

STATISTICAL ANALYSIS PLAN



INCB 39110-301

GRAVITAS-301: A Randomized, Double-Blind, Placebo-Controlled Phase 3 Study of Itacitinib or Placebo in Combination With Corticosteroids for the Treatment of First-Line Acute Graft-Versus-Host Disease

IND Number:	113,428
EudraCT Number:	2017-000538-78
Sponsor:	Incyte Corporation 1801 Augustine Cut-Off Wilmington, DE 19803
Protocol Version:	Protocol Amendment (Version) 4 dated 13 AUG 2018
CRF Approval Date:	12 MAR 2019
SAP Version:	Amendment 1
SAP Author:	██████████ Executive Director, Biostatistics
Date of Plan:	23 JUL 2019

This study is being conducted in compliance with good clinical practice,
including the archiving of essential documents.

TABLE OF CONTENTS

LIST OF ABBREVIATIONS	7
1. INTRODUCTION	9
2. STUDY INFORMATION, OBJECTIVES, AND ENDPOINTS	9
2.1. Protocol and Case Report Form Version	9
2.2. Study Objectives and Endpoints	9
3. STUDY DESIGN	11
3.1. Randomization	12
3.2. Control of Type I Error	12
3.3. Sample Size Considerations	13
3.4. Interim and Final Analyses	14
3.4.1. Interim Analysis 1	15
3.4.2. Interim Analysis 2	15
3.4.3. Final Analysis	17
3.5. Schedule of Assessments	17
4. DATA HANDLING DEFINITIONS AND CONVENTIONS	21
4.1. Scheduled Study Evaluations and Study Periods	21
4.1.1. Day 1	21
4.1.2. Study Day	21
4.1.3. Baseline Value	21
4.1.4. Retreatment Day 1	21
4.1.5. Analysis Visit Definition	21
4.1.6. Last Available Value	22
4.1.7. Date of Last Contact	22
4.1.8. Handling of Missing and Incomplete Data	22
4.2. Variable Definitions	22
4.2.1. Age	22
4.2.2. Body Mass Index	22
4.2.3. Body Surface Area	22
4.2.4. Prior and Concomitant Medication	23
5. STATISTICAL METHODOLOGY	24
5.1. General Methodology	24

5.2.	Treatment Groups	24
5.3.	Analysis Sets.....	24
5.3.1.	Interim Analysis 1.....	24
5.3.1.1.	Interim 1 Analysis Set.....	24
5.3.2.	Interim Analysis 2.....	24
5.3.2.1.	Interim 2 Efficacy Analysis Set	24
5.3.2.2.	Interim 2 Safety Analysis Set	25
5.3.3.	Final Analysis	25
5.3.3.1.	Full Analysis Set.....	25
5.3.3.2.	Per Protocol Analysis Set	25
5.3.4.	Safety Analysis Set	25
5.3.5.	Pharmacokinetic and Translational Research Evaluable Analysis Set	25
6.	BASELINE, EXPOSURE, AND DISPOSITION VARIABLES AND ANALYSES.....	26
6.1.	Demographics and Baseline Characteristics, Disease History, Transplant History, Medical History, and Prior Therapies	26
6.1.1.	Demographics and Baseline Characteristics.....	26
6.1.2.	Hematologic Disease History	26
6.1.3.	Allogeneic Hematopoietic Stem Cell Transplant History	26
6.1.4.	Medical History	27
6.1.5.	Prior Anticancer Therapy or Radiotherapy.....	27
6.2.	Disposition of Subjects	27
6.3.	Protocol Deviations.....	27
6.4.	Transfusions Performed On Study or Within 30 Days Before Enrollment	27
6.5.	Exposure	27
6.6.	Study Drug Compliance.....	28
6.7.	Prior and Concomitant Medication.....	28
7.	EFFICACY	29
7.1.	General Considerations.....	29
7.2.	Efficacy Hypotheses	29
7.3.	Analysis of the Primary Efficacy Endpoint	30
7.3.1.	Primary Efficacy Analysis	30
7.3.1.1.	Day 28 Overall Response Rate	30

8.6.	Electrocardiograms	45
9.	INTERIM ANALYSES AND FINAL ANALYSES.....	46
9.1.	Overview of Interim Analyses	46
9.1.1.	Interim Analysis 1.....	46
9.1.2.	Interim Analysis 2.....	46
9.2.	Final Analyses.....	48
10.	CHANGES AND MODIFICATIONS TO THE ANALYSIS PLAN.....	48
10.1.	Changes to Protocol-Defined Analyses	48
10.2.	Changes to the Statistical Analysis Plan.....	49
10.2.1.	Original to Amendment 1	49
11.	REFERENCES	50
APPENDIX A.	STATISTICAL MODELS FOR EFFICACY ANALYSES	51
APPENDIX B.	PLANNED TABLES, FIGURES, AND LISTINGS.....	53
APPENDIX C.	CLINICALLY NOTABLE EVENTS BASED ON CUSTOMIZED MEDDRA QUERIES	62
APPENDIX D.	VARIOUS UPPER EFFICACY BOUNDARIES FOR TESTING DAY 28 ORR AT BOTH INTERIM 2 AND FINAL ANALYSES USING HWANG-SHIH-DECANI ALPHA SPENDING FUNCTION WITH SHAPE PARAMETER -4	66

LIST OF TABLES

Table 1:	Objectives and Endpoints	10
Table 2:	Operation Characteristics Using Hwang-Shih-DeCani Alpha Spending Function With Shape Parameter -4.....	16
Table 3:	Schedule of Assessments	18
Table 4:	Laboratory Assessments	19
Table 5:	Clinical Laboratory Analytes.....	20
Table 6:	Criteria for Clinically Notable Vital Sign Abnormalities.....	45
Table 7:	Criteria for Clinically Notable ECG Abnormalities	45
Table 8:	Statistical Analysis Plan Versions	48

LIST OF FIGURES

Figure 1:	Study Design: Screening and Randomization	11
Figure 2:	Study Design: Study Flowchart	12
Figure 3:	Upper Efficacy and Lower Futility Boundaries for Day 28 ORR at the Interim Analyses and Critical Value at the Final Analysis.....	14
Figure 4:	Recommended Decision Process at Interim 2	17

LIST OF ABBREVIATIONS

Abbreviation	Term
AE	adverse event
aGVHD	acute graft-versus-host disease
allo-HSCT	allogeneic hematopoietic stem cell transplantation
ALP	alkaline phosphatase
ALT	alanine aminotransferase
AST	aspartate aminotransferase
BMI	body mass index
BOR	best overall response
BSA	body surface area
cGVHD	chronic graft-versus-host disease
CI	confidence interval
CIBMTR	Center for International Blood & Marrow Transplant Research
CL/F	apparent oral dose clearance
C _{max}	maximum observed plasma or serum concentration
CMH	Cochran-Mantel-Haenszel
C _{min}	minimum observed plasma or serum concentration over the dose interval
CMV	cytomegalovirus
CR	complete response
CRF	case report form
CSR	clinical study report
CTCAE	Common Terminology Criteria for Adverse Events (v4.03)
DMC	Data Monitoring Committee
DNA	deoxyribonucleic acid
DOR	duration of response
ECG	electrocardiogram
ECOG	Eastern Cooperative Oncology Group
eCRF	electronic case report form
EOT	end of treatment
FAS	full analysis set
FFS	failure-free survival
GI	gastrointestinal
GVHD	graft-versus-host disease
HBV	hepatitis B virus
HCV	hepatitis C virus
HDL	high-density lipoprotein
HIV	human immunodeficiency virus

Abbreviation	Term
HLA	human leukocyte antigen
IC ₅₀	concentration that results in 50% inhibition
IVRS	interactive voice response system
JAK	Janus kinase
KM	Kaplan-Meier
LDL	low-density lipoprotein
MAGIC	Mount Sinai Acute GHVD International Consortium
MedDRA	Medical Dictionary for Regulatory Activities
MQ	MedDRA query
NCI	National Cancer Institute
NRM	nonrelapse mortality
ORR	overall response rate
OS	overall survival
PK	pharmacokinetic
PR	partial response
PS	performance status
PT	preferred term
PTLD	post-transplant lymphoproliferative disorder
████	██████████
RNA	ribonucleic acid
SAE	serious adverse event
SAP	Statistical Analysis Plan
SD	standard deviation
SE	standard error
████	██████████
SOC	system organ class
STAT	signal transducer and activator of transcription
TEAE	treatment-emergent adverse event
TFLs	tables, figures, and listings
t _{max}	time to maximum concentration
TTR	time to response
TYK	tyrosine kinase
VGPR	very good partial response
WHO	World Health Organization

1. INTRODUCTION

Itacitinib potently inhibits JAK1 ($IC_{50} = 3.6$ nM at 1 mM adenosine triphosphate concentration), with 22- to > 500-fold selectivity over the other JAK family members, JAK2, JAK3, and TYK2. It does not significantly inhibit (< 30% inhibition) a broad panel of approximately 60 other kinases. Itacitinib is also potent (IC_{50} values of approximately 10 nM to 350 nM) in cytokine-driven cell-based assays.

New treatments for the therapy of aGVHD after allo-HSCT are urgently needed. [REDACTED]

[REDACTED]

[REDACTED]

[REDACTED]

[REDACTED]

[REDACTED]

Section 1 of the Protocol provides a detailed description of the investigational products, target patient population, and the potential risks and benefits from the combination treatments. The purpose of this document is to provide a detailed SAP.

2. STUDY INFORMATION, OBJECTIVES, AND ENDPOINTS

2.1. Protocol and Case Report Form Version

This SAP is based on INCB 39110-301 Protocol Amendment (Version) 4 dated 13 AUG 2018 and CRFs approved 12 MAR 2019. Unless superseded by an amendment, this SAP will be effective for all subsequent Protocol amendments and CRFs.

2.2. Study Objectives and Endpoints

[Table 1](#) presents the study objectives and endpoints.

Table 1: Objectives and Endpoints

Primary Objective	Primary Endpoint
Compare the efficacy of itacitinib in combination with corticosteroids versus placebo in combination with corticosteroids in terms of ORR at Day 28 in subjects with aGVHD.	ORR at Day 28, defined as the proportion of subjects demonstrating a CR, VGPR, or PR.
Key Secondary Objective	Key Secondary Endpoint
Compare the efficacy between treatment cohorts at a subsequent key clinical landmark.	NRM at Month 6, defined as the proportion of subjects who died due to causes other than malignancy relapse at Month 6.
Secondary Objectives	Secondary Endpoints
Compare additional response and longer-term efficacy outcomes between treatment cohorts.	ORR, defined as the proportion of subjects demonstrating a CR, VGPR, or PR at Days 14, 56, and 100.
	NRM at Months 9, 12, and 24.
	DOR for responders will be calculated. The DOR is defined from the time of the onset of response to loss of response. Subjects who died or discontinued will be censored at the death date or the previous assessment.
	TTR, defined as the interval from treatment initiation to first response.
	Relapse rate of malignant and nonmalignant hematologic diseases, defined as the proportion of subjects whose underlying hematologic disease relapses.
	Malignancy relapse-related mortality rate, defined as the proportion of subjects whose malignancy relapses and has a fatal outcome.
	FFS, defined as the proportion of subjects who are still alive, have not relapsed, have not required additional therapy for aGVHD, and have not demonstrated signs or symptoms of cGVHD, at Month 6.
OS, defined as the interval from study enrollment to death due to any cause.	
Assess the incidence and severity of AEs and SAEs.	Clinical safety data (eg, AEs, infections) will be tabulated and listed.
Evaluate the PK of itacitinib when administered in combination with corticosteroids.	C_{max} , C_{min} , t_{max} , AUC, and CL/F.
Evaluate the incidence of secondary graft failure.	Incidence rate of secondary graft failure, defined as > 95% recipient cells any time after engraftment with no signs of relapse, OR retransplantation because of secondary neutropenia ($< 0.5 \times 10^9/L$) and/or thrombocytopenia ($< 20 \times 10^9/L$) within 2 months of transplant.
Evaluate the use and discontinuation of corticosteroids.	Average and cumulative corticosteroid dose at Days 28, 56, 100, and 180; proportion of subjects who discontinue corticosteroids at Days 56 and 100.
Evaluate the use and discontinuation of immunosuppressive medications.	Proportion of subjects who discontinue immunosuppressive medications at Days 56 and 100.
Evaluate the incidence of aGVHD flares.	Incidence rate of aGVHD flares through Day 100.
Evaluate the incidence of cGVHD.	Incidence rate of cGVHD at Days 180 and 365.

3. STUDY DESIGN

This is a randomized, double-blind, placebo-controlled, multicenter Phase 3 study of itacitinib or placebo in combination with corticosteroids as first-line treatment of subjects with Grade II to IV aGVHD. Subjects will be randomized 1:1 to itacitinib 200 mg plus corticosteroids or matching placebo plus corticosteroids. Randomization will be stratified by GVHD risk status (ie, standard risk vs high risk; [MacMillan et al 2015](#)). Subjects will receive randomized study treatment until treatment failure (progression of disease, no response, or requiring additional systemic therapy), unacceptable toxicity, or death. Transfusion support and continued use of anti-infective medications, GVHD prophylaxis medications (including calcineurin inhibitors), and topical steroid therapy are permitted.

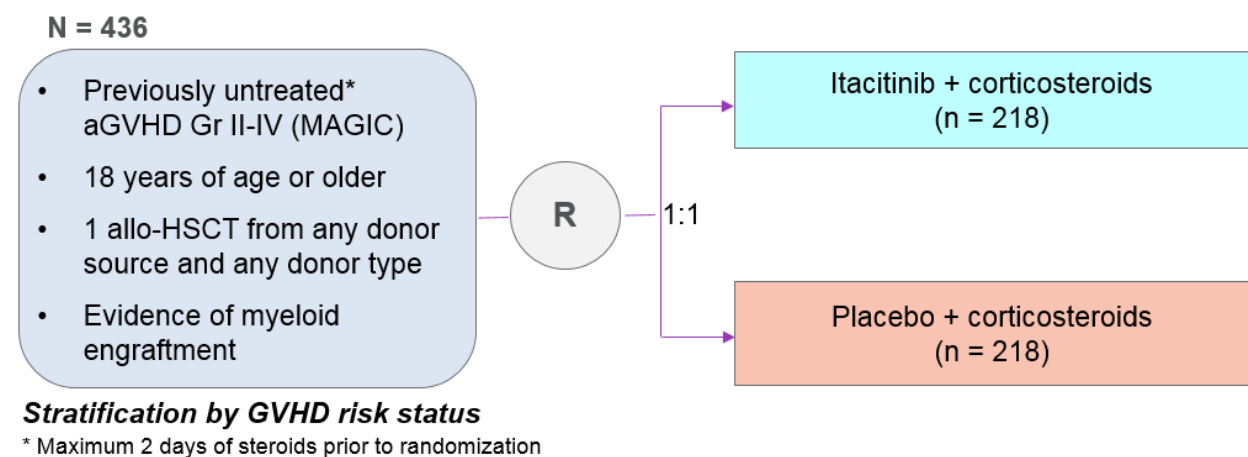
GVHD staging and grading will be assessed for efficacy as per MAGIC criteria ([Harris et al 2016](#)); safety and tolerability will be assessed as per NCI CTCAE v4.03 ([NCI 2010](#)).

