NOVARTIS

Clinical Development

CTL019/tisagenlecleucel/Kymriah®

CCTL019H2301 / NCT03570892

Tisagenlecleucel versus standard of care in adult patients with relapsed or refractory aggressive B-cell non-Hodgkin lymphoma: A randomized, open label, phase III trial (BELINDA)

Statistical Analysis Plan (SAP)

Document type:	SAP Documentation
Document status:	Draft – Addendum 1
Release date:	26-Nov-2021
Number of pages:	59

Document History – Changes compared to previous final version of SAP

Date	Time point	Reason for update	Outcome for update	Section and title impacted (Current)
14-Oct-2021	After the DBL for primary analysis	The primary endpoint was not met. To better understand the results further post-hoc analyses are planned.	Removed reference to a second interim analysis for OS. As the study did not meet the primary endpoint, no statistical test will be performed for OS. Updates align with protocol language. Added a new analysis set, the "Efficacy tisagenlecleucel infused set". This analysis set will be used to assess efficacy post-infusion and excludes patients whose efficacy is not yet evaluable post-infusion. Defined a "modified EFS" and "modified BOR", to be used as a supportive analysis to better understand the impact of study design on the primary endpoint. Added dose-response analyses Other minor updates to specify some additional outputs that are needed for the primary CSR	Section 1 Section 2.1.1.7 Section 2.2.6 Section 2.3.3 Section 2.4.1 Section 2.5.4.3 Section 2.5.4.4 Section 2.7.1 Section 2.7.2 Section 2.13.3 Section 2.14 Section 4
08-June-2021	Prior to DBL for primary analysis	Additional interim OS analysis added. Updates to planned PRO analyses. Other minor updates added for clarification.	Added an additional interim OS analysis to occur 18 months after the randomization of the last patient. Update to time to definitive deterioration analysis, in order to align analysis with the ICH E9 addendum on estimands. Update removes new-ANP and crossover as censoring reasons to target a treatment policy estimand. Clarification on how to derive censoring date and reason for patients with no baseline assessment for PRO. Added separate analysis of FACT-Lym B-symptoms questions. Clarification on definition of baseline. Assessments on the same day as	Section 1 Section 2.1.1.6 Section 2.4.1.2.3 Section 2.5.3 Section 2.7.2 Section 2.11 Section 2.14

Date	Time point	Reason for update	Outcome for update	Section and title impacted (Current)
			randomization/infusion can be considered as baseline assessments. Clarification on how to derive the censoring date and reason for EFS for patients who discontinue study prior to week 12 assessment.	
08-Apr-2021	Prior to DBL for primary analysis	Updates due to COVID-19 pandemic (timing of primary analysis and COVID-19 sensitivity analyses) Updates to align SAP with latest protocol amendment version 3	Added clarification on timing of primary analysis that all patients should have week 12 visit (or discontinued early). Added clarification that this SAP is only for analyses of the global study and any analyses of the China extension cohort will be in a separate SAP. Added secondary objective of time to response Added paragraph about subgroup analysis of patients from Japan Added clarification that concomitant radiotherapy in a palliative setting is allowed in both arms as part of the treatment strategy Added section to address the secondary objective related to RCL testing [redacted] Other minor updates to ensure alignment with protocol	Section 1 Section 2.2.6 Section 2.4.2 Section 2.7 Section 2.8.5.3 Section 2.13 Section 5.6

Table of contents

List of abbreviations

ABC	Activated B-cell
AE	Adverse event
AESI	Adverse event of special interest
ASTCT	American society for transplantation and cellular therapy
ATC	Anatomical therapeutic chemical
AUC	Area under the curve
Bcl-2	B-cell lymphoma 2
Bcl-6	B-cell lymphoma 6
BEAM	Carmustine, etoposide, cytarabine and melphalan
BIRC	Blinded independent review committee
BOR	Best overall response
CAR	Chimeric antigen receptor
CI	Confidence interval
CK	Cellular kinetic
CKAS	Cellular kinetics analysis set
C_{last}	Last concentration
C_{max}	Maximum concentration
C-myc	C-myc proto-oncogene
CR	Complete response
CRO	Contract research organization
CRP	C-reactive protein
CRS	Cytokine release syndrome
CSF	Cerebral spinal fluid
CSR	Clinical study report
CT	Computerized tomography
CTC	Common Toxicity Criteria
CTCAE	Common Terminology Criteria for Adverse Events
CV%	Coefficient of variation (%)
DHAP	Cisplatin, cytarabine and dexamethasone
DI	Dose intensity
DLBCL	Diffuse large B-cell lymphoma
DMC	Data monitoring committee
DOR	Duration of response
DRL	Drug reference listing
ECG	Electrocardiogram
ECHO	Echocardiogram
ECOG	Eastern Cooperative Oncology Group
eCRF	Electronic case report form
eCRS	Electronic case retrieval strategy
EFS	Event-free survival
EOT	End of treatment
EQ-5D	EuroQol 5 dimension

EQ-VAS	EuroQol visual analogue scale
EWB	Emotional well-being
FACT-G	Functional Assessment of Cancer Therapy – General
FACT-Lym	Functional Assessment of Cancer Therapy – Lymphoma
FAS	Full analysis set
FDG	Fluorodeoxyglucose
FL3B	Follicular lymphoma grade 3B
FWB	Functional well-being
GCB	Germinal center B-cell
[redacted]	[redacted]
HDCT	High dose chemotherapy
HR	Hazard ratio
HRQoL	Health related quality of life
HSCT	Hematopoietic stem cell transplant
ICU	Intensive care unit
[redacted]	[redacted]
[redacted]	[redacted]
IPI	International prognostic index
IRT	Interactive response technology
ITT	Intention to treat
KM	Kaplan-Meier
LD	Lymphodepleting
LLOQ	Lower limit of quantification
LPLV	Last patient last visit
LYMS	Lymphoma-specific subscale
MCS	Mental component summary
MedDRA	Medical Dictionary for Drug Regulatory Affairs
mEFS	Modified event-free survival
MRA	Magnetic resonance angiography
NCI	National Cancer Institute
NHL	Non-Hodgkin lymphoma
NGS	Next generation sequencing
NOS	Not otherwise specified
ORR	Overall response rate
OS	Overall survival
PCS	Physical component summary
PD	Progressive disease
PD1	Programmed cell death 1
PET	Positron emission tomography
PFS	Progression-free survival
PK	Pharmacokinetics
PMBCL	Primary mediastinal large B-cell lymphoma
PPS	Per-Protocol Set

PR	Partial response
PRO	Patient-reported outcomes
PWB	Physical well-being
QoL	Quality of life
qPCR	Quantitative polymerase chain reaction
RCL	Replication competent lentivirus
r/r	Relapsed/refractory
R-DHAP	Rituximab plus cisplatin, cytarabine and dexamethasone
R-GDP	Rituximab plus gemcitabine, cisplatin and dexamethasone
R-GemOx	Rituximab plus gemcitabine and oxaliplatin
R-ICE	Rituximab plus ifosfamide, carboplatin, etoposide and mesna
RMST	Restricted mean survival time
SAE	Serious adverse event
SD	Stable disease
SAP	Statistical analysis plan
SF-36 v2	Short Form (36) Health Survey, version 2
SOC	Standard of care
SWB	Social well-being
$T_{1/2}$	Time to half life
TFLs	Tables, Figures, Listings
TIS	Tisagenlecleucel infused set
T_{last}	Time of last concentration
T_{max}	Time to peak concentration
TOI	Trial outcome index
TS	Total score
UNK	Unknown
VSV-g	Vesicular stomatitis virus/glycoprotein
WBC	White blood cells
WHO	World Health Organization

1 Introduction

This Statistical Analysis Plan (SAP) describes the implementation of the statistical analyses planned in the protocol for study CCTL019H2301: Tisagenlecleucel versus standard of care in adult patients with relapsed or refractory aggressive B-cell non-Hodgkin lymphoma: A randomized, open label, phase III trial (BELINDA).

Up to two clinical study reports (CSRs) could result from this SAP:

- Primary analysis: performed after approximately 200 event-free survival (EFS) events have been documented by the blinded independent review committee (BIRC) and all patients have had their week 12 assessment or discontinued prior to week 12 assessment. If statistically significant, this will trigger the first interim analysis for overall survival (OS).

  The timing of the primary analysis was updated compared to the protocol to ensure that all patients have their week 12 assessment (or discontinued early) at the time of primary analysis.

  The week 12 assessment is the first assessment for the primary and secondary efficacy objectives and if patients do not have this assessment they would be censored prior to the week twelve assessment or excluded from the analysis.

  Due to the COVID-19 pandemic, the study suspended screening and enrollment on 31-Mar-2020. Sites were then reopened on a case-by-case basis starting from 12-May-2020. This resulted in a delay in recruitment for a large number of patients, leading to many patients with a shorter follow-up time than simulated in the original sample size calculations. In particular 30 patients (9.3%) were randomized in 2021. The timing of 200 events confirmed by BIRC is expected to happen in April 2021, meaning that some of these patients might not have yet had their week 12 assessment. The week 12 visit of the last patient randomized is expected in early May 2021.
- Final analysis: a final analysis will be performed after the last patient last visit

This SAP is not planned to be used for any other analyses. Other analyses planned to be performed for this study include safety analyses for data monitoring committee (DMC) meetings and an analysis of the China extension cohort. The details of these analyses are contained in separate SAPs.

This SAP was based on version 03 of the study protocol dated 08-January-2021.

1.1 Study design

This is a randomized, open label, multicenter phase III trial comparing the efficacy, safety and tolerability of tisagenlecleucel treatment strategy versus standard of care (SOC) treatment strategy as second line treatment in adult patients with aggressive B-cell non-Hodgkin lymphoma (NHL). Patients must be refractory/relapsed (r/r) within 12 months of last dose of first line immunochemotherapy, which must contain both rituximab and an anthracycline. Refractory disease is defined as absence of complete response (CR) at the end of first line therapy; relapsed disease is defined as CR at the end of first line therapy followed by progressive disease (PD).

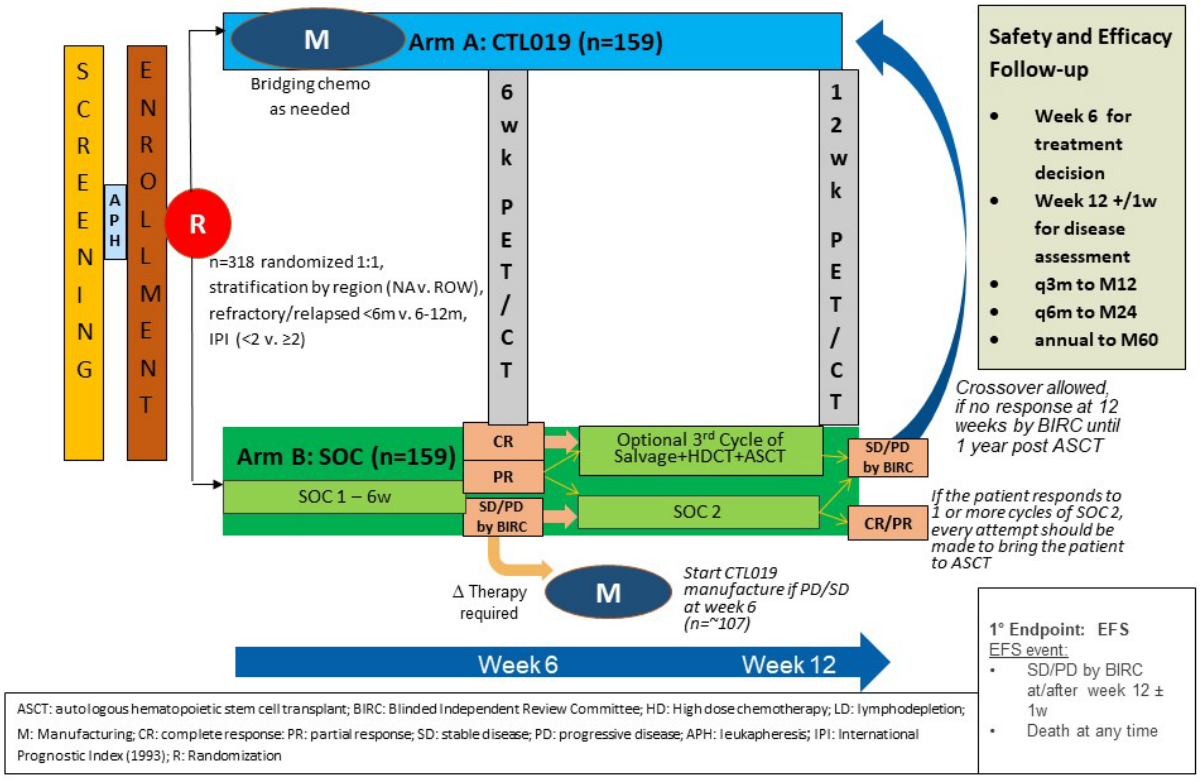
All screened patients will undergo non-mobilized leukapheresis for autologous T cell collection after obtaining informed consent. During the screening period, no lymphoma-specific systemic therapy is allowed prior to randomization.

A subject randomization list has been produced by the Interactive Response Technology (IRT) provider, based on a validated system that automates the random assignment of subject numbers to randomization numbers. Each randomization number is linked to one of the two treatment strategy arms. The randomization is stratified by three binary factors:

- Remission duration: refractory to first line therapy or relapsed < 6 months after last dose of first line therapy versus relapsed 6 - 12 months after last dose of first line therapy
- International Prognostic Index (IPI) score at study entry: < 2 versus ≥ 2
- Region: North America versus Rest of World

The randomization list has been reviewed and approved by a member of the Novartis Randomization Office. The study design is illustrated in Figure 1-1. Eligible patients will be randomized in a 1:1 ratio within each of the eight strata combinations to one of the following arms:

Figure 1-1 Study design



Arm A: tisagenlecleucel treatment strategy, or tisagenlecleucel arm, consisting of optional bridging chemotherapy and lymphodepleting (LD) chemotherapy followed by a single infusion of tisagenlecleucel. After randomization, tisagenlecleucel is manufactured for the patient, which is expected to take 3-4 weeks. During this period the use of one of four platinum-based bridging immunochemotherapies is allowed (R-ICE, R-DHAP, R-GDP and R-GemOx). Once the

tisagenlecleucel product is released by the manufacturing facility, patients should receive LD chemotherapy for 2 or 3 days duration (not required in patients with significant cytopenias), and then receive the tisagenlecleucel infusion about 5 days later. Thus it is expected that tisagenlecleucel infusion will occur approximately 4 to 6 weeks after randomization. Following infusion of tisagenlecleucel, no further anticancer therapies are allowed.

Arm B: standard of care (SOC) treatment strategy, or SOC arm, consisting of standard of care chemotherapy with transplant. Patients will receive one of four platinum-based immunochemotherapies (R-ICE, R-DHAP, R-GDP and R-GemOx) followed in responding patients by high dose chemotherapy (HDCT) and autologous hematopoietic stem cell transplantation (HSCT). Every effort should be made to perform autologous HSCT in patients achieving a PR, if deemed in the patient's best interest by the treating physician. Patients with response that is not sufficient to allow HSCT should change therapy to one of the other immunochemotherapy regimens listed above, at the investigator's discretion in an attempt to achieve a sufficient response and then proceed to HSCT. Only patients who are deemed no longer eligible for HSCT (e.g., due to adverse event [AE], poor tolerance to immunochemotherapy, worsening of performance status) after two cycles of immunochemotherapy may proceed to lenalidomide or ibrutinib treatment at investigator discretion. If the assessment of SD or PD is confirmed by BIRC at the Week 6 assessment, the investigator may request manufacturing of tisagenlecleucel (but not crossover). In addition, patients with PR FDG+ disease, SD or PD per local assessment at the Week 12 assessment may request manufacturing of tisagenlecleucel. Crossover can only occur following confirmation of SD/PD per BIRC at or after the Week 12 assessment, until 1 year after autologous HSCT.

