for all PK parameters by Age group/formulation.

Population PK approach Concentration results from the sparse sampling will be analyzed using nonlinear mixed effects modeling (population PK) or other model-based approaches, as appropriate, and key pharmacokinetic parameters may be derived for the study population from sparse concentration-time data e.g. AUC, Cmax, Ctrough. Details of the analysis method will be developed in an independent 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 Food and Drug Administration (FDA) guidance will be followed (Guidance for Industry: Population Pharmacokinetics; 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.

2.10 PD and PK/PD analyses

Analyses of PK data will not be repeated for the final CSR.

Exposure-Response analysis

A detailed description of exposure-response analysis may be developed in an independent analysis plan if the data permits, and the analysis will be documented in a separate report. Briefly, the objectives are to:

- Characterize the exposure-efficacy relationship of ruxolitinib in terms of concentration-effect and dose-effect (effect: overall response rate at Cycle 7 Day 1; Overall survival at Cycle 4 Day 1, Cycle 7 Day 1 and 12 months). Exposure metrics will be described in further detail in the analysis plan.
- Characterize the exposure-safety relationship of ruxolitinib in terms of concentration-AEs and/or dose-AEs (AEs: frequency of AEs, severity of AEs, AEs of special interest, time to onset of AEs). Exposure metrics will be described in further details in the analysis plan
- Explore exposure-biomarker relationships as appropriate – immune cell subsets, inflammatory cytokine levels and soluble cGvHD biomarkers.

2.11 Patient-reported outcomes

Not applicable.

2.12 Biomarkers

Specified in a separate analysis plan and reported in a separate biomarker report.

2.13 Other Exploratory analyses

Chronic GvHD recurrence following completion of a taper of ruxolitinib

This analysis is performed only for patients who taper off ruxolitinib due to achieving a CR or PR. These subjects will be identified as they have reason “responder” for ruxolitinib permanent discontinuation on the exposure CRF. It’s planned to do this analysis at time of final CSR and only if sufficient number of subjects with cGvHD recurrence after tapering off occurred.

Time to recurrence is defined as the time from completion of a taper of systemic therapy until cGvHD recurrence. Patients without event will be censored at the last contact date. Kaplan-

Meier method and the Kaplan-Meier curves as well as survival probabilities at appropriate time points will be presented with 95% confidence intervals (Brookmeyer and Crowley 1982).

In case of single subjects (less than 5 subjects) with cGvHD recurrence after tapering off, data will be reported in listing only.

Systemic corticosteroid-free response rate

Systemic corticosteroid-free response rate is defined as the proportion of subjects taking no systemic corticosteroid therapy since at least 1 month (i.e. time between date of last corticosteroid drug intake and response assessment data is ≥ 30 days) and demonstrating a complete response (CR) or partial response (PR), at any time post-baseline. Subjects who didn't receive corticosteroids at baseline and still not taking corticosteroids at time of response will be considered as corticosteroid-free responder. Initiation of any new systemic therapy after cGvHD flare or recurrence, regardless when, means treatment failure and the subject is considered as non-responder. The endpoint is similar to the primary endpoint ORR, but considers all subjects who could not taper off systemic corticosteroids as non-responder regardless if they are in CR or PR, and considers all response assessments up to End of Treatment of the subject or data cut-off date.

Systemic corticosteroid-free response rate and its 90% confidence interval will be presented based on all subjects in the FAS and by subgroup of treatment naive cGvHD vs SR-cGvHD.

2.14 Interim analysis

In order to summarize one year of safety data, in addition to the primary efficacy analysis, an interim CSR will be reported when all subjects have completed 1 year of treatment or discontinued earlier. The final analysis will be performed and the final CSR produced after all subjects have discontinued from the study. The same SAP will be used for both the interim and final analysis.

Since the data for the primary efficacy assessment, ORR at Cycle 7 Day 1, will already be final once all subjects have completed 6 months of treatment, and in this clinical trial it is not planned to test specific hypotheses related to the efficacy endpoint(s), but to provide estimates of efficacy endpoints, no alpha adjustment will be made for the analysis of the primary endpoint.

3 Sample size calculation

The sample size for the primary objective of measuring ORR at Cycle 7 Day 1 is approximately 42 subjects, regardless of age.

The sample size calculation is based on the ORR at Cycle 7 Day 1, and the calculation will consider the Saw-Toothed behavior of power waving for single binomial proportion using an exact method (Chernick and Liu 2002). Considering the response rate of children using corticosteroids is 30% to 50% (Wolff et al 2011), we assume the true ORR at Cycle 7 Day 1 of the study population is 70%, and therefore, a minimum sample size of 42 subjects would provide >80% probability to have a 90% CI with lower limit ≥ 50%. Table 3-1 provides

estimates of the probability to have a 90% CI with lower limit ≥ 50%, minimum number of responders required and two-sided 90% Clopper-Pearson CIs for various sample sizes.