An independent DMC will perform an interim analysis for futility once 112 subjects (around 56 per cohort) have completed the Day 28 visit or withdrew early from the study. If the futility boundary is crossed at Interim Analysis 1, the sponsor may consider stopping the study. A second interim analysis for the primary endpoint of Day 28 ORR (both efficacy and futility) will be performed once 240 eligible subjects have completed the Day 180 visit or withdrew early from the study. If the Day 28 ORR efficacy boundary is crossed, the Day 180 NRM will be tested.

The final analysis will be conducted once the last subject completes the Day 180 visit or withdraws from the study.

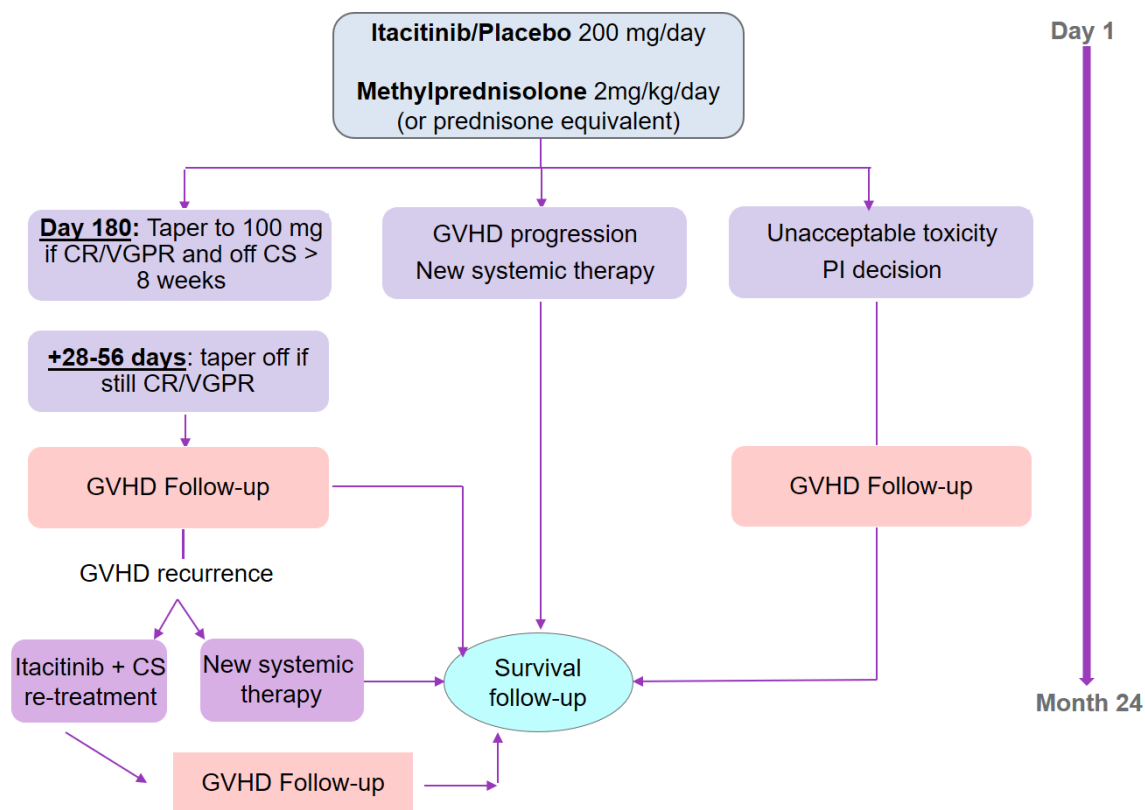
[Figure 1](#) below depicts study design at screening and randomization, and [Figure 2](#) summarizes study flow after randomization.

Figure 1: Study Design: Screening and Randomization



R = randomization.

Figure 2: Study Design: Study Flowchart



Note: Methylprednisolone is given at 2 mg/kg per day or at a dose per local treatment guidelines. Some subjects started with 1 mg/kg per day on Day 1, and this is accepted per the Protocol.

3.1. Randomization

Randomization will occur centrally by the IVRS. Subjects will be randomized in a 1:1 ratio to receive itacitinib plus corticosteroids or matching placebo plus corticosteroids and will be stratified by baseline GVHD risk status (ie, standard-risk versus high-risk). As this is a randomized, double-blind, placebo-controlled study, neither the investigators nor the sponsor will be aware of the treatment to which a subject is randomized.

3.2. Control of Type I Error

For the primary and key secondary endpoints, the overall 1-sided Type 1 error is controlled at 0.025. Early stopping for positive efficacy is permitted in this study. Testing of the primary endpoint will use an alpha-spending approach to control Type I error. Subsequent testing of the key secondary endpoint will control Type I error using the Bonferroni method. See Sections 3.4 and 9 for more details.

No multiplicity adjustment will be applied to other endpoints. For estimation purposes, 95% CIs will be provided for parameter estimates of other endpoints.

3.3. Sample Size Considerations

A sample size of 414 subjects with 1:1 randomization (itacitinib vs placebo) and stratification based on baseline aGVHD risk status (standard risk vs high risk) provides approximately 90% power to test for the primary endpoint (ORR at Day 28) and approximately 80% power to test for the key secondary endpoint (NRM rate at Month 6). The family-wise α -level will be controlled at 0.025 overall for the two comparisons. Specifically, this study will claim to have achieved the efficacy objective when the primary endpoint ORR at Day 28 shows a significant treatment effect at 1-sided $\alpha = 0.025$. Conditional on significance of the primary endpoint, the key secondary endpoint (NRM rate at Month 6) will be tested at 1-sided $\alpha = 0.025$.

An absolute improvement of 16% in the primary endpoint of ORR at Day 28 would be considered a clinically meaningful improvement over standard first-line systemic treatment for aGVHD. A sample size of 414 subjects (207 subjects per treatment group) will provide approximately 90% power to detect a 16% overall difference (72% vs 56%) between treatment groups for the primary endpoint (ORR at Day 28) at a 1-sided alpha level of 0.025.

The assumed Day 28 ORR of 56% is based on MacMillan et al (2015). The stratum-specific response rates (standard risk, 68%; high risk, 44%) assume a standard risk:high risk ratio of 0.50:0.50. It is expected that treatment with itacitinib will result in a 16% increase in the ORR, that is, an expected odds ratio of 2.02 (which corresponds to an overall increase in ORR to 72%), with stratum-specific response rates of 81% and 61% for standard-risk and high-risk aGVHD cohorts, respectively. Power for the CMH test with stratification based on GVHD risk score was calculated using software N-Query 5.0.

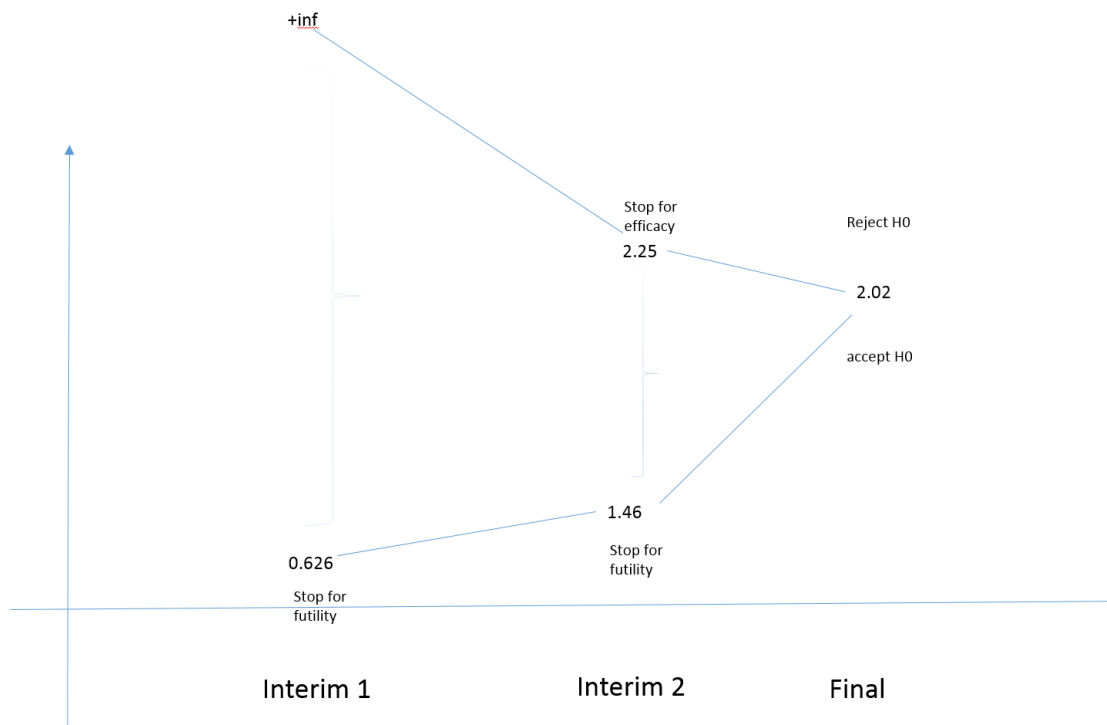
Given its clinical significance, NRM rate at Month 6 will be assessed as the key secondary endpoint. The Day 180 NRM rate difference between 2 groups is tested using a KM estimator of the cumulative distribution function at Day 180. Based on MacMillan et al (2015), the overall NRM at Month 6 in the best available therapy group is 33%, with a stratum-specific NRM rate of 22% for standard risk and 44% for high risk (assuming a standard risk:high risk ratio of 0.50:0.50). A 40% relative reduction in NRM rate will be considered a clinically meaningful improvement. Based on simulation, a sample size of 414 subjects (207 subjects per treatment cohort) will provide approximately 80% power to detect a relative reduction of 40% (or absolute difference of 13.2%, ie, 19.8% vs 33% at Day 180) between treatments at a 1-sided α level of 0.025. NRM at Month 6 will be tested using a CMH test stratified by risk status (standard-risk vs high-risk aGVHD). Stratum-specific NRM rates of 12.4% and 28.2% are expected for standard-risk and high-risk cohorts, respectively. Power for the CMH test with stratification based on GVHD risk score was calculated using software N-Query 5.0.

In order to compensate for an anticipated early withdrawal rate of up to 5% (ie, subjects who do not have any response data before Day 28), a total of 436 subjects (218 per treatment arm) will be randomized into the study to ensure a sample size of 414 evaluable subjects.

3.4. Interim and Final Analyses

There are 2 interim analyses planned in this study. Interim Analysis 1 will be conducted at the time when 112 subjects have completed the Day 28 visit or withdrawn from the study. Only a futility analysis will be conducted based on test statistic of Day 28 ORR using stratified CMH test (see details below). The observed normalized z statistic will be compared with the futility boundary of 0.626 as depicted in Figure 3 (equivalently conditional power equal to 0.2); if the normalized z statistic is less than the futility boundary, it is recommended to stop the study early for futility. Interim Analysis 2 will be conducted when approximately 240 subjects have completed the Day 180 visit or have withdrawn from the study. At Interim Analysis 2, both futility and efficacy will be tested for Day 28 ORR using stratified CMH test with z-score efficacy boundary of 2.25 and futility boundary of 1.46 corresponding to conditional power of 0.2. Critical value at final analysis is 2.02. See details on boundary specifications in Sections 3.4.1 and 3.4.2. Note that boundaries for both the Interim Analysis 2 and the Final Analysis are constructed based on a hypothetical example solely for illustration purposes, whereas the real boundaries used in the study will be based on the observed information fraction at Interim Analysis 2.

Figure 3: Upper Efficacy and Lower Futility Boundaries for Day 28 ORR at the Interim Analyses and Critical Value at the Final Analysis



In any interim analysis, the sponsor may consider stopping the study for futility if the conditional power is less than 20% and stopping for efficacy if both efficacy boundaries are crossed. If neither the efficacy nor the futility boundary for Day 28 ORR is crossed, the study will continue. Additionally, if efficacy for Day 28 ORR is crossed but efficacy for Day 180 NRM rate is not crossed, the study will continue.

3.4.1. Interim Analysis 1

The first interim analysis for futility only will be performed when 112 subjects (25.7% of the planned 436 subjects) have completed the Day 28 visit or withdrew early from study. No alpha will be spent during the first interim analysis. The stopping rule is nonbinding. When the conditional power is < 20%, the sponsor may consider terminating the study if the totality of evidence is warranted. The conditional power is calculated assuming that the estimated treatment effect at the interim analysis is the true effect using the formula (Jennison and Turnbull 1999):

$$\Phi \left(\frac{z}{\sqrt{t(1-t)}} - \frac{z_{\alpha/2}}{\sqrt{1-t}} \right)$$

where Φ is the cumulative distribution function of the observed normalized test statistic z , and t is the information fraction (ie, observed sample at interim divided by the planned total sample size 436); $t = 0.257$ at Interim 1 and $z_{\alpha/2} = 1.96$.

Since CMH test with only 1 stratification factor, aGVHD risk status (standard vs high risk), will be used to test ORR at Day 28, it is thus a chi-square test with degree of freedom equal to 1, that is, $\chi^2_{df=1}$. Taking the square root of the observed CMH test statistic $\chi^2_{df=1}$ will result in normalized test statistic z .

3.4.2. Interim Analysis 2

For Interim Analysis 2, the analysis of both efficacy and futility will use a clinical database cutoff based on when 240 eligible subjects (55.0% of the targeted 436 subjects) have completed the Day 180 visit or withdrew from the study prior to Day 180.

The Hwang-Shih-DeCani alpha spending function with shape parameter -4 (O'Brian-Fleming like boundary) will be used to define upper efficacy boundary for the primary analysis of ORR at Day 28 when approximately 240 eligible subjects reach Day 180 (Hwang et al 1990). Based on enrollment projection, there will be approximately 360 subjects who have completed Day 28.

Table 2 lists efficacy boundary when information fraction (ie, the percentage of subjects) is 0.826 (360 subjects complete Day 28) at Interim 2; if ORR tested at Day 28 crosses the boundary, then the key secondary endpoint, 6-month NRM rate, will be tested by using the Bonferroni method for alpha spending (ie, 0.0125 as compared with 1-sided p-value at both interim and final analyses). By simulation, without considering early stopping for futility, the power of testing 6-month NRM at Interim 2 (with 240 eligible subjects who have completed Day 180) is approximately 0.608.

Table 2: Operation Characteristics Using Hwang-Shih-DeCani Alpha Spending Function With Shape Parameter -4

		No. of Subjects (%)	Alpha Spent /Cumulative	Efficacy Boundary or Critical Value (as Compared With Observed Normalized Test Statistic z)	Nominal Significance /Alpha Level: $1 - \Phi(d)$	Efficacy Boundary or Critical Value at Chi-Square Test Statistic With Degree of Freedom 1
ORR at Day 28	Interim Analysis 1	112 (25.7)	–	–	–	–
	Interim Analysis 2	360 (82.6)	0.0122/0.0122	2.249810	0.0122305	5.061645
	Final	436 (100)	0.0128/0.025	2.015483	0.02192703	4.062172

If, at Interim Analysis 2, the observed number of subjects who have completed Day 28 is different from 360, then the boundaries for efficacy in [Table 2](#) will be recalculated using observed information fraction.

The interim analyses will be performed by an independent statistician and programmer not involved with the conduct of the study as described in the DMC charter.