Tumor response assessments will be performed at baseline (within 2 weeks prior to randomization), and post-baseline at the following times after randomization: weeks 6 (± 2 weeks) and 12 (± 1 week), months 6, 9 and 12 (± 2 weeks), months 18 and 24 (± 2 weeks), and thereafter annually (± 2 weeks) until 5 years after randomization. Efficacy will be assessed using the Lugano criteria (as detailed in Appendix 2 of the study protocol).

It is planned to randomize a total of 318 patients. The primary endpoint is EFS, and the primary analysis is to be performed after approximately 200 EFS events have been documented and all patients have had their week 12 assessment or discontinued prior to week 12 assessment. There is no planned interim analysis for EFS, however, an interim analysis for the secondary endpoint of OS will be conducted at the time of the primary analysis, but only if the primary endpoint EFS is statistically significant. If such an interim analysis takes place and OS is not statistically significant, a second interim analysis for OS will be conducted approximaely 18 months after the last patient was randomized, and a final analysis for OS will be conducted approximately 5 years after the first patient was randomized.

Note that after recruitment of the 318 patients into the global study, the study will continue enrollment into a China extension cohort. This SAP contains details of the analyses planned for the global study and any patients enrolled into the China extension cohort are excluded from these analyses.

1.2 Study objectives and endpoints

The primary objective of this study is to compare tisagenlecleucel treatment strategy to standard of care (SOC) treatment strategy with respect to EFS.

The two treatment strategies to be compared are defined as:

- Tisagenlecleucel treatment strategy: optional bridging chemotherapy and lymphodepleting chemotherapy followed by a single infusion of tisagenlecleucel
- Standard of Care (SOC) treatment strategy: SOC immunochemotherapy followed in responding patients by high dose chemotherapy (HDCT) and autologous hematopoietic stem cell transplantation (HSCT)

A list of study objectives and related endpoints are provided in Table 1-1, reproduced from the study protocol:

Table 1-1 Objectives and related endpoints

Objectives	Endpoints	Section
Primary Objective		
• To compare tisagenlecleucel treatment strategy to SOC treatment strategy with respect to delaying the composite event of disease progression / stable disease at or after the week 12 assessment; or death at any time.	• EFS, defined as time from date of randomization to the date of first documented disease progression or stable disease at or after the week 12 assessment, as assessed by blinded independent review committee (BIRC) per Lugano criteria, or death at any time	• Section 2.5
Secondary Objectives		
• To compare tisagenlecleucel treatment strategy to SOC treatment strategy with respect to EFS as assessed by local investigator.	• EFS as assessed by local investigator	• Section 2.5.4
• To compare tisagenlecleucel treatment strategy to SOC treatment strategy with respect to overall survival (OS).	• OS: defined as the time from randomization to date of death	• Section 2.7
• To compare tisagenlecleucel treatment strategy to SOC treatment strategy with respect to overall response rate (ORR) • To evaluate duration of response (DOR) by BIRC and local investigator. • To compare tisagenlecleucel treatment strategy and SOC treatment strategy with respect to time to response (TTR)	The following endpoints will be evaluated by BIRC and investigator assessment per Lugano criteria: • ORR: overall response rate as per the Lugano criteria • Duration of response: time from the date of first documented response of CR or PR to the date of first documented progression (SD or PD at or after the week 12 assessment will be considered progression) or death due to aggressive B-cell NHL • TTR: time from the date of randomization to the date of a patient first achieved a response of CR or PR on or after the Week 12 assessment	• Section 2.7

• To evaluate safety and tolerability of tisagenlecleucel treatment strategy versus SOC treatment strategy	• Type, frequency and severity of serious and non-serious adverse events and laboratory abnormalities and discontinuations due to adverse events	• Section 2.8
• To compare patient reported outcomes (PRO) of health-related quality of life (HRQoL) in both treatment arms.	• Time to definitive deterioration in SF-36v2, FACT-Lym, and EQ-VAS	• Section 2.11
• Evaluate efficacy and safety of both treatment arms in histological subgroups (e.g., DLBCL NOS, FL3B, other) and molecular subgroups (e.g. GCB, ABC, other)	• EFS, OS and AE	• Section 2.2.6 • Section 2.5 • Section 2.7
• To characterize the in vivo cellular kinetics of tisagenlecleucel transduced cells into target tissues (blood, bone marrow, cerebral spinal fluid and other tissues if available), as measured by qPCR summarized by clinical response in patients receiving tisagenlecleucel therapy in arm A or after crossover	• Summary of qPCR detected tisagenlecleucel transgene concentrations in peripheral blood and bone marrow (and other tissue, if available), and cellular kinetic parameters from peripheral blood profile samples by time point and clinical response status	• Section 2.9
• To characterize the incidence and prevalence of tisagenlecleucel immunogenicity (humoral and cellular) and impact on cellular kinetics, efficacy, and safety in patients receiving tisagenlecleucel therapy in arm A or after crossover	• Summary of pre-existing and treatment induced immunogenicity (cellular and humoral) of tisagenlecleucel • Levels of pre-existing and treatment induced immunogenicity. Cellular kinetic parameters, concentration-time profile by immunogenicity category (positive/negative), and efficacy (Month 3 response)	• Section 2.10
• To assess presence of RCL in patients receiving tisagenlecleucel in arm A or after crossover	• RCL by VSV-qPCR	• Section 2.8.5.3

[redacted]

2 Statistical methods

2.1 Data analysis general information

The data will be analyzed by Novartis and/or a designated Contract Research Organization (CRO), including the possible interim analysis of OS (an external statistician is not needed for this interim analysis because it can only be performed after the primary analysis of EFS).

SAS version 9.4 or later, and R version 3.0.2 or later, will be used to perform all data analyses and to generate tables, figures and listings (TFLs).

Data included in the analysis

The primary analysis is the single planned analysis for the primary efficacy endpoint of EFS. A unique data cutoff date will be established after the targeted number of EFS events for the primary analysis (n=200) has been documented by the BIRC and all patients have had a week 12 assessment or discontinued early. All statistical analyses for the primary analysis will be performed using all data collected in the database up to the data cutoff date. All data with an assessment date or event start date (e.g., vital sign assessment date or start date of an AE) prior to or on the cutoff date will be included in the analysis. Any data collected beyond the cutoff date will not be included in the analysis and will not be used for any derivations.

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

The data cutoff date for the primary analysis will also serve as the data cutoff date for the first possible interim analysis of the secondary endpoint of OS. The data cutoff date for the second interim analysis of OS will be set at 18 months after the randomization of the last patient. The data cutoff date for the final analysis of OS will be set at 5 years after the randomization of the first patient.

General analysis conventions

Pooling of centers: Unless specified otherwise, data from all study centers will be pooled for the analysis. Due to the expected small number of subjects enrolled at centers, no center effect will be assessed.

Qualitative data (e.g., gender, race, etc.) will be summarized by means of contingency tables by treatment group; a missing category will be included as applicable. Percentages will be calculated using the number of subjects in the relevant population or subgroup as the denominator.

Quantitative data (e.g., age, body weight, etc.) will be summarized by appropriate descriptive statistics (i.e., mean, standard deviation, median, q1, q3, minimum and maximum) by treatment group.

2.1.1 General definitions

2.1.1.1 Treatment strategy

The treatment strategies were defined in Section 1.2 above.

2.1.1.2 Completion of treatment strategy

The tisagenlecleucel treatment strategy is considered completed when the subject is infused with tisagenlecleucel. Subjects are considered as discontinued from the tisagenlecleucel treatment strategy if they discontinue the study without tisagenlecleucel infusion.

The SOC treatment strategy is considered completed when the subject has undergone autologous HSCT following SOC chemotherapy. Subjects are considered as discontinued from the SOC treatment strategy if they discontinue the study or crossover to the tisagenlecleucel arm without autologous HSCT.

2.1.1.3 Start date of treatment strategy

The start date of treatment strategy is defined as the randomization date. The start date of treatment is defined as the first date when a non-zero dose of any component of the treatment strategy was administered.

2.1.1.4 End date of treatment strategy

The end date of treatment strategy is defined as the last date when a non-zero dose of any component of study treatment strategy was administered.

In the case of SOC arm patients who do not go to transplant, and who crossover to tisagenlecleucel, the end date of treatment strategy must be on or before the start date of any

bridging chemotherapy or lymphodepleting chemotherapy administered prior to tisagenlecleucel infusion.

2.1.1.5 Study day

The study day describes the day of an event/assessment relative to the randomization date, and is defined as:

- (date of event/assessment − date of randomization + 1) if event/assessment is on or after the date of randomization
- (date of event/assessment − date of randomization) if event/assessment precedes the date of randomization. In this case the study day will be negative.

The study day will be displayed in the data listings.

In addition, days from tisagenlecleucel infusion will be calculated and listed for selected analyses of efficacy, safety and cellular kinetics post-tisagenlecleucel infusion.

2.1.1.6 Baseline

For ***baseline disease evaluations***, the most current assessments (imaging, pathology assessment, bone marrow biopsy or aspirate, CSF cytology, lesions from physical exam findings, etc.) on or prior to the date of randomization will be used as the baseline assessment. (If the assessment is on the date of randomization and assessment time is available, then assessment time must also be before randomization time). Any imaging or disease assessments obtained after randomization cannot be considered for baseline.

For ***safety evaluations*** (i.e., AEs, laboratory abnormalities, vital signs, etc.), the last available assessment on or prior to the date of randomization is taken as baseline. (If the assessment is on the date of randomization and assessment time is available, then assessment time must also be before randomization time).

In addition, the last available assessment on or before the date of tisagenlecleucel infusion is taken as baseline for selected analyses of efficacy, safety and cellular kinetics post-tisagenlecleucel infusion. (If the assessment is on the date of infusion and assessment time is available, then assessment time must also be before infusion start time).

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

In rare cases where multiple measurements meet the baseline definition, with no further flag or label that can identify the chronological order, then the following rule should be applied. If values are from central and local laboratories, the value from central assessment should be considered as baseline.

2.1.1.7 Duration of follow-up

Duration on study defined as the time from the date of randomization until the analysis cutoff date for the primary analysis, or until the LPLV for the final analysis, will be calculated and summarized.

For patients infused with tisagenlecleucel, the duration from infusion defined as the time from the date of infusion until the analysis cutoff date for the primary analysis, or until the LPLV for the final analysis, will also be calculated and summarized.

2.1.1.8 Last contact date

The last contact date will be used for censoring of patients in the analysis of OS. For patients not known to have died as of the analysis cutoff date, the last contact date should be derived as the latest date on or before the data cutoff date from the dates listed in the first column of Table 2-1 below. For each of the sources, specific conditions listed in the table have to be fulfilled to ensure that there was true contact with the patient. No additional dates are allowed to be used, for example, dates coming from concomitant medications, PRO, etc.

Table 2-1 Data sources for last contact date

Source data	Conditions
Date of Randomization	No condition
Last date patient was known to be alive from Survival Follow-up page	No condition
Start/End dates from further antineoplastic therapy	Non-missing medication/procedure term
Start/End dates from drug administration record	Non-missing dose
Any specific efficacy assessment date if available	Evaluation is not missing
Laboratory/cellular kinetics collection dates	Sample collection with non-missing value
Vital signs date	At least one non-missing parameter value
Performance Status date	Non-missing performance status
Start/End dates of AE	Non-missing verbatim term

Note: imputed dates will not be used to derive the last contact date with the exception of partially imputed dates from the Survival Follow-up page.

2.1.1.9 Lost to follow-up

For OS analysis, patients will be considered as lost to follow-up if the time between their last contact date and the analysis cutoff date is greater than or equal to 105 days (i.e., 3 months plus 2 weeks, assuming 1 month = 30.4375 days).

For response-related time-to-event analyses (i.e., EFS and DOR), patients will be considered as lost to follow-up only if they discontinued the study due to loss to follow-up.

2.2 Analysis sets

The analysis sets to be used are defined as below.

2.2.1 Screened Set

The screened set comprises all subjects who have signed informed consent and were screened in the study.

2.2.2 Full Analysis Set

The full analysis set (FAS) comprises all subjects to whom study treatment has been assigned by randomization. According to the intent to treat principle, subjects will be analyzed according to the treatment and strata they have been assigned to during the randomization procedure.

The FAS will be used as the main analysis set for efficacy, demographics and other baseline characteristics.

2.2.3 Safety Set

The safety set comprises all subjects to whom study treatment has been assigned by randomization. Subjects will be analyzed according to randomization.

The safety set will be used for all randomized safety comparisons.

2.2.4 Per-Protocol Set

The per-protocol set (PPS) consists of a subset of subjects in the FAS who are complient with key requirements of the study protocol:

Protocl deviations leading to exclusion rom the PPS include:

- Diagnosis of a disease other than histologically confirmed aggressive B-cell NHL (inclusion criterion 3)
- Relapse/progression more than 12 months after last dose of first line therapy for aggressive B-cell NHL (inclusion criterion 4)

2.2.5 Tisagenlecleucel Infused Set

The tisagenlecleucel infused set (TIS) comprises all subjects who received infusion of tisagenlecleucel (i.e., including crossover patients from the SOC arm).

The TIS will be used for all safety summaries for subjects infused with tisagenlecleucel. Subjects will be analyzed according to the treatment arm they were originally assigned to during the randomization procedure (i.e. tisagenlecleucel treatment strategy, and crossover from SOC treatment strategy).

2.2.6 Efficacy tisagenlecleucel infused set

The efficacy tisagenlecleucel infused set (ETIS) comprises a subset of the TIS who are evaluable for efficacy post-infusion. Patients will be included in the ETIS if they have received an infusion of tisagenlecleucel, and have at least one response assessment (BIRC or local) post-infusion or if they died, discontinued the study or started new anticancer therapy (as defined in Section 2.4.2) post-infusion and prior to any response assessment. This definition is used in order to exclude patients who have not yet had a response assessment post-infusion due to the timing of the analysis cutoff date and the patient's next scheduled response assessment.

The ETIS will be used for all post-infusion efficacy summaries. Subjects will be analyzed according to the treatment arm they were originally assigned to during the randomization procedure (i.e. tisagenlecleucel treatment strategy, and crossover from SOC treatment strategy).

2.2.7 Cellular Kinetic Analysis Set

The cellular kinetic analysis set (CKAS) consists of subjects in the TIS who provide evaluable tisagenlecleucel cellular kinetic data. A subject is considered as having evaluable cellular kinetic data if at least one cellular kinetic parameter can be derived. The CKAS will be used for summaries (tables and figures) of cellular kinetic data. The TIS will be used for listings of cellular kinetic data.

Note that subjects will be removed from the estimation of certain CK parameters on an individual basis depending on the number of available samples. These subjects will be identified at the time of the analyses.

