NOVARTIS

Clinical Development

CTL019 (tisagenlecleucel)

CCTL019B2202 / NCT02435849

A Phase II, single arm, multicenter trial to determine the efficacy and safety of CTL019 in pediatric patients with relapsed and refractory B-cell acute lymphoblastic leukemia

Statistical Analysis Plan
Amendment 4

Author: [redacted]

Document type: RAP Documentation

Document status: Final 6.0

Release date: 27-May-2019

Number of pages: 45

Document History

Version	Date	Changes
1.0	17-May-2015	Draft for CTT review
1.1	21-May-2015	Draft for sDLT review, incorporating comments received from CTT
2.0	29-May-2015	Final version
3.0	16-Jul-2016	Amendment 1 Change file name to “Statistical Analysis Plan” following new RAP document naming convention. Protocol Amendment 2,3,4 Related updates: • Update the primary endpoint to be ORR within 3 months post CTL019 infusion instead of 6 months. • Add interim analysis with first 50 infused patients • Add analysis related to CTL019 manufactured at [redacted] • Add key secondary endpoints related to patients infused with CTL019 from US manufacturing facility • Revise the order of main and additional analysis for DOR to be consistent with the protocol • Update samples size calculation • Update details w.r.t. PRO analysis and healthcare resource utilization • Update the visit schedule and primary follow-up length; added summary of duration of study follow-up • Update efficacy baseline definition • Update response status category at study entry • Update AESI search criteria • Update analysis time window • Add summary plan for relapsed patients: site of initial relapse; CD19 status at initial relapse. • Add analysis about CR/CRI with MRD negative at day 28 • Update subgroup analysis plan Related to CSPD discussion • Update analysis for hematopoietic cytopenia • Update analysis of CRS and anti-cytokine therapies • Update PK analysis plan • Update analysis plan for B- and T- cells • Update analysis plan for apheresis product Details/Clarification for Programming • Update growth data analysis plan • Update partial data imputation rule

Version	Date	Changes
		• Update reference table • Definition of treatment failure • Definition of Enrolled set: add 'clinical' to the inclusion/exclusion criteria • Update PK language about imputation of non-quantifiable values; and values with status showing not reliable • Algorithm for censoring after missing two scheduled assessment • Add baseline B cell phenotype and CD19 expression summary • Other minor editorial changes and clarifications
4.0	15-Oct-2016	Amendment 2 • Add analysis of other safety events of interest • Other minor editorial changes and clarifications
5.0	19-Jun-2017	Amendment 3 • Add efficacy subgroup analysis by median time since enrollment to CTL019 infusion, and by number of previous relapses • Add OS sensitivity analysis censoring post infusion SCT • Revise the definition of relapsed disease by grouping chemorefractory into relapsed disease per investigator's feedback • Add analysis of tocilizumab PK • Revise analysis plan of humoral immunogenicity and add analysis of cellular immunogenicity • Revise definition of B cell recovery to take into account additional assay available in secondary follow-up • Other minor editorial changes and clarifications
6.0	27-May-2019	Amendment 4 Updates made to align with analysis plan for rest of world submission • Add efficacy analysis on Enrolled set (for EFS and OS) and in patients who achieved BOR of CR or CRi within Month 3 • Updated AESI section to include summaries based on identified and potential risks and reference to eCRS form. • Add analysis of serious neurological adverse reactions • Add KM analysis of time to resolution for cytopenia Grade 1-4 to Grade 0 • Update KM estimates from % unresolved to % resolved cases based time to resolution based on lab results • Remove analysis of time to first CRS onset using KM method. • Add analysis of fibrinogen • Add analysis of febrile neutropenia per lab exam • Remove analysis of CTL019 product by category of EFS

Table of contents

List of tables

List of figures

List of abbreviations

AE	Adverse Event
AESI	Adverse Event of Special Interest
ALL	Acute Lymphoblastic Leukemia
ATC	Anatomic Therapeutic Chemical (Classification)
AUC	Area Under the Curve
CDC	Centers for Disease Control and Prevention
CI	Confidence Interval
CIF	Cumulative Incidence Function
Cmax	Maximum concentration
CNS	Central Nervous System
CR	Complete remission
CRi	Complete remission with incomplete blood count recovery
CRO	Contract Research Organization
CRP	C-Reactive Protein
CRS	Cytokine Release Syndrome
CSF	Cerebral Spinal Fluid
CTC	Common Toxicity Criteria
DNA	Deoxyribonucleic Acid
DOR	Duration of Remission
ECG	Electrocardiogram
eCRF	electronic Case Report Form
EFS	Event Free Survival
FAS	Full Analysis Set
GVHD	Graft versus Host Disease
IEAS	Interim efficacy analysis set
IL	Interleukin
IRC	Independent Review Committee
KM	Kaplan Meier
LLOQ	Lower Limit of Quantification
LOQ	Limit of Quantification
MRD	Minimal Residual Disease
ORR	Overall Remission Rate
OS	Overall Survival
PD	Pharmacodynamics
PK	Pharmacokinetics

PPS	Per-Protocol Set
q-PCR	Quantitative Polymerase Chain Reaction
RFS	Relapse Free Survival
SCT	Stem Cell Transplantation
SDS	standard deviation score
Tmax	Time to peak concentration
ULOQ	Upper Limit of Quantification

1 Introduction

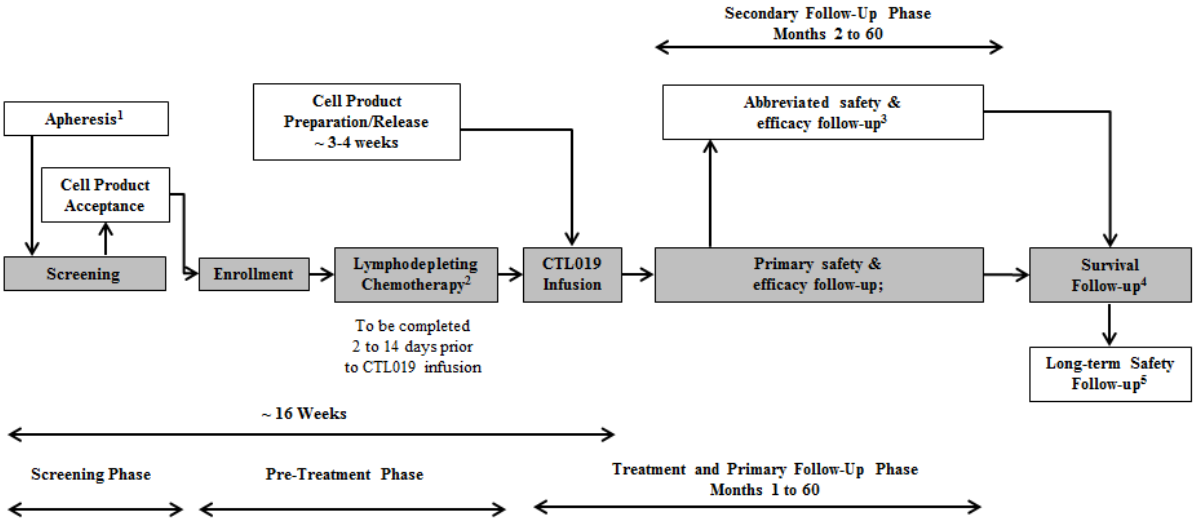
This document describes the detailed statistical methodology for the study CTL09B2202: A Phase II, single arm, multicenter trial to determine the efficacy and safety of CTL019 in pediatric patients with relapsed and refractory B-cell acute lymphoblastic leukemia. The data will be analyzed by Novartis and/or a designated CRO. It is planned that the data from all centers that participate in this protocol will be used.

2 Study design, objectives and endpoints

2.1 Study Design

The target population for this study consists of pediatric patients with B-cell ALL who are refractory, relapsed after allogeneic SCT, or are otherwise ineligible for allogeneic SCT. The study will have the following sequential phases for all patients (see Figure 2-1): Screening, Pre-Treatment (cell product preparation and lymphodepleting chemotherapy), Treatment and Primary Follow-up (60 months), Secondary Follow-up, and Survival Follow-up. The total duration of the study is 5 years.

Figure 2-1 Study design



1 Performed prior to Study Entry
2 As indicated per protocol
3 Only for patients who drop out of the Primary Follow-up before Month 60.
4 Patients will be followed for survival until the end of trial, or until they are enrolled in the long-term follow-up.
5 Long term safety follow-up conducted per health authority guidance under a separate protocol

Only following informed consent/assent and confirmation of all clinical eligibility criteria will the patient's apheresis product be shipped to the manufacturing facility. The manufacturing facility will then evaluate the patient's apheresis product for acceptance. Enrollment is defined as the point at which a patient meets all clinical inclusion/exclusion criteria and the patient's apheresis product is received and accepted by the manufacturing facility.

Following enrollment, lymphodepleting chemotherapy may be started approximately one to three weeks before CTL019 infusion. The purpose of the chemotherapy is to induce lymphopenia in order to facilitate engraftment and homeostatic expansion of CTL019 cells. If patients have a White Blood Cell (WBC) count ≤ 1,000 cells/μL within one week prior to CTL019 infusion, lymphodepleting chemotherapy is not required.

CTL019 infusion will be given 2 to 14 days after completion of lymphodepleting chemotherapy, if lymphodepleting chemotherapy is required. A single dose will be administered. The target cell dose range is 2.0 to 5.0 x 10^6 autologous CTL019 transduced viable T cells per kg body weight (for patients ≤ 50 kg) and 1.0 to 2.5 x 10^8 autologous CTL019 transduced viable T cells (for patients > 50 kg). The allowable infused cell dose range of CTL019 transduced cells have been defined as 0.2 to 5.0 x 10^6 autologous CTL019 transduced viable T cells per kg body weight (for patients ≤ 50 kg) and 0.1 to 2.5 x 10^8 autologous CTL019 transduced viable T cells (for patients > 50 kg). CTL019 products below these minimum transduced cell doses will not be released for infusion.

After CTL019 infusion, efficacy will be assessed monthly for the first 6 months, then quarterly up to 2 years and semi-annually afterwards up to 5 years, or until patient relapse based on the Novartis guidelines for response assessment in ALL (Appendix 1 of protocol), which is based on NCCN version 1.2013 guidelines, Cheson et al (2003) and Appelbaum et al (2007). For patients who discontinued from primary follow-up while in remission, relapse status will be obtained every 3 months in the secondary follow-up until first relapse (if applicable). Safety will be assessed throughout the study. A post-study follow-up for lentiviral vector safety will continue under a separate destination protocol for 15 years post infusion per health authority guidelines.

The end of study is defined as the last patient's last visit (LPLV), which is the last patient's Month 60 evaluation (End of Treatment and Primary Follow-Up or End of Secondary Follow-up visit), or the time of premature withdrawal.

Patients may continue to be followed under the current protocol for relapse and survival until LPLV or until they choose to enroll into the 15 year long term follow-up protocol, whichever occurs first. Once a discontinued patient relapses, the patient will only be followed for survival. The relapse and survival follow-ups can be conducted in the form of telephone contact.

2.2 Study objectives and endpoints

Objectives and related endpoints are provided in Table 2-1 and detailed in the study protocol.

Table 2-1 Study objectives and endpoints

Objective	Endpoint
Primary	
Evaluate the efficacy of CTL019 therapy from all manufacturing facilities as measured by overall remission rate (ORR) during the 3 months after CTL019 administration, which includes CR and CR with incomplete blood count recovery (CRi) as determined by IRC assessment	ORR (= CR + CRi); See Protocol Appendix 1 for response definition
Key secondary	
Evaluate the efficacy of CTL019 therapy from US manufacturing facility as measured by overall remission rate (ORR) during the 3 months after CTL019 administration, which includes CR and CR with incomplete blood count recovery (CRi) as determined by IRC assessment	ORR (= CR + CRi) assessment; See Appendix 1 for response definition
Evaluate the percentage of patients who achieve a best overall response (BOR) of CR or CRi with a MRD negative bone marrow by central analysis using flow cytometry among all patients who receive CTL019 from all manufacturing facilities	Percentage of patients with BOR of CR or CRi with MRD negative bone marrow by flow cytometry during the 3 months after CTL019 infusion among all patients who are infused with CTL019 from all manufacturing facilities
Evaluate the percentage of patients who achieve a best overall response (BOR) of CR or CRi with a MRD negative bone marrow by central analysis using qPCR among patients who receive CTL019 from US manufacturing facility	Percentage of patients with BOR of CR or CRi with MRD negative bone marrow by qPCR during the 3 months after CTL019 infusion among all patients who are infused with CTL019 from US manufacturing facility
Other secondary	
Evaluate the percentage of patients who achieve CR or CRi at Month 6 without stem cell transplant (SCT) between CTL019 infusion and Month 6 response assessment	Percentage of patients who achieve CR or CRi at Month 6 without SCT between CTL019 infusion and Month 6 response assessment
Evaluate the percentage of patients who achieve CR or CRi and then proceed to SCT while in remission before Month 6 response assessment Evaluate the duration of remission (DOR)	• Percentage of patients who achieve CR or CRi and then proceed to SCT while in remission prior to Month 6 response assessment • In addition, all patients that proceed to SCT after CTL019 infusion will be described • DOR, i.e. the time from achievement of CR or CRi, whichever occurs first, to relapse or death due to ALL • Site of involvement of subsequent relapse will be summarized