At the Interim Analysis 2 clinical data cutoff date, the date when the 240th eligible subject randomized in the study has reached Day 180, 360 subjects are expected to have been randomized. Testing of Day 28 ORR will be conducted for subjects who have completed the Day 28 visit or withdrew from the study by the clinical data cutoff date. This means that subjects who were ongoing but had not yet reached the Day 28 visit by the Interim Analysis 2 clinical data cutoff date will be excluded from both the Day 28 ORR and NRM efficacy analyses.

[Figure 4](#) depicts the recommended decision process at Interim Analysis 2: first, testing Day 28 ORR using stratified CMH and assuming 360 subjects have completed Day 28 or withdrew from the study by the cutoff date, if observed normalized test statistic is ≤ 1.46 , it is recommended to stop the study for futility; if the observed normalized test statistic of Day 28 ORR is ≥ 2.25 , then the NRM rate at Day 180 will be tested using the Bonferroni method. If both the primary efficacy analysis and the key secondary efficacy analysis pass their corresponding efficacy boundary, the superiority of itacitinib over placebo for both the Day 28 ORR and Day 180 NRM endpoints can be claimed at Interim Analysis 2.

Table 3: Schedule of Assessments

Visit Day Item	Screening -28 to -1	Treatment ^a												EOT ^d	Retreatment ^e	Follow-Up		
		D1	D7	D14	D21	D28	D35	D42	D49	D56 ^b	D100	D180	D365 ^c			Safety ^f	GVHD ^g	Survival ^h
Informed consent	X																	
I/E criteria	X	X																
Contact IVRS	X	X					X				X	X		X	X	X		
Demography/disease history	X																	
Prior/concomitant medications	X	X												X	X	X		
Supportive care medications	X	X												X	X	X		
AE assessment	X	X												X	X	X		
Physical examination	X	X	X	X	X	X	X	X	X	X	X		X	X	X	X		
Vital signs	X	X	X	X	X	X	X	X	X	X	X		X	X	X	X		
ECOG PS	X	X	X	X	X	X	X	X	X	X	X		X	X	X	X		
12-lead ECG	X	As indicated												X				
aGVHD grading and response	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	
cGVHD assessment	X	As indicated												X	X	X	X	
Chimerism/graft failure assessment	X	As indicated																
PTLD assessment		As indicated												X	X	X	X	
Underlying disease relapse assessment		As indicated												X	X	X	X	X
Dispense study drug		X				X				X			X		X			
Study drug compliance		X				X				X			X	X	X			
Steroid dose monitoring	X	X												X	X			
Survival follow-up																		X
New aGVHD therapies																		X

ECG = electrocardiogram; I/E = inclusion/exclusion; PS = performance status; PTLD = post-transplant lymphoproliferative disorder.

^a A ± 3-day window is permitted to facilitate scheduling during the treatment phase.

^b After Day 56, visits will occur every 28 days and will include all Day 56 assessments.

^c The Day 365 visit will also serve as the regularly scheduled visit occurring every 28 days after Day 56 (ie, Day 364).

^d A second EOT should occur if the subject restarts itacitinib and subsequently discontinues treatment.

^e Retreatment assessments occur every 28 days ± 3 days.

^f 30 to 35 days after EOT.

^g Every 28 days ± 7 days, for participants who completed study treatment or discontinued early for reasons other than GVHD progression.

^h Every 8 weeks ± 7 days.

Table 4: Laboratory Assessments

Visit Day Item	Screening	Treatment ^a											EOT ^e	Retreatment ^f	GVHD Follow-Up	Safety Follow-Up ^g	
	-28 to -1	D1 ^b	D7	D14	D21	D28	D35	D42	D49	D56 ^c	D100	D180					D365 ^d
Chemistry panel	X	X	X	X	X	X	X	X	X	X	X		X	X	X	X ^h	X
Hematology	X	X	X	X	X	X	X	X	X	X	X		X	X	X		X
Hepatitis screening	X																
HIV screening	X																
Serum pregnancy test (childbearing females only)	X													X			
Urine pregnancy test ⁱ (childbearing females only)						X				X							
PK assessment ^j		X	X			X											

^a A ± 3-day window is permitted to facilitate scheduling during the treatment phase.
^b Day 1 laboratory assessments do not need to be repeated if screening assessments were performed within the preceding 7 days.
^c After Day 56, visits will occur every 28 days and will include all Day 56 assessments.
^d The Day 365 visit will also serve as the regularly scheduled visit occurring every 28 days after Day 56 (ie, Day 364).
^e A second EOT should occur if the subject restarts itacitinib and subsequently discontinues treatment.
^f Retreatment assessments occur every 28 days.
^g 30 to 35 days after EOT.
^h Liver testing including total bilirubin data
ⁱ Urine pregnancy tests are required at Day 28 and every 28 days.
^j Subjects 1-112: samples collected on Days 1, 7, and 28 at predose and at 1 hour ± 15 minutes, 2 hours ± 30 minutes, and 5 hours ± 60 minutes postdose. Subjects 113-436: Day 7 and 28 at predose only.
^k Day 56 only.

Table 5: Clinical Laboratory Analytes

Serum Chemistries	Hematology ^a	Other
Albumin ALP ALT AST Bicarbonate or CO ₂ Blood urea nitrogen Calcium Chloride Creatinine Glucose Lactate dehydrogenase Lipid profile (total cholesterol, HDL, LDL, and triglycerides) Phosphorus Potassium Sodium Total bilirubin Total protein	Hematocrit Hemoglobin Mean corpuscular volume Reticulocytes Platelet count Red blood cell count White blood cell count White blood cell differential (5 parts): <ul style="list-style-type: none"> • Basophils • Eosinophils • Lymphocytes • Monocytes • Neutrophils 	Serum pregnancy test Urine pregnancy test Hepatitis Screening Tests Hepatitis B surface antigen Hepatitis B surface antibody Hepatitis B core antibody HCV antibody HCV-RNA HBV-DNA

ALP = alkaline phosphatase; ALT = alanine aminotransferase; AST = aspartate aminotransferase;

HDL = high-density lipoprotein; LDL = low-density lipoprotein.

^a Hematology and chemistry assessments will be performed locally.

4. DATA HANDLING DEFINITIONS AND CONVENTIONS

4.1. Scheduled Study Evaluations and Study Periods

4.1.1. Day 1

Day 1 is the date of first dose of itacitinib/placebo in the treatment phase.

4.1.2. Study Day

If a visit/reporting date is on or after Day 1, then the study day at the visit/reporting date will be calculated as

$$\text{Day \#} = (\text{Visit/Reporting Date} - \text{Day 1 date} + 1).$$

If the visit/reporting date is before Day 1, then the study day at the visit/reporting date will be calculated as

$$\text{Day \#} = (\text{Visit/Reporting Date} - \text{Day 1 date}).$$

A study day of -1 indicates 1 day before Day 1.

4.1.3. Baseline Value

Baseline is the last nonmissing measurement obtained before or on the first day of itacitinib or placebo administration (in terms of date and time, if available).

4.1.4. Retreatment Day 1

Retreatment Day 1 is the date of first dose of itacitinib in the retreatment phase.

4.1.5. Analysis Visit Definition

In order to summarize data collected over time (including unscheduled visits and EOT visit), the following general rule will be applied in creating analysis visits:

- Step 1: The EOT assessment will be assigned to the next available scheduled assessment if it is measured within 3 days of the target day of next scheduled visit. For example, if a subject has an EOT assessment after scheduled visit Day 21, and the next scheduled visit is the Day 28 visit, then the assessment at EOT will be assigned to Analysis Visit Day 28 for this subject if they are assessed within 3 days of the target day of next scheduled visit.
- Step 2: If the scheduled assessment is available (including mapped EOT assessments in Step 1), then the scheduled assessment should be used; otherwise, the first nonmissing unscheduled assessment should be used.

Note that laboratory data were not scheduled to be collected per protocol at the Day 180 visit but were collected in study execution to support GVHD staging.

4.1.6. Last Available Value

The last available value is the last nonmissing measurement obtained after starting study treatment.

4.1.7. Date of Last Contact

The date of last contact is defined at the subject level as the latest date of discontinuation, laboratory assessment, GVHD/ECOG assessment, vital sign, ECG, scheduled or unscheduled visit, and any available date in the clinical database showing that the subject is still alive prior to or on the clinical data cutoff date.

4.1.8. Handling of Missing and Incomplete Data

In general, values for missing data will not be imputed unless methods for handling missing data are specified in this section or relevant sections.

Partial disease/cancer diagnosis date will be handled as follows:

- If only the day is missing, then the imputed day will be the first of the month.
- If both the month and day are missing, then the imputed day and month will be 01 JAN.
- No imputation will be done if the date is completely missing.

Missing date of last dose in the treatment period will be handled as follows:

- If the date is completely missing, then the imputed date of the last dose will be date of EOS, date of death, date of clinical data cutoff, or retreatment Day 1 – 1, whichever is the earliest.

4.2. Variable Definitions

The following variables will only be calculated if not reported on the eCRF.

4.2.1. Age

Subject age was collected before enrollment.

Age category 1 is defined as < 65 years and ≥ 65 years, and age category 2 is defined as < 65 years, 65 to 74 years, 75 to 84 years, and ≥ 85 years.

4.2.2. Body Mass Index

Body mass index will be calculated as follows:

$$\text{BMI (kg/m}^2\text{)} = [\text{weight (kg)}] / [\text{height (m)}]^2.$$

4.2.3. Body Surface Area

Body surface area will be calculated based on the Mosteller (1987) formula as follows:

$$\text{BSA (m}^2\text{)} = \{[\text{weight (kg)} \times \text{height (cm)}] / 3600\}^{1/2}.$$

Sites will also record the BSA calculated per institutional standards.

4.2.4. Prior and Concomitant Medication

Prior medication is defined as any nonstudy medication taken within 30 days before the first dose of study drug.

Concomitant medication is defined as any nonstudy drug that is started accordingly:

- Before the date of first administration of study drug and is ongoing throughout the study or ends on/after the date of first study drug administration.
- On/after the date of first administration of study drug and is ongoing or ends during the course of study drug.

A prior medication could also be classified as "both prior and concomitant medication" if the end date is on or after first dose of study drug.

For the purposes of analysis, all medications will be considered concomitant medications unless the medications can unequivocally be defined as not concomitant.

The start/stop dates recorded in the eCRF by the investigator and his or her research staff will be used to identify when a concomitant medication was taken during the study. Unresolved missing start dates will be handled as follows:

- For the starting date, impute as follows:
 - If the day is missing, then impute as the first day of the month.
 - If the month is missing, then impute as the first day of the year.
 - If the year is missing, then the medication will be considered both prior and concomitant.
- For the ending date, impute as follows:
 - If the day is missing, then impute as the last day of the month.
 - If the month is missing, then impute as the last day of the year.

If the year is missing, then the medication will be considered as ongoing until the end of the study.

Note that the above imputation rules are for the purpose of defining prior and concomitant medications in the study. Imputed start/end dates are not saved in the analysis dataset.

5. STATISTICAL METHODOLOGY

5.1. General Methodology

Unless otherwise noted, SAS® software (SAS Institute Inc, Cary, NC; Version 9.4 or later) will be used for the generation of all tables, graphs, and statistical analyses. Descriptive summaries for continuous variables will include, but not be limited to, the number of observations, mean, standard deviation, median, minimum, and maximum. Descriptive summaries for categorical variables will include the number and percentage of subjects in each category.

Interim analyses are planned for this study as defined in Section 9.

5.2. Treatment Groups

This is a randomized, double-blinded, placebo-controlled study with itacitinib or placebo in combination with corticosteroids. There are 2 treatment groups. Subjects will be randomized 1:1 to itacitinib 200 mg plus corticosteroids or matching placebo plus corticosteroids.

5.3. Analysis Sets

5.3.1. Interim Analysis 1

The Interim Analysis 1 clinical data cutoff date was defined as the date when the first 112 randomized and dosed subjects have either completed the Day 28 visit or withdrawn from the study. This occurred on 25 JUN 2018.

5.3.1.1. Interim 1 Analysis Set

The Interim 1 analysis set contains all subjects who have been randomized and have received at least 1 dose of study drug as of the Interim 1 cutoff date. The Interim 1 analysis set will be used for both planned efficacy and safety analyses at Interim Analysis 1. Note that subjects randomized after the 112th subject in the Interim 1 analysis set are not included in Interim 1 analyses, even though they might have safety and efficacy data in the clinical database as of the Interim 1 cutoff date. Because Interim Analysis 1 includes limited numbers of analyses, only 1 analysis set is defined, which is used for both efficacy and safety analyses.

5.3.2. Interim Analysis 2

The Interim Analysis 2 cutoff date is when the 240th eligible subject randomized in the study reaches Day 180 in the treatment phase. This occurred on 17 APR 2019.

5.3.2.1. Interim 2 Efficacy Analysis Set

The Interim 2 efficacy analysis set contains all subjects who have been randomized as of the Interim Analysis 2 clinical data cutoff date but excludes subjects who are ongoing at the clinical data cutoff date and have not finished the Day 28 visit. The Interim 2 efficacy analysis set will be used for all planned efficacy analyses at Interim Analysis 2.

5.3.2.2. Interim 2 Safety Analysis Set

The Interim 2 safety analysis set contains all subjects who have been randomized and have received at least 1 dose of study drug as of the Interim 2 cutoff date. The Interim 2 Safety Analysis Set will be used for all planned safety analyses at Interim 2. All safety tables, figures, and listings will be presented by actual treatment group.

5.3.3. Final Analysis

All efficacy analyses and demographic and baseline summaries will be presented by planned treatment group, that is, by either itacitinib or placebo.

5.3.3.1. Full Analysis Set

All subjects who are randomized constitute the FAS.

The FAS will be used for the summary of demographics, baseline characteristics, subject disposition, and analyses of all efficacy data according to intent-to-treat principle.

5.3.3.2. Per Protocol Analysis Set

The per protocol analysis set comprises the subset of subjects in the FAS who are compliant with requirements of the clinical study Protocol, excluding the following:

- Did not receive any dose of itacitinib/placebo after randomization
- Did not meet eligibility criteria
- Had significant deviation(s) from the Protocol in terms of missed study visits, noncompliance with randomized therapy or prohibited concomitant medications
- Subjects later considered as not having aGVHD
- Subjects with MAGIC Criteria Grade 0 or I at Baseline

5.3.4. Safety Analysis Set

The safety analysis set includes subjects randomized into the study who received at least 1 dose of study drug. All safety tables, figures, and listings will be presented by actual treatment group.

5.3.5. Pharmacokinetic and Translational Research Evaluable Analysis Set

Pharmacokinetic and translational research evaluable analysis set includes subjects who receive at least 1 dose of study drug and provide at least 1 plasma sample (1 corresponding PK or translational research measurement) after study drug administration will be considered as potential PK evaluable subjects. The study pharmacokineticist will review data listings of study drug administration and sample records to identify subjects to be excluded from analyses of PK data.

6. BASELINE, EXPOSURE, AND DISPOSITION VARIABLES AND ANALYSES

[Appendix B](#) provides a list of data displays.

In the event that the study terminates at Interim Analysis 1 or 2, the interim safety set will be used in place of the safety analysis set, the interim efficacy set will be used in place of FAS, and any per protocol populations will use subsets of the appropriate interim safety or efficacy set.