2.2.8 Subgroup of interest

Efficacy

If the primary efficacy analysis based on the FAS is statistically significant, the primary efficacy endpoint of EFS will be summarized by the following subgroups to examine the homogeneity of the treatment effect:

- Stratification factors, based on the data from eCRF:
  - Remission duration (Refractory or relapsed within 6 months, Relapsed at 6-12 months)
  - IPI at study entry (< 2, ≥ 2 risk factors)
  - Geographical region (North America, Rest of the World)
- Age (< 65 years, ≥ 65 years)
- Sex (Male, Female)
- Race (Asian, Black or African American, White, Other)
- Ethnicity (Hispanic or Latino, Not Hispanic or Latino)
- ECOG performance status (0, 1)
- Elevated LDH (Yes, No)
- Prior response status (Primary refractory, relapsed)
- Histology (e.g., DLBCL NOS, FL3B, PMBCL, High grade B-cell lymphoma, Other)
- Rearrangements in MYC/BCL2/BCL6 genes (Double/Triple hits, Other)
- Stage of disease at study entry (I/II, III/IV)
- Molecular DLBCL cell of origin (Germinal Center B cell [GCB], non-GCB)

No formal statistical test of hypotheses will be performed for the subgroups, only point estimates of the treatment effect and 95%-confidence intervals will be provided. The objective of the efficacy subgroup analysis is to demonstrate homogeneity of treatment effect in the above subgroups. Subgroup analyses will only be performed if adequate number of events are observed.

Safety

Key safety analyses on AESIs, deaths, SAEs and AEs leading to discontinuation will be repeated in the safety set, and in the tisagenlecleucel infused set, in the following subgroups:

- Age (< 65 years, ≥ 65 years)

- Sex (Male, Female)
- Race (Asian, Black or African American, White, Other)
- Ethnicity (Hispanic or Latino, Not Hispanic or Latino)
- Histology (e.g., DLBCL NOS, FL3B, PMBCL, High grade B-cell lymphoma, Other)
- Molecular: DLBCL cell of origin (GCB, non-GCB)

The objective for carrying out these subgroup analyses is to identify potential safety issues that may be limited to a subgroup of patients, or safety issues that are more commonly observed in a subgroup of patients. Summary tables will only be performed if at least 5 patients are present in each subgroup. Some grouping of classes will be considered.

In addition, key summary tables and/or figures for efficacy, safety and others will be developed for the Japan subgroup, defined as all patients who have been randomized in the study from the investigational sites located in Japan (i.e., Country = Japan), to address the regulatory submission in Japan. No statistical hypothesis testing will be conducted for the Japan subgroup. Summary tables and/or figures developed for the Japan subgroup will be specified in the TFL shell document.

Note that histology, rearrangements in MYC/BCL2/BCL6 genes and molecular DLBCL cell of origin are reported locally on CRFs, and also tested using central labs. Analyses for the CSR will consistently use locally reported data.

2.3 Patient disposition, demographics and other baseline characteristics

Unless specified otherwise, the FAS will be used for all baseline and demographic summaries. Summaries will be reported by treatment arm and for all subjects. The FAS will also be used for listings, where subjects will be presented by treatment arm.

2.3.1 Subject disposition

Subject disposition will be summarized as follows:

- Screening disposition for the screened set
- Treatment disposition for the FAS by treatment arm
- Study disposition for the FAS by treatment arm

For each disposition, subject status including completed, ongoing or discontinued (with reason for discontinuation) will be summarized based on the number and percentage of subjects as listed on the disposition eCRF pages.

Study follow-up will be summarized numerically as well as by categories: <6 months, 6 months to <12 months, 12 months to <24 months, >=24 months etc. for the FAS.

All disposition data will be listed using the screened set.

2.3.2 Analysis Sets

The number (%) of subjects in each analysis set (defined in Section 2.2) will be summarized by treatment group and stratum.

2.3.3 Basic demographic and background data

All demographic and baseline disease characteristics data will be summarized and listed by treatment arm. A summary of crossover patients will also be provided separately. The following grouping will be applied:

- Age: <65 vs. ≥ 65 years
- ECOG performance status: 0 vs. 1

Baseline stratification factors

The number (%) of subjects in each stratum, based on data obtained from the eCRF, will be summarized overall and by treatment arm for the FAS. Subgroup analysis using descriptive statistics based on stratification factors will also use data from the eCRF. These analyses may be repeated using data from the IRT system, but only if strata membership differs between IRT and eCRF for at least one patient. For statistical models that are stratified by one or more of the baseline stratification factors, the stratification per IRT will be used.

2.3.4 Medical history

Medical history and ongoing conditions, including cancer-related conditions and symptoms entered on eCRF will be summarized by treatment arm.. The summaries will be presented by primary system organ class (SOC), preferred term (PT) and treatment arm. Medical history and current medical conditions will be coded using the Medical Dictionary for Regulatory Activities (MedDRA) terminology.

2.3.5 Diagnosis and extent of cancer

The summary of diagnosis and extent of cancer (i.e., disease history) will include predominant histological subgroup (e.g., DLBCL NOS, FL3B, PMBCL, etc.), transformation of prior lymphoma, stage at initial diagnosis, stage at time of study entry, lines of therapy for prior lymphoma, lines of therapy for current lymphoma, IPI at study entry, DLBCL cell of origin subtype (local), rearrangements in MYC/BCL2/BCL6 genes (e.g. double/triple hits etc.) (local), time (in months) from initial diagnosis of current lymphoma to start of treatment strategy, time (in months) since most recent relapse/progression to start of treatment strategy, time (in months) from initial diagnosis of current lymphoma to most recent relapse/progression, and time (in months) from diagnosis of prior lymphoma to initial diagnosis of current lymphoma (only for patients with prior lymphoma).

Subjects will be classified by their prior treatment response as:

- Refractory: defined as subjects who did not achieve CR on first line therapy to current lymphoma
- Relapsed: defined as subjects who had CR on first line therapy to current lymphoma and relapsed prior to the study

Note: For subjects who received first line treatment for a prior NHL within the subtypes allowed by the protocol which then transformed into the current lymphoma, remission duration is derived considering the last treatment for the prior lymphoma.

2.3.6 Protocol deviations

The number (%) of subjects in the FAS with any protocol deviation will be tabulated by deviation category (as specified in the study Edit Check document) overall and by treatment group. All protocol deviations will be listed.

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

2.4.1 Study treatment / compliance

2.4.1.1 Tisagenlecleucel arm

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

2.4.1.1.1 Tisagenlecleucel

The total viable cell count (cells) and the total CAR-positive viable T cell count (cells) will be listed and summarized using descriptive statistics. Subjects will be categorized as below, within or above the target dose ranges.

The time between screening, leukapheresis, randomization, receipt of apheresis product by manufacturing, shipment date of tisagenlecleucel from manufacturing and tisagenlecleucel infusion will be summarized using descriptive statistics.

2.4.1.1.2 Lymphodepleting chemotherapy

The lymphodepleting chemotherapies, received after randomization but prior to tisagenlecleucel infusion will be listed. The number and percentage of subjects who received lymphodepleting chemotherapy will be summarized by therapy type, i.e., fludarabine/cyclophosphamide, bendamustine or other. Duration of exposure (Section 5.4), actual cumulative dose (in mg/m^2) and reason for therapy discontinuation will also be summarized by therapy type.

2.4.1.1.3 Bridging chemotherapy

Bridging chemotherapies are defined as chemotherapies received after randomization but prior to lymphodepleting chemotherapies. The number and percentage of subjects who received bridging chemotherapy will be summarized by therapy type, i.e., R-ICE, R-DHAP, R-GDP, R-GemOx or other. Duration of exposure (Section 5.4), dose reduction, interruption or discontinuation and corresponding reasons will also be summarized by therapy type; actual cumulative dose (in mg/m^2) and dose intensity (mg/m^2/day) will be summarized by each component of the bridging chemotherapy regimen.

The number of cycles of bridging chemotherapy received as part of the treatment strategy will also be summarized. Regimens with a total duration of exposure less than or equal to 10 days will be considered as one cycle of chemotherapy and regimens with a total duration of exposure greater than 10 days will be considered as more than one cycle, i.e. patients will be classified as receiving 0, 1 or >1 cycle of bridging chemotherapy.

2.4.1.2 SOC arm

2.4.1.2.1 Immunochemotherapy

The number and percentage of subjects who received chemotherapy will be summarized by therapy type, i.e., R-ICE, R-DHAP, R-GDP or R-GemOx. Duration of exposure (Section 5.4), number of cycles, dose reduction, interruption or discontinuation and corresponding reasons will also be summarized by therapy type; actual cumulative dose (in mg/m^2) and dose intensity (mg/m^2/day) will be summarized by each component of the immunochemotherapy regimen.

The number of cycles of chemotherapy received as part of the treatment strategy will also be summarized. Regimens with a total duration of exposure less than or equal to 10 days will be considered as one cycle of chemotherapy and regimens with a total duration of exposure greater than 10 days will be considered as more than one cycle, i.e. patients will be classified as receiving 0, 1 or >1 cycle of chemotherapy.

2.4.1.2.2 Conditioning chemotherapy

The number and percentage of subjects who received conditioning high dose chemotherapy will be summarized by therapy type, i.e., BEAM or other. Duration of exposure (Section 5.4) and reason for therapy discontinuation will also be summarized by therapy type; actual cumulative dose (in mg/m^2) and dose intensity (mg/m^2/day) will be summarized by each component of the conditioning chemotherapy.

2.4.1.2.3 Autologous HSCT

The number and percentage of subjects who become ineligible for autologous HSCT during the treatment strategy will be summarized. The reason for ineligibility will also be summarized.

The number and percentage of subjects who underwent autologous HSCT will be summarized.

2.4.2 Prior, concomitant and post therapies

Medications will be coded using the latest version of World Health Organization (WHO) Drug Reference Listing (DRL) dictionary that employs the WHO Anatomical Therapeutic Chemical (ATC) classification system at the time of analysis; surgical and medical procedures will be coded using MedDRA. The versions of the WHO-DRL and the MedDRA that will be used will be footnoted in all relevant outputs.

Prior anti-cancer therapy

Prior anti-cancer therapy refers to all anti-cancer interventions (therapeutic treatments and procedures) for aggressive B-cell NHL that are administered prior to randomization. The number and percentage of subjects in the FAS who received any prior anti-cancer medications, prior anti-cancer radiotherapy or prior anti-cancer surgery will be summarized by treatment arm, and listed separately

New anti-cancer therapy

New anti-cancer therapy consists of anti-cancer therapy administered on or after randomization, excluding therapies given as part of the randomly assigned treatment strategy. As such, new anti-cancer therapy is defined separately for each treatment arm as follows:

In the tisagenlecleucel arm:

- any anti-CD19 or gene therapy other than tisagenlecleucel
- conditioning therapy (high dose chemotherapy) with intention of HSCT
- any anti-neoplastic therapy other than optional bridging chemotherapy (regardless of protocol-specified or not) or lymphodepleting chemotherapy prior to tisagenlecleucel infusion (including for patients who do not go on to receive tisagenlecleucel)
- any anti-neoplastic therapy at any time after tisagenlecleucel infusion

In the SOC arm:

- any anti-CD19 or gene therapy including tisagenlecleucel
- any anti-neoplastic therapy prior to HSCT except SOC treatment options (including patients who do not go to HSCT)
- any anti-neoplastic therapy at any time after HSCT

Please note that in both arms radiation therapy in a palliative setting during the treatment strategy will not be considered as new anti-cancer therapy. Any radiation therapy given after the end of the treatment strategy would be considered a new anti-cancer therapy.

New anti-cancer therapies will be listed and summarized by ATC class, preferred term, overall and by treatment arm by means of frequency counts and percentages using the FAS.

Concomitant therapies

Concomitant therapy is defined as all interventions (therapeutic treatments and procedures) given to a subject during the study other than those specified as study treatment. Concomitant therapy includes medications (other than study drugs) starting on or after randomization or medications starting prior to start date of randomization and continuing after the start date of randomization.

Concomitant medications will be summarized by highest ATC class and preferred term using frequency counts and percentages by treatment arm. Surgical and medical procedures will be summarized by SOC and preferred term.

All concomitant therapies will be listed. Any concomitant therapies starting and ending prior to the start of study treatment will be flagged in the listing. The FAS will be used for all concomitant medication tables and listings.

Anti-cytokine medications are given for severe CRS due to tisagenlecleucel cells. The number of subjects administered with anti-cytokine medications, type of anti-cytokine medications received, and number of tocilizumab doses given for the management of CRS will be summarized using the TIS.

2.5 Analysis of the primary objective

The primary aim of the study is to compare two second line treatment strategies in adult patients with aggressive B-cell non-Hodgkin lymphoma who are refractory to or relapsed after frontline standard of care and are eligible for stem cell transplantation. The treatment strategies will be compared based on their effect on delaying the composite event of PD / stable disease (SD) at or after the Week 12 assessment or death at any time. These two treatment strategies will be compared based on all randomized patients, irrespective of whether the patient received all or some of the components of the randomized treatment. Intercurrent events preventing compliance with these strategies such as initiation of alternative cancer therapies prior to the composite event of interest, will be handled accordingly. A Cox regression model stratified by randomization stratification factors will be used to estimate the hazard ratio (HR) of EFS.

2.5.1 Primary endpoint

The primary endpoint of the study is EFS, defined as the time from the date of randomization to the date of the first documented PD or SD at or after the Week 12 assessment, as assessed by BIRC per Lugano criteria (see Appendix 2 of the study protocol), or death due to any cause, at any time. The protocol-allowed time window for the Week 12 assessment is +/- 1 week; however, as an analysis convention, response assessments as early as week 10 will be taken into account as valid Week 12 assessments (i.e., on or after study day 71, where study day 1 is the date of randomization). This approach is taken in case some patients have an early Week 12 assessment, for example to avoid delaying future study treatment options, e.g., HSCT in the SOC arm, tisagenlecleucel infusion in the tisagenlecleucel arm. Censoring conventions are provided in Section 2.5.3.

2.5.2 Statistical hypothesis, model, and method of analysis

The primary efficacy analysis will be the comparison of the distribution of EFS between the two treatment strategies in the FAS. Assuming proportional hazards for EFS, the following statistical hypothesis will be tested to address the primary efficacy objective:

$$H_0: \theta_1 \geq 1 \quad \text{vs.} \quad H_{A1}: \theta_1 < 1$$

where θ_1 is the EFS HR (tisagenlecleucel arm versus standard-of-care arm). The analysis to test this hypothesis will consist of a stratified log-rank test at an overall one-sided 2.5% level of significance. Stratification will be based on the randomization stratification factors, as assigned in IRT, of remission duration (refractory or relapsed within 6 months vs. relapsed at 6-12 months), IPI score at study entry (< 2 vs. ≥ 2) and region (North America vs. Rest of the World).

There will be no interim analysis for EFS. The final analysis for EFS will be performed on the data observed in the FAS up to the data cutoff date, after approximately 200 EFS events have been documented by the BIRC and all patients have had their week 12 assessment or discontinued prior to week 12 assessment. The study will be considered positive if the stratified log-rank test performed at the final analysis for EFS has a one-sided p-value ≤ 0.025.

The survival distribution of EFS will be estimated using the Kaplan-Meier (KM) method and will be plotted graphically by treatment arm. The median EFS along with its 95% confidence interval will be presented by treatment arm. The survival probabilities at 3, 6, 9 and 12 months, and the associated 95% confidence intervals, will be summarized by treatment arm. A stratified

Cox regression model will be used to estimate the HR of EFS, along with its 95% confidence interval, using the same strata as for the primary efficacy comparison.

2.5.3 Handling of missing values/censoring/discontinuations

The analysis of EFS will be based on all randomized patients, regardless of whether the patient received all or some of the components of the randomized treatment, and the amount of dose received.