Objective	Endpoint
Evaluate the relapse-free survival (RFS)	RFS, i.e. the time from achievement of CR or CRi whichever occurs first to relapse or death due to any cause during CR or CRi
Evaluate the event-free survival (EFS)	EFS, i.e. the time from date of CTL019 infusion to the earliest of death, relapse or treatment failure
Evaluate the overall survival (OS)	OS, i.e. the time from date of CTL019 infusion to the date of death due to any reason
Evaluate the response at Day 28 +/- 4 days	Proportion of patients attaining CR or CRi at Day 28 +/- 4 days post CTL019 infusion
Evaluate the impact of baseline tumor burden on response	Response as a function of baseline tumor burden (tumor load) (MRD, extramedullary disease, etc)
Evaluate the quality of response using MRD disease assessments before treatment, and at day 28 +/- 4 days after treatment using central assessment by qPCR and before SCT by local assessment (flow or PCR)	MRD quantitative result (% leukemic cells) and qualitative result (positive/negative)
Evaluate the safety of CTL019 therapy	Type, frequency and severity of adverse events and laboratory abnormalities
Characterize the *in vivo* cellular pharmacokinetic (PK) profile (levels, persistence, trafficking) of CTL019 cells in target tissues (blood, bone marrow, cerebral spinal fluid, and other tissues if available)	- q-PCR detected DNA encoding anti-CD19 chimeric antigen receptor (CTL019) in blood, bone marrow and CSF - Cmax, Tmax, AUCs and other relevant PK parameters of CTL019 in blood, bone-marrow, CSF if available - Persistence of CTL019 in blood, bone marrow and CSF (if available) (e.g. Mean Residence Time [MRT] last)
Describe the incidence of newly acquired and confirmed immunogenicity to CTL019	- Incidence of newly acquired and confirmed immunogenicity and anti-CTL019 assay titers
Describe the effect of CTL019 therapy on Patient Reported Outcomes (PRO)	PRO as measured by PedsQL and EQ-5D questionnaires
Derivation of a score to predict cytokine release syndrome	Develop a score utilizing clinical and biomarker data and assess its ability for early prediction of cytokine release syndrome
Describe the profile of soluble immune factors that may be key to cytokine release syndrome	Frequent monitoring of concentrations of soluble immune factors in blood
Describe the levels of B and T cells (peripheral blood and bone marrow) prior to and following CTL019 infusion for safety monitoring	Lymphocyte subsets of B and T cells and description of associated safety events
Assess the efficacy, safety and in vivo cellular pharmacokinetics of patients infused with CTL019 manufactured by [redacted]	- ORR and MRD negative remission - Type, frequency and severity of adverse events and laboratory abnormalities - CTL019 transgene levels by qPCR in blood, bone marrow and CSF if available
Exploratory	

Objective	Endpoint
[redacted]	[redacted]
T cell trafficking (CTL019 immunophenotyping)	- CTL019 positive T cells and other leukocyte subsets
Describe the effect of anti-cytokine therapy on CRS, CTL019 PK/PD, and disease response	- Clinical CRS adverse events and laboratory measures of CRS (e.g. IL-6, , CRP, and ferritin concentrations) by anti-cytokine therapy - CTL019 concentrations and B cell depletion by anti-cytokine therapy - Disease response by anti-cytokine therapy
Quantify the relationship between 1) CTL019 cell product/apheresis product [redacted] 2) other cell product/apheresis product characteristics and clinical endpoints (efficacy, safety, PK)	- [redacted] - Apheresis and cell product characteristics [redacted] - Clinical response (CR, CRi, relapse) - MRD and B cell recovery assay results - PK parameters - CRS status - Cytokine response
To explore the relationship between CRS and initial tumor burden, clinical tumor response, and PK/PD parameters	- CRS occurrence, CRS grade, need for anti-cytokine therapies - Baseline tumor burden - Maximum clinical response - CTL019 concentrations and B cell depletion
[redacted]	[redacted]
Describe hospital resource utilization	- Number of patients with hospitalized infusion, total number of hospitalizations, and length of stay

3 Definitions and general methodology

3.1 Definitions

3.1.1 Study follow-up

Study follow-up consists of treatment and primary follow-up and secondary follow-up (if applicable).

After CTL019 infusion, patients are followed in the "treatment and primary follow-up phase" for up to 60 months for disease response assessment (blood, bone marrow and extramedullary disease etc.) and safety. When patient is unable to be followed in the primary follow-up phase, the patient will enter a secondary follow-up phase with reduced data collection schedule to collect remission/relapse information and health authority requested data (e.g. delayed adverse events, etc.). It is anticipated that patients may leave the primary follow-up and move to secondary follow-up due to reasons including: treatment failure, relapse after remission, pursuing SCT while in remission, or withdrawal from the primary follow-up (see protocol Section 7.1).

The end of primary follow-up will be refer to completion or discontinuation date of the treatment and primary follow-up phase. The study follow-up completion or discontinuation date will be refer to the completion/discontinuation date of the last phase (i.e. treatment and primary follow-up or secondary follow-up) the patient has entered.

3.1.2 Study drug and study treatment

Study drug is defined as CTL019 transduced cells.

Study treatment includes not only the study drug, i.e., CTL019 transduced cells, but also lymphodepleting chemotherapy.

3.1.3 Date of first administration of lymphodepleting chemotherapy

The date of first administration of lymphodepleting chemotherapy is defined as the first date when a non-zero dose of chemotherapy was administered and recorded on the “Concomitant Antineoplastic Therapy” electronic Case Report Form (eCRF) for the indication “Lymphodepleting”.

3.1.4 Date of infusion of study drug

The date of infusion of study drug is defined as the date when a non-zero dose of study drug (CTL019 transduced cells) was administered and recorded on the “Dosage administration record” eCRF.

3.1.5 Date of first study treatment

For patients who received lymphodepleting chemotherapy, the date of first study treatment is the date of first administration of lymphodepleting chemotherapy (as defined in Section 3.1.3);

for patients who did not receive lymphodepleting chemotherapy, the date of first study treatment is the date of infusion of study drug (as defined in Section 3.1.4).

3.1.6 Study day

The study day will be calculated as the difference between the date of the assessment and the date of first infusion of CTL019 (**Day 1**) plus 1 for assessments on or after the date of first infusion. For assessment before the date of first infusion, the study day will be calculated as the difference between the date of the assessment and the date of first infusion of CTL019 (**Day 1**) *(Note: if an event happens before the first day of CTL019 infusion then the study day will be negative.)* For patients who did not receive CTL019 infusion, their study days will not be calculated.

The study day will be displayed in all relevant data listings.

3.1.7 Baseline

For ***baseline disease evaluations***, the most current assessments (bone marrow, blood count, CSF, physical exam, etc.) on or prior to the date of enrollment will be used as the baseline assessment.

In the case that both bone marrow aspirate and biopsy morphological results are available the highest blasts value will be considered, and the corresponding assessment date will be used as reference for other assessments.

For ***safety evaluations*** (i.e. laboratory and vital signs), the last available assessment before CTL019 infusion is taken as 'baseline' values.

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

3.1.8 Last contact date

The last contact date will be used for censoring of patients in the analysis of overall survival.

For patients not known to have died as of the analysis cut-off date, the last contact date should be derived as the latest date on or before the data cut-off date from the dates listed in the first column of Table 3-1Last contact date data sources. For each of the sources specific conditions listed in the second column of Table 3-1 have to be fulfilled to ensure that there was true contact with the patient.

No additional dates are allowed to be used, e.g. dates coming from concomitant medications, PRO, etc.

Table 3-1 Last contact date data sources

Source data	Conditions
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.

Source data	Conditions
Any specific efficacy assessment date if available	Evaluation is not missing.
Laboratory/PK 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: completely imputed dates will not be used to derive the last contact date. Partial date's imputation is allowed to be used for event (death) and for censoring date only if coming from Survival Follow-up eCRF page (see Section 5.5.6 for details).

3.1.9 Lost to follow-up

For overall survival 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 analysis (i.e. DOR, RFS and EFS), patients will be considered as lost to follow-up if the patient discontinued the study due to lost to follow-up.

3.2 Data Included in the analysis

Data from all participating centers will be combined.

An interim analysis will be performed when the first 50 patients who receive CTL019 have completed 3 months follow-up from study day 1 infusion or discontinued earlier. At the time of this interim analysis, assessment of all endpoints will be based only on patients who receive CTL019 manufactured from US manufacturing facility because there will be no patients treated with CTL019 manufactured from other manufacturing facilities.

The final analysis of the primary endpoint will be performed after all patients infused with CTL019 have completed 3 months from study day 1 infusion or discontinued earlier.

Selected efficacy and safety analysis will be updated annually afterwards. A final Clinical Study Report (CSR) will be produced once all patients complete or discontinue from the study.

3.3 Definitions of analysis sets

The analysis sets to be used are defined as below. The Interim efficacy analysis set (IEAS) and the Full analysis set (FAS) will be used as the primary efficacy analysis set for the interim and final analysis respectively. The Safety Set will be used for all safety analyses, unless otherwise specified. The Pharmacokinetic analysis set (PAS) will be used for pharmacokinetics analyses.

All tables and listings will be presented by one treatment arm of CTL019, unless otherwise specified.

Screened Set

The Screened Set comprises all patients who have signed informed consent/assent and screened in the study.

Enrolled Set

The Enrolled Set comprises all patients who are enrolled in the study. Enrollment date is defined as the point at which the patient meets all clinical inclusion/exclusion criteria, and the patients' leukapheresis product is received and accepted by the manufacturing facility. In case of protocol deviation such that patients are enrolled without meeting all inclusion/exclusion criteria, such patients will still be considered in the Enrolled Set, if the patients' leukapheresis product is received and accepted by the manufacturing facility.

Full Analysis Set (FAS)

The Full Analysis Set comprises all patients who received infusion of CTL019.

Interim Efficacy Analysis Set (IEAS)

At the time of interim analysis, the Interim Efficacy Analysis Set comprises the first 50 patients who receive CTL019 infusion.

Safety Set

The Safety Set comprises all patients who received infusion of CTL019. Note that the Safety Set and FAS are the same for this study.

Per-Protocol Set (PPS)

The Per-Protocol Set consists of a subset of the patients in the IEAS or FAS (at time of interim and final analysis respectively) who are compliant with major requirements of the study protocol.

Major protocol deviations leading to exclusion from the PPS include:

- Diagnosis of disease other than ALL at baseline;
- Prior therapy does not match with study protocol requirements in terms of number and types of previous therapy regimens;
- Missing or incomplete documentation of disease at baseline;
- CTL019 T-cells was infused to patients without fulfilling either of the following two conditions: (A) meeting all approved manufacturing release criteria; (B) released through exceptional release.

In addition, patients who receive a dose less than the minimum target dose of 2.0×10^6/kg (for patients ≤50 kg) or 1.0×10^8 (for patients >50 kg) CTL019 transduced viable T cells will also be excluded.

Pharmacokinetic Analysis Set (PAS)

The pharmacokinetic analysis set consists of a subset of IEAS or FAS (at time of interim and final analysis respectively) patients who have at least one sample providing evaluable PK data (i.e., samples not flagged for exclusion by the clinical pharmacologist) for CTL019. The PAS will be used for summaries (tables and figures) of PK data, and listings will be provided based on FAS. For any correlation analysis between PK data and other efficacy/safety endpoints, the PAS will be used.

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

Tocilizumab Pharmacokinetic Analysis Set (TPAS)

The Tocilizumab Pharmacokinetic Analysis set (TPAS) consists of patients in FAS who have taken at least one dose of tocilizumab and provided at least one tocilizumab PK concentration.

3.4 Response evaluation for ALL

3.4.1 Response criteria

The ALL response guideline is outlined in the Protocol Appendix 1 - Guidelines for efficacy evaluation in Acute Lymphoblastic Leukemia studies.

The overall disease response is determined at a given evaluation using the criteria described in Table 3-1Last contact date data sources below.

Table 3-2 Overall disease response classification at a given evaluation time

Response category	Definition
Complete remission (CR)	All the following criteria are met: **Bone marrow** • < 5% blasts **Peripheral blood** • Neutrophils > 1.0 x 10^9/L, and • Platelets > 100 x 10^9/L, and • Circulating blasts < 1% **Extramedullary disease** • No clinical evidence of extramedullary disease (by physical exam and central nervous system (CNS) symptom assessment), and • If additional assessments (e.g. CSF assessment by lumbar puncture (LP), CNS imaging, biopsy, etc.) are performed, results must show remission status **Transfusion independency** • No platelet and/or neutrophil transfusions less than or equal to 7 days before peripheral blood sample for disease assessment

Response category	Definition
Complete remission with incomplete blood count recovery (CRi)	All criteria for CR as defined above are met, except that the following exist: • Neutrophils ≤ 1.0 x 10^9/L, and/or • Platelets ≤ 100 x 10^9/L, and/or • Platelet and/or neutrophil transfusions less than or equal to 7 days before peripheral blood sample for disease assessment
No response	Failure to attain the criteria needed for any response categories or relapse
Relapsed Disease	Only in patients with a CR or CRi and who have: • Reappearance of blasts in the blood (≥ 1%), or • Reappearance of blasts in bone marrow (≥ 5%), or • (Re-)appearance of any extramedullary disease after CR or CRi
Unknown	"Unknown" is assigned in case the baseline assessment or the response assessment is not done, incomplete, indeterminate, or not performed within the respective time frame. If there is evidence of relapse, the overall response will assessed as "relapsed disease" with the relapsed component alone.

3.4.2 Establishing CR/CRi and subsequent maintenance of CR/CRi with no clinical evidence of relapse

A full response evaluation, including assessments of peripheral blood, bone marrow, CNS symptoms, physical exam and CSF assessment by LP, is required at the first time a CR or CRi is demonstrated. If the patient is not in CR/CRi at Month 1, then a bone marrow biopsy/aspirate and CSF assessment by LP are also required at the first time clinical evidence of remission is seen by peripheral blood and extramedullary disease assessment (physical exam and CNS symptom assessment) to establish that a patient has achieved CR/CRi for the first time. Additional bone marrow biopsies/aspirates and CSF assessments by LP may be recommended in the protocol.