6.1. Demographics and Baseline Characteristics, Disease History, Transplant History, Medical History, and Prior Therapies

Demographic and baseline characteristics, disease history, transplant history, medical history, and prior therapies will be summarized for the FAS and listed.

6.1.1. Demographics and Baseline Characteristics

Demographic and baseline characteristics include the following: age, sex, race, ethnicity, weight, height, BMI, BSA, and ECOG performance status.

6.1.2. Hematologic Disease History

Underlying hematologic disease, diagnosis/type of non-malignant and malignant disease, time since initial diagnosis of underlying hematologic disease, and disease risk status will be summarized.

Time since diagnosis of underlying hematologic disease will be calculated as follows:

Time since diagnosis of underlying hematologic disease (days) = Day 1 – date of diagnosis of underlying malignancy + 1.

6.1.3. Allogeneic Hematopoietic Stem Cell Transplant History

Type of allogeneic transplant, time since transplant, donor source, ex vivo T-cell depletion of the allograft (Yes/No), CD34+ positive selection of the allograft (Yes/No), donor CMV type, subject CMV status, HLA matching, HLA score, best response of hematologic disease at the time of transplant, days of diagnosis of aGVHD since transplantation, GVHD organ staging categories, and type of conditioning treatment regimens will be summarized in the FAS by treatment group.

Time since transplant will be calculated as follows:

Time since transplant (days) = Day 1 – date of transplant + 1.

Days of diagnosis of aGVHD since transplantation will be calculated as follows:

Days of diagnosis of aGVHD (days) since transplantation = date of diagnosis of GVHD – date of transplantation + 1.

6.1.4. Medical History

For subjects in the FAS, medical history will be summarized by treatment group. This summation will include the number and percentage of subjects with significant medical history for each body system/organ class as documented on the eCRF.

6.1.5. Prior Anticancer Therapy or Radiotherapy

Number of subjects who received prior anticancer therapy will be summarized for the FAS. Regimen name, component drugs, start and stop dates, purpose of the regimen, best response, reason for discontinuation, and date of relapse/progression will be listed.

Number of subjects who received prior radiation will be summarized for the FAS. Radiotherapy type, body site, start and stop dates, total dose, and best response will be listed.

6.2. Disposition of Subjects

The number and percentage of subjects who were randomized, discontinued study treatment with a primary reason for discontinuation, and withdrew from the study with a primary reason for withdrawal will be summarized for the FAS by treatment group. Subjects entering and discontinuing retreatment will also be summarized.

6.3. Protocol Deviations

Protocol deviations recorded on the eCRF will be presented in the subject data listings. Incidence of different types of protocol deviations will be summarized by treatment group.

6.4. Transfusions Performed On Study or Within 30 Days Before Enrollment

Within 30 days before enrollment, blood component, time since transfusion, quantity, and unit, given AE for transfusion, and the reason for the PRBC/platelet transfusion will be listed. Transfusion performed at any time after randomization will be listed together with those performed within 30 days before enrollment.

6.5. Exposure

For subjects in the safety analysis set, exposure to study drug will be summarized descriptively as follows:

- **Duration of treatment with study drug:**

Duration of treatment (days) = Date of last dose of study drug – Date of first dose of study drug + 1

- **Average daily dose of study drug:**

Average daily dose of study drug (mg/day) = [total actual study drug dose taken (mg)] / [duration of treatment with study drug]

Frequency distribution of itacitinib/placebo duration category (days) will be tabulated by concatenated intervals: Day 1 through Day 28, Day 29 through Day 100, Day 101 through Day 180, and > Day 180.

Similarly, duration of corticosteroid and average daily dose of corticosteroid before EOT as the background therapy will be summarized by treatment group. Frequency distribution of corticosteroid duration category (days) will be tabulated by concatenated intervals.

Frequency distribution of itacitinib/placebo dose modification in the treatment phase (ie, dose decreased due to toxicity, dose increased, dose tapered, dose temporarily interrupted and with previous dose resumed) will be tabulated by concatenated intervals: Day 1 through Day 28, Day 29 through Day 56, Day 57 through Day 84, Day 85 through Day 112, Day 112 through Day 140, Day 141 through Day 168, Day 169 through Day 196, Day 197 through Day 224, Day 225 through Day 252, Day 253 through Day 280, Day 281 through Day 308, Day 381 through Day 336, Day 337 through Day 365, and > Day 365.

In order to summarize corticosteroid use over time in the treatment phase, for days of interest Day X (ie, X = 14, 28, 56, 100, and 180, separately), average corticosteroid dose (mg/day) during the week ending on Day X, relative corticosteroid dose (%) during the week ending on Day X will be summarized by treatment group. Number and percentage of subjects ongoing with itacitinib/placebo and ongoing with corticosteroid on Day X respectively will be summarized by treatment group.

6.6. Study Drug Compliance

For subjects in the safety analysis set, overall compliance (%) for study drug will be calculated for all subjects as follows:

$$\text{Overall compliance (\%)} = 100 \times [\text{total dose actually taken (mg)}] / [\text{total prescribed dose (mg)}]$$

The total prescribed dose is defined as the sum of the doses prescribed by the investigator accounting for dose modifications. The total dose actually taken (mg) is derived by sum of daily dose taken in a specific duration of treatment.

6.7. Prior and Concomitant Medication

Prior medications and concomitant medications will be coded using the WHO Drug Dictionary. For subjects in the safety analysis set, the number and percentage of subjects with prior and concomitant medications will be summarized by WHO drug class and WHO drug term. In the data listing, each medication will be recorded as prior, concomitant, or both prior and concomitant.

The number and percentage of subjects receiving prior GVHD prophylaxis medications will be summarized by treatment group.

The number and percentage of subjects who have ever taken immunosuppressive medications at baseline will be summarized overall and by treatment group. Subjects who took immunosuppressive medications in the 7 days before or on Days 1, 14, 28, 56, 100, and 180 during the treatment phase will be summarized by treatment group.

The number and percentage of subjects undergoing each type of conditioning regimen will be summarized by treatment group.

7. EFFICACY

[Appendix B](#) provides a list of planned data displays.

7.1. General Considerations

In all efficacy analyses, initiation of new, systemic anti-GVHD therapy includes 2 scenarios:

1) initiation of new anti-GVHD therapy including a higher dose of corticosteroid than that used at Day 1 and 2) initiation of new, systemic anti-GVHD therapy when subjects who were treated with placebo in the treatment phase were re-treated with itacitinib in the retreatment phase.

Initiation of new, systemic anti-GVHD therapy will be considered when defining the following:

1) onset of response and 2) loss of response. When defining responder for Day 28 ORR; overall response at Days 14, 56, 100, and 180; and BOR, the onset of response will be censored by initiation of new, systemic anti-GVHD therapy. Responders who did not have GVHD progression or death after response will be censored at the time of either study withdrawal, initiation of new, systemic anti-GVHD therapy, or the date of last assessment of GVHD before or on the cutoff date, whichever is earliest. In the survival analysis of NRM, all subjects in the analysis set in question without a death due to causes other than malignancy relapse will be censored at either study withdrawal, death due to malignancy relapse, or the date of last contact before or on the cutoff date, whichever is earliest.

Duration of response will be calculated for every subject who has a response in the study; DOR will also be summarized for any subgroups of interest.

7.2. Efficacy Hypotheses

The primary efficacy analysis is the comparison of ORR at Day 28 between the 2 treatment groups. The following statistical hypotheses will be tested to address the primary efficacy objective:

$$H_0: ORR_{itacitinib} = ORR_{pbo} \text{ versus } H_1: ORR_{itacitinib} > ORR_{pbo}$$

where $ORR_{itacitinib}$ and ORR_{pbo} are the ORRs at Day 28 in the itacitinib and placebo groups, respectively. The CMH chi-square test, stratified by the randomization stratification factor (ie, standard risk and high risk), will be used to compare ORR between the 2 treatment groups, at the 1-sided 2.5% level of significance.

For the primary efficacy analyses, Day 28 ORR, as well as ORR for scheduled visits at Days 14, 56, 100, and 180, if available, will be used; if a scheduled assessment is not available but an unscheduled visit is available (within 3 days of the scheduled date), then the unscheduled assessment closest to the scheduled date will be used.

For summary analyses over time, analysis visits are defined Section 4.1.5, where analysis visits will be used in both tables and figures. The listings will include all collected data from the eCRF regardless of whether the visits are scheduled or unscheduled.

7.3. Analysis of the Primary Efficacy Endpoint

For all of the following regression analyses specified, regardless of logistic regression, Cox regression, or cause-specific hazard/subdistribution hazard regression in competing-risks analyses, the covariates in the final model can differ from what is specified, depending on the unblinded data.

7.3.1. Primary Efficacy Analysis

7.3.1.1. Day 28 Overall Response Rate

The primary efficacy endpoint, ORR at Day 28 (ie, with either a CR, VGPR, or PR at Day 28 assessment or other response assessments within 3 days of Day 28, before start of new anti-GVHD therapy), will be analyzed using a stratified CMH test. Note that the EOT visit must be accounted for in the Day 28 ORR calculation, because when the EOT and Day 28 visits occur at the same time, per protocol, EOT will supersede the Day 28 visit. That is, if the EOT visit occurs when the Day 28 visit is planned, data from the EOT visit will be used in lieu of the Day 28 visit.

Responders ([CIBMTR 2009](#), [Harris et al 2016](#); refer to Protocol Appendix B):

- (CR) Complete response
- (VGPR) Very good partial response
- (PR) Partial response

Non-responders:

- (MR) Mixed response
- (PD) Progression of disease
- (NR) No response
- (NA) Not applicable
- Other
- Missing (primary reason)
 - Death prior to the Day 28 visit
 - Withdrawal from study prior to the Day 28 visit
 - Initiation of new anti-GVHD therapy prior to the Day 28 visit
 - Missing visit

Day 28 ORR will be summarized using descriptive statistics (N, %) by treatment group with 2-sided exact binomial 95% CIs. P-value, mean proportion difference, odds ratio, and 95% Wald confidence limits calculated from stratified CMH test will be also presented.

7.3.1.2. Sensitivity Analyses for Day 28 Overall Response Rate

Sensitivity analyses for Day 28 ORR will consist of the following:

- Stratified CMH test for Day 28 ORR in the full and per-protocol analysis sets
- Logistic regression method with both treatment group and baseline GVHD risk status as cofactors in the model for the efficacy population
- Separate CMH test by baseline GVHD risk status

7.3.2. Duration of Response for Day 28 Responders

The DOR is defined from the time of the first response to the loss of response (ie, either disease progression or death). Responders who did not progress or die after onset of response will be censored, and response will be ended at either study withdrawal, initiation of new anti-GVHD therapy, or last response assessment before or on the cut-off date, whichever is earliest.

Duration of response will be calculated for every responder as well as Day 28 responders.

The DOR for the Day 28 responders will be assessed using the KM method only for subjects who achieve Day 28 response. The KM curves; medians; 1-, 2-, 6-, 12-, 18- and 24-month (equivalently at Days 30, 60, 180, 365, 547, and 730) survival estimates and 95% CIs will be estimated ([Brookmeyer and Crowley 1982](#)). The DOR will be assessed for each treatment group and also by stratum and Day 28 response status.

7.3.3. Alternative Ways of Calculating Duration of Response for Day 28 Responders

There are 2 other ways of calculating DOR for Day 28 responders:

- **DOR alternate approach 1** – DOR is defined as the interval from Day 28 to either death or initiation of new anti-GVHD therapy, including a higher corticosteroid dose than that administered at Day 1, whichever is earlier.
- **DOR alternate approach 2** – DOR is defined as the interval from Day 28 to the earliest of death, worsening from best response at or after Day 28 in any organ stage, or initiation of new anti-GVHD therapy (including a higher corticosteroid dose than that administered at Day 1).

Summary statistics and KM estimate will be provided by treatment group. Kaplan-Meier plots will also be provided.

7.3.4. Subgroup Analyses for Primary Endpoint

The primary endpoint will be tabulated and forest plots for odds ratios will be generated for the following subgroups:

- Age group (18-65 vs > 65 years)
- MacMillan ([2015](#)) age group: (18-40 years vs >40 years)
- Sex
- Race (White vs non-White)

- Region (North America vs Western Europe vs Rest of world)
 - **North America:** Canada, United States
 - **Western Europe:** Belgium, Denmark, France, Germany, Italy, Netherlands, Spain, Sweden, Switzerland, United Kingdom, Austria, Finland, Greece, Portugal, Switzerland, Ireland
 - **Rest of world:** Australia, Chile, Israel, Japan, South Korea, Mexico, New Zealand, Poland, Russia, South Africa, Czech Republic, Hungary, Singapore, Taiwan
- HLA/donor source category (Identical, Matched related, Matched unrelated, Mismatched related, Mismatched unrelated, Other, Missing)
- Type of Allogeneic Transplant (Bone Marrow, Peripheral Blood Stem Cell, Cord Blood, Other)
- Baseline aGVHD grade
- Skin involvement at Baseline (Yes or No)
- Lower GI involvement at Baseline (Yes or No)
- Upper GI involvement at Baseline (Yes or No)
- Liver involvement at Baseline (Yes or No)
- Baseline GVHD risk status
- Underlying hematologic disease (malignant or non-malignant)
- Days from allo-HSCT to Day 1 (< 28 days and ≥ 28 days)
- Subjects aGVHD status determined after randomization (had aGVHD, did not have aGVHD)

7.4. Analysis of the Secondary Efficacy Endpoints

7.4.1. Key Secondary Endpoint Analysis

Nonrelapse mortality is defined as death due to causes other than malignancy relapse. For the primary analysis, the NRM rate will be estimated using the proportion of subjects with a death due to causes other than malignancy relapse within the subjects who were known alive at Day 180 or who had such a death. Reasons for not including a subject in this proportion include:

- Subject was alive and had not reached Day 180
- Death due to malignancy relapse
- Subject ended study participation prior to Day 180 (summarized by reason for study withdrawal)

A sensitivity analysis will also be conducted estimating the NRM rate using the KM estimator for each treatment.

7.4.1.1. Primary Analysis of Nonrelapse Mortality Rate at Month 6 (Day 180)

For the primary analysis, NRM will be analyzed using a stratified CMH test. Month 6 NRM will be summarized using descriptive statistics by treatment group with 2-sided exact binomial 95% CIs. P-value, mean proportion difference, odds ratio, and 95% Wald confidence limits calculated from stratified CMH test will also be presented.

Let $\hat{p}_{pbo,i}$ be the estimated proportion of evaluable subjects in the placebo group in stratum i who died due to causes other than malignancy relapse prior to or on Day 180 (ie, excluding those subjects who are either ongoing and have not reached Day 180, who withdrew from study prior to Day 180, or who died due to relapse) with corresponding variance estimator $\hat{V}(\hat{p}_{pbo,i})$. Let $\hat{p}_{trt,i}$ and $\hat{V}(\hat{p}_{trt,i})$ be the same for the itacitinib group in stratum i .