If no EFS event is observed prior to the earliest censoring event, EFS will be censored. Censoring events include loss to follow-up, withdrawal of consent, study discontinuation, data cutoff date and initiation of new anticancer therapy (as defined in Section 2.4.2). If the earliest censoring event occurs before the Week 12 assessment, then EFS will be censored at the date of the censoring event (since only death counts as an EFS event prior to the Week 12 assessment). If the earliest censoring event occurs after the Week 12 assessment, then EFS will be censored at the date of the last assessment with CR/PR prior to the earliest censoring event and on or after the Week 12 assessment. In the case where the Week 12 assessment has a response of "unknown" and the censoring event occurs before any further response assessment with CR/PR, EFS will be censored on the day before the Week 12 assessment. If no assessments are available after day 70 and before censoring event, EFS will be censored on day 70 (i.e. the last day when only death is considered an event).

2.5.4 Supportive analyses

Patients in the tisagenlecleucel arm are only expected to receive tisagenlecleucel infusion between 4 and 6 weeks after randomization (see Section 1.1), and before the infusion they may be receiving bridging chemotherapy, consisting of one of the four chemotherapy regimens planned for the SOC arm. As such, any efficacy benefit to the tisagenlecleucel arm is only anticipated to emerge after 4-6 weeks. This delayed treatment effect is an example of non-proportional hazards, and can lead to a loss of power for the stratified log-rank test, the primary analysis in the study. To support the analyses of EFS under non-proportional hazards assumptions, the following supportive analyses will be used to compare EFS per BIRC assessment between the tisagenlecleucel and SOC arms.

2.5.4.1 Weighted log-rank test

Considering the delayed effect of the tisagenlecleucel treatment, piece-wise weighted log-rank test [Xu et al, 2017]: assigning weights of 0 to event times in the first 6 weeks, and weights of 1 thereafter, in both treatment arms.

The test will be conducted at the one-sided 2.5% level, and be stratified by the randomization stratification factors of remission duration, IPI score at study entry and region. In addition, the treatment effect will be summarized by the average HR and 95% CI obtained from a weighted stratified Cox model, stratified by the randomization stratification factors.

2.5.4.2 Restricted mean survival time

The restricted mean survival time (RMST) [Zhang, 2013] analysis assesses the treatment difference in expected survival time between two arms with the restriction of a certain cutoff

time point (τ). The choice of τ should take into the following considerations: (1) τ should not exceed the minimum of the largest follow-up time for both arms so that the RMST(τ) of both arms can be adequately estimated [Zhao et al, 2015]; (2) τ should be large enough to cover the majority of patients' outcomes so that the RMST(τ) provides a meaningful assessment of the treatment effect; (3) τ is desired to be clinically meaningful. Given the situation of the patients' enrollment and scheduled assessments, the τ_0= Month 12 is selected as an appropriate cut-off point. In case the minimum of the largest follow-up time for both arms does not extend beyond Month 12, τ_1= minimum of largest follow-up time of both arms will be used as the cut-off point.

The RMST (for EFS, by the chosen cutoff point) will be estimated based on the KM method for both tisagenlecleucel arm and SOC arm. RMST will be summarized descriptively by treatment group. The RMST difference between the two arms will be calculated, along with the associated 95% CI.

2.5.4.3 Further supportive analyses

Further supportive analyses will include:

- HR and 95% CI of EFS per BIRC from an unstratified Cox model without any covariate adjustment.
- HR and 95% CI of EFS per BIRC from a stratified Cox model, stratified by the randomization stratification factors of remission duration, IPI at study entry and region, and adjusted for the following possibly prognostic covariates: age, gender, race, ECOG performance status, histological subgroup, stage of disease at study entry and DLBCL subtype.
- HR and 95% CI of EFS per BIRC from an unstratified Cox model, adjusted for the following possibly prognostic covariates: age, gender, race, ECOG performance status, histological subgroup, stage of disease at study entry, DLBCL subtype, remission duration (per CRF), IPI at study entry (per CRF) and region.
- EFS per BIRC review in the PPS, using the same analysis method as in the primary efficacy analysis (with the exception of the log rank test, which will not be performed).
- EFS per local investigator review, using the same analyses as in the primary efficacy analysis, with the exception of the log-rank test which will not be performed (this was one of the secondary objectives of the study, see Section 1.2).

Also depending on the amount of missing assessments, further sensitivity analyses maybe undertaken, for example, censoring EFS after two or more missing response assessments.

If the primary analysis is statistically significant, subgroup analyses to assess the homogeneity of the treatment effect across demographic and baseline disease characteristics will be performed per BIRC. The list of subgroups is provided in Section 2.2.6. The analyses will include KM summaries of median EFS with 95% CI by treatment arm, and stratified Cox models will be used to estimate EFS HR and their 95% CIs. Only subgroups with an adequate number of events (at least 5) will be included. Note that even if the primary analysis is not statistically significant, the above subgroup analyses will be undertaken for both the histological and molecular subgroups, because this is a stated secondary objective of the study (see Section 1.2).

The number of subjects censored and reasons for censoring will be summarized by treatment arm using descriptive statistics, presented separately for EFS per BIRC and EFS per local investigator review.

Stratified Cox regression models will be repeated for the following alternative definitions of EFS:

1. EFS per BIRC irrespective of new anti-cancer therapy for lymphoma, i.e., EFS events will be counted even if occurring after start of a new anti-cancer therapy. This corresponds to a fully intention-to-treat approach for both treatment strategies
    - This analysis targets the Treatment Policy estimand, i.e. the comparison between tisagenlecleucel treatment strategy and SOC treatment strategy including possible future treatments

2. EFS per BIRC considering new anti-cancer therapy for lymphoma at any time as an EFS event.
    - This analysis considers a Composite estimand, where new anti-cancer therapies are considered the same as PD/SD or death

2.5.4.4 Post-hoc supportive analyses

The primary definition of EFS considers any SD/PD after day 70 to be an event. In order to assess the impact of this definition on the primary analyses, a supportive analysis comparing the two treatment arms will be performed with a modified definition of EFS (mEFS). The modified definition of EFS is the same as the primary definition of EFS except that:

- For patients on both arms, an assessment of SD/PD will be ignored as an event if it is later followed by CR/PR and no new anticancer therapy was started inbetween,
- For patients in the tisagenlecleucel arm who received infusion, any SD/PD assessments prior to infusion will be ignored.

This modified definition has been derived due to the fact that the week 12 assessment (as early as 71 days post-randomization) sometimes occured too early in the treatment strategy for a response to be observed, i.e. patients had not yet been infused on Arm A, or patients were still undergoing salvage chemotherapy with the intent to proceed to HSCT. The modified definition accounts for the fact that some of these patients did ultimately respond to the randomized treatment strategy.

Due to the study design, BIRC assessments may not be available for these patients after the initial SD/PD assessment, and in such cases local assessments will be used instead. For patients in the tisagenlecleucel arm, if no assessment is available post-infusion, the patient will be censored at the date of infusion.

The mEFS definition will also be used in a separate analysis to assess EFS post-infusion in patients randomized to the tisagenlecleucel treatment strategy by their disease response status pre-infusion.

2.6 Analysis of the key secondary objective

There is no key secondary objective in this study.

2.7 Analysis of secondary efficacy objectives

The secondary efficacy objectives are to compare tisagenlecleucel treatment strategy to SOC treatment strategy with respect to:

- EFS as assessed by local investigator (already discussed in Section 2.5.4 as supportive analysis for primary objective)
- OS
- Overall response rate (ORR), both by BIRC and local investigator assessment
- Duration of response (DOR), both by BIRC and local investigator assessment
- EFS as assessed by BIRC and OS in histological subgroups (DLBCL NOS, FL3B, PMBCL, etc) and molecular subgroups (e.g., GCB, ABC, other)
- Time to response (TTR), both by BIRC and local investigator assessment

2.7.1 Secondary endpoints

Overall survival

OS is defined as the time from date of randomization to date of death due to any cause. A cutoff date will be established for each analysis of OS. All deaths occurring on or before the cutoff date will be used in the OS analysis. If a patient is not known to have died at the data cutoff date, OS will be censored at the date of last contact.

Overall response rate

ORR is defined as the proportion of subjects with best overall response (BOR) of CR or partial response (PR) according to the Lugano criteria (see Appendix 2 of the study protocol for details).

BOR is defined as the best overall disease response from the sequence of overall disease responses, observed between the Week 12 assessment and the first to occur between the data cutoff date, the start date of a new anticancer therapy (as defined in Section 2.4.2) and the date of EFS event. That is, response assessments before the Week 12 assessment are not used in the calculation of BOR, in order to maintain consistency with the definition of EFS used in this study. For example, a patient with overall disease response of PD at Week 6 followed by CR at Week 12, would have a BOR of CR. Response assessments as early as week 10 (study day 71) will be taken into account as valid Week 12 assessments.

Similar to the modified EFS definition specified in Section 2.5.4.4, a modified BOR (mBOR) will also be defined where the date of EFS event is not a censoring reason, i.e. responses of CR/PR after SD/PD will be considered for mBOR, as long as no new anticancer therapies were started inbetween.

A *separate* definition “BOR post-infusion” will also be used to assess efficacy post infusion. This refers to the best overall disease response considering efficacy assessments post-infusion and before the data cutoff date, the start date of a new anticancer therapy and the date of

progressive disease). BOR post-infusion will also be used to assess any differences by response in PK (Section 2.9), immunogenicity (Section 2.10) [redacted]

Duration of response

Duration of response (DOR) only applies to patients whose BOR is CR or PR according to the Lugano criteria. It is defined as the time from the date of first documented response of CR or PR, to the date of the first subsequent documented progression or death due to aggressive B-cell NHL. In this study, "documented progression" refers to a response of SD or PD on or after the Week 12 assessment, and assessments on or after week 10 (study day 71) will be considered as valid Week 12 assessments. Censoring conventions are provided in Section 2.7.3.

Time to response

Time to overall disease response (CR or PR) is defined as time from the date of randomization to the date of first documented overall disease response of PR or CR according to Lugano criteria based on disease response data per BIRC on or after the Week 12 assessment.

2.7.2 Statistical hypothesis, model, and method of analysis

Overall survival

OS will only be tested if the primary endpoint (EFS as assessed by BIRC) is statistically significant at the primary analysis for EFS. In that case, and assuming proportional hazards for OS, the following statistical hypotheses will be tested in the FAS:

$$H_{02}: \theta_2 \geq 1 \quad \text{vs.} \quad H_{A2}: \theta_2 < 1$$

where θ_2 is the OS HR (tisagenlecleucel arm versus standard-of-care arm). The analysis to test this hypothesis will consist of a stratified log-rank test at an overall one-sided 2.5% level of significance. Stratification will be based on the randomization stratification factors: remission duration, IPI score at study entry and region.

If the EFS primary endpoint is statistically significant, OS will be analyzed using a group sequential design with two looks, the first at the time of the primary analysis for EFS, , and the final look at 5 years after the randomization of the first patient. A Haybittle-Peto boundary will be used, where the one-sided significance level is 0.05% at theinterim analysis and 2.5% at the final analysis. Analyses will be based on the FAS according to the randomized treatment group and stratum assigned at randomization.

Irrespective of whether the EFS primary endpoint is statistically significant or not, the following analyses will be undertaken. The distribution of OS will be estimated using the KM method, and KM curves will be presented. The median OS and the proportion of patients alive at 6 and 12 weeks, and at 6, 12, 18, 24, 36, 48 and 60 months, with 95% confidence intervals, will be presented by treatment arm. The HR for OS will be calculated, along with its 95% confidence interval, using the following models:

- a stratified Cox model without any covariate adjustment, stratified by the randomization stratification factors of remission duration, IPI at study entry and region,
- an unstratified Cox model without any covariate adjustment,

- a stratified Cox model, stratified by the randomization stratification factors of remission duration, IPI at study entry and region, and adjusted for the following possibly prognostic covariates: age, gender, race, ECOG performance status, histological subgroup, stage of disease at study entry and DLBCL subtype,
- an unstratified Cox model, adjusted for the following possibly prognostic covariates: age, gender, race, ECOG performance status, histological subgroup, stage of disease at study entry, DLBCL subtype, remission duration (per CRF), IPI at study entry (per CRF) and region.

Subgroup analyses will be undertaken for histological subtype and molecular subgroup, because these are a stated secondary objective of the study (see Section 1.2). The analyses will include KM summaries of median OS with 95% CI by treatment arm.

The number of subjects censored for OS and reasons for censoring will be summarized by treatment arm using descriptive statistics.

Overall response rate

ORR based on BIRC assessment will be summarized using descriptive statistics (N, %) by treatment arm, along with two-sided standard Wald asymptotic (i.e., normal approximation) 95% CIs. As a supportive analysis, ORR will also be summarized based on the local investigator assessment of response data. In addition, comparative summary of BOR between BIRC assessment and local assessment will be provided to evaluate the consistency of the two results.

Furthermore, a descriptive summary of response status at the week 6 assessment will be provided for all patients, and a summary of response status pre-infusion will be provided for all infused patients.The time between the pre-infusion response assessment and the date of infusion will also be summarized.

Duration of response

DOR will be summarized by treatment arm for all patients in the FAS with BOR of CR or PR. The distribution of DOR will be estimated using the KM method, and KM curves will be presented. The median DOR and the proportion of patients remaining relapse-free at 3, 6, 12, 18, 24, 36, 48 and 60 months after first response, with 95% confidence intervals, will be presented by treatment arm. The HR for DOR will be calculated, along with its 95% confidence interval, using a Cox model stratified by the randomization stratification factors of remission duration, IPI at study entry and region.

Time to response

The TTR analysis will be based on the FAS. TTR will be estimated using the KM method and the median TTR will be presented along with a 95% confidence interval. As a sensitivity analysis, TTR as per investigator assessment will be presented by treatment group, along with 95% confidence intervals.

Descriptive analyses for patients who achieved CR or PR (e.g. mean, SD, median, minimum and maximum) will be presented by treatment group.

2.7.3 Handling of missing values/censoring/discontinuations

Overall survival

If a patient is not known to have died at the data cutoff date, OS will be censored at the date of last contact (see Section 2.1.1.8 for definition of last contact date).

Overall response rate

Patients with unknown or missing BOR will be counted as non-responders. If there is no baseline response assessment, all post-baseline overall disease responses are expected to be Unknown. If no valid post-baseline response assessments are available, the BOR will be Unknown unless progression is reported. For the computation of ORR, these patients will be counted as non-responders.

Duration of response

If no DOR event (i.e., SD/PD on or after the Week 12 assessment or death due to aggressive B-cell NHL) is observed prior to the earliest censoring event, DOR will be censored. The same censoring reasons used in the primary EFS analysis will be used for DOR, with the addition of "death due to reason other than aggressive B-cell NHL". The censoring date will be the date of the last assessment with response of CR or PR on or prior to the earliest censoring event.

Time to response

TTR will be censored for patients who did not achieve a PR or CR:

- At maximum follow-up (i.e. FPFV to LPLV used for the analysis) for patients who died due to any cause, progressed on or after the week 12 assessment or initiated new anticancer therapy (as defined in Section 2.4.2),
- At the date of the last adequate assessment on or after the week 12 assessment otherwise.
    - If patient has no adequate assessments on or after the week 12 assessment, patient will be censored at day 1.

2.8 Safety analyses

The main focus of the safety analyses are:

- to compare the safety of the two treatment strategies as defined in Section 1.2 in the safety set (randomized safety comparison)
- to evaluate safety post-tisagenlecleucel infusion in the tisagenlecleucel infused set.

The safety analyses will be based on the analysis sets specified above unless specified otherwise.

2.8.1 Analysis and reporting periods

Note that following the definitions in Table 2-2 below, the **safety comparison period** will be the main safety reporting period for the randomized safety comparison, and the **post-infusion period** will be the main safety reporting period for the evaluation of safety after tisagenlecleucel infusion.