Complete remissions in patients with ALL have been observed to take place within 1 month after infusion with CTL019. The onset of complete remissions is rapid and dramatic, and patients quickly regain a normal performance status. ALL relapse in the bone marrow is rapidly followed by signs or symptoms of disease recurrence as well as abnormalities in the peripheral blood.

Therefore, following initial achievement of CR/CRi, patients will be considered to have maintained a clinical CR/CRi if the patient has no evidence of extramedullary disease (by physical exam and CNS symptom assessment) and circulating blasts in peripheral blood are <1%.

In order for the best overall disease response to be categorized as CR or CRi, there must be no clinical evidence of relapse as assessed by peripheral blood and extramedullary disease assessment (physical exam and CNS symptom assessment) at a minimum of 4 weeks (28 days) after the initial achievement of CR or CRi. Please note, if additional assessments (e.g. bone marrow, CSF assessment by LP, CNS imaging, biopsy, etc.) are performed in the same evaluation for disease response evaluation purpose, they will also need to show remission status.

The onset date of CR or CRi will then be derived as the evaluation date of the initial CR or CRi assessment.

3.4.3 Date of overall disease response evaluation

A complete evaluation of response includes at the minimum the assessments of peripheral blood, CNS symptoms and physical exam. In addition, bone marrow and CSF assessment may be required. All components of disease assessments must be performed within the specified time frame (See Protocol Appendix 1) to be qualified as the same response evaluation.

If the overall disease response is CR, CRi or Unknown, the evaluation date (i.e. for one evaluation number) is defined as the latest of all dates of required measurements at that evaluation number. This rule applies also in case of multiple measurements of the same variable.

Relapse or No response can be assessed based on a partial evaluation (e.g. a relapse is assessed from blood alone). The assessment date for relapse or no response is calculated as the earliest date of all assessments that reveal a relapse or lack of response.

3.5 Time-to-event definitions

General rule for the calculation of the time to event interval is:

Time to event = event date - start date + 1 (in days)

When no post-baseline assessments of the event are available, the date of CTL019 infusion will be used as end date when time is to be censored at last post-baseline assessment of event, i.e. time to event variables will never be negative.

Often censoring time is determined based on date of adequate response assessment. Any response assessment is considered to be adequate if the assessment was performed and the outcome of the assessment was other than “unknown” or “not done”.

4 Statistical methods used in reporting

4.1 General presentation of descriptive summaries

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

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

4.2 Patient disposition

Patient disposition will be summarized for the following: screening phase for the Screened Set, pre-treatment phase for the Enrolled Set, treatment and primary follow-up phase and secondary follow-up phase for the FAS. The patient disposition for each phase will be summarized for all patients who entered that phase. The number and percentage of patients in each of the categories as listed for “End of Phase Disposition eCRF” pages will be tabulated and listed. Patients who have entered any study phase but have not completed/discontinued will be listed as appropriate.

For the screening phase, the clinical eligibility criteria that were not met by patients will also be tabulated. In addition, the number and percentage of patients who enrolled in the long term follow-up study will be summarized.

In addition, a high level disposition summary including all phases will be provided for all screened patients.

Duration of primary follow-up and total duration of 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.

4.3 Background and demographic characteristics

The FAS (as well as the IEAS at interim analysis) will be used for all baseline disease characteristics and demographic summaries. The Enrolled Set will be used for listings, where patients will be presented by whether they have received CTL019 or not.

4.3.1 Basic demographics data

Demographic and other baseline data will be listed by patient and/or summarized descriptively.

4.3.2 Medical history and ALL disease characteristics

Medical history and ongoing conditions, including cancer-related conditions and symptoms at the time of informed consent will be summarized and listed. Ongoing and historical medical conditions will be flagged separately in the listing. The summaries will be presented by primary system organ class and preferred term. Medical histories are coded using the medical dictionary for regulatory activities (MedDRA) terminology.

The CD19 status, MRD status by central assessment, local morphologic blast count, CNS classification and other extramedullary disease status prior to enrollment will be summarized.

Number and percentage of patients with CNS involvement by ALL at any time prior to enrollment will be summarized.

Other CNS disease history (usually non-leukemic, see Section 5.8) and CNS related prior radiotherapy (e.g. to the brain or cranial spinal axis) will also be summarized.

4.3.3 Prior anti-neoplastic therapy

Number and percentage of patients with prior anti-neoplastic medications/therapies (including medications for hematological disease, radiotherapy and SCT) will be summarized. Number of previous complete remissions, number of previous lines of therapies, setting of last medication (induction, consolidation, maintenance, salvage, conditioning for SCT), best response (including MRD status) of last medication and locations of last radiotherapy will also be summarized.

Prior anti-neoplastic medications for hematological disease will be summarized by anatomic therapeutic chemical (ATC) class, and preferred term.

Patients will also be classified and summarized by their response status at study entry:

- Primary refractory: If patient never had a morphologic CR prior to the study

- Relapsed disease: If patient had a CR from other therapy and relapsed prior to the study

All prior anti-neoplastic medications, radiotherapy and SCT will be listed. The number of previous complete remissions and number of previous lines of therapies will also be listed.

4.3.4 Cytogenetic abnormalities

Number and percentage of patients with cytogenetic abnormalities (yes/no) and those with complex karyotypes (≥5 unrelated abnormalities) at study entry will be summarized. All cytogenetic abnormalities will be listed.

4.3.5 Others

All other data collected at baseline will be listed.

4.4 Protocol deviation summaries

The number and percentage of patients in the Full Analysis Set with any protocol deviation will be tabulated by the deviation category. Major protocol deviations leading to exclusion from the PPS will be summarized.

All protocol deviations will be listed.

4.5 Treatments (study treatment, rescue medication, other concomitant therapies, compliance)

The total cells infused (both cells and cells/kg) and total transduced CTL019 cells infused (both cells and cells/kg) will be listed and summarized using descriptive statistics. Weight provided to the manufacturing facility for CTL019 product manufacturing is used in calculating the weight adjusted doses.

Patients will be categorized as below, within or above the prescribed dose range. Patients with dose interruptions, as recorded in the dosage administration record eCRF, will be listed. Because the study drug of CTL019 is administered via one time infusion, no specific compliance will be summarized other than the CTL019 dose administration.

Prior and concomitant medications and significant non-drug therapies prior to and after the start of infusion will be listed by patient and summarized by ATC class and preferred term.

Antineoplastic therapies, including the lymphodepleting chemotherapies, received after enrollment but prior to infusion will be listed. Patients will also be summarized by the types of lymphodepleting chemotherapies received (i.e. fludarabine based lymphodepleting therapy, non-fludarabine based lymphodepleting therapy and no lymhodepleting therapy).

Transfusions collected per protocol requirement during the study will be listed.

Anti-cytokine medications are given for severe CRS due to CTL019 cells. Number of patients requiring anti-cytokine medications for the management of CRS will be summarized. The frequency and dose of rescue medications will also be summarized by preferred term.

4.6 Efficacy evaluation

4.6.1 Primary efficacy endpoint

The primary objective of the study is to evaluate the efficacy of CTL019 therapy as measured by overall remission rate (ORR) during the 3 months after CTL019 administration, which includes CR and CRi in the FAS.

The primary analysis will be based on the IRC assessment.

In addition, sensitivity analysis will be performed using the local investigator's response assessment instead of the IRC's assessment.

4.6.1.1 Variable

The primary endpoint is the ORR during the 3 months after CTL019 administration as determined by IRC assessment. The ORR is defined as the proportion of patients with a best overall disease response of CR or CRi. The best overall disease response is the best disease response recorded from first CTL019 infusion until start of new anticancer therapy (including SCT).

Best overall response will be assigned according to the following order:

1. CR
2. CRi
3. No response
4. Unknown

The best overall disease response for a patient is always calculated, based on the sequence of overall disease responses.

For the best overall disease response to be categorized as CR or CRi, there must be no clinical evidence of relapse as assessed by peripheral blood and extramedullary disease assessment (physical exam and CNS symptom assessment) at a minimum of 4 weeks (28 days) after the initial achievement of CR or CRi. Please note, if additional assessments of bone marrow and/or CSF are performed in the same evaluation, they will also need to show remission status (Section 3.4.2).

If a patient achieved CR or CRi once, without maintaining for at least 28 days, the best overall response for this patient will be considered as 'No response'. If a patient achieved CR or CRi once, but did not perform any subsequent response assessment, the best overall response for this patient will be considered as 'Unknown'.

See also the Section 3.4 for details regarding the definition of overall disease response.

4.6.1.2 Statistical hypothesis, model, and method of analysis

The primary efficacy analysis will be performed by testing whether the ORR within 3 months is greater than 20% at overall one-sided 2.5% level of significance, i.e.,

$$H_0: p \leq 0.2 \text{ vs. } H_a: p > 0.2.$$

The primary efficacy endpoint, ORR within 3 months, will be analyzed at the interim look and final look following a group sequential design. The ORR will be summarized along with the 2-sided exact Clopper-Pearson confidence intervals (CI) with coverage level determined by the O'Brien-Fleming type α-spending approach according to Lan-DeMets as implemented in East 6.3 (Lan and DeMets, 1983). The study will be considered successful if the lower bound of the 2-sided exact confidence interval for ORR is greater than 20%, so that the null hypothesis that the ORR is less than or equal to 20% can be rejected.

The primary efficacy endpoint, ORR will be analyzed based on the data observed in the IEAS and FAS at interim and final analysis respectively.

In addition, time to response (CR or CRi) will also be summarized descriptively for responders.

4.6.1.3 Handling of missing values/censoring/discontinuations

Patients in the study who are of unknown clinical response will be treated as non-responders.

In case of missing data for the full evaluation required to qualify for a certain response category, the overall evaluation "unknown" will be assigned unless at least one observation was made which qualifies for relapse. Relapse can be determined by the relapsed component alone.

Other missing data are simply noted as missing on appropriate tables/listings.

The censoring rules for time to event endpoints are specified in the corresponding sections in Section 4.6.3.

4.6.1.4 Supportive analyses

The analysis of the primary endpoint will be performed among all patients in the PPS using the same methodology as outlined at interim and final analysis, respectively.

The analysis of primary endpoint will also be performed among all patients in the IEAS or FAS (at interim and final analysis respectively) plus enrolled patients who have discontinued prior to CTL019 infusion.

In addition, the analysis of the primary endpoint will also be performed using all patients in the IEAS or FAS (at interim and final analysis respectively) plus those who satisfy all clinical eligibility criteria and have discontinued prior to CTL019 infusion.

Proportion of patients attaining CR or CRi at Day 28 +/- 4 days post CTL019 infusion, among all patients in the IEAS or FAS (at time of interim and final analysis respectively), will be summarized along with exact 95 % CI.

4.6.2 Key secondary efficacy endpoint

4.6.2.1 ORR within 3 months in all patients infused with CTL019 from US manufacturing facility

The first key secondary objective of the study is to evaluate the efficacy of CTL019 therapy from US manufacturing facilities as measured by overall remission rate (ORR) during the 3 months after CTL019 administration by IRC assessment among patients who receive CTL019

therapy from US manufacturing facility in IEAS and FAS at interim and final analysis respectively.

The hypothesis testing will be performed to test whether the ORR within 3 months is less than or equal to 20% against the alternative hypothesis that ORR is greater than 20%.

This hypothesis testing will only be performed when the primary objective is met, so that the family-wise type I error rate will be controlled at one-sided 2.5% level under this hierarchical testing scheme. The type I error probability will be controlled by using a Lan-DeMets (O'Brien-Fleming) alpha spending function at 2.5% level of significance.

This key secondary endpoint will be summarized along with the 2-sided exact Clopper-Pearson confidence intervals with coverage level according to the above alpha spending function. This key secondary objective will be considered successfully met if the lower bound of the 2-sided exact confidence interval is greater than 20%, so that the null hypothesis above can be rejected.

4.6.2.2 Remission with MRD negative bone marrow in patients infused with CTL019 from all manufacturing facilities

The second key secondary objective of the study is to evaluate the percentage of patients who receive CTL019 from all manufacturing facilities and achieve a BOR of CR or CRi with a MRD negative bone marrow by central analysis using flow cytometry during the 3 months after CTL019 administration. The main analysis of this key secondary objective will be performed among all patients in the IEAS and FAS population at time of interim and final analysis respectively. See Protocol Appendix 1 for details of determination of MRD negativity.

The key secondary efficacy analysis will be performed by testing whether the percentage of MRD negative responder among all patients in IEAS or FAS as defined above is less than or equal to 15% against the alternative hypothesis that it is greater than 15% at overall one-sided 2.5% level of significance, i.e.,

$$H_0: p \leq 0.15 \text{ vs. } H_a: p > 0.15.$$

This hypothesis testing will only be performed when both the primary endpoint and the first key secondary endpoint are met, so that the family-wise type I error rate will be controlled at one-sided 2.5% level under this hierarchical testing scheme. The type I error probability will be controlled by using a Lan-DeMets (O'Brien-Fleming) alpha spending function at 2.5 % level of significance.

This key secondary endpoint will be summarized along with the 2-sided exact Clopper-Pearson confidence intervals with coverage level according to the above alpha spending function. This key secondary objective will be considered successfully met if the lower bound of the 2-sided exact confidence interval is greater than 15%, so that the null hypothesis above can be rejected.

The key secondary endpoint will also be summarized among those who achieve a BOR of CR or CRi during the 3 months after CTL019 administration.

Additional analysis will be done using the qPCR MRD analysis instead of flow cytometry.