The overall estimate of treatment difference in Day 180 NRM between placebo and itacitinib is defined as the following:

$$w * [\hat{p}_{pbo,standard} - \hat{p}_{trt,standard}] + (1 - w) * [\hat{p}_{pbo,high} - \hat{p}_{trt,high}]$$

Its 95% CI is the above estimate of difference

$$\pm 1.96 * \text{sqrt} (w^2 * (\hat{V}(\hat{p}_{pbo,standard}) + \hat{V}(\hat{p}_{trt,standard})) + (1 - w)^2 * (\hat{V}(\hat{p}_{pbo,high}) + \hat{V}(\hat{p}_{trt,high})))$$

With weight w defined by inverse of variance in 2 baseline GVHD risk groups. That is:

$$w = \frac{1/[\hat{V}(\hat{p}_{pbo,standard}) + \hat{V}(\hat{p}_{trt,standard})]}{1/[\hat{V}(\hat{p}_{pbo,standard}) + \hat{V}(\hat{p}_{trt,standard})] + 1/[\hat{V}(\hat{p}_{pbo,high}) + \hat{V}(\hat{p}_{trt,high})]}$$

7.4.1.2. Sensitivity Analysis of Nonrelapse Mortality Rate at Month 6 (Day 180) Using Kaplan-Meier Approach

A sensitivity analysis will also be conducted where NRM at Day 180 will be estimated by 1 minus the survival probability at Day 180 using the KM estimator.

A stratified test statistic for NRM rate at Day 180 is proposed as:

$$TS_{Day180NRM} = \frac{\left[\sum_{j=1}^2 (\widehat{CDF}_{pbo,j}(t=D180) - \widehat{CDF}_{trt,j}(t=D180)) \right]^2}{\sum_{j=1}^2 Var(\widehat{CDF}_{pbo,j}(t=D180) - \widehat{CDF}_{trt,j}(t=D180))}$$

$$= \frac{\left[\sum_{j=1}^2 (\widehat{CDF}_{pbo,j}(t = D180) - \widehat{CDF}_{trt,j}(t = D180)) \right]^2}{\sum_{j=1}^2 (SE(\widehat{CDF}_{pbo,j}(t = D180))^2 + SE(\widehat{CDF}_{trt,j}(t = D180))^2)}$$

$\widehat{CDF}_{pbo,j}(t = D180)$ and $SE(\widehat{CDF}_{pbo,j}(t = D180))^2$ are the cumulative distribution function estimate at Day 180 in the placebo group in the jth stratum and its variance, respectively.

$\widehat{CDF}_{trt,j}(t = D180)$ and $SE(\widehat{CDF}_{trt,j}(t = D180))^2$ are defined similarly in the itacitinib group for jth stratum subjects. $TS_{Day180NRM}$ will be asymptotically distributed as a Chi-square test statistic with degree of freedom equal to 1.

The overall estimate of treatment difference in Day 180 NRM between placebo and itacitinib is thus defined as the following:

$$w * [\widehat{CDF}_{pbo,standard}(t = D180) - \widehat{CDF}_{trt,standard}(t = D180)] + (1 - w) * [\widehat{CDF}_{pbo,high}(t = D180) - \widehat{CDF}_{trt,high}(t = D180)]$$

Its 95% CI is the above estimate of difference

$$\pm 1.96 * \sqrt{w^2 * (SE(\widehat{CDF}_{pbo,standard}(t = D180)))^2 + (SE(\widehat{CDF}_{trt,standard}(t = D180)))^2 + (1 - w)^2 * (SE(\widehat{CDF}_{pbo,high}(t = D180)))^2 + SE(\widehat{CDF}_{trt,high}(t = D180))^2)}$$

with weight w defined by inverse of variance in two baseline GVHD risk groups. That is:

$$w = \frac{1/[SE(\widehat{CDF}_{pbo,standard}(t=D180))^2 + SE(\widehat{CDF}_{trt,standard}(t=D180))^2]}{1/[SE(\widehat{CDF}_{pbo,standard}(t=D180))^2 + SE(\widehat{CDF}_{trt,standard}(t=D180))^2] + 1/[SE(\widehat{CDF}_{pbo,high}(t=D180))^2 + SE(\widehat{CDF}_{trt,high}(t=D180))^2]}$$

In order to conduct above hypothesis testing for NRM at Day 180, time to NRM, defined as the interval between randomization and death due to causes other than malignancy relapse, will be derived. All subjects in the FAS without a death due to causes other than malignancy relapse will be censored at either study withdrawal, death due to malignancy relapse, or the date of last contact before or on the cutoff date, whichever is earliest.

7.4.1.3. Nonrelapse Mortality Rate Using Kaplan-Meier Approach

The survival distribution of NRM in each treatment group overall [REDACTED] will be estimated using the KM method, and the KM curves; medians; 1-, 2-, 3-, 6-, 9-, 12-, 18-, and 24-month survival estimates; and 95% CIs will be presented (Brookmeyer and Crowley 1982).

7.4.2. Survival Analyses of Nonrelapse Mortality

Time to death due to causes other than malignancy relapse (NRM) will be analyzed according to the randomized treatment group and strata assigned at randomization.

The hazard ratio of itacitinib relative to placebo overall [REDACTED] will be calculated, along with its 95% CI, using a stratified Cox model.

Stratified log-rank test will also be used to test survival curve of NRM in the treatment group relative to placebo group using the FAS.

Subgroup KM analyses of NRM will also be conducted by Day 28 response status [REDACTED].

7.4.3. Competing-Risks Analyses of Nonrelapse Mortality

In the literature, analyzing NRM is normally conducted in the context of a competing-risks situation. Time from randomization to first occurrence of NRM or competing event is observed and used for analysis; otherwise, the subject is censored. Nonrelapse mortality and a competing-risk event are mutually exclusive. This is because a subject either dies due to malignancy relapse, dies due to an event other than malignancy relapse, or is censored at last follow-up. There are 2 types of competing-risks analyses:

- I: Time to NRM competing with time to malignancy relapse
Time to first occurrence of either death due to reasons other than malignancy relapse (NRM) or malignancy relapse is defined as an event of interest and competing risk event, respectively. Subjects without an event (neither event of interest nor competing event) will be censored at the date of study withdrawal or date of last contact before or on the cutoff date, whichever is the earlier.
- II: Time to NRM competing with time to death due to malignancy relapse
Time to first occurrence of either death due to reasons other than malignancy relapse (NRM) or death due to malignancy relapse is defined as an event of interest and competing risk event, respectively. Subjects without an event (neither event of interest nor competing event) will be censored at the date of study withdrawal or date of last contact before or on the cutoff date, whichever is the earlier.

For each of the competing-risks analyses (ie, I and II), there are 2 ways of modeling conducted in this study. The first one is cause-specific hazard modeling and the second one is subdistribution hazard modeling.

In cause-specific hazard modeling, competing events will be censored, and a stratified log-rank test [REDACTED] will be used to test 2 KM curves with KMC estimate equals 1 minus KM estimate. In cause-specific hazard modeling, stratified Cox regression model [REDACTED] with a cofactor of treatment group will be used to estimate cause-specific hazard ratio (itacitinib vs placebo) and 95% CI.

In subdistribution hazard modeling, competing-risk events will be considered to remain in the risk set as described by Gray (1988). Stratified Gray's test [REDACTED] will be used to test equality of 2 cumulative incidence curves. Cumulative incidence function at Day 180 (SE) will be estimated. Stratified subdistribution hazard regression (Fine and Gray 1999) with cofactor of treatment group (itacitinib vs placebo) will be performed, and subdistribution hazard ratio (95% CI) on treatment effect (itacitinib vs placebo) will be estimated.

Cumulative incidence curve for each competing-risk analysis will be plotted overall by Day 28 response status [REDACTED]

7.4.4. Overall Response Rate at Days 14, 56, 100, and 180

Similar to Day 28 ORR, ORR at Days 14, 56, 100, and 180, respectively, is defined as the proportion of subjects demonstrating a CR, VGPR, or PR at a defined analysis visit/window (ie, ± 3 days of scheduled visit date).

The ORR on Days 14, 56, 100, and 180 will be estimated with a 95% CI. The proportion of responders will be estimated with 95% CIs. Confidence intervals will be calculated based on the exact method for binomial distributions. Summary statistics will be provided.

7.4.5. Best Overall Response

Best overall response for each subject is defined as best response during the course of the study (best in the categories of CR, VGPR, or PR), including responses at unscheduled visits.

Subjects without a BOR (ie, subjects not achieving a response of either CR, VGPR, or PR) in the treatment phase will be considered as nonresponders for BOR in the responder analysis. Number and percentage of subjects in each best response category will be tabulated by treatment group in the FAS.

7.4.6. Time to Response

Time to response is defined as time from randomization to the first response of either CR, VGPR, or PR before initiation of new anti-GVHD therapy. In the TTR analysis, subjects without a response after randomization will be censored at either study withdrawal, initiation of new, systemic anti-GVHD therapy, death, or last response assessment prior to or on the cutoff date, whichever is earliest.

The KM curves; medians; 1-, 2-, 6-, 12-, 18-, and 24-month (at Days 30, 60, 180, 365, 547, and 730, respectively) survival estimates, and 95% CIs will be estimated (Brookmeyer and Crowley 1982).

7.4.7. Relapse Rate of Malignant and Nonmalignant Hematologic Diseases

Relapse rate of malignant and nonmalignant hematologic diseases is defined as the proportion of subjects whose underlying hematologic diseases relapse. Incidence of such events will be tabulated with summary statistics.

7.4.8. Malignancy Relapse-Related Mortality Rate

Malignancy relapse-related mortality rate is defined as the proportion of subjects whose malignancies relapse and have a fatal outcome. Incidence of such events will be tabulated with summary statistics.

7.4.9. Failure-Free Survival

Failure-free survival is defined as the proportion of subjects who are still alive, have not relapsed of underlying hematologic disease, have not required additional therapy for aGVHD, and have not had signs or symptoms of cGVHD. Subjects without an event of relapse of underlying hematologic disease, death due to relapse, death due to causes other than relapse, initiation of new anti-GVHD therapy, or signs or symptoms of cGVHD will be censored at the date of study withdrawal or the date of last contact before or on the data cutoff date, whichever is earlier. After randomization, the FFS event date will be the first occurrence of one of the following:

- Initiation of anti-GVHD therapy, or
- Relapse/recurrence of underlying hematologic disease, or
- Death due to malignancy relapse, or
- Death due to causes other than malignancy relapse, or
- Signs or symptoms of cGVHD.

Failure free survival will be analyzed according to the randomized treatment group and strata assigned at randomization.

The survival distribution of FFS in each treatment group overall [REDACTED] will be estimated using the KM method, and the KM curves; medians; 1-, 2-, 6-, 12-, 18-, and 24-month (at Days 30, 60, 180, 365, 547, and 730, respectively) survival estimates, and 95% CIs will be presented (Brookmeyer and Crowley 1982). The hazard ratio of treatment relative to placebo overall [REDACTED] will be calculated, along with its 95% CI, using a stratified Cox model. Stratified log-rank test will also be used to test FFS survival curve in the itacitinib group against the placebo group.

Subgroup KM analyses of FFS will also be conducted by Day 28 response status [REDACTED]

7.4.10. Overall Survival

Overall survival is defined as the time from randomization to death due to any cause during the study. The KM method will be used to estimate the survival time distribution and the median survival. Subjects with no observed death or loss to follow-up will be treated as censored at either the date of study withdrawal or the date of last contact prior to or on the data cutoff date, whichever is later.

Overall survival will be analyzed according to the randomized treatment group and strata assigned at randomization.

The survival distribution of OS in each treatment group overall [REDACTED] will be estimated using the KM method, and the KM curves; medians; 1-, 2-, 6-, 12-, 18-, and 24-month (at Days 30, 60, 180, 365, 547, and 730, respectively) survival estimates, and 95% CIs will be presented (Brookmeyer and Crowley 1982). The hazard ratio of itacitinib relative to placebo overall [REDACTED] will be calculated, along with its 95% CI, using a stratified Cox model. Stratified log-rank test will also be used to test the OS survival curve in the itacitinib group against the placebo group.

Subgroup KM analyses of OS will also be conducted by Day 28 response status [REDACTED].

7.5. Other Secondary Efficacy Analyses

Secondary graft failure will be tabulated with summary statistics.

Incidence rate of aGVHD flares through Day 100 will be summarized. Summary statistics will be provided.

Average and cumulative corticosteroid dose during the weeks ending on Days 28, 56, 100, and 180 will be tabulated. Proportions of subjects who discontinue corticosteroids will be summarized.

Proportions of subjects who discontinue immunosuppressive medication on Days 56 and 100 will be tabulated. Summary statistics will be provided.

Incidence of cGVHD (Days 180 and 365 and overall) will be estimated. Summary statistics will be provided.

[REDACTED]

Subgroup analyses will be conducted by underlying hematologic disease (malignant vs nonmalignant), as appropriate.

7.6. Analysis of Exploratory Efficacy Variables

The following exploratory endpoints will be summarized by treatment group using the FAS:

- [REDACTED]
- [REDACTED]
- Summary of primary and contributing causes of death.
- Summary of donor chimerism.
- Summary of criteria and action taken for relapse of underlying disease.

7.7. Pharmacokinetic Analyses

A population PK analysis for itacitinib will be performed. Pharmacokinetic data from clinical Study INCB 39110-105 will be combined with data from this study, INCB 39110-301, and will aid in identification of the structural model and improve parameter estimation.

The population PK data preparation will be performed using SAS version 9.1 or later (SAS Institute Inc., Cary, NC). Exploratory data analyses and presentations of data will be performed using R version 3.2.0 or later (www.r-project.org), SPlus version 7.0 or later (Tobco Software Inc., Palo Alto, CA), and/or SAS version 9.1 or later. The PK analyses will use NONMEM version 7.1.0 or later (ICON Development Solutions, Dublin, Ireland) or Monolix version 2018R1 or later (Lixoft, Antony, France). The individual post hoc PK model parameters (CL/F, V/F, if applicable) will be determined and will be used to simulate individual exposures. Observed PK parameters such as C_{max} , t_{max} , and C_{min} will also be presented and summarized.

Simulation may be used to evaluate potential dose modifications if any covariates, for example, renal or hepatic impairment, are noted to have a clinically significant impact on exposure.

Additional details regarding this analysis may be found in the data analysis plan (DMB-18.155.1).

[REDACTED]

[REDACTED]

[REDACTED]

[REDACTED]

[REDACTED]

[REDACTED]

[REDACTED]

8. SAFETY AND TOLERABILITY

[Appendix B](#) provides a list of data displays.

8.1. General Considerations

All safety analyses will be conducted using the safety analysis set and summarized by actual treatment received on Day 1. Summary tables may be replaced with listings when appropriate. For instance, an AE frequency table may be replaced with a listing if it only contains a few unique PTs reported on relatively few subjects.

For summary analyses over time, analysis visits for laboratory assessments/vital signs/ECGs are defined in Section [4.1.5](#).

Listings will include all collected visits from the eCRF regardless of whether the visits are scheduled or unscheduled.

8.2. Adverse Events

8.2.1. Adverse Event Definitions

A TEAE in the treatment phase is either an AE reported for the first time or worsening of a pre-existing condition after first dose of itacitinib/placebo and before or at 30 days after last dose in the treatment phase. Analysis of AEs will be limited to TEAEs that occur during the treatment phase, but data listings will include all AEs, regardless of their timing with regard to study treatment administration.

A TEAE in the retreatment phase is either an AE reported for the first time or worsening of a pre-existing condition on or after Day 1 of retreatment with itacitinib. Treatment-emergent AEs in the retreatment phase will be listed.