Table 2-2 Safety reporting periods

Period	Definition	Subjects to be included
For both arms		
Screening period *	From the day of subject's informed consent to the day before randomization	Screened set
Safety comparison period	From the day of randomization to the earlier day of starting a new anticancer therapy, or 56 days after last study treatment administration	Safety set
Post safety-comparison period	From the day after the earlier day of starting a new anticancer therapy, or 56 days after last study treatment administration, until end of study	Safety set
For tisagenlecleucel arm and subjects in SOC arm with a crossover visit		
Pre-lymphodepleting period **	From day of randomization (subjects in tisagenlecleucel arm) or day of crossover visit (subjects in SOC arm with a crossover visit) to the day before first lymphodepleting chemotherapy dose or the day before infusion of tisagenlecleucel if lymphodepleting chemotherapy is not given	Full analysis set
Lymphodepleting period ***	From the first day of lymphodepleting chemotherapy to • the day before infusion of tisagenlecleucel, for subjects who received infusion, or • the earlier of date of discontinuation and 30 days after last dose of lymphodepleting chemotherapy for subjects who didn't receive infusion of tisagenlecleucel	All subjects who received lymphodepleting chemotherapy
Post-infusion period	Starting on the day of the first tisagenlecleucel infusion until end of study	Tisagenlecleucel infused set

* If a subject was not randomized, all the AEs for the subject are considered to be in the screening period.

** If a subject did not receive lymphodepleting chemotherapy or tisagenlecleucel infusion, all the AEs for the subject are considered to be in the pre-lymphodepleting period.

*** This period only applies to subjects who received lymphodepleting chemotherapy.

2.8.2 Adverse events (AEs)

Reporting of AEs follows the modified safety reporting rules described in Protocol Appendix 3.

Reporting of AEs (except for CRS) will be based on MedDRA (latest version per database lock) and the Common Terminology Criteria for Adverse Events (CTCAE) version 5.0. The grading of CRS will be primarily based on the Lee criteria [Lee et al, 2014].

- For the randomized safety comparison, AE summaries will include all AEs that started or worsened during the safety comparison period, i.e. ***post-randomization*** AEs.
- For the safety evaluation post-tisagenlecleucel infusion, AE summaries will include all AEs that started or worsened during the post-infusion period, i.e. ***tisagenlecleucel-treatment-emergent*** AEs.

AEs will be summarized by number and percentage of subjects with at least one AE, at least one AE in each primary system organ class and for each preferred term. A subject with multiple occurrences of an AE will be counted only once in the respective AE category. A subject with multiple CTCAE grades for the same preferred term will be summarized under the maximum CTCAE grade recorded for the event. AE with missing CTCAE grade will be included in the all grades column of the summary tables. The frequency of AEs of grade 3 or above will be summarized together.

In AE summary tables, the primary system organ class will be presented alphabetically and preferred terms will be sorted within the primary system organ class in descending frequency. The sort order for the preferred terms will be based on their frequency in the 'All grades' column as reported in the tisagenlecleucel arm.

For the **randomized safety comparison**, the following AE summaries will be produced by treatment arm in subjects from the safety set for the safety comparison period:

- Overview of adverse events, deaths, and other serious or clinically significant AEs
- Adverse events, regardless of study treatment relationship, by primary system organ class, preferred term and maximum grade
- Adverse events, suspected to be study treatment related, by primary system organ class, preferred term and maximum grade
- Serious adverse events, regardless of study treatment relationship, by primary system organ class and preferred term and maximum grade
- Serious adverse events, suspected to be study treatment related, by primary system organ class and preferred term and maximum grade
- Adverse events leading to study treatment discontinuation, regardless of study treatment relationship, by primary system organ class and preferred term
- Adverse events leading to study treatment interruption/adjustment, regardless of study treatment relationship, by primary system organ class and preferred term
- Non-serious adverse events, regardless of study treatment relationship, by primary system organ class and preferred term

For the **safety evaluation post-tisagenlecleucel infusion**, the AE summaries listed below will be produced separately for subjects in Arm A and subjects crossed over from Arm B to Arm A from the tisagenlecleucel infused set for the post-infusion period by timing of onset: within 8 weeks of first tisagenlecleucel infusion, >8 weeks to 1 year after the first tisagenlecleucel infusion, >1 year after the first tisagenlecleucel infusion, and any time after the first tisagenlecleucel infusion. The denominator for each time period will be the number of subjects still remaining in the study at the start of the corresponding time period.

- Overview of adverse events, deaths, and other serious or clinically significant AEs

- Adverse events, regardless of tisagenlecleucel relationship, by primary system organ class, preferred term and maximum grade
- Adverse events, suspected to be tisagenlecleucel related, by primary system organ class, preferred term and maximum grade
- Serious adverse events, regardless of tisagenlecleucel relationship, by primary system organ class and preferred term and maximum grade
- Serious adverse events, suspected to be tisagenlecleucel related, by primary system organ class and preferred term and maximum grade
- Non-Serious Adverse events, regardless of tisagenlecleucel relationship, by primary system organ class and preferred term

2.8.2.1 Adverse events of special interest / grouping of AEs

AESIs include all important identified and potential risks of tisagenlecleucel, and may also include additional relevant safety topics (e.g., missing information [redacted]). The list of AESIs and their search criteria are updated on a regular basis at program level in the electronic Case Retrieval Strategy (eCRS) form. The most recent version of the eCRS form will be used for the reporting activity.

For the **randomized safety comparison**, the AESIs will be summarized by treatment arm in the safety comparison period, in subjects from the safety set.

For the **safety evaluation post-tisagenlecleucel infusion**, AESIs will be summarized separately for subjects in Arm A and subjects crossed over from Arm B to Arm A by drug relationship, group term, preferred term, maximum grade and timing of onset in subjects from the tisagenlecleucel infused set for the post-infusion period by timing of onset: within 8 weeks of first tisagenlecleucel infusion, >8 weeks to 1 year after the first tisagenlecleucel infusion, >1 year after first tisagenlecleucel infusion, and any time after first tisagenlecleucel infusion.

AESI based on important identified risks and potential risks will be summarized intables by timing of onset:

- AESI post-tisagenlecleucel infusion based on important **identified** risks, regardless of study drug relationship, by group term, preferred term and maximum CTC grade
- AESI post-tisagenlecleucel infusion based on important **potential** risks, regardless of study drug relationship, by group term, preferred term and maximum CTC grade

2.8.2.1.1 Cytokine release syndrome (CRS)

Detailed information regarding the first episode of CRS, including [redacted] time to onset of CRS, time to resolution of CRS, time to grade 3/4/5 CRS, concurrent infections, timing and duration of ICU stay, selected complications and use of anti-cytokine therapies, will be summarized by treatment arm. Time to resolution of the first CRS episode will be summarized for subjects with CRS using KM methodology. In case the end date of an episode of CRS is missing, it will be censored as the minimum of the cutoff date, end of study evaluation date and death date (if applicable).

[redacted]

2.8.2.1.2 Neurological events

Neurological events refer to a group of neurological AEs defined in the AESI search criteria form. A neurological event episode may include multiple overlapping or consecutive neurological AEs as long as the end date and the start date of two consecutive AEs are no more than 3 days apart (i.e., current AE start date – previous AE end date ≤ 3). The onset day of a neurological event episode is the start date of the first neurological AE in the episode. The resolution date is the end day of the last AE in the episode. If there are multiple AEs with the same last end date and one or more of these AEs are unresolved, the entire episode will be considered unresolved. Time to onset of the first neurological event episode will be summarized descriptively. Time to resolution of all neurological event episodes from all subjects will be summarized using KM methodology, without taking into account that multiple episodes might be clustered by subject. For example, for a subject with two episodes, both episodes will be included in the analysis. Although these two episodes from one subject are not independent, they will be treated as if they are from two subjects (each with one episode).

Neurological events will be reported according to the CTCAE v5.0. Additionally, neurological events will also be assessed using other grading scales (e.g., ASTCT consensus grading) and a sensitivity analysis will be conducted.

2.8.2.1.3 Hematopoietic cytopenias

In addition to the analysis of AEs reported by the investigator, analysis of laboratory results will also be performed for hematopoietic cytopenias not resolved by week 4 (defined as 35 days after the end of study treatment [28 days +7 day window])., (Section 2.8.4).

2.8.3 Deaths

Summary tables for deaths will be produced by treatment arm, system organ class and preferred term.

For the **randomized safety comparison**, summary tables for deaths will be provided by treatment arm for all deaths occurring in the safety set during the safety comparison period, i.e., from randomization until the earlier date of starting a new anticancer therapy and 56 days after last study treatment administration.

For the **safety evaluation post-tisagenlecleucel infusion**, summary tables for deaths will be provided separately for subjects in Arm A and subjects crossed over from Arm B to Arm A for all deaths in the tisagenlecleucel infused set that occurred after tisagenlecleucel infusion, by timing of death: within 30 days of tisagenlecleucel infusion, >30 days after tisagenlecleucel infusion and any time after tisagenlecleucel infusion.

All deaths will be listed, and study period as defined in Section 2.8.1 will be flagged in the listings.

2.8.4 Laboratory data

For the analysis of laboratory abnormalities, data from central and local laboratories will be combined.

For laboratory tests covered by the CTCAE, Novartis will assign the appropriate CTC grade to each laboratory value. For laboratory tests where grades are not defined by CTCAE, results will be graded by the low/normal/high classification based on laboratory normal ranges.

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

- Shift tables using CTC grades to compare baseline to the worst post-infusion or safety comparison period value
- For laboratory tests where CTC grades are not defined, shift tables using the low/normal/high classification to compare baseline to the worst post-infusion or safety comparison period value.

For the **randomized safety comparison**, the shift tables will be generated for the safety comparison period in subjects from the safety set.

For the **safety evaluation post-tisagenlecleucel infusion**, the shift tables for the tisagenlecleucel arm will be generated separately for subjects in Arm A and subjects crossed over from Arm B to Arm A from the tisagenlecleucel infused set by timing: within 8 weeks of first tisagenlecleucel infusion, >8 weeks to 1 year after first tisagenlecleucel infusion, >1 year after first tisagenlecleucel infusion, and any time after first tisagenlecleucel infusion.

In addition, for the safety evaluation, the percentage of subjects with hematopoietic cytopenia 4 weeks after end of study treatment of grade 3 or above will be summarized. (Note: 4 weeks is defined as 35 days: to allow for visits happening within 7 day time window allowed for visit). Among these subjects, the timing of resolution to grade 2 or below will be presented via KM methodology. The grading of cytopenias will be derived using laboratory results in absolute lymphocytes (hypo), absolute neutrophils (hypo), hemoglobin (hypo), platelet count (hypo) and WBC (hypo) according to CTCAE 5.0. If a subject did not achieve resolution at the last laboratory assessment, timing of resolution will be censored at that time. Patients will also be censored when receiving a new antineoplastic therapy. The KM median time to resolution and estimates of the percentage of unresolved cases at different time points (e.g., month 2, month 3, etc.) will be presented. The analysis will be presented for both the Safety Set and the Tisagenlecleucel Infused Set.

Liver function parameters of interest are total bilirubin (BILI), ALT, AST and alkaline phosphatase (ALP). The number (%) of subjects with worst post-baseline values as per Novartis Liver Toxicity guidelines will be summarized:

- ALT or AST > 3xULN
- ALT or AST > 5xULN
- ALT or AST > 8xULN
- ALT or AST > 10xULN
- ALT or AST > 20xULN
- BILI > 2xULN
- BILI > 3xULN

- ALT or AST > 3xULN & BILI > 2xULN
- ALT or AST > 3xULN & BILI > 2xULN & ALP ≥ 2xULN
- ALT or AST > 3xULN & BILI > 2xULN & ALP < 2xULN

2.8.5 Other safety data

2.8.5.1 ECG and cardiac imaging data

All ECG data will be listed by treatment arm, subject and visit, and abnormalities will be flagged.

2.8.5.2 Vital signs

All vital signs data will be listed by treatment group, subject and visit, and if ranges are available, abnormalities will be flagged. Summary statistics will be provided by treatment arm and visit for subjects in the safety set, and by visit for subjects in the tisagenlecleucel infused set.

2.8.5.3 Replication competent lentivirus

The presence of detectable replication competent lentivirus (RCL) will be tested by VSV-G at protocol scheduled assessments and listed.

2.9 Cellular kinetics endpoints

Tisagenlecleucel concentrations in peripheral blood (and bone marrow and CSF if available) will be listed, graphed and summarized (individual profiles, arithmetic and geometric means, standard deviation, CV%, CV% geometric mean, minimum, median and maximum) by time points, BOR (after infusion) and treatment arm (tisagenlecleucel arm or crossover from SOC arm) for subjects in the cellular kinetic analysis set (CKAS) as assessed by the following:

- Tisagenlecleucel transgene levels as measured by q-PCR

[redacted]

The cellular kinetics parameters listed in Table 2-3 along with other relevant cellular kinetics parameters will be estimated, if feasible, from the individual concentration versus time profiles using a non-compartmental approach within the modeling program Phoenix® (Pharsight, Mountain View, CA) and reported by BOR category. The non-quantifiable concentrations will be imputed to zero for concentration summaries, and will not be included for estimation of cellular kinetics parameters. Results reported but deemed unreliable will be flagged and excluded from the summaries and cellular kinetics parameter derivations.

Table 2-3 Non compartmental cellular kinetics parameters

Parameter	Definition
AUC 0 - 28d and/or AUC0-84d	The AUC from time zero to day 28 and/or day 84 and or other disease assessment days, in peripheral blood (days*copies/ µg)
Cmax	The maximum (peak) observed in peripheral blood or other body fluid drug concentration after single dose administration (copies/ µg)
Tmax	The time to reach maximum (peak) peripheral blood or other body fluid drug concentration after single dose administration (days)
T1/2	The half-life associated with the elimination phase slope of a semi logarithmic concentration-time curve (days) in peripheral blood
Clast	The last observed quantifiable concentration in peripheral blood (copies/ug)
Tlast	The time of last observed quantifiable concentration in peripheral blood (days)

Descriptive statistics of cellular kinetics parameters (arithmetic and geometric means, standard deviation, CV%, CV% geometric mean, minimum, median and maximum) will be summarized by BOR (after infusion) for subjects in the CKAS. For T_{max} and T_{last} only minimum, median and maximum will be presented. The summary of parameters will be presented separately by Arm A and crossover patients.

For subjects who receive tocilizumab for management of CRS, the cellular kinetic parameters based on qPCR data will be summarized by use of tocilizumab [redacted]

2.10 Immunogenicity

The humoral immunogenicity assay measures the antibody titers specific to the tisagenlecleucel molecule prior to and following infusion. Data will be further fractionated to determine proportion of subjects who make transient versus sustained antibody responses. The cellular immunogenicity assay assesses the presence of T lymphocytes activated by the tisagenlecleucel protein. Data will be reported as summary statistics of pre and post-dose levels of activated T lymphocytes.

The analysis of immunogenicity will be performed separately for patients received tisagelecleucel in Arm A, or patients crossed over from Arm B to Arm A and infused with tisagelecleucel.