The quality of response (i.e. proportion of patients with MRD negative disease response) at day 28 +/- 4 days after treatment using central assessment by flow cytometry will also be summarized. For patients who proceed to SCT in remission, the MRD status before SCT by

local assessment (flow or PCR) will be listed. Both quantitative MRD result and qualitative results (positive/negative) will be analyzed if available.

4.6.2.3 Remission with MRD negative bone marrow in patients infused with CTL019 from US manufacturing facility

The third key secondary objective of the study is to evaluate the percentage of patients who achieve a BOR of CR or CRi with a MRD negative bone marrow by central analysis using flow cytometry during the 3 months after CTL019 administration among all patients who receive CTL019 from US manufacturing facility.

The hypothesis testing will be performed to test whether the above rate is less than or equal to 15% against the alternative hypothesis that it is greater than 15%.

This hypothesis testing will only be performed when both the primary objective and the first two secondary endpoints are met, so that the family-wise type I error rate will be controlled at one-sided 2.5% level under this hierarchical testing scheme. The type I error probability will be controlled by using a Lan-DeMets (O'Brien-Fleming) alpha spending function at 2.5% level of significance.

This key secondary endpoint will be summarized along with the 2-sided exact Clopper-Pearson confidence intervals with coverage level according to the above alpha spending function. This key secondary objective will be considered successfully met if the lower bound of the 2-sided exact confidence interval is greater than 15%, so that the null hypothesis above can be rejected.

4.6.3 Other secondary efficacy endpoints

No formal hypothesis testing is planned other than for the primary objective and key secondary objectives. The other secondary efficacy objectives are outlined in the following sections. IRC assessment will be used in the main analysis of secondary endpoints that involve disease response. Note that IRC assessment is only performed during treatment and primary follow-up phase. Additional relapse information is collected for patients entering secondary follow-up in remission. The relapse information collected during secondary follow-up will be used in all time to event analysis that involves disease response assessment.

4.6.3.1 Percentage of patients who achieve CR or CRi at Month 6 without SCT between CTL019 infusion and Month 6 response assessment

This analysis will be conducted when all patients have completed 6 months post CTL019 infusion or have discontinued earlier, and hence will not be conducted at the interim analysis.

The percentage of patients who are in CR or CRi at Month 6 without SCT (post CTL019 infusion) between CTL019 infusion and Month 6 response assessment, among all patients in FAS, will be summarized with exact 95% CI. In addition, the percentage will also be summarized among all patients who achieved CR or CRi.

The patient will be considered to be in CR or CRi at Month 6 if there is at least one CR or CRi assessment after day 167 (i.e. >30.4375 × 5.5) without any relapse prior to this CR or CRi assessment. If such patient does not have SCT prior to Month 6, this patient is considered as

having achieved CR or CRi at Month 6 without SCT between CTL019 infusion and Month 6 response assessment.

Here the time of proceeding to SCT is defined as the time of commencing the conditioning regimen as required for hematopoietic SCT. This definition applies to all analyses involving SCT.

4.6.3.2 Percentage of patients who achieve CR or CRi and then proceed to SCT while in remission before Month 6 response assessment

This analysis will be conducted when all patients have completed 6 months post CTL019 infusion or have discontinued earlier, and hence will not be conducted at the interim analysis.

The percentage of patients who achieve CR or CRi and then proceed to SCT during remission before Month 6 response assessment, among all patients in FAS will be summarized with exact 95% CI. In addition, the percentage will also be summarized among all patients who achieved CR or CRi. All patients that proceed to SCT post CTL019 infusion will be listed.

For patients who discontinue and undergo SCT before the scheduled Month 6 evaluation, they will be considered to have met this secondary endpoint if the patients are still in morphologic remission, i.e. the DOR is not lost or censored.

The "Month 6" evaluation is as defined in Section 4.6.3.1.

4.6.3.3 Duration of remission (DOR)

Duration of remission is defined as the duration from the date when the response criteria of CR or CRi is first met to the date of relapse or death due to underlying cancer.

In the main analysis of DOR (Method 1), in case a patient does not have relapse or death due to underlying cancer prior to data cutoff, DOR will be censored at the date of the last adequate disease assessment on or prior to the earliest censoring event (except for SCT). The censoring reason could be

- Ongoing without event
- Lost to follow-up
- Withdrew consent
- New anticancer therapy (also see below for handling SCT)
- Adequate assessment no longer available
- Event after at least two missing scheduled disease assessments

In addition, if there are any patients who respond to CTL019 but experience death due to any reason other than ALL, death due to reason other than ALL will be considered as a competing risk event to other events of interest (relapse or death due to ALL). Sensitivity analyses will be performed in which death due to reason other than ALL will be censored.

As SCT may be a further treatment option in responding patients, it is appropriate to consider the date of SCT as censoring date, instead of censoring at the last disease assessment date.

A sensitivity analysis will be performed in which the date of relapse or death (if due to the underlying cancer) after SCT will be used for the calculation of DOR as a sensitivity analysis.

If a patient receives SCT after a CR or CRi, relapse or survival status after SCT will be recorded on the corresponding follow-up eCRFs, although data on individual disease response components (e.g. bone marrow) will not be collected. Censoring due to SCT (Method 1) will overestimate the rate of relapse and therefore may be considered inappropriate for the main analysis when a substantial number of patients choose to receive SCT (CHMP 2010). Therefore the above described sensitivity analysis will be performed if there is at least 1 patient with SCT after CTL019 infusion while in remission.

The proposed analyses for DOR are summarized in Table 4-1 below.

Table 4-1 Analyses of duration of remission

	Death due to reason other than underlying cancer	SCT after remission
Method 1	Competing risk analysis	Censor at time of SCT
Method 2	Censor at last adequate disease assessment	Censor at time of SCT
Method 3	Competing risk analysis	Ignore SCT
Method 4	Censor at last adequate disease assessment	Ignore SCT

DOR will be assessed only in patients with the best overall response of CR or CRi. The estimated percentage of relapsed patients (at 6 months, 12 months, etc.) will be presented with 95% confidence intervals using the cumulative incidence function (CIF) or the Kaplan-Meier (KM) method.

For Method 1 and Method 3, the CIF is used to estimate the probability of the event of interest in the presence of the competing risks (Kim 2007). These analyses will only be performed if there is at least 1 patient with competing risk event.

For Method 2 and Method 4, the distribution function of DOR will be estimated using the KM method. The median DOR along with 95% confidence intervals will be presented if appropriate.

If a considerable number of patients receive SCT while in remission after CTL019 infusion, then exploratory analyses may be performed on patients who achieve CR/CRi after CTL019 infusion to assess the effect of SCT on DOR. Baseline disease characteristics and post-baseline factors (e.g. time to CR/CRi, minimal residual disease) that may be correlated with the decision to receive SCT and with DOR will be identified. A Cox model with SCT as a time dependent covariate and potential confounding factors as additional covariates may then be explored in patients who achieve CR/CRi after CTL019 infusion. The hazard ratio (SCT v/s No SCT after CR/CRi) estimate along with its 95% confidence interval will be provided. Additional exploratory analyses may be considered to account for the confounding factors.

For relapse patients, the following characteristics of the initial relapse will be summarized:

- Site of initial relapse:
  - Bone marrow or peripheral blood relapse
    - With extramedullary relapse
    - Without extramedullary relapse
    - Unknown extramedullary status
  - Extramedullary only relapse

- CD19 status of initial bone marrow or peripheral blood relapse: Determined by ALL phenotyping from bone marrow or peripheral blood flow cytometry assessment:
  - CD19 positive
  - CD19 dim
  - CD19 negative
  - CD19 positive/negative
  - Unknown

If CD19 status is obtained from both bone marrow and peripheral blood, the bone marrow result will be used.

4.6.3.4 Relapse free survival (RFS)

Relapse free survival is measured by the time from achievement of CR or CRi whatever occurs first to relapse or death due to any cause during CR or CRi.

In case a patient does not have relapse or death due to any cause prior to data cutoff, RFS will be censored at the date of the last adequate disease assessment on or prior to the earliest censoring event (except for SCT). The censoring reason could be

- Ongoing without event
- Lost to follow-up
- Withdrew consent
- New anticancer therapy (see below for handling SCT)
- Adequate assessment no longer available
- Event after at least two missing scheduled disease assessments

In the main analysis of RFS, patients who proceed to SCT after CTL019 infusion will be censored at the time of SCT (see Section 4.6.3.3 for the rationale). In addition, a sensitivity analysis of RFS will be performed without censoring SCT, if there is at least 1 patient with SCT after CTL019 infusion while in remission.

RFS will be assessed only in patients with the best overall response of CR or CRi. The distribution function of RFS will be estimated using the KM method. The median RFS along with 95% confidence intervals will be presented if appropriate.

4.6.3.5 Event free survival (EFS)

Event free survival is the time from date of first CTL019 infusion to the earliest of the following:

- Death from any cause after remission
- Relapse
- Treatment failure: Defined as no response in the study and discontinuation from the study due to any of the following reasons:
  - Death
  - Adverse event

- Lack of efficacy
- New anticancer therapy

In case of treatment failure, the event date will be set to study Day 1 (CHMP 2010). In addition, a sensitivity of EFS will be performed by considering time of discontinuation from the study as the event time for treatment failure, instead of setting to study Day 1.

In case a patient does not have relapse, death due to any cause or treatment failure (e.g. discontinuation as a result of withdrawal of consent, lost to follow-up, protocol violation or administrative problems) prior to data cutoff, EFS is censored at the last adequate disease assessment date on or prior to the earliest censoring event (except for SCT). The censoring reason could be

- Ongoing without event
- Lost to follow-up
- Withdrew consent
- New anticancer therapy (see below for handling SCT)
- Adequate assessment no longer available
- Event after at least two missing scheduled disease assessments

In the main analysis of EFS, patients who proceed to SCT while in remission after CTL019 infusion will be censored at the time of SCT (see Section 4.6.3.3 for the rationale). In addition, a sensitivity analysis of EFS will be performed without censoring SCT, if there is at least 1 patient with SCT after CTL019 infusion while in remission.

EFS will be assessed in all patients (IEAS, FAS and Enrolled Set) and in those who achieve BOR of CR or CRi within 3 months after CTL019 administration. The distribution function of EFS will be estimated using the KM method. The median EFS along with 95% confidence intervals will be presented if appropriate.

4.6.3.6 Overall survival (OS)

Overall survival is the time from date of first CTL019 infusion to the date of death due to any reason.

Patients not known to have died at the data cut-off date are censored at their last contact date, which is defined as the latest date they were known to be alive. Patients were followed for survival also in case of SCT. OS will be summarized censoring for SCT as sensitivity analysis.

OS will be assessed in all patients (IEAS, FAS and Enrolled Set) and in those who achieve BOR of CR or CRi within 3 months after CTL019 administration. The distribution function of OS will be estimated using the KM method. The median OS along with 95% confidence intervals will be presented if appropriate.

4.6.3.7 Efficacy in patients infused with CTL019 manufactured by [redacted]

The ORR and MRD negative remission rate will be summarized with 95% exact confidence intervals for patients infused with CTL019 manufactured by [redacted].

This analysis is not applicable at interim analysis because there is no patient infused with CTL019 from [redacted] at time of interim analysis.

4.7 Safety evaluation

4.7.1 Analysis set and reporting periods for the analyses

Table 4-2 summarizes the mutually exclusive safety reporting periods as well as the patients to be included in each of the segments. Note that the post-infusion period will be the main period of safety reporting (see Section 4.7.2 for details).

Table 4-2 Safety reporting periods

Period	Definition	Patients to be included
Pre-treatment period	From day of patient's informed consent to the day before first lymphodepleting chemotherapy dose or the pre-infusion visit if the lymphodepleting chemotherapy is not given	Screened patients
Lymphodepleting period (note: this period only applies to patients who received lymphodepleting chemotherapy)	From the first day of lymphodepleting chemotherapy • to the day before infusion of CTL019, for patients who received infusion, or • to the earlier of date of discontinuation and 30 days after last dose of lymphodepleting chemotherapy for patients who didn't receive infusion of CTL019	All patients who received lymphodepleting chemotherapy
Post-infusion period	Starting at day of first CTL019 infusion until end of study (60 months from CTL019 infusion)	Safety Set

4.7.2 Adverse events

The adverse events reporting follows a modified safety reporting rule as described in Protocol Appendix 3.

Reporting of AEs (except for CRS and graft versus host disease (GVHD)) will be based on MedDRA (latest version per database lock) and Common Terminology Criteria for Adverse Events (CTCAE) version 4.03. For the analysis purpose, latest version of Meddra at time of the analysis will be used.

The grading of CRS and GVHD will be based on protocol specific grading scales (Protocol section 6.2.4.2, Table 6-1 and Table 6-3, respectively).

Summary tables for AEs will be provided for AEs that started or worsened during the post-infusion period, i.e. the ***CTL019-treatment-emergent*** AEs. However, all safety data (including all observation periods as defined in Section 4.7.1) will be listed and with the period (as defined in Section 4.7.1) flagged for the starting date of the AE.

The incidence of CTL019-treatment-emergent AEs (new or worsening during the post-infusion period) will be summarized by system organ class, preferred term, severity (based on CTCAE grades), and relation to study drug. A patient with multiple CTC grades for an AE will be

summarized under the maximum CTC grade recorded for the event. The frequency of CTC grade 3 and 4 AEs will be summarized separately.

4.7.2.1 Adverse events of special interest (AESI)

Adverse events of special interest (AESI) include all important identified and potential risks of tisagenlecleucel. The list of AESIs and their search criteria will be 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.

AESI will be summarized by timing of onset: Within 8 weeks post CTL019 infusion, 8 weeks to 1 year post CTL019 infusion, >1 year post CTL019 infusion and any time post first tisagenlecleucel infusion.