Adverse events will be tabulated by MedDRA PT and SOC. Severity of AEs will be described and graded using the NCI CTCAE v4.03 ([NCI 2010](#)). The CTCAE reporting guidelines and grading details are available on the Cancer Therapy Evaluation Program website.

A grading (severity) scale is provided for each AE term. If the toxicity is not included in the CTCAE v4.03 criteria, then it will be rated on a scale of 1 to 5 as follows: 1 = mild, 2 = moderate, 3 = severe, 4 = life-threatening, and 5 = fatal. All toxicities will be graded based on the worst level reached, not the level they may have reached if they had not been treated. When the intensity of an AE changes over time for a reporting period (eg, between visits), then each change in intensity will be reported as an AE until the event resolves.

The subset of AEs considered by the investigator to be related to study drug will be considered to be treatment-related TEAEs. If the investigator does not specify the relationship of the AE to study drug, then the AE will be considered to be treatment-related. The incidence of TEAEs and treatment-related TEAEs will be tabulated. Serious TEAEs will also be tabulated.

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

- An unresolved missing causality will be considered treatment-related.
- An unresolved missing severity will be identified as an unknown severity.

For purposes of analysis, all AEs will be considered TEAEs unless the AE can unequivocally be defined as not treatment-emergent.

8.2.2. Adverse Event Summaries

An overall summary of AEs in the treatment phase by treatment group will include the following:

- Number (%) of subjects reporting any TEAEs
- Number (%) of subjects reporting any serious TEAEs
- Number (%) of subjects reporting any Grade 3 or higher TEAEs
- Number (%) of subjects reporting any TEAEs related to itacitinib/placebo
- Number (%) of subjects reporting any TEAEs related to corticosteroids
- Number (%) of subjects who temporarily interrupted itacitinib/placebo because of TEAEs
- Number (%) of subjects with itacitinib/placebo dose reduction because of TEAEs
- Number (%) of subjects who permanently discontinued itacitinib/placebo because of TEAEs
- Number (%) of subjects who permanently discontinued itacitinib/placebo and corticosteroids because of TEAE
- Number (%) of subjects who had a fatal TEAE
- Number (%) of subjects who withdrew from the study because of a TEAE

Overall summaries of AEs (as described above) will also be provided by age group, sex, race, and region.

The following summaries will be produced for TEAEs occurring in the treatment phase by MedDRA term (if 10 or fewer subjects appear in a table, a listing may be appropriate):

- Number (%) of subjects reporting TEAEs by SOC and PT
- Number (%) of subjects reporting TEAEs by PT in decreasing order of frequency
- Number (%) of subjects reporting TEAEs by SOC, PT, and highest grade
- Number (%) of subjects reporting Grade 3 or higher TEAEs by SOC and PT
- Number (%) of subjects reporting Grade 3 or higher TEAEs by PT in decreasing order of frequency

- Number (%) of subjects reporting itacitinib/placebo treatment-related TEAEs by SOC and PT
- Number (%) of subjects reporting itacitinib/placebo treatment-related TEAEs by PT in decreasing order of frequency
- Number (%) of subjects reporting itacitinib/placebo treatment-related TEAEs by SOC, PT, and highest grade
- Number (%) of subjects reporting Grade 3 or higher itacitinib/placebo treatment-related TEAEs by SOC and PT
- Number (%) of subjects reporting fatal TEAEs by SOC and PT
- Number (%) of subjects reporting serious TEAEs by SOC and PT
- Number (%) of subjects with serious TEAEs by PT in decreasing order of frequency
- Number (%) of subjects reporting itacitinib/placebo treatment-related serious TEAEs by SOC and PT
- Number (%) of subjects reporting TEAEs leading to itacitinib/placebo dose interruption by SOC and PT
- Number (%) of subjects reporting TEAEs leading to itacitinib/placebo dose reduction by SOC and PT
- Number (%) of subjects reporting TEAEs leading to discontinuation of itacitinib/placebo by SOC and PT
- Number (%) of subjects reporting corticosteroid treatment-related TEAEs by SOC and PT
- Number (%) of subjects reporting corticosteroid treatment-related TEAEs by PT in decreasing order of frequency
- Number (%) of subjects reporting corticosteroid treatment-related TEAEs by SOC, PT, and highest grade
- Number (%) of subjects reporting corticosteroid treatment-related Grade 3 or higher TEAEs by SOC and PT
- Number (%) of subjects reporting corticosteroid treatment-related serious TEAEs by SOC and PT
- Number (%) of subjects reporting TEAEs leading to corticosteroid dose interruption by SOC and PT
- Number (%) of subjects reporting TEAEs leading to corticosteroid dose reduction by SOC and PT
- Number (%) of subjects reporting TEAEs leading to discontinuation of corticosteroids by SOC and PT
- Number (%) of subjects reporting TEAEs leading to withdrawal from the study by SOC and PT

The following summaries will be produced for TEAEs occurring in the treatment phase by MedDRA term based on baseline hepatic impairment status in combination with postbaseline azole usage:

- Number (%) of subjects reporting serious TEAEs by SOC and PT
- Number (%) of subjects reporting Grade 3 or higher TEAEs by SOC and PT
- Number (%) of subjects reporting fatal TEAEs by SOC and PT
- Number (%) of subjects reporting itacitinib/placebo treatment-related TEAEs by SOC and PT
- Number (%) of subjects reporting Grade 3 or higher thrombocytopenia TEAEs by SOC and PT based on customized MQ
- Number (%) of subjects reporting Grade 3 or higher neutropenia TEAEs by SOC and PT based on customized MQ
- Number (%) of subjects reporting Grade 3 or higher anemia TEAEs by SOC and PT based on customized MQ

Severity of hepatic impairment at baseline is defined as follows:

- Normal: Total bilirubin \leq ULN and AST \leq ULN
- Mild: Total bilirubin > 1.0 to $1.5 \times$ ULN and AST = any value or total bilirubin \leq ULN and AST $>$ ULN
- Moderate: Total bilirubin > 1.5 to $3 \times$ ULN and AST = any value
- Severe: Total bilirubin $> 3.0 \times$ ULN and AST = any value

If multiple azoles were used postbaseline, the most potent azole will be used for classification according to the following potency order (listed from strongest to weakest):

1. Ketoconazole
2. Itraconazole
3. Voriconazole
4. Posaconazole
5. Fluconazole

8.3. Clinically Notable Treatment-Emergent Adverse Events Based on Customized MedDRA Queries

The following categories of clinically notable events will be summarized based on customized MQs (see [Appendix C](#)):

- Anemia
- Neutropenia
- Thrombocytopenia
- QT prolongation
- Hyperlipidemia
- EBV infection
- CMV infection

Within each category, the number and percentage of subjects with at least 1 TEAE occurring during the treatment phase tabulated by MedDRA PT will be summarized. Summaries of clinically notable TEAEs based on customized MQs will be provided by treatment group, category, PT, and the highest grade or outcome.

- Overall summary of clinically notable TEAEs in the treatment phase based on customized MQs
- Number (%) of subjects with clinically notable TEAEs in the treatment phase based on customized MQs by PT and outcome
- Number (%) of subjects with clinically notable TEAEs in the treatment phase based on customized MQs by PT and highest grade

8.4. Clinical Laboratory Tests

Laboratory test values outside the normal range will be assessed for severity based on the normal ranges for the clinical reference laboratory. The incidence of abnormal laboratory values and shift tables relative to baseline will be tabulated.

Laboratory data will be classified into CTC grades for AEs (CTCAE v4.03). The following summaries will be produced for the laboratory data:

- Shift tables using CTC grades to compare baseline with the worst postbaseline value will be produced with CTC grade.
- For laboratory parameters where CTC grades are not defined, shift tables to the worst postbaseline value will be produced using the low/normal/high classifications based on laboratory reference ranges.

Numeric laboratory values will be summarized descriptively in SI units, and non-numeric test values will be tabulated when necessary. In addition, line graphs will be provided for platelets, neutrophils, cholesterol, and triglycerides over time in the treatment phase, plotted by treatment group.

8.5. Vital Signs

Descriptive statistics and mean change from baseline will be summarized for vital signs (blood pressure, pulse, respiratory rate, and body temperature) at each assessment time by treatment group. Vital sign results will be reviewed for clinically notable abnormalities (see [Table 6](#)), and subjects exhibiting clinically notable vital sign abnormalities will be listed. A value will be considered an "alert" value if it is outside the established range and shows a > 25% change from baseline.

Table 6: Criteria for Clinically Notable Vital Sign Abnormalities

Parameter	High Threshold	Low Threshold
Systolic blood pressure	> 155 mmHg	< 85 mmHg
Diastolic blood pressure	> 100 mmHg	< 40 mmHg
Pulse	> 100 bpm	< 45 bpm
Temperature	> 38°C	< 35.5°C
Respiratory rate	> 24 breaths/min	< 8 breaths/min

8.6. Electrocardiograms

Electrocardiogram results will be reviewed for clinically notable abnormalities according to predefined criteria (see [Table 7](#)). Subjects exhibiting clinically notable ECG abnormalities will be listed and summarized by treatment group. Adverse events will be reported for clinically notable abnormalities that are considered clinically significant in the judgment of the investigator.

Descriptive statistics (N, mean [SD], minimum, median, and maximum) at baseline for each parameter will be summarized by treatment group. At EOT, summary statistics of mean change (SD) and percentage change will also be tabulated by parameter and treatment group.

Table 7: Criteria for Clinically Notable ECG Abnormalities

Parameter	High Threshold	Low Threshold
QTcF	> 460 ms	< 295 ms
PR	> 220 ms	< 75 ms
QRS	> 120 ms	< 50 ms
QT	> 500 ms	< 300 ms
RR	> 1330 ms	< 600 ms

QTcF = Fridericia correction.

9. INTERIM ANALYSES AND FINAL ANALYSES

9.1. Overview of Interim Analyses

There are 2 planned interim analyses. The first interim analysis is for futility checking and second interim analysis is for efficacy and futility.

9.1.1. Interim Analysis 1

The first interim analysis for futility only will be performed when 112 (25.7% of the planned 436 subjects) have completed the Day 28 visit or withdrew early from study. No alpha will be spent during the first interim analysis. The stopping rule is nonbinding. When the conditional power is < 20%, the sponsor may consider terminating the trial if the totality of evidence is warranted.

Interim Analysis 1 on ORR at Day 28 will be carried out as follows:

F1: Obtain CMH test statistic (chi-square test with degree of freedom 1);

F2: Obtain observed normalized test statistic z by taking square root of CMH in F1;

F3: Calculate conditional power based on $\Phi\left(\frac{z}{\sqrt{t(1-t)}} - \frac{z_{\alpha/2}}{\sqrt{1-t}}\right)$ (Jennison and Turnbull 1999),

where Φ is the cumulative distribution function of the observed normalized test statistic z . t is the information fraction and (ie, observed sample at interim divided by planned total sample size 436); $t = 0.257$ at Interim 1 and $z_{\alpha/2} = 1.96$.

F4: If conditional power is less than 20%, the sponsor may consider stopping the study for futility; otherwise, the study will continue.

9.1.2. Interim Analysis 2

For Interim Analysis 2, the analysis of both efficacy and futility will use a clinical database cutoff based on when 240 eligible subjects (55.0% of the targeted 436 subjects) have completed the Day 180 visit or withdrawn from the study prior to Day 180.

The Hwang-Shih-DeCani alpha spending function with shape parameter -4 (O'Brien Fleming like boundary) will be used to define nonbinding upper efficacy boundary for the primary analysis of ORR at Day 28 when approximately 240 eligible subjects reach Day 180. Based on enrollment projection, there will be approximately 360 subjects who have completed Day 28.

The futility analysis based on test of ORR at Day 28 at Interim Analysis 2 will be carried out the same way as in Interim Analysis 1 (Steps F1-F4) but using the observed information fraction and observed normalized test statistic from Interim Analysis 2.

[Table 2](#) lists efficacy boundary when information fraction (ie, the percentage of subjects) is 0.826 (360 subjects complete Day 28 or withdraw from the study) at Interim Analysis 2. If the test on ORR at Day 28 crosses the boundary, then the key secondary endpoint, 6-month NRM rate, will be tested using a separate boundary using the Bonferroni method for alpha-spending (ie, 0.0125 as compared with 1-sided p-value at both interim and final analyses). Note that [Table 2](#) will be re-derived when the observed information for ORR at Day 28 is not 0.826.

Therefore, efficacy analysis at Interim 2, assuming information fraction being 0.826 (ie, 360/436) will be conducted as follows:

- E1: Obtain CMH test statistic on ORR at Day 28 (chi-square test with degree of freedom 1),
- E2: CMH test statistic will be compared with the critical value in Column 7 of [Table 2](#) (Interim 2 chi-square test critical value),
- E3: If positive on ORR at Day 28, NRM rate at Day 180 will be tested using the Bonferroni method,
- E4: If neither efficacy nor futility on Day 28 ORR is crossed, study will continue.

Assuming that at the Interim Analysis 2 data cutoff date (Day 180 of the 240th randomized eligible subject) there are approximately 360 subjects based on estimation who have been randomized, then testing of Day 28 ORR will be conducted for subjects who have completed Day 28 or withdrawn from the study upon the cutoff date. This means that, by the Interim Analysis 2 data cutoff date, subjects who were ongoing but had not reached the Day 28 visit will be excluded for both the Day 28 ORR and Day 180 NRM analyses.

Analysis of NRM rate at Day 180 will use the same method as described in Section [7.4.1](#) for final analysis but using the Interim Analysis 2 cutoff data instead. The decision process is described in [Figure 4](#).

At the Interim Analysis 2 data cutoff date, if the total number of randomized subjects in the Interim 2 analysis set is different from the predicted number 360, the efficacy boundaries will be recalculated based on the actual observed information fraction at Interim Analysis 2.

An independent DMC will be charged with evaluating interim results. The DMC will consist of clinicians and an independent statistician. The sponsor will remain blinded, and DMC decisions will be communicated through sponsor management as dictated in the DMC charter. Additional operational details of the interim analyses, including tables, figures, and listings provided to the DMC, will be provided in the DMC Charter.

9.2. Final Analyses

Final analyses are for efficacy of both ORR at Day 28 and NRM rate at Day 180. Assuming the information fraction at Interim Analysis 2 is 0.826 (360 subjects complete Day 28), efficacy boundaries in [Table 2](#) will be used; if not, the efficacy boundaries will be regenerated using observed information fraction at Interim Analysis 2.

The test steps are as follows:

- E1: Obtain CMH test statistic on ORR at Day 28 (chi-square test with degree of freedom 1),
- E2: CMH test statistic will be compared with the critical value in Column 7 of [Table 2](#) (final analysis chi-square test critical value),
- E3: If positive on ORR at Day 28, NRM rate at Day 180 will be tested using the Bonferroni method.

Note that if there is substantial over-running at the end of the study (ie, total number of subjects is greater than 446), the final critical value will be adjusted in order to control overall Type I error rate.

The final analyses will be conducted by the sponsor.

10. CHANGES AND MODIFICATIONS TO THE ANALYSIS PLAN

All versions of the SAP are listed in [Table 8](#).

Table 8: Statistical Analysis Plan Versions

SAP Version	Date
Original	14 MAY 2018
Amendment 1	23 JUL 2019

10.1. Changes to Protocol-Defined Analyses

Note that the FAS population is equivalent to the efficacy evaluable population as defined in the protocol.