2.10.1 Humoral immunogenicity

The proportion of humoral immunogenicity positive and negative patients will be summarized by time points. Summary statistics will be presented for tisagelecleucel cellular kinetic parameters for qPCR by anti- tisagelecleucel antibody post-infusion status (positive or negative). A strip plot of anti-tisagenlecleucel antibodies by time points will be presented. Regarding the boosted/induced humoral immunogenicity, a patient is only defined as positive for tisagenlecleucel treatment-induced or -boosted anti-mouse CAR19 (antimCAR19) antibodies when the anti-mCAR19 antibody MFI at any time post-infusion was at least 2.28-fold higher than pre-infusion levels. The summary of cellular kinetic parameters will be presented by categories of patients with and without treated induced boosted/induced anti-CTL019 antibodies.

A scatter plot of baseline anti-tisagelecleucel antibodies versus qPCR AUC0-28d and Cmax will be presented along with the appropriate regression line. In addition boxplots of anti-

tisagelecleucel antibodies at enrollment by BOR (after infusion) will be presented. The same response categories will be used for a similar boxplot summarizing the maximum fold change (based on post-infusion/baseline MFI at various time points) of anti-tisagelecleucel post-infusion.

2.10.2 Cellular immunogenicity

All the analyses described in this section will be performed separately for both CD4 and CD8 T cell responses related to both Pool 1 and Pool 2 peptides. The cellular immunogenicity will be summarized by time points and will be presented as strip plots with time points on the x-axis and net responses on y-axis. The strip plot of maximum net response of cellular immunogenicity by BOR (after infusion) will be presented. The scatter plot of maximum net response post baseline versus qPCR AUC0-28d and Cmax will also be presented along with the appropriate regression line. In addition, a summary table will be presented for net response by BOR (after infusion).

2.11 Patient-reported outcomes

Three separate patient-reported outcome (PRO) instruments will be used in the study:

- FACT-Lym: Functional Assessment of Cancer Therapy – Lymphoma, version 4, to assess lymphoma-specific quality of life. It is composed of the FACT-General (FACT-G), a 27-item questionnaire of general questions, and the FACT-Lym lymphoma-specific subscale (FACT-Lym LymS), an additional 15 items that assess patient concerns relating to lymphoma. All questions are scored on a 5-point scale ranging from 0 = “not at all” to 4 = “very much”; negatively worded questions are reverse-scored so that higher scores are always reflective of better health-related quality of life (HRQoL). FACT-G items are divided into four primary domains: PWB (Physical Well-Being; seven items; range, 0–28), SWB (Social Well-Being; seven items; range, 0–28), EWB (Emotional Well-Being; six items; range, 0–24), and FWB (Functional Well-Being; seven items; range, 0–28). The FACT-Lym LymS consists of common lymphoma disease and/or treatment-related symptoms (e.g., pain, fever, swelling, night sweats, insomnia, itching, weight loss, fatigue and loss of appetite). Three summary scales: FACT-Lym TOI (Trial Outcome Index; range, 0–116; composed of the PWB, FWB, FACT-Lym LymS); FACT-G (range, 0–108; composed of the PWB, FWB, SWB, EWB), and the FACT-Lym TS (Total Score; range, 0–168; composed of all of the scales) are also calculated.
- EQ-5D-5L: EuroQol 5D, to assess health utility for the purpose of economic evaluation. It is composed of the EQ-VAS, a visual analogue scale from 0-100 with higher values indicating better HRQoL, and the EQ-5D assessing five dimensions of health state (mobility, self-care, usual activities, pain/discomfort, anxiety/depression) on an ordinal scale with five categories: “no”, “slight”, “moderate”, “severe” and “extreme”.
- SF-36v2: Short Form (36) Health Survey, version 2, to assess general health and quality of life. It is composed of 36 questions making 8 domains (physical functioning, role – physical, bodily pain, general health, vitality, social functioning, role – emotional, mental health), from which two overall component scores are obtained: PCS (Physical Component Summary) and MCS (Mental Component Summary). Domain and component summary scores are converted to norm-based scores based on the general population, with

higher scores indicative of better HRQoL. The scoring of the questionnaire will be provided by the vendor for this instrument.

For missing items within any questionnaire, prorated scores will be calculated according to developer guidance.

All summary scores from the FACT-Lym, the EQ-VAS and the SF-36v2 will be tabulated for each treatment arm in the FAS, presented over time by scheduled visit. Change from baseline in the scores at the time of each assessment will also be summarized, where baseline is defined as the last PRO assessment on or prior to randomization.

The individual scores and change from baseline for B-symptoms from the FACT-Lym over time by scheduled visit will also be summarized separately. The three B-symptom questions are: bothered by fevers (FACTLY31), night sweats (FACTLY32), and losing weight (FACTLY36).

For each of the five dimensions of the EQ-5D-5L, the proportions of patients in each treatment arm in the FAS having reported “no”, “slight”, “moderate”, “severe” and “extreme” problems will be tabulated at each time-point.

Change from baseline item scores means and 95% CIs will be plotted by visit and by treatment arm.

Time to definitive deterioration

A secondary objective of the study is to compare the treatment arms with respect to the PRO scores. The FACT-Lym TOI will be the primary score for this objective. In addition, the FACT-Lym LymS, FACT-Lym TS and FACT-G will be analyzed as secondary scores, as will the EQ-VAS from the EQ-5D-5L, and the PCS and MCS from the SF-36v2. Note that treatment comparisons based on hypothesis tests will be considered as descriptive only, and no adjustment for multiple testing will be performed.

The primary PRO endpoint will be time to definitive deterioration, defined as the time from randomization to the date of definitive deterioration. For FACT-Lym TOI, a clinically meaningful deterioration is defined as a decrease from baseline of at least 5.5 points, and the decrease is considered definitive if there is no later improvement above this threshold. The date of definitive deterioration is the earliest such post-baseline assessment. In the absence of an earlier definitive deterioration, patients are censored at the date of the last assessment before the data cutoff date. Patients with no baseline data will be censored at day 1, due to no baseline assessment.

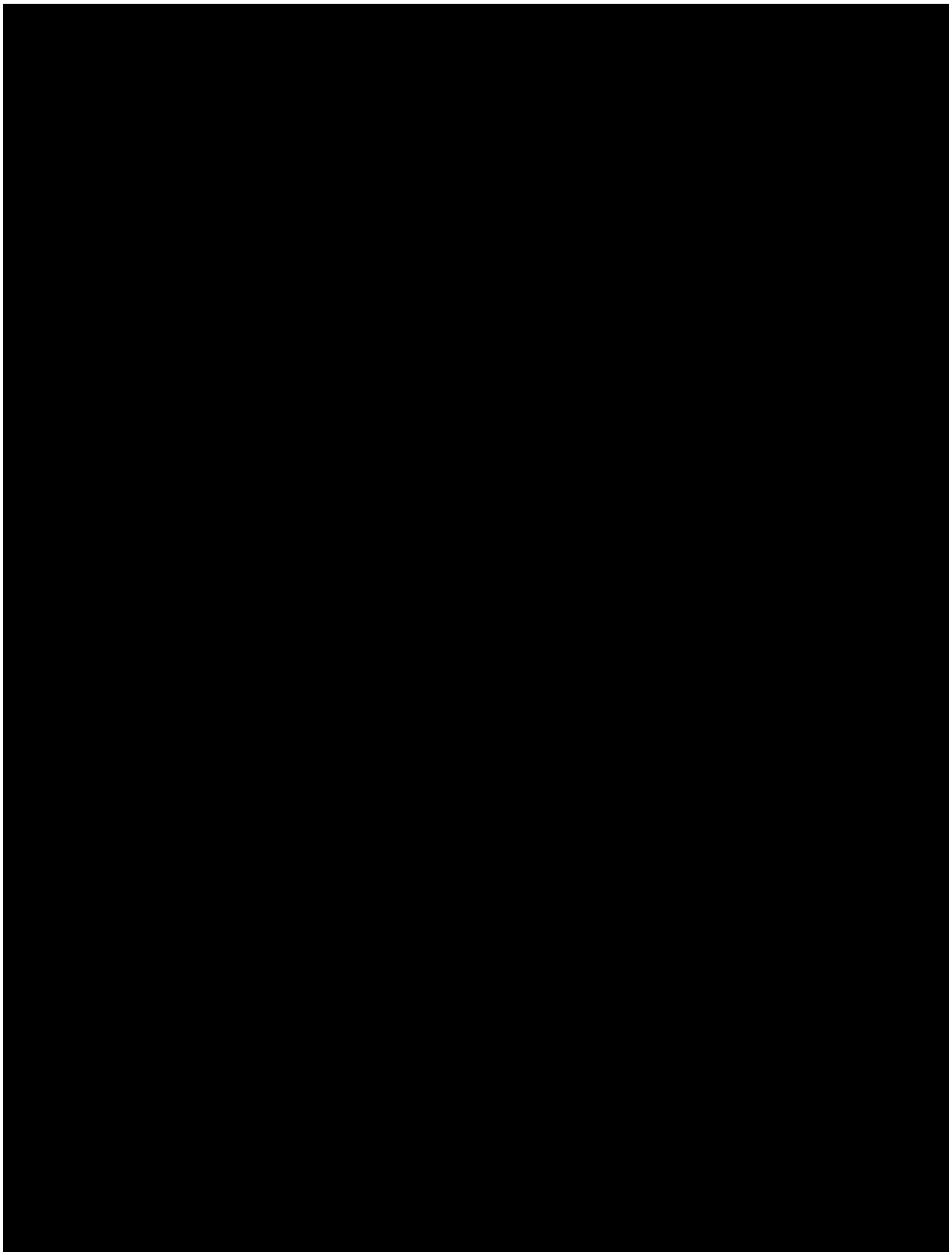
Death is considered as a definitive deterioration when it occurs within a period of time defined by twice the period between two assessments as planned in the study protocol. This avoids overestimating the time to definitive worsening in patients dying after an irregular assessment scheme. Patients who die after more than twice the planned period between two assessments since the last assessment are censored at the date of their last available questionnaire. For patients with no baseline assessment, death is considered as definitive deterioration if the death happens on or before study day 85 (twice the planned assessment period of 42 days), otherwise patient is censored on day 1 due to no baseline assessment.

Time to definitive deterioration in the FACT-Lym TOI will be compared between the two treatment arms in the FAS using a stratified log-rank test at the one-sided 2.5% level of significance. The test will be stratified by the randomization stratification factors of remission duration, IPI score at study entry and region, as assigned at randomization in IRT. The survival distributions will be presented descriptively using KM curves.

Similar analyses of time to definitive deterioration will be performed on the secondary PRO scales of interest as listed above, with deterioration defined as a decrease from baseline of ≥ 2.9 points for FACT-Lym LymS, ≥ 6.5 points for FACT-Lym TS, ≥ 3 points for FACT-G, ≥ 7 points for EQ-VAS, and ≥ 3 points for both PCS and MCS. Sensitivity analysis may be conducted for PCS and MCS using ≥ 5 as the MCID.

A supportive analysis of time to first deterioration may also be conducted, because of the possibility that first deterioration is not definitive due to crossover and new anti-neoplastic therapies.

[redacted]



2.14 Interim analysis

No interim analysis is planned for this trial for the primary endpoint of EFS. A hierarchical testing procedure will be adopted and the statistical test for OS will be performed only if the primary efficacy endpoint, EFS is statistically significant.

Two analyses are planned for OS: 1) an interim analysis at the time of the primary analysis for EFS (provided EFS is significant, as outlined in Section 2.7.2) in the FAS, and 2) a final analysis for OS at approximately 5 years from the first patient randomized. A Haybittle–Peto boundary will be used for testing OS, where the one-sided significance level is 0.05% at the interim analysis and 2.5% at the final analysis.

3 Sample size calculation

Based on the data from the ORCHARRD study (Novartis unpublished analyses), EFS time for patients who were randomized to receive salvage chemotherapy (DHAP plus Rituximab or DHAP plus Ofatumumab), who never reached CR before or relapsed within 12 months from response to previous therapy, or had a response of PR, SD or PD to previous therapy was considered as a reference for SOC. In ORCHARRD study, for these patients, who continue to

be in SD status at the end of cycle 2/3 (which is earlier than the 12 week assessment, each cycle: 21 days) or had progressed earlier than the 12 week assessment, based on the definition of EFS endpoint used in BELINDA (where documented SD/PD at the Week 12 assessment is considered an EFS event), EFS event time was adjusted to 12 weeks, to account for these earlier events.

The 9 month EFS rate is estimated to be 22.32% in SOC arm and is assumed to be 40% in tisagenlecleucel arm. Due to delayed tisagenlecleucel infusion of 6 weeks and ignoring SD/PD at week 6, we assume piecewise HR in both treatment arms. The HR between two arms during the first 6 weeks is assumed to be 1 and 0.61 after 6 weeks. Given those assumptions, a total of 200 EFS events are required to have 92% power at an one-sided 2.5% level of significance to reject the null hypothesis (i.e. the survival functions for EFS in the two arms are identical), using a stratified log-rank test with equal weights. The sample size calculation was conducted via simulation with software package East 6.4.

Based on a recruitment period of approximately 21 months using staggered enrollment rates of 2, 10, and 16 patients in the 1st 3 months followed by 17 patients thereafter, and assuming 15% drop out rate, a total of approximately 318 patients will need to be randomized to the two treatment arms in a 1:1 ratio. Given the above assumptions, it is estimated that the 200th EFS event will be observed at approximately 23 months from the date of first patient randomized.

4 Change to protocol specified analyses

1. In Section 2.4.2, definition of new anti-cancer therapy for SOC arm: protocol specified that:

- any anti-neoplastic therapy prior to HSCT except protocol-allowed-SOC treatment options (including patients who do not go to HSCT).

After discussion with the clinical team, it was decided that SOC treatments even if not protocol allowed, if taken prior to HSCT or for those who do not go to HSCT but are still eligible, is a protocol deviation but should not be considered as new anti-cancer therapy. Accordingly, the following updates have been made in the SAP:

- any anti-neoplastic therapy prior to HSCT except SOC treatment options (including patients who do not go to HSCT but are still eligible)

2. Added an efficacy subgroup analysis based on ECOG performance status (0, 1) and Elevated LDH (Yes, No).

[redacted]

5. A secondary endpoint of the study is time to definitive deterioration in SF-36v2, FACT-Lym and EQ-VAS. Further PRO analyses described in the protocol, e.g. mixed-effects models and improvement rates, will be performed outside the CSR, because they are not endpoints of the study.

5 Appendix

5.1 Imputation rules

5.1.1 Study drug

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

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

The subject should be treated as on-going and the cutoff date should be used as the dose end date.

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

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

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

Use Dec31yyyy

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

Use EOT date

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

Use last day of the Month (mm)

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

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

Use the treatment start date

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

5.1.2 AE, ConMeds and safety assessment date imputation

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

Missing Element	Rule
day, month, and year	• No imputation will be done for completely missing dates
day, month	• If available year = year of study treatment start date then

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

Table 5-2 Imputation of end dates (AE)

Missing Element	Rule
day, month, and year	• Completely missing end dates (incl. ongoing events) or with end date on or after the cutoff date will be imputed by min(cutoff date, end of study evaluation, date of death, withdrawal of consent date)
day, month	• If partial end date contains year only, set end date = min(31DEC, cutoff date, end of study evaluation, date of death, withdrawal of consent date)
day	• If partial end date contains month and year, set end date = min(last day of the month, cutoff date, end of study evaluation, date of death, withdrawal of consent date)

Partial or missing ConMeds end dates will not be imputed.

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

Any AEs and ConMeds which are continuing as per data cutoff will be shown as 'ongoing' rather than the end date provided.

Note that if the imputed AE start date is after the AE end date (regardless of if the end date is imputed or not), use the AE end date as the imputed AE start date.