The following summaries will be produced for the Safety Set:

- Adverse events of special interest based on identified risks, regardless of study drug relationship, by group term, preferred term and maximum grade
- Adverse events of special interest based on identified risks, suspected to be study drug related, by group term, preferred term and maximum grade
- Adverse events of special interest (AESI) based on potential risks, regardless of study drug relationship, by group term, preferred term and maximum grade
- Adverse events of special interest (AESI) based on potential risks, suspected to be study drug related, by group term, preferred term and maximum grade

4.7.2.2 Summaries of adverse events

Post-infusion period:

The following AE summaries will be produced for the Safety Set:

- Adverse events, regardless of study drug relationship, by primary system organ class, preferred term and maximum grade
- Adverse events, suspected to be study drug related, by primary system organ class, preferred term and maximum grade
- Deaths post infusion, by primary system organ class and preferred term
- Serious adverse events, regardless of study drug relationship, by primary system organ class and preferred term and maximum grade
- Serious adverse events, suspected to be study drug related, by primary system organ class and preferred term and maximum grade
- Adverse events leading to study discontinuation, regardless of study drug relationship, by primary system organ class and preferred term
- Non-Serious Adverse events, regardless of study drug relationship, by primary system organ class and preferred term

Adverse events will be summarized by timing of onset: Within 8 weeks post CTL019 infusion, 8 weeks to 1 year post CTL019 infusion, >1 year post CTL019 infusion and any time post first tisagenlecleucel infusion.

Lymphodepleting period:

In addition, AEs that started or worsened during the lymphodepleting period will be summarized for all patients in the Enrolled Set who received lymphodepleting chemotherapy. The following tables will be produced:

- Adverse events, regardless of study treatment relationship by primary system organ class and preferred term
- Serious adverse events, regardless of study treatment relationship by primary system organ class and preferred term
- Adverse events, with suspected study treatment relationship by primary system organ class and preferred term
- Serious adverse events, with suspected study treatment relationship by primary system organ class and preferred term

Pre-treatment period:

AEs that started or worsened during the pre-treatment period will be separately summarized for the Enrolled Set:

- Adverse events, by primary system organ class, preferred term and maximum grade
- Serious adverse events, by primary system organ class and preferred term

4.7.2.3 Safety in patients infused with CTL019 manufactured by [redacted]

Key safety summaries for adverse events regardless of relationship to study drug by System Organ Class (SOC) and PT, and adverse events of special interest will be performed on the Safety Set.

4.7.3 Laboratory abnormalities

For laboratory tests covered by the CTCAE, the study's biostatistics and reporting team will grade laboratory data accordingly. For laboratory tests covered by CTCAE, a Grade 0 will be assigned for all non-missing values not graded as 1 or higher. Grade 5 will not be used.

For laboratory tests where grades are not defined by CTCAE, results will be graded by the low/normal/high classifications based on laboratory normal ranges.

The following summaries will be generated separately for hematology and biochemistry laboratory tests for Safety Set:

- Shift tables using CTCAE grades to compare baseline to the worst post-infusion value.
    - for laboratory tests where CTCAE grades are not defined, shift tables using the low/normal/high/(low and high)
- Change from baseline to the worst post-infusion value, with descriptive statistics of baseline value, worst post-infusion value and the change.

The shift tables will be generated by timing: Within 8 weeks post CTL019 infusion, >8 weeks to 1 year post CTL019 infusion, >1 year post CTL019 infusion.

In addition, percentage of patient with Grade 3 or 4 hematopoietic cytopenias 28 days post CTL019 infusion will be summarized. Among patients with Grade 3 or 4 hematopoietic cytopenias 28 days post CTL019 infusion, the timing of resolution to Grade 2 or below will be summarized via Kaplan-Maier method. Similarly, among patients with Grade 1 to Grade 4 hematopoietic cytopenias 28 days post CTL019 infusion, the timing of resolution to Grade 0 will be summarized via Kaplan-Maier method. Grading of cytopenias will be derived using lab results in absolute lymphocytes (hypo), absolute neutrophils (hypo), hemoglobin (hypo), platelet count (hypo) or WBC (hypo) according to CTCAE 4.03. If a patient did not achieve resolution at the last lab assessment, timing of resolution will be censored at the last assessment. The median time to resolution and KM estimates of % resolved cases at different time point (month 2, month 3 and etc.) will be summarized.

The following listings will be provided for Enrolled Set.

- Listing of patients with laboratory abnormalities of CTC grade 3 or 4 with the corresponding CTC grades and the classifications relative to the laboratory reference ranges.
- Listing of all laboratory data with values flagged to show the corresponding CTC grades and the classifications relative to the laboratory reference ranges.

4.7.4 Immunogenicity

4.7.4.1 Humoral immunogenicity

Humoral immunogenicity assessment will include prevalence of immunogenicity (patients with pre-existing antibodies that bind to CTL019) and incidence of immunogenicity (patients with treatment-induced or treatment-boosted antibodies that bind to CTL019), together with antibody titers. Data will be further fractionated to determine proportion of patients who make transient versus sustained antibody responses. The assay for humoral immunogenicity will be a cell-based assay, detecting antibodies that bind to a Jurkat cell line transfected with the CTL019 construct. This cell line stably expresses the complete CTL019 sequence and can be used to detect antibodies that bind to any epitope on the extracellular domain of the protein.

The proportion of humoral immunogenicity positive and negative patients will be summarized by time points. Summary statistics will be presented for CTL019 cellular kinetic parameters for qPCR by anti-CTL019 antibody post-infusion status (positive or negative).

A scatter plot of baseline anti-CTL019 antibodies versus qPCR AUC0-28d and Cmax will be presented along with the appropriate regression line and equation. In addition boxplots of anti-CTL019 antibodies at enrollment by day 28 disease response will be presented. The same response categories will be used for a similar boxplot summarizing the maximum fold change of anti-CTL019 post-infusion.

4.7.4.2 Cellular immunogenicity

The cellular immunogenicity will be summarized by time points separated for CTL019 Pool 1 Peptides and CTL019 Pool 2 peptides. The boxplot of maximum fold change of cellular immunogenicity by month 3 disease response will be presented. The scatter plot of maximum

fold change versus qPCR AUC0-28d and Cmax will also be presented along with the appropriate regression line and equation.

4.7.5 Cytokine release syndrome and anti-cytokine therapies

To explore the relationship between CRS and other endpoints, the goal of this statistical analysis should be considered as the generation of new scientific hypotheses and observing new trends, since the studies are not adequately powered to propose a scoring system.

Clinical and biomarker data will be analyzed to potentially identify an early predictive score which reflects the risk of developing severe cytokine release syndrome. Only parameters that can be potentially utilized in clinical setting by treating physicians will be considered for the score development.

Detailed information regarding the CRS will be summarized by day 28 disease response from IRC assessment. Information summarized includes: maximum CRS grade, time to onset of CRS; duration of CRS; time to Grade 3/4 CRS, concurrent infections, timing and duration of ICU stay, selected complications, and use of anti-cytokine therapies, etc.

In addition, time to resolution of the first CRS will also be summarized using KM method for patients with CRS. In case the end date of a CRS is missing, it will be censored / imputed as the minimum of the following dates: the cut-off date, end of study evaluation (i.e., completion of the last phase of the study), date of death (if applicable).

Peak cytokine level, time to high fever onset, CTL019 PK parameters (e.g. Cmax and AUC0-d28), baseline tumor burden, CTL019 product characteristics (i.e. CD3+CD45+ [%], transduction efficiency [%], vector DNA sequence for CTL019 PCR [copies/cell]) and CTL019 dose administered will be plotted against the maximum CRS grade using strip plot as appropriate. The relationship between maximum CRS grade of the overall study vs CTL019 dose will also be explored using strip plots.

In addition, fibrinogen level together with the use of fibrinogen concentrate or cryoprecipitate will be summarized by CRS grade during CRS.

Individual patient time-profile for key inflammatory markers and cytokine parameters up to month 1 will be plotted, with annotation of tocilizumab and siltuximab usage.

4.7.6 Serious neurological adverse reactions

Serious neurological adverse reactions (SNARs) refer to a group of neurological adverse events defined in the AESI search criteria form. A SNAR episode may include multiple overlapped or consecutive SNARs as long as the end day and the start day of two consecutive events are no more than 3 days apart (i.e., current SNAR Start date – previous SNAR End date ≤ 3 days). The onset day of a SNAR episode is the start date of the first SNAR in the episode. The resolution date is the end day of the last SNAR in the episode. If there are multiple SNARs 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 SNAR episode will be summarized descriptively. Time to resolution of all SNAR episodes from all patients will be summarized using KM method by ignoring the fact that multiple episodes might be clustered by patient. That is, for one patient, if there are 2 episodes, both episodes are included in the KM analysis. Though the 2 episodes

for one patient are not completely independent, they are treated as if they are from two patients (each with 1 episode).

4.7.7 Febrile neutropenia

In addition the adverse event of febrile neutropenia reported by investigators, the number and percentage of patients with neutropenia with concurrent high fever, defined as grade 3 or above neutropenia per lab exam with temperature >38.3 °C within +/- 1 day, will be summarized.

4.7.8 Growth data

For patients under 18 years of age at the time of CTL019 infusion, height and weight will be summarized at semi-annual intervals before and after starting CTL019, using the standard deviation score (SDS), velocity and velocity SDS. The relevant height and weight values for each semi-annual period are defined using time windows, as defined in Section 5.4.

SDS is calculated using the formulae (provided by Centers for Disease Control and Prevention (CDC)):

$$\text{SDS} = \frac{\left(\frac{X}{M}\right)^{L} - 1}{LS} \text{ if } L \neq 0, \qquad \text{or} \qquad \text{SDS} = \frac{log\left(\frac{X}{M}\right)}{S} \text{ if } L = 0,$$

where X is height in centimeters or weight in kilograms, and L, M and S are height-, weight-, sex- and age-specific reference values from the CDC Growth Charts (http://www.cdc.gov/growthcharts/percentile_data_files.htm). The files for height and weight are named STATAGE and WTAGE for children older than 2 years (see Appendix). Age is listed at the half month point for the entire month; for example, 1.5 months represents 1.0 month up to but not including 2.0 months of age. SDS is actually a Z score that measures the distance from the population mean in units of standard deviations. That is, SDS < 0 refers to values lower than the population mean, and for example SDS ≤ -1.645 refers to values in the lowest 5%. (The usual percentile more commonly used in the clinical practice can be derived from the Z-score by a normal distribution).

Height velocity is defined as follows:

Height velocity (cm/6-months) = (height in time window *k* – height in time window *k-1*)

÷ ([assessment date in time window *k* – assessment date in time window *k-1*] ÷ [365.25/2]),

and similarly for weight velocity.

Velocity SDS is calculated as (velocity – mean) / SD, where mean and SD are obtained as the height-, weight-, sex- and age-specific values in Tables 5 to 8 in Baumgartner (1986), where the age category immediately above the patient's exact age (at the assessment date in time window *k*) should be used. Velocity SDS will only be calculated for time window *k* if data also exists for time window *k-1*, since calculating across multiple units of 6 months requires more than one reference value to be taken into account.

Height/weight SDS and velocity SDS will be summarized using descriptive statistics (mean, standard deviation, range) for each time window, as well as by presenting number of patients with SDS values lower/higher than 5^{th}/95^{th} percentiles respectively. Box plots will also be plotted for each time window. All height/weight SDS, velocity and velocity SDS data will be

listed, and values of SDS and velocity SDS outside of the central 95% of population values will be flagged as either High (SDS ≥ 1.645) or Low (SDS ≤ -1.645).

Depending on the actual enrolled population (e.g. country, race, etc.), adjustment of the method may be made if appropriate.

4.7.9 Puberty Stage

Puberty stage will only be analyzed among pre-pubescent patients, i.e., using patients from the Safety Set who were classified as Tanner Stage 4 or lower at the latest assessment prior to the infusion of CTL019.

Tanner Stage includes two components for boys, namely testis and pubic hair, and two components for girls: breast development and pubic hair. It is expected that data will become available during the trial on a proportion of patients as they go through puberty attaining higher levels of the Tanner Stage. For the age at which Tanner Stages 2-5 are achieved, age at thelarche (females), age at menarche (females) and age at adrenarche (males), summary statistics from Kaplan-Meier distributions will be determined, including the median age and the proportions of patients reaching these milestones at some given ages. The statistics will be given as point estimates with 95% confidence intervals.

Delayed puberty in girls is defined as failure to attain Tanner Stage 2 (for both breast development and pubic hair) by age 13, or absence of menarche by age 15 or within 5 years of attainment of Tanner Stage 2 (Fenichel et al. 2012). Delayed puberty in boys is defined as failure to attain Tanner Stage 2 (for both testis and pubic hair) by age 14 (Crowley et al. 2012). Rates of delayed puberty will be presented for boys and girls separately, along with 95% confidence intervals, among the patients who did not have delayed puberty at baseline.

4.7.10 Other safety data

Vital signs will be collected as clinically needed. Findings supportive of GVHD will be listed for patients who have received prior allogeneic SCT.

Karnofsky/Lansky performance scores will be listed by subject.

4.8 Pharmacokinetic analysis

4.8.1 CTL019 cellular kinetics

PAS will be used for all CTL019 cellular kinetics summaries (tables and figures). FAS will be used for CTL019 cellular kinetics data listings.

CTL019 concentrations in peripheral blood and bone marrow (and CSF if available) will be listed, graphed, and summarized by time point as assessed by the following:

- CTL019 transgene levels as measured by q-PCR
- CTL019 transduced cells measured by flow cytometry of CD3-positive, CD3-positive/ CD4-positive and CD3-positive/CD8-positive CTL019 transduced cells.