10.2. Changes to the Statistical Analysis Plan

10.2.1. Original to Amendment 1

The Type 1 error control approach for the key secondary endpoint tested at Interim Analysis 2 and the Final Analysis was modified to use the Bonferroni method. Portions of Sections 3 and 9 were modified to reflect that change. In addition, the definition of endpoint for and analysis of Month 6 NRM was refined to reflect Health Authority feedback.

- Section 2 was updated to align with Protocol Amendment 4.
- Section 3 was revised [REDACTED] to provide detail on the interim and final analyses. Figures were added to help visualize the study design and analyses, and the schedule of assessments tables were updated according to Protocol Amendment 4.
- Section 4 was updated to provide additional clarification and definitions.
- Section 5.3 was updated to add the analysis sets used at Interim Analyses 1 and 2. “FAS” replaced “efficacy evaluable population,” and the definition of per protocol analysis set was refined.
- Section 6 was updated to provide clarification and more details on analyses. Disease history was updated according to the current approved CRF, and sections were added for allo-HSCT history and transfusions.
- Section 7 was revised to provide additional clarification and details on efficacy analyses. Alternative ways of calculating DOR were added. Subgroup analyses by [REDACTED], [REDACTED], aGVHD status determined after randomization, MacMillan age group, HLA/donor source category, aGVHD status determined after randomization, and Day 28 response status were added. Timing of 12-, 18-, and 24-month analyses was specified.
- Section 8 was updated for clarification and to add TEAE in the retreatment phase, AE Grade 5, and more AE summaries. The clinically notable AEs were rewritten based on customized MQs. Analyses for EBV in post-transplant lymphoproliferative disorder evaluation were also added.
- Section 9.1 was edited to correct the conditional power formula and to provide more clarification on analyses at Interim Analysis 2.
- [Appendix A](#) was added.
- [Appendix B](#) was updated to redefine TFLs.
- [Appendix C](#) was updated to redefine customized MQs.
- [Appendix D](#) was added for other possible efficacy boundaries for testing Day 28 ORR at both the Interim 2 and final analyses.

In addition, other minor, administrative changes have been incorporated throughout the SAP and are noted in the redline version of the amendment.

11. REFERENCES

- Armitage P, McPherson CK, Rowe BC. Repeated significance tests on accumulating data. *Journal of the Royal Statistical Society* 1969;132:235-244.
- Brookmeyer R, Crowley J. A confidence interval for the median survival time. *Biometrics* 1982;38:29-41.
- Center for International Blood & Marrow Transplant Research. Clinical trial endpoints for patients with acute GVHD. 2009. <https://www.cibmtr.org/Meetings/Materials/GVHDworkshop/pages/index.aspx>. Accessed February 5, 2015.
- Collett D. *Modeling Survival Data in Medical Research*. London: Chapman & Hall; 1994.
- Fine JP, Gray RJ. A proportional hazards model for the subdistribution of a competing risk. *Journal of the American Statistical Association* 1999;94:496 -509.
- Gray RJ. A Class of k-sample tests for comparing the cumulative incidence of a competing risk. *The Annals of Statistics* 1988;16:1141-1154.
- Harris AC, Young R, Devine S, et al. International, multicenter standardization of acute graft-versus-host disease clinical data collection: a report from the Mount Sinai Acute GVHD International Consortium. *Biol Blood Marrow Transplant* 2016;22:4-10.
- Hwang IK, Shih WJ, De Cani JS. Group sequential designs using a family of type I error probability spending functions. *Stat Med* 1990;9:1439-1445.
- Jennison C, Turnbull BW. *Group Sequential Methods With Applications to Clinical Trials*. 1st ed. Boca Raton, FL: Chapman and Hall; 1999.
- MacMillan ML, Robin M, Harris AC, et al. A refined risk score for acute graft-versus-host disease that predicts response to initial therapy, survival, and transplant-related mortality. *Biol Blood Marrow Transplant* 2015;21:761-767.
- Mantel N, Fleiss JL. Minimum expected cell size requirements for the Mantel-Haenszel one-degree-of-freedom chi-square test and a related rapid procedure. *Am J Epidemiol* 1980;112:129-134.
- Mosteller RD. Simplified calculation of body-surface area. *N Engl J Med* 1987;317:1098.
- National Cancer Institute. *Common Terminology Criteria for Adverse Events (CTCAE) Version 4.03*. 2010. http://evs.nci.nih.gov/ftp1/CTCAE/CTCAE_4.03/CTCAE_4.03_2010-06-14_QuickReference_5x7.pdf. Accessed December 1, 2017.

APPENDIX A. STATISTICAL MODELS FOR EFFICACY ANALYSES

Analysis for Overall Response

Cochran-Mantel-Haenszel test

The null hypothesis of equality of response rate in the 2 treatment groups will be tested against 1-sided alternative. The statistical hypotheses are:

$$H_0: \text{ORR}_{\text{trt}} \leq \text{ORR}_{\text{pbo}} \text{ versus } H_1: \text{ORR}_{\text{trt}} > \text{ORR}_{\text{pbo}}$$

where ORR_{trt} is the probability of response in treatment group and ORR_{pbo} is the probability of response in placebo group. The CMH chi-square test X2CMH (implemented again via SAS procedure FREQ with CMH option in the TABLES statement) will be used to test the difference in response rates between the treatment groups. The p-value corresponding to the CMH test for “general association” will be used, which follows a Chi-square distribution with 1 degree of freedom.

If the sampling assumptions for chi-square test is not met, the exact CMH test will be used (implemented via SAS procedure MULTTEST). The test is performed by running a stratified version of the Cochran-Armitage permutation test (Armitage et al 1969). In studies with stratified randomization, the chi-square approximation is considered appropriate for the X2CMH statistics if the rule of Mantel and Fleiss (1980) is satisfied.

Logistic Regression

Odds ratio will be used as a measure of association between treatment and response. The odds ratio will be derived from the logistic regression model (implemented using SAS procedure LOGISTIC, with both treatment and stratification factor specified as an explanatory variable in the CLASS statement), which allows for including not only the stratification factor but also for adjustments for other covariates (both categorical and continuous). The odds ratio will be presented with 95% Wald confidence limits.

In cases where an exact test has been used to compare response rates, the odds ratio should be determined using exact logistic regression, and the odds ratio presented with exact 95% confidence limits. In these cases, SAS PROC LOGISTIC with EXACTONLY option will be used.

Analyses for Survival Data

Analyses for Survival Data in this section use OS as an example.

Stratified Log-Rank Test

The null hypothesis stating that OS distributions of the 2 treatment groups are equivalent will be tested against 1-sided alternative

Assuming proportional hazards for OS, the following statistical hypotheses will be tested:

$$H_{02}: \theta_2 \geq 1 \text{ versus } H_{A2}: \theta_2 < 1$$

where θ_2 is the OS hazard ratio (treatment versus placebo). The analysis to test these hypotheses will consist of a stratified log-rank test at an overall 1-sided 2.5% level of significance. The stratification will be based on the randomization stratification factors (ie, aGVHD high vs standard).

The stratified log-rank test adjusting for the strata used in the randomization will be implemented as follows: In each of the K strata separately, the LIFETEST procedure with **STRATA** statement including only the treatment group variable and with the TIME statement will be used to obtain the rank statistic S_k and variance $\text{var}(S_k)$ where $k=1, 2, \dots, K$. The final test statistics will then be reconstructed as follows:

$Z = [S_1 + \dots + S_K] / \sqrt{[\text{var}(S_1) + \dots + \text{var}(S_k)]}$. One-sided P-value will be obtained from the standard normal distribution using the Z statistic.

(One-sided will be obtained using Z statistic / Two-sided p-value will be obtained using Z^2 statistic.)

Kaplan-Meier Estimates

An estimate of the OS survival function in each treatment group at distinct event times will be constructed using KM (product-limit) method as implemented in PROC LIFETEST with METHOD=KM option. The PROC LIFETEST statement will use the option CONFTYPE=LOGLOG. Stratification factor (high vs standard) will be use in **STRATA** statement to get KM estimates by both treatment group and stratum.

Median survival for each treatment group will be obtained along with 95% CIs calculated from PROC LIFETEST output using the method of ([Brookmeyer and Crowley 1982](#)). Kaplan-Meier estimates of the survival function with 95% CIs at specific time points will be summarized. The standard error of the KM estimate will be calculated using Greenwood's formula ([Collett 1994](#)).

Hazard Ratio

Hazard ratio will be estimated by fitting the Cox proportional hazards model using SAS procedure PHREG.

A stratified unadjusted Cox model will be, that is, the MODEL statement will include the treatment group variable as the only covariate and the **STRATA** statement will include stratification variable(s).

Hazard ratio with 2-sided 95% CI will be based on Wald test.

The above survival analyses (ie, log-rank test, KM method and Cox proportional hazards model) do not consider competing risk to be an issue.

Table No.	Title	Population	Standard	In-Text
2.2.31	Cox Regression Analysis of Overall Survival	FAS		X
████	██	████		█
████	██	████		█
2.2.34	Summary of Post-Transplant Lymphoproliferative Disorder Assessment	FAS		X
2.2.35	Summary of ECOG Status in the Treatment Phase	FAS		X
2.2.36	Summary of Death	FAS		X
2.2.37	Summary of Relapse of Underlying Hematologic Disease	FAS		X
Exposure				
3.1.1	Summary of Itacitinib Dose Modifications in the Treatment Phase	Safety		X
3.1.2	Summary of Corticosteroid Exposure Prior to Day 1	Safety		X
3.1.3	Summary of Corticosteroid Exposure in the Treatment Phase	Safety		X
3.1.4	Summary of Corticosteroid Exposure Between Day 1 and Start of New Anti-GVHD Therapy	Safety		X
3.1.5	Summary of Corticosteroid Use Over Time in the Treatment Phase	Safety		X
3.1.4.1	Summary of Itacitinib/Placebo Exposure and Compliance in the Treatment Phase	Safety		X
3.1.4.2	Summary of Itacitinib/Placebo Exposure and Compliance in the Treatment Phase by Age Group	Safety		X
3.1.4.3	Summary of Itacitinib/Placebo Exposure and Compliance in the Treatment Phase by Sex	Safety		X
3.1.4.4	Summary of Itacitinib/Placebo Exposure and Compliance in the Treatment Phase by Race	Safety		X
3.1.4.5	Summary of Itacitinib/Placebo Exposure and Compliance in the Treatment Phase by Region	Safety		X
Adverse Events				
3.2.1.1	Overall Summary of Treatment-Emergent Adverse Events in the Treatment Phase	Safety	X	X
3.2.1.2	Overall Summary of Treatment-Emergent Adverse Events in the Treatment Phase by Age Group	Safety	X	X
3.2.1.3	Overall Summary of Treatment-Emergent Adverse Events in the Treatment Phase by Sex	Safety	X	X
3.2.1.4	Overall Summary of Treatment-Emergent Adverse Events in the Treatment Phase by Race	Safety	X	X
3.2.1.5	Overall Summary of Treatment-Emergent Adverse Events in the Treatment Phase by Region	Safety	X	X
3.2.2	Summary of Treatment-Emergent Adverse Events in the Treatment Phase by MedDRA System Organ Class and Preferred Term	Safety	X	X
3.2.3	Summary of Treatment-Emergent Adverse Events in the Treatment Phase by MedDRA Preferred Term in Decreasing Order of Frequency	Safety	X	X
3.2.4	Summary of Treatment-Emergent Adverse Events in the Treatment Phase by MedDRA System Organ Class, Preferred Term, and Highest Grade	Safety	X	X
3.2.5.1	Summary of Grade 3 or Higher Treatment-Emergent Adverse Events in the Treatment Phase by MedDRA System Organ Class and Preferred Term	Safety	X	X

Table No.	Title	Population	Standard	In-Text
3.2.5.2	Summary of Grade 3 or Higher Treatment-Emergent Adverse Events in the Treatment Phase by MedDRA Preferred Term in Decreasing Order of Frequency	Safety	X	X
3.2.6	Summary of Itacitinib/Placebo Treatment-Related Treatment-Emergent Adverse Events in the Treatment Phase by MedDRA System Organ Class and Preferred Term	Safety	X	X
3.2.7	Summary of Itacitinib/Placebo Treatment-Related Treatment-Emergent Adverse Events in the Treatment Phase by MedDRA Preferred Term in Decreasing Order of Frequency	Safety	X	X
3.2.8	Summary of Itacitinib/Placebo Treatment-Related Treatment-Emergent Adverse Events in the Treatment Phase by MedDRA System Organ Class, Preferred Term, and Highest Grade	Safety	X	X
3.2.9	Summary of Grade 3 or Higher Itacitinib/Placebo Treatment-Related Treatment-Emergent Adverse Events in the Treatment Phase by MedDRA Preferred Term in Decreasing Order of Frequency	Safety	X	X
3.2.10	Summary of Treatment-Emergent Adverse Events With a Fatal Outcome in the Treatment Phase by MedDRA System Organ Class and Preferred Term	Safety	X	X
3.2.11	Summary of Serious Treatment-Emergent Adverse Events in the Treatment Phase by MedDRA System Organ Class and Preferred Term	Safety	X	X
3.2.12	Summary of Serious Treatment-Emergent Adverse Events in the Treatment Phase by MedDRA Preferred Term and in Decreasing Order of Frequency	Safety	X	X
3.2.13	Summary of Itacitinib/Placebo Treatment-Related Serious Treatment-Emergent Adverse Events in the Treatment Phase by MedDRA System Organ Class and Preferred Term	Safety	X	X
3.2.14	Summary of Treatment-Emergent Adverse Events Leading to Itacitinib/Placebo Dose Interruption in the Treatment Phase by MedDRA System Organ Class and Preferred Term	Safety	X	X
3.2.15	Summary of Treatment-Emergent Adverse Events Leading to Itacitinib/Placebo Dose Reduction in the Treatment Phase by MedDRA System Organ Class and Preferred Term	Safety	X	X
3.2.16	Summary of Treatment-Emergent Adverse Events Leading to Discontinuation of Itacitinib/Placebo in the Treatment Phase by MedDRA System Organ Class and Preferred Term	Safety	X	X
3.2.17	Summary of Corticosteroid Treatment-Related Treatment-Emergent Adverse Events in the Treatment Phase by MedDRA System Organ Class and Preferred Term	Safety	X	
3.2.18	Summary of Corticosteroid Treatment-Related Treatment-Emergent Adverse Events in the Treatment Phase by MedDRA Preferred Term in Decreasing Order of Frequency	Safety	X	
3.2.19	Summary of Corticosteroid Treatment-Related Treatment-Emergent Adverse Events in the Treatment Phase by MedDRA System Organ Class, Preferred Term, and Highest Grade	Safety	X	
3.2.20	Summary of Corticosteroid Treatment-Related Grade 3 or Higher Treatment-Emergent Adverse Events in the Treatment Phase by MedDRA System Organ Class, Preferred Term, and Highest Grade	Safety	X	