5.1.3 Dates of prior lymphoma, initial diagnosis of cancer, first relapse and most recent recurrence/relapse

If the day or month of prior lymphoma, initial diagnosis, first relapse or most recent relapse is missing, it will be imputed to the minimum of the informed consent date -1 and the following:

- Missing day: 15th day of the month and year
- Missing day and month: July 1st of the year

5.1.4 Response assessments

All investigation dates for response (e.g., dates of PET scan, CT scan, bone marrow biopsy) must be completed with day, month and year. At any response assessment, if one or more investigation dates are incomplete, the response assessment date is calculated from the complete investigation dates as follows: if the overall disease response at that assessment is CR/PR/SD/UNK, then the assessment date is assigned as the latest complete investigation date, otherwise if overall disease response is PD, then the assessment date is the earliest complete investigation date at that assessment. If no investigation dates at a particular response assessment have day recorded, the 1st of the month is used. If month is not recorded at any of the investigations at a particular response assessment, the response assessment date will be imputed to the mid-point between the previous and following assessments. If a previous and following assessment are not available, this response assessment will not be used for any calculation.

5.1.5 Anti-neoplastic therapies

5.1.5.1 Prior therapies

Start date:

The same rule which is applied to the imputation of AE/concomitant medication start date will be used with the exception that for scenario (B) will be replaced to be ‘randomization date -1’.

End date:

Imputed date = min (randomization date, last day of the month), if day is missing;

Imputed date = min (randomization date, 31DEC), if month and day are missing.

If the end date is not missing and the imputed start date is after the end date, use the end date as the imputed start date.

If both the start date and the end date are imputed and if the imputed start date is after the imputed end date, use the imputed end date as the imputation for the start date.

5.1.5.2 Post therapies

Start date:

Imputed date = max (randomization date + 1, first day of the month), if day is missing;

Imputed date = max (randomization date + 1, 01JAN), if day and month are missing.

End date: No imputation.

5.1.6 Date of hospitalization imputation

Missing hospitalization end date or end date after data cutoff will be imputed following the same conventions as for AE end date imputation.

5.1.7 Incomplete date for death or last known date subject alive

If the day or month of death is missing from the death CRF, death will be imputed to the maximum of the full (non-imputed) last contact date (Section 2.1.1.8) and the following:

- Missing day: 15th day of the month and year of death
- Missing day and month: July 1st of the year of death

If the day or month of last known date subject alive is missing in the survival CRF, it will be first imputed with the following:

- Missing day: minimum of the date of assessment and 15th day of the month and year of last known date subject alive
- Missing day and month: minimum of the date of assessment and July 1st of the year of last known date subject alive

Then the above imputed last know date subject alive will be used to calculate the last contact date as defined in Section 2.1.1.8.

5.2 AEs coding/grading

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

AEs will be assessed according to CTCAE version 5.0, except for CRS, where grading of CRS will be primarily based on the Lee criteria [Lee et al, 2014]

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

5.3 Laboratory parameters derivations

Grade categorization of laboratory values will be assigned programmatically as per NCI CTCAE version 5.0. The calculation of CTCAE grades will be based on the observed laboratory values only, clinical assessments will not be taken into account. The criteria to assign CTCAE grades are given in Novartis internal criteria for CTCAE grading of laboratory parameters per Table 5-3.

For laboratory tests where grades are not defined by CTCAE version 5.0, results will be graded by the low/normal/high (or other project-specific ranges, if more suitable) classifications based on laboratory normal ranges.

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

Table 5-3 CTC grades for laboratory values based on CTCAE v5

CTC grades for laboratory values in Novartis Oncology (based on CTCAE v5 – Nov 2017) **Page 1**

				CTC Grades[1]				
Lab test (toxicity)	**SI unit**	**Lab test (NCDS)**	**Normal ranges (Merck manual, July 2015) and *conversion factors***	**0**	**1**	**2**	**3**	**4**
Hematology								
WBC ↓	10^9/L	WBC	**3.9 – 10.7 x 10^9/L**	≥ LLN	< LLN - 3.0 x 10^9/L	< 3.0 – 2.0 x 10^9/L	< 2.0 – 1.0 x 10^9/L	< 1.0 x 10^9/L
WBC (Leukocytosis)	10^9/L	WBC			-	-	> 100 x 10^9/L	-
Hemoglobin (Anemia)	g/L	HGB	**120 - 160 g/L** or 7.4 - 9.9 mmol/L (F) **140 - 170 g/L** or 8.7 – 10.6 mmol/L (M) *(16.113 x mmol/L = g/L)*	≥ LLN	< LLN - 100 g/L < LLN - 6.2 mmol/L	< 100 - 80 g/L < 6.2 - 4.9 mmol/L	< 80 g/L < 4.9 mmol/L	-
Hemoglobin ↑	g/L	HGB			Increase >0-20 g/L above ULN	Increase >20-40 g/L above ULN	Increase >40 g/L above ULN	-
Platelets ↓	10^9/L	PLAT	**150 - 350 x 10^9/L**	≥ LLN	< LLN - 75.0 x 10^9/L	< 75.0 - 50.0 x 10^9/L	< 50.0 - 25.0 x 10^9/L	< 25.0 x 10^9/L
Neutrophils ↓	10^9/L	NEUT		≥2x10^9/L	< 2.0 - 1.5 x 10^9/L	< 1.5 - 1.0 x 10^9/L	< 1.0 - 0.5 x 10^9/L	< 0.5 x 10^9/L
Lymphocytes↓	10^9/L	LYM		≥1.5x10^9/L	< 1.5 - 0.8 x 10^9/L	< 0.8 - 0.5 x 10^9/L	< 0.5 - 0.2 x 10^9/L	< 0.2 x 10^9/L
Lymphocytes ↑	10^9/L	LYM			-	> 4 - 20 x 10^9/L	> 20 x 10^9/L	-
Biochemistry								
AST ↑	U/L	AST	**0 - 35 U/L** or 0 – 0.58 ukat/L *(60 x ukat/L = U/L)*	≤ ULN	> ULN – 3.0 x ULN	> 3.0 - 5.0 x ULN	> 5.0 - 20.0 x ULN	> 20.0 x ULN
ALT ↑	U/L	ALT	**0 - 35 U/L** or 0 – 0.58 ukat/L *(60 x ukat/L = U/L)*	≤ ULN	> ULN – 3.0 x ULN	> 3.0 - 5.0 x ULN	> 5.0 - 20.0 x ULN	> 20.0 x ULN
Total bilirubin ↑	umol/L	BILI	**5.1 – 20.5 umol/L** or 0.3 – 1.2 mg/dL *(17.1 x mg/dL = umol/L)*	≤ ULN	> ULN - 1.5 x ULN	> 1.5 - 3.0 x ULN	> 3.0 - 10.0 x ULN	> 10.0 x ULN
Alk. Phosphatase ↑	U/L	ALP	**36 - 92 U/L** or 0.5 - 1.5 ukat/L *(60 x ukat/L = U/L)*	≤ ULN	> ULN - 2.5 x ULN	> 2.5 - 5.0 x ULN	> 5.0 - 20.0 x ULN	> 20.0 x ULN
Creatinine ↑	umol/L	CREAT	**61.9 - 115 umol/L** or 0.7 – 1.3 mg/dL *(88.4 x mg/dL = umol/L)*	≤ ULN	> ULN - 1.5 x ULN	> 1.5 - 3.0 x ULN	> 3.0 - 6.0 x ULN	> 6.0 x ULN
Creatinine kinase↑	U/L	CK	**30 - 170 U/L** or 0.5 – 2.83 ukat/L *(60 x ukat/L = U/L)*	≤ ULN	> ULN - 2.5 x ULN	> 2.5 - 5.0 x ULN	> 5.0 - 10.0 x ULN	> 10.0 x ULN
Albumin (Hypoalbuminemia)	g/L	ALB	**35 - 55 g/L** or 3.5 to 5.5 g/dL	≥ LLN	< LLN - 30 g/L	< 30 - 20 g/L	< 20 g/L	-
Total Cholesterol ↑	mmol/L	CHOL	**3.88 – 5.15 mmol**/L or 150 - 199 mg/dL *(38.67 x mg/dL = mmol/L)*	≤ ULN	> ULN - 7.75 mmol/L > ULN - 300 mg/dL	> 7.75 -10.34 mmol/L > 300 – 400 mg/dL	>10.34-12.92 mmol/L > 400 – 500 mg/dL	>12.92 mmol/L > 500 mg/dL
Lipase ↑	U/L	LIPASE	**<95 U/L** or <1.58 ukat/L *(60 x ukat/L = U/L)*	≤ ULN	> ULN - 1.5 x ULN	> 1.5 - 2.0 x ULN	> 2.0 - 5.0 x ULN	> 5.0 x ULN
Amylase ↑	U/L	AMYLASE	**0 - 130 U/L** or 0 – 2.17 ukat/L *(60 x ukat/L = U/L)*	≤ ULN	> ULN - 1.5 x ULN	> 1.5 - 2.0 x ULN	> 2.0 - 5.0 x ULN	> 5.0 x ULN
Uric acid (Hyperuricemia)	umol/L	URATE	**150 - 470 umol/L** or 2.5 – 8 mg/dL *(59.48 x mg/dL = umol/L)*	Defined by clinical criteria only in CTCAE V5				

ULN = Upper Limit of Normal range; LLN = Lower Limit of Normal range

CTC grades for laboratory values in Novartis Oncology (based on CTCAE v5 – Nov 2017) **Page 2**

				CTC Grades[1]				
Lab test (toxicity)	**SI unit**	**Lab test (NCDS)**	**Normal ranges (Merck manual, July 2015) and *conversion factors***	**0**	**1**	**2**	**3**	**4**
Phosphorus (Hypophosphatemia)	mmol/L	PHOS	**0.97 – 1.45 mmol/L** or 3.0 - 4.5 mg/dL *(0.32 x mg/dL = mmol/L)*	Defined by clinical criteria only in CTCAE V5				
Calcium (corrected) (Hypercalcemia)	mmol/L	CACALC	**2.2 - 2.6 mmol/L** or 9 - 10.5 mg/dL *(0.2495 x mg/dL = mmol/L)*	≤ ULN	> ULN - 11.5 mg/dL > ULN - 2.9 mmol/L	> 11.5 - 12.5 mg/dL > 2.9 - 3.1 mmol/L	> 12.5 - 13.5 mg/dL > 3.1 - 3.4 mmol/L	> 13.5 mg/dL > 3.4 mmol/L
Calcium (corrected) (Hypocalcemia)	mmol/L	CACALC		≥ LLN	< LLN - 8.0 mg/dL < LLN - 2.0 mmol/L	< 8.0 - 7.0 mg/dL < 2.0 - 1.75 mmol/L	< 7.0 - 6.0 mg/dL < 1.75 - 1.5 mmol/L	< 6.0 mg/dL < 1.5 mmol/L
Magnesium (Hypermagnesemia)	mmol/L	MG	**0.62 – 0.99 mmol/L** or 1.5 – 2.4 mg/dL *(0.4114 x mg/dL = mmol/L)*	≤ ULN	> ULN - 3.0 mg/dL > ULN - 1.23 mmol/L	-	> 3.0 – 8.0 mg/dL > 1.23 – 3.3 mmol/L	> 8.0 mg/dL > 3.3 mmol/L
Magnesium (Hypomagnesemia)	mmol/L	MG		≥ LLN	< LLN - 1.2 mg/dL < LLN - 0.5 mmol/L	< 1.2 - 0.9 mg/dL < 0.5 - 0.4 mmol/L	< 0.9 - 0.7 mg/dL < 0.4 - 0.3 mmol/L	< 0.7 mg/dL < 0.3 mmol/L
Glucose (non-fasting) (Hyperglycemia)	mmol/L	GLUCSN	**<7.8 mmol/L** or <140 mg/dL *(0.05551 x mg/dL = mmol/L)*	Defined by clinical criteria only in CTCAE V5				
Glucose (fasting) (Hyperglycemia)	mmol/L	GLUCSF	**3.9 – 5.8 mmol/L** or 70 - 105 mg/dL *(0.05551 x mg/dL = mmol/L)*					
Glucose (Hypoglycemia)	mmol/L	GLUCSN/ GLUCSF		≥ LLN	< LLN - 55 mg/dL < LLN - 3.0 mmol/L	< 55 - 40 mg/dL < 3.0 - 2.2 mmol/L	< 40 - 30 mg/dL < 2.2 - 1.7 mmol/L	< 30 mg/dL < 1.7 mmol/L
Potassium (Hyperkalemia)	mmol/L	K	**3.5 - 5.0 mmol/L** *(0.2558 x mg/dL = mEq/L = mmol/L)*	≤ ULN	> ULN - 5.5 mmol/L	> 5.5 - 6.0 mmol/L	> 6.0 - 7.0 mmol/L	> 7.0 mmol/L
Potassium (Hypokalemia)	mmol/L	K		≥ LLN	< LLN - 3.0 mmol/L	-	< 3.0 - 2.5 mmol/L	< 2.5 mmol/L
Sodium (Hypernatremia)	mmol/L	SODIUM	**136 - 145 mmol/L** *(0.435 x mg/dL = mEq/L = mmol/L)*	≤ ULN	> ULN - 150 mmol/L	> 150 - 155 mmol/L	> 155 - 160 mmol/L	> 160 mmol/L
Sodium (Hyponatremia)	mmol/L	SODIUM		≥ LLN	< LLN - 130 mmol/L	< 129 - 125 mmol/L	< 124 - 120 mmol/L	< 120 mmol/L
Triglyceride ↑	mmol/L	TRIG	**< 2.82 mmol/L** or < 250 mg/dL *(0.01129 x mg/dL = umol/L)*	< 150 < 1.71	≥ 150 - 300 mg/dL ≥ 1.71 – 3.42 mmol/L	> 300 - 500 mg/dL > 3.42 – 5.7 mmol/L	> 500 - 1000 mg/dL > 5.7 – 11.4 mmol/L	> 1000 mg/dL > 11.4 mmol/L
Coagulation								
INR↑	1	INR	**0.8 – 1.2**	≤ 1.2	> 1.2- 1.5	> 1.5- 2.5	> 2.5	-
Activated partial thromboplastin time↑	sec	APTT	**25 - 35 sec**	≤ ULN	> ULN - 1.5 x ULN	> 1.5 - 2.5 x ULN	> 2.5 x ULN	-
Fibrinogen ↓	g/L	FIBRINO	**1.5 – 3.5 g/L** or 150 – 350 mg/dL *(0.01 x mg/dL = g/L)*	≥ LLN	< LLN - 0.75 x LLN	< 0.75 - 0.5 x LLN	< 0.5 - 0.25 x LLN	< 0.25 x LLN

ULN = Upper Limit of Normal range; LLN = Lower Limit of Normal range

(1) LAB CTC grades 1, 2, 3, 4 overrule the study specific (central or local) normal range criteria, e.g. if ULN of Sodium is 151 mmol/L and the value is 151 mmol/L, CTC grade 2 is assigned although the value is ≥ ULN.
Clinical criteria such as 'asymptomatic' or 'Life-threatening consequences' are <u>not</u> considered for determination of LAB CTC grades. Concomitant usage of therapy is also not considered.
Values and LNRs for blood differentials can be given as %, absolute values should then be calculated using WBC. Generally, ≥ 1.5 x 109 /L (lymphocytes) and ≥ 2 x 109 /L (neutrophils) are considered as LAB CTC grade 0
The **comparison with baseline is <u>not</u> considered** for derivation of LAB CTC grades

5.3.1.1 Imputation Rules

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

If laboratory values are provided as '<X' (i.e., below limit of detection) or '>X', prior to conversion of laboratory values to SI unit, then these numeric values are set equal to X.