The PK parameters listed in Table 4-3 will be estimated from the individual concentration versus time profiles using a non-compartmental approach within Phoenix® (Pharsight,

Mountain View, CA). The non-quantifiable concentrations will be imputed to zero for PK concentration summaries, and will not be included for estimation of PK parameters. Results reported but deemed unreliable will be flagged and excluded from the summaries and PK parameter derivations.

Table 4-3 Noncompartmental pharmacokinetic parameters

Parameter	Definition
AUC 0 - Tmax	The AUC from time zero to Tmax in peripheral blood (% or copies/µg x days)
AUC Tmax - 28d and 84d	The AUC from time Tmax to day 28 and 84 or other disease assessment days, in peripheral blood (% or copies/µg x days)
AUC 0 - 28d and 84d	The AUC from time zero to day 28 and 84 or other disease assessment days, in peripheral blood (% or copies/µg x days)
Cmax	The maximum (peak) observed in peripheral blood drug concentration after single dose administration (% or copies/µg)
Tmax	The time to reach maximum (peak) peripheral blood drug concentration after single dose administration (days)
T1/2	The half-life associated with the disposition phase slopes (alpha, beta, gamma etc.) of a semi logarithmic concentration-time curve (days) in peripheral blood
Clast	The last observed quantifiable concentration in peripheral blood (% or copies/ug)
Tlast	The time of last observed quantifiable concentration in peripheral blood (days)

Descriptive statistics of CTL019 cellular kinetics parameters (mean, standard deviation, coefficient of variation, geometric mean, CV% geometric mean, median, min and max) will be summarized by day 28 disease response from IRC assessment. When a geometric mean will be presented, it will be stated as such. A range of values will be presented for selected variables. For Tmax median values and ranges only will be given.

The relationship between anti-cytokine treatment, use of steroids, occurrence of immunogenicity or other relevant covariates and CTL019 cellular kinetics might be explored. Population and/or mechanistic PK / PD models may also be generated.

For patients who were treated with tocilizumab during CRS, the tocilizumab concentrations will be summarized by time points (depending upon sample availability) relative to time of first tocilizumab dose.

CTL019 cellular kinetics parameters will be summarized by tocilizumab usage to investigate the effect of tocilizumab on CTL019 cellular kinetics.

4.8.2 CTL019 cellular kinetics in patients infused with CTL019 manufactured by [redacted]

The CTL019 PK parameters for CTL019 transgene levels as measured by q-PCR will also be summarized. The CTL019 PK parameters as measured by flow cytometry (exploratory only) will also be summarized, as appropriate.

4.8.3 Tocilizumab and siltuximab PK

Tocilizumab and siltuximab concentrations will be presented by time points relative to the time of tocilizumab and siltuximab dosing by dose reference ID, showing n, m (number of non-zero values), mean, standard deviation (SD), CV%, median, minimum, maximum, geometric mean, and geometric CV%. The plot of concentration-time profile will also be provided.

The descriptive statistics (n, mean, CV%, standard deviation (SD), median, geometric mean, geometric CV%, minimum and maximum) will be presented by dose reference ID for relevant PK parameters.

All individual PK parameters and PK concentration data will be listed.

4.9 Biomarkers analyses

As a project standard, Novartis Oncology BDM will analyze only biomarkers collected in the clinical database. For exploratory markers, since the studies are not adequately powered to assess specific biomarker–related hypotheses, the goal of these exploratory statistical analyses should be considered as the generation of new scientific hypotheses. These hypotheses may be compared with results found in literature as well as verified with data derived from subsequent clinical trials. No adjustment for multiple comparisons is usually planned for exploratory analyses. Furthermore, additional post hoc exploratory assessments are expected and may be performed.

There may be circumstances when a decision is made to stop sample collection, or not perform or discontinue the analysis of blood / archival tumor samples / fresh tumor biopsies / fine needle aspirates due to either practical or strategic reasons (e.g. issues related to the quality and/or quantity of the samples or issues related to the assay). Under such circumstances, the number of samples may be inadequate to perform a rigorous data analysis and the available data will only be listed and potentially summarized.

The analyses to be performed for the CSR are outlined below. Additional analyses that may be performed after the completion of the end-of-study CSR will be documented in separate reports. These analyses may include but are not limited to the meta-analysis of data from this study combined with data from other studies or the analysis of biomarkers generated from samples collected during the study but analyzed after the database lock and completion of the CSR. The data analysis will be described in a stand-alone analysis plan document, as appropriate.

4.9.1 Biomarker Data Analysis Set

The FAS will be used for all biomarker analysis. Assessment of associations between biomarker and safety data will be conducted using the Safety Set.

4.9.1.1 Data Handling of Serum Cytokine Data

Serum cytokine data represent quantitative soluble protein measurements that tend to follow a log normal distribution. Thus, a log10 transformation of the data is typically required for normalization prior to performing any statistical modeling. Values below the lower limit of quantitation (which may be reported with the label “LLOQ” or have a numerical value below the assay’s lower limit of quantification) will be imputed / replaced as 0.5×LLOQ, which will

be specified by the performing lab and is assay and analyte specific. In some cases a value, although below LLOQ, is reported, this value should not be used and the data should be imputed as 0.5×LLOQ.

For values above the upper limit of quantification (either reported as "ULOQ" or a numerical value greater than the assays upper limit of quantification), the values will be set to the ULOQ threshold of the assay.

4.9.2 Basic Tables, Listings and Figures

4.9.2.1 B-cell and T-cell level

The levels (%) of CD19+ total B cells amongst viable WBC in peripheral blood will be summarized by day 28 response of IRC assessment and time point (see Section 5.4). The levels (%) of T cells amongst mono-nuclear cells (lymphocytes and monocytes with the exclusion of granulocytes) in peripheral blood and bone marrow will be described.

CD19 phenotype determined by bone marrow flow cytometry assessment (CD19 positive, CD19 dim, CD19 negative, CD19 positive/negative, Unknown) and CD19+ intensity level among B-ALL cells in bone marrow at baseline and time of bone marrow or blood relapse will be summarized.

It is anticipated that all patients who achieve complete remission will exhibit B-cell aplasia. Time to B-cell recovery will be summarized. Here B cell recovery is defined as the time from onset of remission date to the earliest time when the percentage of CD19+ total B cell among viable WBC in blood is at least 1%, or the percentage of CD19+ total B cell among lymphocyte in blood is at least 3%. If no B cell recovery is observed, time to B cell recovery is censored at the last B cell result. Note that if CD19+ ALL tumor cells are also present in the blood (recurrence), total B cells are affected by the malignant B cells and hence should be interpreted with caution.

Data may also be summarized by response status and potentially graphed using strip plots. Patient level and average longitudinal plots of the cell counts and percent changes from baseline may be generated.

For abnormal T cell or B cell results, associated safety events such as infections and use of associated therapies (i.e. antibiotics, immunoglobulin replacement) will be investigated using patient listings.

4.9.2.2 Soluble immune factors

Soluble immune and inflammatory cytokines (e.g. IL-10, interferon gamma, IL-6, CRP, and Ferritin) will be listed and summarized by patient and time point. If both the baseline and post baseline values are below LLOQ, absolute, percent and fold change from baseline will not be imputed and reported as missing. Summaries of baseline and change from baseline (absolute change, percent change and fold change) at each time point will be summarized in tables that include sample size, mean, standard deviation, %CV, median, minimum and maximum. Optionally the number and percent of missing values or the values below LLOQ for each time point will be reported.

CRP and ferritin results assessed by local lab will be used for summary.

Baseline levels may also be summarized by clinical response status and severity of CRS and potentially graphed using strip plots. In addition, the maximum change from baseline measure for each cytokine may also be graphed against clinical response status and severity of CRS response using strip plots. Patient level and averaged cytokine measures and change from baseline may be displayed using longitudinal plots.

[redacted]

4.11 Patient reported outcome (PRO) and healthcare resource utilization

4.11.1.1 Patient reported outcome

Patient Reported Outcomes (PRO) will be assessed using PedsQL and EQ-5D. PedsQL™ and EQ-5D will be completed by patients aged 8 and above. Descriptive statistics (e.g. mean, median, and frequency) and change from baseline of the summary scores for each post baseline time-point/window of assessment will be provided based on all available data at the time of analysis. IEAS or FAS will be used for all analysis at interim and final analysis respectively.

No imputation will be applied if the total or subscale scores are missing at a visit.

Separate summaries will be provided for month 3 and 6 EQ-5D and PedsQL results for those patients who achieve best overall response CR/CRi.

Subgroup analysis by age group may be performed as ad hoc analysis if there is sufficient number of patients within each age subgroup.

EQ-5D

The EQ-5D health state is represented by the answer to the 5 dimensions of the questionnaire (mobility, self-care, usual activity, pain/discomfort, anxiety/depression in this order). Each question has 3 levels answers:

- level 1: no problems
- level 2: some problems
- level 3: a lot of problems

For each question, the level at baseline, as well as the change of level in each post baseline time-point/window of assessment will be summarized.

In addition, there is a general question about the overall health (EQ-VAS) with range 0-100 (the larger number indicates better health). The EQ-VAS values as well as the change from baseline will be summarized for each post baseline time-point/window of assessment.

PedsQL

The PedsQL questionnaire composes of 4 subscales: emotional, social, school and physical functioning scales. Each subscale contains 5-8 questions each with 5 choices indicating the frequencies. The items in each question will first be scored as the following: “Never”=100, “Almost Never”=75, “Sometimes”=50, “Often”=25, and “Almost Always”=0.

The mean scores in following categories will then be calculated:

- each of the 4 subscales (i.e. emotional, social, school and physical),
- the psychosocial health summary score (combining emotional, social and school functioning scales),
- the total score (combining all 4 subscales).

Descriptive statistics will be used to summarize the raw and change from baseline of the above summary scores for each post baseline time-point/window of assessment.

If more than 50% of the items are missing in a subscale for PedsQL, the score for this subscale will be considered missing for this assessment. Otherwise, the average of the non-missing items in the subscale will be used to impute for the missing items when calculate the score for the subscale.

4.11.1.2 Healthcare resource utilization

Data relating to resource utilization (described in Section 7.2.5 of study protocol) will be used to support health economic evaluations.

Number of CTL019 inpatients and outpatients infusions will be summarized. Descriptive statistics of hospitalizations, including the total and average number and duration of hospitalizations, timing and duration of ICU stay, etc., will be provided.

4.12 Subgroup analyses

4.12.1 Efficacy subgroup analyses

Subgroup analyses for ORR, MRD and DOR will be performed on the following based on the patient's baseline status:

- Age: <10 years, ≥10 years to <18 years, ≥18 years
- Gender: Male, Female
- Race: White, Asian, Other
- Ethnicity: Hispanic or Latino, Other
- Response status at study entry: Primary refractory, Relapsed disease
- Prior SCT therapy: Yes, No
- Eligibility for SCT: Eligible for SCT, ineligible for SCT
- Baseline bone marrow tumor burden: Low (defined as either morphologic or MRD result is <50% and neither is ≥50%), High (defined as either morphologic or MRD result is ≥50%)
- Baseline extramedullary disease presence: Yes, No
- Philadelphia chromosome/BCR-ABL: Positive, Negative
- Mixed-Lineage Leukemia (MLL) rearrangement: Yes, No
- Hypodiploidy: Yes, No
- BCR-ABL1-like: Yes, No
- Complex Karyotypes (≥5 unrelated abnormalities): Yes, No
- Down syndrome: Yes, No
- Time since enrollment to CTL019 infusion: ≤Median, >Median
- Number of previous relapses: 0, 1, 2, ≥3

The rationale for performing subgroup analyses are as follows:

- Age, gender, race and ethnicity are demographic factors that are typically requested by health authorities to assess internal consistency of the study results and also have been shown to impact ALL outcome in first line and first relapse settings.
- Prior response status is a key prognosis factor due to potentially different rates of treatment related morbidity in patients who have relapsed following allogeneic SCT vs those who have not undergone SCT.
- Baseline bone marrow tumor burden and extramedullary disease presence are important indicators of overall disease burden, which is a potential predictive factor.
- BCR-ABL, MLL rearrangement, Hypodiploidy, BCR-ABL1-like gene signatures and complex karyotype (≥5 unrelated abnormalities) are high risk factors for ALL outcome in the first line and first relapse settings. Patients with these high risk factors have poorer

diagnosis (Harrison et al 2010; van der Veer et al 2013; NCCN v6 2013). In case there are very few patients with these high risk features individually, analysis may be performed for patients with any of these high risk features versus those who do not.

- Patients with Down syndrome are known to have increased ALL treatment related morbidity and mortality rates. Because of increased risk, stem cell transplant is often not recommended in this population. Therefore, the experience with CTL019 in this rare population may address an unmet medical need.

Subgroup analyses will only be performed if at least 5 patients are present in each subgroup. Some grouping of classes will be considered if there are too few patients in some subgroups.

Efficacy analyses in subgroups will generally be purely exploratory and are intended to explore the intrinsic consistency of any treatment effects found overall.

Subgroup analyses of the primary endpoint (ORR) will be performed on the FAS by presenting the point estimates in the subgroup with the exact 95% CIs. Summary tables and forest plots will be presented.

4.12.2 Safety subgroup analyses

Key safety summaries for adverse events regardless of relationship to study drug by SOC and PT, and adverse events of special interest will be repeated on the Safety Set in the following subgroups:

- Age: <10 years, ≥10 years to <18 years, ≥18 years
- Gender: Male, Female
- Race: White, Asian, Other
- Ethnicity: Hispanic or Latino, Other
- Response status at study entry: Primary refractory, Relapsed disease
- Prior SCT therapy: Yes, No
- Down syndrome: Yes, No

Summary tables will only be performed if at least 5 patients are present in each subgroup. Some grouping of classes will be considered.