Table No.	Title	Population	Standard	In-Text
3.2.21	Summary of Corticosteroid Treatment-Related Serious Treatment-Emergent Adverse Events in the Treatment Phase by MedDRA System Organ Class and Preferred Term	Safety	X	
3.2.22	Summary of Treatment-Emergent Adverse Events Leading to Corticosteroid Dose Interruption in the Treatment Phase by MedDRA System Organ Class and Preferred Term	Safety	X	X
3.2.23	Summary of Treatment-Emergent Adverse Events Leading to Corticosteroid Dose Reduction in the Treatment Phase by MedDRA System Organ Class and Preferred Term	Safety	X	X
3.2.24	Summary of Treatment-Emergent Adverse Events Leading to Discontinuation of Corticosteroid in the Treatment Phase by MedDRA System Organ Class and Preferred Term	Safety	X	X
3.2.25	Summary of Treatment-Emergent Adverse Events Leading to Withdrawal From the Study in the Treatment Phase by MedDRA System Organ Class and Preferred Term	Safety	X	X
3.2.26	Overall Summary of Clinical Notable Treatment-Emergent Adverse Events in the Treatment Phase Based on Customized MQs	Safety	X	X
3.2.27	Summary of Clinically Notable Treatment-Emergent Adverse Events in the Treatment Phase Based on Customized MQs by MedDRA Preferred Term and Outcome	Safety	X	X
3.2.28	Summary of Clinically Notable Treatment-Emergent Adverse Events in the Treatment Phase Based on Customized MQs by MedDRA Preferred Term and Highest Grade	Safety	X	X
3.2.29	Summary of Serious Treatment-Emergent Adverse Events in the Treatment Phase by MedDRA System Organ Class, Preferred Term, Hepatic Impairment Status at Baseline, and Postbaseline Azole Usage	Safety	X	X
3.2.30	Summary of Grade 3 or Higher Treatment-Emergent Adverse Events in the Treatment Phase by MedDRA System Organ Class, Preferred Term, Hepatic Impairment Status at Baseline, and Postbaseline Azole Usage	Safety	X	X
3.2.31	Summary of Treatment-Emergent Adverse Events With a Fatal Outcome in the Treatment Phase by MedDRA System Organ Class, Preferred Term, Hepatic Impairment Status at Baseline, and Postbaseline Azole Usage	Safety	X	X
3.2.32	Summary of Itacitinib/Placebo Treatment-Related Treatment-Emergent Adverse Events in the Treatment Phase by MedDRA System Organ Class, Preferred Term, Hepatic Impairment Status at Baseline, and Postbaseline Azole Usage	Safety	X	X
3.2.33	Summary of Grades 3 or Higher Thrombocytopenia Treatment-Emergent Adverse Events in the Treatment Phase by MedDRA System Organ Class, Preferred Term, Hepatic Impairment Status at Baseline, and Postbaseline Azole Usage	Safety	X	X
3.2.34	Summary of Grades 3 or Higher Neutropenia Treatment-Emergent Adverse Events in the Treatment Phase by MedDRA System Organ Class, Preferred Term, Hepatic Impairment Status at Baseline, and Postbaseline Azole Usage	Safety	X	X
3.2.35	Summary of Grades 3 or Higher Anemia Treatment-Emergent Adverse Events in the Treatment Phase by MedDRA System Organ Class, Preferred Term, Hepatic Impairment Status at Baseline, and Postbaseline Azole Usage	Safety	X	X

Table No.	Title	Population	Standard	In-Text
Labs, Vital Signs, and ECGs				
3.3.1	Summary of Laboratory Values – Hematology	Safety	X	X
3.3.2	Shift Summary of Hematology Values – to the Worst Abnormal Value	Safety	X	X
3.3.3	Shift Summary of Hematology Laboratory Values in CTC Grade – to the Worst Abnormal Value	Safety	X	X
3.3.4	Summary of Laboratory Values – Chemistry	Safety	X	X
3.3.5	Shift Summary of Chemistry Values – to the Worst Abnormal Value	Safety	X	X
3.3.6	Shift Summary of Chemistry Laboratory Values in CTC Grade – to the Worst Abnormal Value	Safety	X	X
3.4.1	Summary of Systolic Blood Pressure	Safety	X	
3.4.2	Summary of Diastolic Blood Pressure	Safety	X	
3.4.3	Summary of Pulse	Safety	X	
3.4.4	Summary of Respiratory Rate	Safety	X	
3.4.5	Summary of Body Temperature	Safety	X	
3.4.6	Summary of Body Weight	Safety	X	
3.4.7	Summary of QRS Interval (ms) From 12-Lead ECG	Safety	X	
3.4.8	Summary of QT Interval (ms) From 12-Lead ECG	Safety	X	
3.4.9	Summary of QTcB Interval (ms) From 12-Lead ECG	Safety	X	
3.4.10	Summary of QTcF Interval (ms) From 12-Lead ECG	Safety	X	
3.4.11	Summary of RR Interval (ms) From 12-Lead ECG	Safety	X	
3.4.12	Summary of PR Interval (ms) From 12-Lead ECG	Safety	X	
3.4.13	Summary of Clinically Significant ECG Abnormalities	Safety	X	

Listings

Listing No.	Title
2.1.1	Subject Disposition Status
2.1.2	Subject Enrollment and Site Information
2.1.3	Subject Retreatment Disposition Status
2.2.1	Protocol Deviations
2.2.2	Randomized Subjects With Inclusion/Exclusion Criteria Violations
2.3	Analysis Populations
2.4.4	Demographic and Baseline Characteristics
2.4.5	Medical History
2.4.6	Hematologic Disease History
2.4.7	Prior Anticancer Therapy
2.4.8	Prior Radiotherapy
2.4.9	Prior and Concomitant Medications
2.4.10	Prophylaxis Medications
2.4.11	Conditioning Regimens
2.4.12	PRBC/Platelet Transfusions Performed Within 30 Days Before Enrollment or On Study
2.4.13	Allogeneic Hematopoietic Stem Cell Transplant History
2.5.1	Itacitinib Exposure and Compliance
2.5.2	Corticosteroid Exposure in the Treatment Phase
2.5.3	Corticosteroid Usage After End-of-Treatment
2.6.1	aGVHD Symptoms and Response Assessment
2.6.2	cGVHD Symptoms and Response Assessment
2.6.3	Relapse of Malignant and Nonmalignant Hematologic Diseases
2.6.4	Failure-Free Survival Events and Assessment
2.6.5	Secondary Graft Failure
2.6.6	Acute GVHD Flares
2.6.7	Discontinuation of Immunosuppressive Medication
2.6.8	Post-Transplant Lymphoproliferative Disorder Assessment
████	████████████████████
████	████
2.6.11	Deaths
2.6.12	ECOG Status
2.6.13	PK Blood Sampling Times
2.7.1	Adverse Events
2.7.2	Serious Adverse Events
2.7.3	Grade 3 and Higher Adverse Events
2.7.4	Fatal Adverse Events
2.7.5	Itacitinib/Placebo Treatment-Related Adverse Events
2.7.6	Corticosteroids Treatment-Related Adverse Events
2.7.8	Adverse Events Leading to Itacitinib/Placebo Dose Interruption
2.7.9	Adverse Events Leading to Corticosteroid Dose Interruption
2.7.10	Adverse Events Leading to Itacitinib/Placebo Dose Reduction
2.7.11	Adverse Events Leading to Corticosteroid Dose Reduction
2.7.12	Adverse Events Leading to Discontinuation of Itacitinib/Placebo

Listing No.	Title
2.7.13	Adverse Events Leading to Discontinuation of Corticosteroids
2.7.14	Adverse Events Occurring in the Retreatment Phase
2.8.1	Abnormal Clinical Laboratory Values – Hematology
2.8.2	Clinical Laboratory Values – Hematology
2.8.3	Abnormal Clinical Laboratory Values – Chemistry
2.8.4	Clinical Laboratory Values – Chemistry
2.9.1	Abnormal Vital Sign Values
2.9.2	Vital Sign Values
2.10.1	Abnormal 12-Lead ECG Values
2.10.2	12-Lead ECG Values

APPENDIX C. CLINICALLY NOTABLE EVENTS BASED ON CUSTOMIZED MEDDRA QUERIES

This appendix provides the categories of clinically notable events based on customized MQs.

Clinically Notable Hematological Events

Category	Preferred Term
Anemia	Anaemia macrocytic
Anemia	Aplasia pure red cell
Anemia	Aplastic anaemia
Anemia	Erythroblast count decreased
Anemia	Erythropenia
Anemia	Hypoplastic anaemia
Anemia	Microcytic anaemia
Anemia	Red blood cell count decreased
Anemia	Anaemia
Anemia	Haematocrit decreased
Anemia	Haemoglobin decreased
Anemia	Leukoerythroblastic anaemia
Neutropenia	Agranulocytosis
Neutropenia	Band neutrophil count decreased
Neutropenia	Febrile neutropenia
Neutropenia	Granulocyte count decreased
Neutropenia	Granulocytopenia
Neutropenia	Neutropenia
Neutropenia	Neutropenic infection
Neutropenia	Neutropenic sepsis
Neutropenia	Neutrophil count decreased
Thrombocytopenia	Acquired amegakaryocytic thrombocytopenia
Thrombocytopenia	Platelet count decreased
Thrombocytopenia	Thrombocytopenia
Thrombocytopenia	Plateletcrit decreased

Hematological Infection

Category	Preferred Term
EBV infection	Epstein Barr virus positive mucocutaneous ulcer
EBV infection	Epstein-Barr viraemia
EBV infection	Epstein-Barr virus associated lymphoma
EBV infection	Epstein-Barr virus associated lymphoproliferative disorder
EBV infection	Epstein-Barr virus infection
EBV infection	Hepatitis infectious mononucleosis
EBV infection	Infectious mononucleosis
EBV infection	Oral hairy leukoplakia
EBV infection	Post transplant lymphoproliferative disorder
EBV infection	X-linked lymphoproliferative syndrome
EBV infection	Epstein-Barr virus test positive
EBV infection	Epstein-Barr virus antibody positive
Cytomegaloviral infections	Congenital cytomegalovirus infection
Cytomegaloviral infections	Cytomegalovirus chorioretinitis
Cytomegaloviral infections	Cytomegalovirus colitis
Cytomegaloviral infections	Cytomegalovirus duodenitis
Cytomegaloviral infections	Cytomegalovirus enteritis
Cytomegaloviral infections	Cytomegalovirus enterocolitis
Cytomegaloviral infections	Cytomegalovirus gastritis
Cytomegaloviral infections	Cytomegalovirus gastroenteritis
Cytomegaloviral infections	Cytomegalovirus gastrointestinal infection
Cytomegaloviral infections	Cytomegalovirus gastrointestinal ulcer
Cytomegaloviral infections	Cytomegalovirus hepatitis
Cytomegaloviral infections	Cytomegalovirus infection
Cytomegaloviral infections	Cytomegalovirus mononucleosis
Cytomegaloviral infections	Cytomegalovirus mucocutaneous ulcer
Cytomegaloviral infections	Cytomegalovirus myelomeningoradiculitis
Cytomegaloviral infections	Cytomegalovirus myocarditis
Cytomegaloviral infections	Cytomegalovirus nephritis
Cytomegaloviral infections	Cytomegalovirus oesophagitis
Cytomegaloviral infections	Cytomegalovirus pancreatitis
Cytomegaloviral infections	Cytomegalovirus pericarditis
Cytomegaloviral infections	Cytomegalovirus syndrome
Cytomegaloviral infections	Cytomegalovirus urinary tract infection
Cytomegaloviral infections	Cytomegalovirus viraemia
Cytomegaloviral infections	Encephalitis cytomegalovirus
Cytomegaloviral infections	Disseminated cytomegaloviral infection
Cytomegaloviral infections	Pneumonia cytomegaloviral
Cytomegaloviral infections	Cytomegalovirus test positive

Hyperlipidemia

Category	Preferred Term
Hyperlipidemia	Acquired lipoatrophic diabetes
Hyperlipidemia	Acquired mixed hyperlipidaemia
Hyperlipidemia	Apolipoprotein B/Apolipoprotein A-1 ratio increased
Hyperlipidemia	Autoimmune hyperlipidaemia
Hyperlipidemia	Blood cholesterol abnormal
Hyperlipidemia	Blood cholesterol esterase increased
Hyperlipidemia	Blood cholesterol increased
Hyperlipidemia	Blood triglycerides abnormal
Hyperlipidemia	Blood triglycerides increased
Hyperlipidemia	Diabetic dyslipidaemia
Hyperlipidemia	Dyslipidaemia
Hyperlipidemia	High density lipoprotein abnormal
Hyperlipidemia	High density lipoprotein decreased
Hyperlipidemia	Hypercholesterolaemia
Hyperlipidemia	Hyperlipidaemia
Hyperlipidemia	Hypertriglyceridaemia
Hyperlipidemia	Hypo HDL cholesterolaemia
Hyperlipidemia	Intermediate density lipoprotein increased
Hyperlipidemia	LDL/HDL ratio increased
Hyperlipidemia	Lipids abnormal
Hyperlipidemia	Lipids increased
Hyperlipidemia	Lipoprotein (a) abnormal
Hyperlipidemia	Lipoprotein (a) increased
Hyperlipidemia	Lipoprotein abnormal
Hyperlipidemia	Lipoprotein increased
Hyperlipidemia	Low density lipoprotein abnormal
Hyperlipidemia	Low density lipoprotein increased
Hyperlipidemia	Non-high-density lipoprotein cholesterol increased
Hyperlipidemia	Remnant-like lipoprotein particles increased
Hyperlipidemia	Total cholesterol/HDL ratio abnormal
Hyperlipidemia	Total cholesterol/HDL ratio increased
Hyperlipidemia	Very low density lipoprotein abnormal
Hyperlipidemia	Very low density lipoprotein increased

QT Prolongation

Category	Preferred Term
QT prolongation	Electrocardiogram QT interval abnormal
QT prolongation	Electrocardiogram QT prolonged
QT prolongation	Long QT syndrome
QT prolongation	Long QT syndrome congenital
QT prolongation	Torsade de pointes
QT prolongation	Ventricular tachycardia

APPENDIX D. VARIOUS UPPER EFFICACY BOUNDARIES FOR TESTING DAY 28 ORR AT BOTH INTERIM 2 AND FINAL ANALYSES USING HWANG-SHIH-DECANI ALPHA SPENDING FUNCTION WITH SHAPE PARAMETER -4

No. of Subjects in the Interim 2 Efficacy Analysis Set	Information Fraction at IA2 (Column 1/436)	Efficacy Boundary/Critical Value at Interim 2 (as Compared With Observed Normalized Test Statistic z)	Final Critical Value
410	0.940367	2.062174	2.036718
411	0.942661	2.058304	2.037092
412	0.944954	2.054429	2.037452
413	0.947248	2.050549	2.037794
414	0.949541	2.046665	2.038117
415	0.951835	2.042776	2.038417
416	0.954128	2.038882	2.038691
417	0.956422	2.034984	2.038935
418	0.958716	2.03108	2.039144
419	0.961009	2.027172	2.039312
420	0.963303	2.023259	2.039433
421	0.965596	2.019341	2.039498
422	0.96789	2.015418	2.039497
423	0.970183	2.01149	2.039418
424	0.972477	2.007557	2.039247
425	0.974771	2.003619	2.038963
426	0.977064	1.999677	2.038541
427	0.979358	1.995729	2.037951
428	0.981651	1.991776	2.037147
429	0.983945	1.987817	2.036068
430	0.986239	1.983854	2.034627
431	0.988532	1.979885	2.03269
431	0.990826	1.975912	2.030039
433	0.993119	1.971933	2.026284
434	0.995413	1.967948	2.020608
435	0.997706	1.963959	2.010698