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

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

Further derivation of laboratory parameters might be required for CTCAE grading. For instance, corrected calcium can be derived using the reported total calcium value and albumin at the same assessment using the following formula:

Corrected Calcium (mg/dL) = Calcium (mg/dL) – 0.8 [Albumin (g/dL)-4]

In order to apply the above formula, albumin values in g/L will be converted to g/dL by multiplying by 0.1, and calcium values in mmol/L will be converted to mg/dL by dividing by 0.2495. For calculation of laboratory CTC grades 0 and 1, the normal range for derived corrected calcium is set to the same limits (in mg/dL) as for calcium.

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

5.4 Derivation of treatment exposure endpoints

5.4.1 Duration of exposure to chemotherapy

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

The last date of exposure to chemotherapy isthe date of the last administration of a non-zero dose of the drug.

5.4.2 Cumulative dose

Cumulative dose of chemotherapy is defined as the total dose given during the treatment exposure and will be summarized for each of the treatment components.

The **actual cumulative dose** for a chemotherapy refers to the total actual dose administered, over the duration for which the subject is on that treatment.

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

5.4.3 Dose intensity and relative dose intensity

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

DI (dosing unit / unit of time) = Actual Cumulative dose (dosing unit) / Duration of exposure to *chemotherapy* (unit of time).

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

DI will be summarized for combination chemotherapy regimens separately for each of the treatment components, but using the duration of exposure of each of the components.

5.4.4 Dose reductions, interruptions or permanent discontinuations

The number of subjects who have dose reductions, permanent discontinuations or interruptions, and the reasons, will be summarized separately for each of the treatment components.

'Dose changed', 'Dose interrupted', and 'Dose permanently discontinued' fields from the Study treatment chemotherapy CRF pages will be used to determine the dose reductions, dose interruptions, and permanent discontinuations, respectively.

The corresponding field 'Reason for change/interruption/discontinued' will be used to summarize the reasons.

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

Reduction: A dose change where the prescribed dose level is lower than the previous prescribed dose level or where the actual dose administered/total daily dose is lower than the calculated dose amount based on the prescribed dose. Therefore any dose change to correct a dosing error will not be considered a dose reduction. Only dose change is collected in the CRF, number of reductions will be derived programmatically based on the change and the direction of the change.

5.5 Statistical models

5.5.1 Analysis of time-to-event data

5.5.1.1 Hypothesis testing

The following one-sided hypothesis will be tested using the stratified log-rank test to address the primary efficacy objectives EFS, and secondary objective OS:

$H_{01}: \theta \geq 1$ vs. $H_{A1}: \theta < 1$

where θ is the HR for tisagenlecleucel arm vs. SOC arm.

The stratified log-rank test can be implimented using the example SAS codes attached below

```
proc lifetest data = dataset notable plots=none;
time time * censor(1);
strata str_1 str_2 str 3/ group = trt test=LOGRANK
diff=control('SOC');
run;
```

Please note that the above SAS codes provides the two-sided p-value pchisq. The following SAS codes/algorithm may be used to calcuate the corresponding one-sided p-value upchisq:

```
if estimate < 0 then upchisq = pchisq / 2;
else if estimate > 0 then upchisq = 1 - pchisq / 2;
upchisq = min(1, upchisq);
```

where **estimate** is the coefficient of treatment from PROC PHREG, **pchisq** is the two-sided p-value from PROC LIFETEST, and **upchisq** is the corresponding one-sided p-value.

5.5.1.2 Kaplan-Meier estimates

For time-to-event endpoints (EFS, OS, DOR), an estimate of the survival function in each treatment group will be constructed using the KM (product-limit) method as implemented in PROC LIFETEST with METHOD=KM option. The PROC LIFETEST statement will use the option CONFTYPE=LOGLOG.

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

5.5.1.3 Hazard ratio

The HR will be estimated by fitting a Cox proportional hazards model using the SAS procedure PHREG (with TIES=EXACT option in the MODEL statement).

A stratified unadjusted Cox model will be used, i.e. the MODEL statement will include the treatment group variable as the only covariate, and the STRATA statement will include the stratification variables.

The HR will be presented along with its two-sided 95% confidence interval (based on the Wald test).

In addition, a stratified and covariate adjusted Cox model will also be performed, i.e. the MODEL statement will include the treatment group, and covariates age, gender, race, ECOG performance status at baseline, histological subgroup, stage of disease at study entry and DLBCL subtype, and the STRATA statement will include the stratification variables.

5.5.1.4 Treatment of ties

The STRATA statement in LIFETEST procedure will be used to analyze time to event data with ties. The PHREG procedure in SAS with option TIES=EXACT will be used to fit the Cox proportional hazards model.

5.5.1.5 Checking proportionality of hazard assumption

Plots (SURVIVAL LOGSURV LOGLOGS) generated by the LIFETEST procedure in SAS will be used to provide visual checks of the proportional hazards assumption.

- SURVIVAL plots the estimated survivor functions: the curves should be similar if hazards are proportional
- LOGSURV plots the cumulative hazard functions : the larger cumulative hazard should be a multiple of the smaller cumulative hazards if hazards are proportional
- LOGLOGS plots log (cumulative hazard). The LOGLOG plot will show parallel curves if the hazards are proportional.

5.5.2 Analysis of binary data

ORR will be summarized in terms of percentage rates with 95% CIs. An exact binomial confidence interval (implemented using SAS procedure FREQ with EXACT statement for one-way tables) will be calculated [Clopper & Pearson, 1934].

SAS procedure FREQ will be used to estimate the proportion of responders (binary outcome =1 or "Yes"), along with the associated 95% (=100 × (1 – *two-sided alpha level*)) two-sided Pearson-Clopper CI. These estimates are obtained using the following example SAS codes:

```
PROC FREQ DATA = dataset;
TABLE binary event / binomial(level = "Yes") alpha = two-sided alpha level;
EXACT binomial;
RUN;
```

When there are no responders, SAS does not produce a CI by default. To obtain a CI in this situation, PROC FREQ is used as specified above except changing **level**="No". From the results of this modified procedure, the values in percent of the LCL and UCL of a 0% response rate are calculated as follows:

$LCL_{LEVEL="Yes"}(\%) = 100\% - UCL_{LEVEL="No"}(\%)$

$UCL_{LEVEL="Yes"}(\%) = 100\% - LCL_{LEVEL="No"}(\%)$

5.6 Time windows

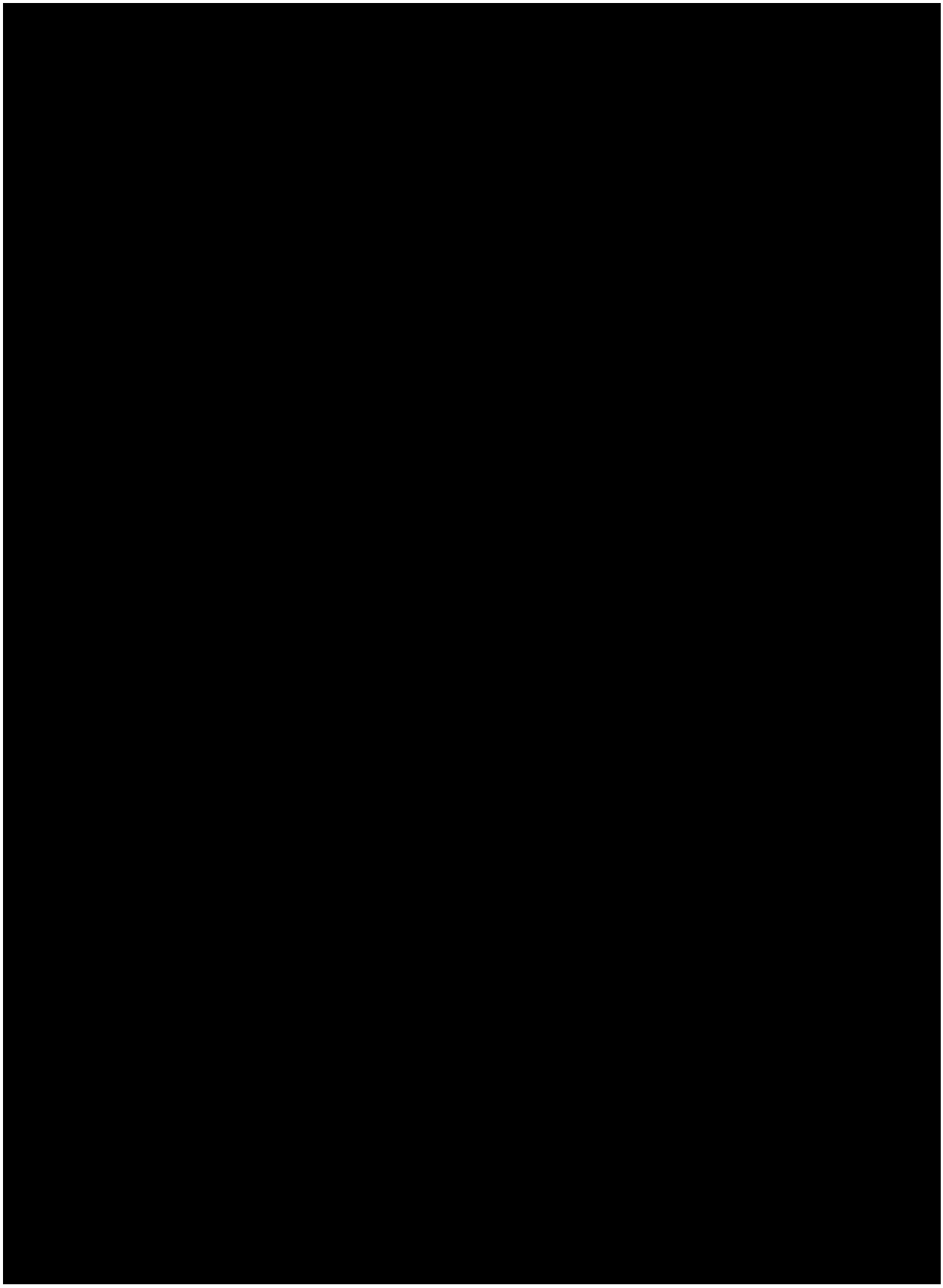
In order to summarize [redacted] immunogenicity, PK, and PRO data over time, assessments will be time-slotted using the following time windows. These windows will be based on the study evaluation schedule and should comprise a set of days around the nominal visits. As a general rule, the following steps are followed to determine the cutoffs for post-baseline time windows:

- Transform all scheduled assessment time points into study days, assuming 1 month = 30.4375 days. Middle points of scheduled assessments are determined.
- The time window associated with the previous assessment ends prior to the middle point; the time window associated with the latter assessment begins after the middle point. In case the middle point is an exact study day, it will belong to the previous assessment.
- The time window of first post-baseline assessment starts with Day 2, unless otherwise indicated.

If more than one assessment is done within the Baseline time window, the last assessment in the baseline time window will be used. For all other time windows, the assessment closest to the planned assessment date will be used; if two or more assessments are equidistant from the planned date, then the mean value will be used.

For each analysis, baseline is either defined as the most recent sample/assessment prior to randomization or the most recent sample/assessment prior to tisagenlecleucel infusion. Visit timing and time window definitions can similarly be defined either based on study day (i.e. from randomization) or on days since tisagenlecleucel infusion.

[redacted]



[redacted]

Table 5-6 shows the defined time windows for immunogenicity samples based on the TIS with timings based on the infusion date.

Table 5-6 Time windows for immunogenicity analyses based on tisagenlecleucel infused set

Time Window	Planned visit timing (infusion day)	Time Window Definition (Infusion day)
Humoral & Cellular Immunogenicity		
Infusion -1d	-1	< Day 1 pre-infusion
Infusion +13d ±3d	14	Day 1 post infusion to 21
Infusion +28d ±7d	29	22 to 54
M4 ±14d*	80	55 to 110
M6 ±14d*	141	111 to 232
M12 ±14d*	323	233 to 414
M18 ±14d*	506	415 to 597
M24 ±14d*	689	598 to 1236
M60 ±14d*	1784	≥ 1237

Infusion Day 1 = start date of tisagenlecleucel infusion

* Sample collection based on randomization date, therefore planned timing in relation to infusion is calculated assuming that infusion happens 6 weeks after randomization

Table 5-7 shows the defined time windows for PK samples based on the CKAS with timings based on the infusion date.

Table 5-7 Time windows for PK analysis based on CKAS

Time Window	Planned visit timing (Infusion day)	Time Window Definition (Infusion day)
CTL019 pharmacokinetics by q-PCR in peripheral blood		
Infusion (pre-infusion)	1	< Day 1 pre-infusion
Infusion +1d	2	Day 1 post-infusion to 3
Infusion +3d ±1d	4	4 to 5
Infusion +6d ±1d	7	6 to 9
Infusion +10d ±3d	11	10 to 12
Infusion +13d ±3d	14	13 to 18
Infusion +21d ±3d	22	19 to 25
Infusion +28d ±7d	29	26 to 35
Week 12 ±7d*	42	36 to 61
M4 ±14d*	80	62 to 110
M6 ±14d*	141	111 to 186

M9 ±14d*	232	187 to 277
M12 ±14d*	323	278 to 414
M18 ±14d*	506	415 to 597
M24 ±14d*	689	598 to 871
M36 ±14d*	1054	872 to 1236
M48 ±14d*	1419	1237 1601
M60 ±14d*	1784	≥ 1602
[redacted]		
CTL019 pharmacokinetics by q-PCR in bone marrow aspirate		
Week 12 ±7d*	42	≥ Day 1 post infusion

Infusion Day 1 = start date of tisagenlecleucel infusion

* Sample collection based on randomization date, therefore planned timing in relation to infusion is calculated assuming that infusion happens 6 weeks after randomization

Table 5-8 shows the defined time windows for PRO assessments for analyses based on the FAS with timings based on randomization date.

Table 5-8 Time windows for PRO analyses based on the full analysis set

Time Window	Planned visit timing (study day)	Time Window Definition (Study day)
SF-36v2, FACT-LYM, EQ-5D-5L		
D1 Randomization	1	≤ Study Day 1
Week 6 ±14d	42	2 to 63
Week 12 ±7d	84	64 to 133
M6 ±14d	183	134 to 228
M9 ±14d	274	229 to 319
M12 ±14d	365	320 to 456
M18 ±14d	548	457 to 639
M24 ±14d	731	640 to 913
M36 ±14d	1096	914 to 1278
M48 ±14d	1461	1279 to 1643
M60 ±14d	1826	≥ 1644

Study Day 1 = randomization date

6 Reference

1. Clopper CJ and Pearson ES (1934). The use of confidence or fiducial limits illustrated in the case of the binomial. Biometrical, 26, 404-413.
2. Brookmeyer R and Crowley J (1982). A Confidence Interval for the Median Survival Time. Biometrics, 38, 29 - 41.
3. Collet D (1994). Modelling survival data in medical research. London, Chapman & Hall.
4. Lee D, Gardner R, Porter DL, et al (2014) Current concepts in the diagnosis and management cytokine release syndrome. Blood; 124 (2): 188-195.

5. Xu Z, Zhen B, Park Y, Zhu B (2017). Designing therapeutic cancer vaccine trials with delayed treatment effect. Stat Med, 36(4): 592-605.
6. Zhang X (2013). Comparison of restricted mean survival times between treatments based on a stratified Cox model. Bio-Algorithms and Med-Systems; 9(4): 183-189.
7. Zhao L, Claggett B, Tian L, et al (2015). On the restricted mean survival time curve in survival analysis. Biometrics. 72, 215-221.