4.13 Determination of sample size

In a previous study of clofarabine in patients with r/r B-cell ALL who have had 2 or more prior regimens, the reported ORR was 20% (95% CI [10%, 34%]; Jeha et al. 2006). Hence, an ORR of 45% that excludes a 20% ORR at the 0.025 significance level would indicate meaningful efficacy in this highly refractory population.

The final analysis of the primary endpoint will be performed after all patients infused with CTL019 have completed 3 months follow-up from study day 1 infusion or discontinued earlier. The sample size for the final analysis of the primary endpoint will be up to 76 patients.

Based on the null hypothesis of ORR ≤ 20% and alternative hypothesis of ORR >20%, 76 patients in the FAS will provide more than 95% power to demonstrate statistical significance at one-sided cumulative 0.025 level of significance, if the underlying ORR is 45% and taking into

account the interim analysis as described in Section 4.14. In this setting, an ORR of 23/76=30% will be needed to claim success.

Within the expected sample size of 76 patients with CTL019, at least 10 patients will be treated with CTL019 manufactured by the [redacted]. If there are at least 6 patients among them who achieved best overall response of CR or CRi, the lower bound of the 95% confidence interval will be higher than 20%. The probability of observing at least 6 CR or CRi among the 10 patients will be 26% if the true ORR is 45%, and will be 84% if the true ORR is 70%.

Table 4-4 Confidence intervals for ORR in patients infused with CTL019 manufactured by the [redacted]

Total number of patients	CR + CRi	95% Exact CI
10	5	(18.7%, 81.3%)
	6	(26.2%, 87.8%)
	7	(34.8%, 93.3%)
	8	(44.4%, 97.5%)
	9	(55.5%, 99.7%)
	10	(69.2%, 100%)

The actual number of patients to be enrolled will depend on the pre-infusion dropout rate. Limited data are available so far to provide robust estimate on the pre-infusion dropout rate. Assuming 20% to 25% enrolled patients will not be infused due to reasons such as manufactory failure, worsening of patient's condition, etc., approximately 95 patients are estimated to be enrolled to reach the number of patients required.

4.13.1 Power for analysis of key secondary variables

4.13.2 ORR within 3 months in patients infused with CTL019 from US manufacturing facility

The same efficacy is assumed for patients infused with CTL019 in US manufacturing facility vs other manufacturing facilities. Under this assumption and conditional on the statistical significance of the primary endpoint, the overall power of this endpoint will be greater than 95%, taking in account an interim analysis will be performed with first 50 patients, and then a final analysis will be performed with up to 66 patients infused with CTL019 from US manufacturing facility.

4.13.3 Remission with MRD negative bone marrow in patients who received CTL019 from all manufacturing facilities

In previous studies in the r/r ALL setting, 67% to 82% patients achieved MRD negative status among patients who achieved CR or CRi (Topp et al 2015, O'Brien et al 2012). Considering that an ORR of 45% that excludes 20% at the 0.025 significance level would indicate meaningful efficacy for ORR, 34% of patients achieving MRD negative bone marrow that excludes 15% at the 0.025 significance level would indicate meaningful efficacy (i.e. 75% among complete responders) for the key secondary objective.

Based on the above assumptions, conditional on the statistical significance of the primary endpoint and the first key secondary endpoint, and taking into account the interim analysis with

first 50 patients as described above, 76 patients in the FAS will provide greater than 95% power to demonstrate statistical significance for the key secondary endpoint at one-sided 0.025 level of significance, if the underlying percentage of patients who achieve BOR or CR or CRi with MRD negative bone marrow is 34%.

4.13.4 Remission with MRD negative bone marrow in patients who received CTL019 from US manufacturing facility

The same efficacy is assumed for patients infused with CTL019 in US manufacturing facility vs other manufacturing facilities. Under this assumption and conditional on the statistical significance of the primary and first 2 key secondary endpoints, the power of this endpoint will be 94%, taking into account an interim analysis will be performed with first 50 patients, and then a final analysis will be performed with up to 66 patients with CTL019 from US manufacturing facility.

4.14 Interim analyses

4.14.1 Interim analysis for the primary endpoint

An interim analysis is planned when the first 50 patients infused have completed 3 months from study day 1 infusion or discontinued earlier. The interim analysis will be performed by testing the null hypothesis of ORR within 3 months being less than or equal to 20% against the alternative hypothesis of ORR within 3 months being greater than 20% at overall one-sided 2.5% level of significance.

The study will not be stopped for outstanding efficacy at the interim analysis regardless of the interim result.

An α-spending function according to Lan-DeMets (O'Brien-Fleming), as implemented in East 6.3, will be used to construct the efficacy stopping boundaries (Lan and DeMets 1983). Based on the choice of α-spending functions described above, if the interim analysis is performed exactly with 50 patients and final analysis will include up to 76 patients (i.e. 50/76=65.8% information fraction), the lower bound of the 2-sided 98.9% exact CI of the ORR will need to be greater than 20% to declare statistical significance. As a result, an ORR of 19/50 = 38% is needed to claim success at interim. If the interim efficacy boundary is not crossed, 2-sided 95.4% exact CI will be used at final analysis correspondingly. As a result, an ORR of 23/76 = 30% will be needed to claim success at final analysis.

The efficacy boundary at the final analysis will be based on the actual number of patients and the alpha already spent at the interim analysis. If the number of patients in the final analysis deviates from the expected number of patients, the final analysis criteria will be determined so that the overall significance level across all analyses is maintained at one-sided 0.025.

4.14.2 Interim analysis for the key secondary endpoints

If the primary endpoint is met at the interim analysis, the key secondary endpoints will also be assessed following hierarchical sequence using an α-spending function according to Lan-DeMets (O'Brien-Fleming).

4.14.2.1 ORR within 3 months in all patients infused with CTL019 from US manufacturing facility

Based on the choice of α-spending functions described above, if the interim analysis is performed exactly with 50 patients and final analysis will include up to 66 patients (i.e. 50/66=75.8% information fraction), the lower bound of the 2-sided 98.0% exact CI of the ORR will need to be greater than 20% to declare statistical significance. As a result, an ORR of 18/50 = 36% is needed to claim success at interim. If the interim efficacy boundary is not crossed, 2-sided 95.6% exact CI will be used at final analysis correspondingly. As a result, an ORR of 21/66 = 32% will be needed to claim success at final analysis.

4.14.2.2 Remission with MRD negative bone marrow in patients infused with CTL019 from all manufacturing facilities

Based on the choice of α-spending functions described above, if the interim analysis is performed exactly with 50 patients and final analysis will include up to 76 patients (i.e. 50/76=65.8% information fraction), the lower bound of the 2-sided 98.9% exact CI will need to be greater than 15% to declare statistical significance. As a result, a MRD negative rate of 15/50 = 30% is needed to claim success at interim. If the interim efficacy boundary is not crossed, 2-sided 95.4% exact CI will be used at final analysis correspondingly. As a result, an ORR of 19/76 = 25% will be needed to claim success at final analysis.

4.14.2.3 Remission with MRD negative bone marrow in patients infused with CTL019 from US manufacturing facility

Based on the choice of α-spending functions described above, if the interim analysis is performed exactly with 50 patients and final analysis will include up to 66 patients (i.e. 50/66=75.8% information fraction), the lower bound of the 2-sided 98.0% exact CI of the ORR will need to be greater than 20% to declare statistical significance. As a result, an ORR of 15/50 = 30% is needed to claim success at interim. If the interim efficacy boundary is not crossed, 2-sided 95.6% exact CI will be used at final analysis correspondingly. As a result, an ORR of 17/66 = 26% will be needed to claim success at final analysis.

5 Additional analysis definitions and conventions

5.1 Response rate analyses

For the analyses of response rate (e.g, ORR), the rates will be summarized along with a 2-sided 95% exact Clopper-Pearson confidence interval. Sample code is provided below.

```
PROC FREQ data=dataset;

EXACT BINOMIAL;

TABLE outcome/binomial(p=0.2) ALPHA=0.xxx;

RUN;

/* outcome is the variable to indicate response or not, note that if the outcome is
dichotomous variable, then the proportion of outcome=0 will be calculated.*/
```

5.2 Time-to-event analyses

For time-to-event analyses (DOR, RFS, EFS and OS), the survival function will be estimated using the Kaplan-Meier (product-limit) method as implemented in PROC LIFETEST (see examples below). Median survival will be obtained along with 95% confidence intervals calculated from PROC LIFETEST output using the loglog option available within PROC LIFETEST, Kaplan-Meier estimates with 95% confidence intervals at specific time points will be summarized.

```
PROC LIFETEST data=dataset METHOD=KM conftype=loglog;

TIME survtime*censor(1);

RUN;

/* survtime represents variable containing event/censor times;

censor represents censoring variable (1=censored, 0=event); */
```

The time points can be expressed in weeks or in months depending on the time-to-event variable (e.g. overall survival might require a different scale than duration of response). If 'months' is used it should be noted that 1 month is defined as (365.25/12) =30.4375 days, which is not equal to 4 weeks.

In completing risk analysis, the cumulative incidence function (CIF) can be estimated following macro:

```
%CIF(data=dataset, out=est, time=survtime, status=status, event=1);

/* survtime represents variable containing event/censor times;

status represents status variable (0=censored, 1= event of interest, 2= competing
events); */
```

5.3 Duration of follow-up

The follow up duration (in months) for time to event endpoints (EFS and OS) is calculated as (Date of event or censoring – Date of first CTL019 infusion + 1)/30.4375.

The follow up duration (in months) for time to event endpoints (DOR, RFS and time to B cell recovery) is calculated as (Date of event or censoring – Date of onset of remission + 1)/30.4375.

Primary follow up duration (in months) will be calculated as (min(Analysis cut-off date, treatment and primary follow-up phase completion or discontinuation date) – Date of first CTL019 infusion + 1)/30.4375).

The total study follow up duration (in months) will be calculated as (min(Analysis cut-off date, Study follow-up completion or discontinuation date) – Date of first CTL019 infusion + 1)/30.4375). Here the study follow-up completion or discontinuation date will be refer to the completion/discontinuation date of the last phase (i.e. treatment and primary follow-up or secondary follow-up) the patient has entered.

5.4 Time windows

In order to summarize the patient reported outcome (PRO), growth data, PK and biomarker 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.

For PK, Biomarker and growth data, 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.

Table 5-1 shows the defined time windows for biomarker sample.

Table 5-1 Time windows for biomarker

Time Window	Planned visit timing (study day)	Time Window Definition (Study days)
Peripheral blood for serum cytokine analyses		
W-16 to D-1 Enrollment/Pre-Chemotherapy*	Before Study Day -1	< first day of Lymphodepleting (LD) chemotherapy
D -1 Pre-infusion**	-1	Day of LD chemo to day 1 pre infusion
D7±1d	7	Day 1 post infusion to 10
D14±3d	14	11 to 17
D21±3d	21	18 to 24
D28±4d	28	25 to 59
M3±14d	91	60 to 136
M6±14d	183	137 to 273
M12±14d	365	≥274
CTL019 Immunophenotyping; B cell; T cell (peripheral blood)		
W-16 to D-1 Enrollment/Pre-Chemotherapy *	Before Study Day -1	< first day of Lymphodepleting (LD) chemotherapy
D -1 Pre-infusion**	-1	Day of LD chemo to day 1 pre infusion
D7±1d	7	Day 1 post infusion to 10
D14±3d	14	11 to 17
D21±3d	21	18 to 24
D28±4d	28	25 to 59
M3±14d (Primary follow-up only)	91	60 to 136
M6±14d (Primary follow-up only)	183	137 to 228
M9±14d (Primary follow-up only)	274	229 to 319
M12±14d (Primary follow-up only)	365	320 to 574
M24±14d (Primary follow-up only)	731	575 to 913
M36±14d (Primary follow-up only)	1096	≥ 914

[redacted]		
W-16 to D-1 Enrollment/Pre-Chemotherapy	Before Study Day -1	< first day of Lymphodepleting (LD) chemotherapy
D28±4d	28	21 to 59
M3±14d	91	60 to 136
M6±14d	183	137 to 273
M12±14d	365	≥274
[redacted]		
CTL019 Immunophenotyping; B cell; T cell (bone marrow aspirate)		
W-16 to W-12 Screening*	Before Study Week -12	< first day of Lymphodepleting (LD) chemotherapy
D28±4d	28	21 to 59
M3±14d (recommended but not required)	91	60 to 136
M6±14d (recommended but not required)	183	≥137

Study Day 1 = start date of CTL019
* for patients who didn't receive LD chemotherapy, this window is ≤-2
**for patients who didn't receive LD chemotherapy, this window is -1 to 1 pre-infusion

As it is critical to understand the change of cytokine level during the first month of study drug and to capture the likely unscheduled assessments, a time window (Table 5-2) more frequent than protocol scheduled assessment is defined for this purpose.

Table 5-2 Time windows for serum cytokine in peripheral blood analyses within 28 days

Time Window	Time Window Definition (Study days)
W-16 to D-1 Enrollment/Pre-Chemotherapy *	< first day of Lymphodepleting (LD) chemotherapy
Pre-infusion**	Day of LD chemo to day 1 pre infusion
D4	Day 1 post infusion*** to Day 5
D7	6 to 9
D11	10 to 12
D14	13 to 15
D17	16 to 19
D21	20 to 24
D28	25 to 35

* for patients who didn't receive LD chemotherapy, this window is ≤-2
** for patients who didn't receive LD chemotherapy, this window is -1 to 1 pre infusion
*** all samples on day 1 are scheduled to be taken pre-infusion. Samples will be considered as post infusion only if time of collection is after CTL019 infusion.