Table 3-1 Probability to have a 90% CI with lower limit ≥ 50% and 90% confidence intervals for different number of subjects

Sample size	Minimum No. of responders	Response Rate	Probability to have a 90% CI with lower limit ≥ 50%	90% CI	
30	20	0.67	0.730	0.501	0.807
31	21	0.68	0.688	0.515	0.813
32	22	0.69	0.644	0.528	0.820
33	22	0.67	0.733	0.509	0.801
34	23	0.68	0.693	0.522	0.807
35	23	0.66	0.773	0.504	0.789
36	24	0.67	0.737	0.517	0.795
37	24	0.65	0.807	0.500	0.778
38	25	0.66	0.774	0.512	0.784
39	26	0.67	0.740	0.523	0.790
40	26	0.65	0.807	0.508	0.774
41	27	0.66	0.776	0.519	0.780
42	27	0.64	0.836	0.505	0.765
43	28	0.65	0.808	0.515	0.771
44	28	0.64	0.861	0.501	0.757
45	29	0.64	0.836	0.511	0.763
46	30	0.65	0.809	0.521	0.768
47	30	0.64	0.860	0.508	0.755
48	31	0.65	0.836	0.517	0.760
49	31	0.63	0.881	0.505	0.747
50	32	0.64	0.859	0.514	0.753

4 Change to protocol specified analyses

One exploratory analysis endpoints is added:
Chronic GvHD recurrence following completion of a taper of systemic therapy (see Section 2.13).

5 Appendix

5.1 Imputation rules

5.1.1 Study drug

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

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

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

5.1.2 AE date imputation

Table 5-5-1 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 5-5-2 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.

5.1.2.1 Other imputations

Incomplete date of initial diagnosis of cGvHD

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

5.2 AEs coding/grading

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

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

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

5.3 Laboratory parameters derivations

Grade categorization of lab values will be assigned programmatically as per NCI Common Terminology Criteria for Adverse Events (CTCAE) version 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)

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

5.4 Statistical models

5.4.1 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]).

5.4.2 Key secondary analysis

Not applicable.

5.4.3 Definition of new or additional systemic treatment of cGvHD

The data source to search for the new or additional systemic therapy would be from eCRFs of "Prior and Concomitant Medication" (CONMED) or eCRFs "Prior or Concomitant non-drug therapies/procedures".

Systemic therapies are identified using the Route (if CMROUTE is Intramuscular, Intravenous, Oral, Subcutaneous, or Parenteral). Therapies with CMROUTE = 'ORAL' will be considered as systemic only if in addition CMADMIN = "SYSTEMIC".

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

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

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

6 Reference

1. Armitage P, McPherson CK and Rowe BC (1969). Repeated significance tests on accumulating data. J .R. Statist. Soc. A, 132, 232-244.
2. Brookmeyer R and Crowley J (1982). A Confidence Interval for the Median Survival Time. Biometrics, 38, 29 - 41.
3. 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.
4. Clopper CJ and Pearson ES (1934). The use of confidence or fiducial limits illustrated in the case of the binomial. Biometrical, 26, 404-413.
5. Collet D (1994). Modelling survival data in medical research. London, Chapman & Hall.
6. Wolff D, Gerbitz A, Ayuk F, et al (2010) Consensus conference on clinical practice in chronic graft-versus-host disease (GVHD): first-line and topical treatment of chronic GVHD. Biology of Blood and Marrow Transplantation; 16(12), pp.1611-1628.
7. Wolff D, Schleuning M, von Harsdorf S, et al (2011) Consensus Conference on Clinical Practice in Chronic GVHD: Second-Line Treatment of Chronic Graft-versus-Host Disease. Biology of Blood and Marrow Transplantation; 17(1), pp.1-17.
8. Zeiser R, Burchert A, Lengerke C, et al (2015) Ruxolitinib in corticosteroid-refractory graft-versus-host disease after allogeneic stem cell transplantation: a multicenter survey. Leukemia; 29(10), p.2062.
9. Baumgartner RN, Roche AF, Himes JH. Incremental growth tables: Supplementary to previously published charts. Am J Clin Nutr 1986;43:711-22.
10. Lee SJ, Wolff D, Kitko C, et al (2015) Measuring therapeutic response in chronic graft-versus-host disease. National Institutes of Health consensus development project on criteria for clinical trials in chronic graft-versus-host disease: IV. The 2014 Response Criteria Working Group report. Biology of Blood and Marrow Transplantation; 21(6), pp.984-999.
11. Rauchenzauner M, Schmid A, Heinz-Erian P, Kapelari K, Falkensammer G, Griesmacher A, Finkenstedt G, Högler W. 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. 2007. Feb;92(2):443-449. Epub 2006 Nov 14.
12. Gooley TA, Leisenring W, Crowley J, Storer BE. Estimation of failure probabilities in the presence of competing risks: new representations of old estimators. Stat Med. 1999;18(6):695–706.