Table 5-3 shows the defined time windows for CTL019 PK sample.

Table 5-3 Time windows for CTL019 PK

Time Window	Planned visit timing (Study day)	Time Window Definition (Study day)
CTL019 pharmacokinetics by q-PCR in peripheral blood		

W-16 to D-1 Enrollment/Pre-Chemotherapy	Before Study Day -1	≤ day 1 pre-infusion
D1 10 min ± 5 min post-infusion	1	Day 1 post-infusion to 2
D4±1d	4	3 to 5
D7±1d	7	6 to 9
D11±1d	11	10 to 12
D14±3d	14	13 to 17
D21±3d	21	18 to 24
D28±4d	28	25 to 59
M3±14d	91	60 to 136
M6±14d	183	137 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
M30±14d	913	914 to 1004
M36±14d	1096	1005 to 1187
M42±14d	1278	1188 to 1369
M48±14d	1461	1370 to 1552
M54±14d	1644	1553 to 1734
M60±14d	1826	≥ 1735
CTL019 pharmacokinetics by flow cytometry in peripheral blood		
W-16 to D-1 Enrollment/Pre-Chemotherapy	Before Study Day -1	≤ day 1 pre-infusion
D4±1d	4	Day 1 post-infusion to 5
D7±1d	7	6 to 9
D11±1d	11	10 to 12
D14±3d	14	13 to 17
D21±3d	21	18 to 24
D28±4d	28	25 to 59
M3±14d	91	60 to 136
M6±14d	183	137 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
M30±14d	913	914 to 1004
M36±14d	1096	1005 to 1187
M42±14d	1278	1188 to 1369
M48±14d	1461	1370 to 1552
M54±14d	1644	1553 to 1734
M60±14d	1826	≥ 1735
CTL019 pharmacokinetics by q-PCR in bone marrow aspirate **CTL019 pharmacokinetics by flow cytometry in bone marrow aspirate**		
W-16 to W-12 Screening	Before Study Week -12	≤-1
D28±4d	28	1 to 59

M3±14d (recommended but not required)	91	60 to 136
M6±14d (recommended but not required)	183	≥137
CTL019 pharmacokinetics by q-PCR in CSF		
W-16 to W-12 Screening	Before Study Week-12	≤-1
D28±4d	28	≥ 1

Table 5-4 shows the defined time windows for tocilizumab PK sample.

Table 5-4 Time windows for tocilizumab PK

Time Window	Time Window Definition
First tocilizumab dose:	
D1 (5-15 minutes post infusion)	First toci admin to <30 minutes post first toci
D1 1 hour ± 15 min post infusion	30 minutes post first toci to <12 hours post first toci
D2 ± 2 hours	12 hours post first toci to <36 hours post first toci
D3 ± 4 hours	36 hours post first toci to <96 hours post first toci
D7 ± 1d	96 hours post first toci to <192 hours post first toci
Second tocilizumab dose:	
D1 (pre-dose; second infusion)	24 hours prior to second toci admin to <second toci admin
D1(5-15 minutes post second infusion)	Second toci admin to <12 hours post second toci
D2 ± 2 hours from second infusion	12 hours post second toci to <36 hours post second toci
* Concentration on or after second tocilizumab administration will not be summarized for first tocilizumab PK.	

Table 5-5 shows the defined time windows for CTL019 PK sample.

Table 5-5 Time windows for immunogenicity

Time Window	Planned visit timing (Study day)	Time Window Definition (Study day)
W-16 to D-1 Enrollment/Pre-Chemotherapy	Before Study Day -1	≤-1
D14±3d	14	1 to 21
D28±4d	28	22 to 59
M3±14d	91	60 to 136
M6±14d	183	137 to 273
M12±14d	365	274 to 574
M24±14d	731	575 to 913
M36±14d	1096	≥ 914
Study Day 1 = start date of CTL019		

Table 5-6 shows the defined time windows for growth data and Tanner staging.

Table 5-6 Time windows for growth data and Tanner staging

Time Window	Planned visit timing (Study day)	Time Window Definition (Study day)
Height and Tanner Stage		
Baseline	Before Study Week -12	≤1
M6±14d	183	>1 to 273

M12±14d	365	274 to 456
M18±14d	548	457 to 639
M24±14d	731	640 to 913
M30±14d	913	914 to 1004
M36±14d	1096	1005 to 1187
M42±14d	1278	1188 to 1369
M48±14d	1461	1370 to 1552
M54±14d	1644	1553 to 1734
M60±14d	1826	≥ 1735
Weight		
Baseline	Before Study Week -12	≤-3
D-1±1d	-1	-2 to -1
D28±4d	28	21 to 59
M3±14d	91	60 to 136
M6±14d	183	137 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
M30±14d	913	914 to 1004
M36±14d	1096	1005 to 1187
M42±14d	1278	1188 to 1369
M48±14d	1461	1370 to 1552
M54±14d	1644	1553 to 1734
M60±14d	1826	≥ 1735
Study Day 1 = start date of CTL019		

Table 5-7 shows the defined time windows for patient reported outcome.

For PRO, if more than one assessment is done within the Baseline time window, the assessment closest to and before the first day of study treatment will be used. For all other time windows, if two assessments are obtained with the same time difference compared to the scheduled visit day, the assessment obtained prior to the scheduled visit will be considered.

Table 5-7 Time windows for patient reported outcome and performance status

Time Window	Planned visit timing (study day)	Time Window Definition (Study days)
W-16 to D-1	Before Study Day -1	Last one on or before first day of study treatment
D28±4d	28	21 to 59
M3±14d	91	60 to 136
M6±14d	183	137 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 = start date of CTL019 infusion		

5.5 Handling of missing or partial dates

Missing or partial date imputation will be conducted according to the logic described in this section. The imputed dates will be used for the calculation of duration of events. However, in the listings only the original reported dates will be listed.

5.5.1 AE date imputation

Date imputation is the creation of a new, complete date from a partial one according to an agreed and acceptable algorithm. Missing date for AE will be handled according to rules specified below. A partial date is simply an incomplete date e.g. DDOCT2001: the days are missing from this DDMMMYYYY date.

Partial AE start dates, if left partial, would ultimately mean the following:

It would not be possible to place the AE in time.

Therefore the treatment/dosage at the time of the event would be unknown.

Therefore the event could not be reported/summarized appropriately – if at all.

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

AE start date will be imputed as follows:

The following Table 5-8 explains the abbreviations used.

Table 5-8 AE/treatment date abbreviations

	Day	Month	Year
Partial Adverse Event Start Date	<not used>	AEM	AEY
Treatment Start Date (TRTSTD)	<not used>	TRTM	TRTY

The following matrix Table 5-9 describes the possible combinations and their associated imputations. In the table body the upper text indicates the imputation and the lower text the relationship of the AE start date to the treatment start date (TRTSTD).

Table 5-9 AE partial date imputation algorithm

	AEM MISSING	AEM < TRTM	AEM = TRTM	AEM > TRTM
AEY MISSING AEY < TRTY	NC Uncertain (D)	NC Uncertain (C)	NC Uncertain (C)	NC Uncertain (C)

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

The following Table 5-10 is the legend to the above table.

Table 5-10 AE/treatment date relationship and imputation legend

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

The following Table 5-11 gives a few examples.

Table 5-11 AE imputation example scenarios

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

Note, it may happen that the imputed AE start is after AE end date, in that case, imputed AE start=AE end date.

There **will be no** attempt to impute the following:

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

Partial AE end date will be imputed as follows:

- Imputed date = min (date of death if applicable, last day of the month), if day is missing;
- Imputed date = min (date of death if applicable, 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.

Missing AE end date or AE end date after data cutoff will be imputed as follows:
All events with start date before or on the cut-off date, and with end date missing or after the cut-off date will have the end date imputed as the minimum of the cut-off date, end of study evaluation (i.e. completion of the last phase of the study) or date of death (if applicable). For these events, the imputed end date will not appear in the listings, instead, they will be reported as "continuing".

5.5.2 Concomitant medication date imputation

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

5.5.3 Incomplete date for anti-neoplastic therapies

Prior therapies

Start date:

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

End date:

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

Imputed date = min (start date of study treatment, 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.

Post therapies

Start date:

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

Imputed date = max (last date of study treatment + 1, 01JAN), if day and month are missing.

End date: No imputation.

5.5.4 Incomplete assessment dates for tumor assessment

All investigation dates (e.g. peripheral blood, bone marrow) must be completed with day, month and year.

If one or more investigation dates are incomplete but other investigation dates are available, the incomplete date(s) are not considered for calculation of the assessment date and assessment date is calculated as the latest of all investigation dates (e.g. peripheral blood, bone marrow) if the overall disease response at that assessment is CR/CRi/UNK. Otherwise, if overall lesion response is relapsed disease or no response, the assessment date is calculated as the earliest date of all investigation dates at that evaluation number that reveals a relapse/no response. If all measurement dates have no day recorded, the 1st of the month is used.

If the month is not completed, for any of the investigations, the respective assessment will be considered to be at the date which is exactly between the previous and the following assessment. If both a previous and following assessments are not available, this assessment will not be used for any calculations.

5.5.5 Incomplete date for relapse or last known date subject in remission

The "Remission/Relapse Information" CRF will be used to track the relapse status for those patients who enter the secondary follow up phase while in remission.

If the day or month of date of relapse or last known date subject in remission is missing, it will be imputed to the minimal of date of assessment and the following:

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

5.5.6 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 3.1.8) 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 3.1.8.

5.5.7 Incomplete date for initial diagnosis, first relapse and most recent relapse

If the day or month of initial diagnosis, first relapse or most recent relapse is missing, the date of initial diagnosis 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

5.5.8 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.

5.6 Determination of missing scheduled disease assessments

For some time-to-event endpoints (i.e. DOR, RFS, EFS), classification of censoring or event can depend on the number of missing scheduled disease assessments.

The protocol defined schedule of disease assessments is every month for the first 6 months, every 3 months thereafter until Month 24, and every 6 months thereafter until Month 60. Each assessment is expected to be performed at the scheduled time point plus or minus 2 weeks in general, i.e. the window is 4 weeks or 1 month.

An event is considered as after 2 or more missing scheduled disease assessments if the distance between the last adequate non-relapse assessment and the event is larger than the threshold, defined as two times the protocol specified interval between the disease assessments plus the protocol allowed window around the assessments.

More specifically, an event is considered as having occurred after 2 or more missing scheduled disease assessments if the distance between the last adequate non-relapse assessment and the event is:

- >91 days (i.e. 1+1+1 months), if the last adequate non-relapse assessment occurs on or before Day 136 (i.e. middle point of Month 4 and Month 5)
- >152 days (i.e. 1+3+1 months), if the last adequate non-relapse assessment occurs after Day 136 and on or before Day 167 (i.e. middle point of Month 5 and Month 6)
- >213 days (i.e. 3+3+1 months), if the last adequate non-relapse assessment occurs after Day 167 and or on before Day 593 (i.e. middle point of Month 18 and Month 21)
- >304 days (i.e. 3+6+1 months), if the last adequate non-relapse assessment occurs after Day 593 and on or before Day 684 (i.e. middle point of Month 21 and Month 24)
- >395 days (i.e. 6+6+1 months), if the last adequate non-relapse assessment occurs after Day 684

5.7 EFS category

Patients will be categorized into EFS categories for exploratory analysis (e.g. EFS category vs manufactured product characteristics, etc.):

- EFS ≥ 6 (or 3) months: if patient achieved remission, and EFS event or censoring day ≥ 6 (or 3) months
- EFS event < 6 (or 3) months: if patient achieved remission, and EFS event day < 6 (or 3) months.
- EFS censor < 6 (or 3) months: if patient achieved remission, and EFS censor day < 6 (or 3) months
- Treatment failure: if patient is treatment failure
- Other: if patient does not satisfy any of above (i.e. pending patients who have not achieved CR nor have been classified as treatment failure; will NOT be included in the analysis)

Here EFS ≥ 6 months is determined by whether EFS (in days) is ≥ 167 days (i.e. 5.5 months). Similarly, EFS ≥ 3 months is determined by whether EFS (in days) is ≥ 76 days (i.e. 2.5 months).

5.8 CNS disease history search

CNS disease history is defined by the following MedDRA terms as collected in medical history:

- Neurological disorders congenital (HLGT)
- Congenital and peripartum neurological conditions (HLGT)
- Central nervous system haemorrhanges and cerebrovascular accidents (HLT)
- Noninfectious encephalopathy/delirium (SMQ) (broad)

6 References

Baumgartner RN, Roche AF, Himes (1986). Incremental growth tables: supplementary to previously published charts. American Journal of Clinical Nutrition, 43, 711-722.

CHMP (2010) Appendix 2 to the Guidance on the evaluation of anticancer medicinal products in man (CPMP/EWP/205/95 Rev. 3) on confirmatory studies in haematological malignancies

Fenichel P (2012). Delayed puberty. Endocr Dev, 22, 138-59.

Harrison CJ, Haas O, Harbott J, et al. (2010) Detection of prognostically relevant genetic abnormalities in childhood B-cell precursor acute lymphoblastic leukaemia: recommendations from the Biology and Diagnosis Committee of the International Berlin-Frankfürt-Münster study group. Brit J Hematol; 151(2):132-142.

Jeha S, Gaynon PS, Razzouk BI, et al. (2006) Phase II study of clofarabine in pediatric patients with refractory or relapsed acute lymphoblastic leukemia. J Clin Oncol, 24:1917–23

Kim H (2007) Cumulative Incidence in Competing Risks Data and Competing Risks Regression Analysis. Clinical Cancer Research 2007;13:559-565

National Comprehensive Cancer Network (NCCN) Guidelines (NCCN, 2013 vl), Acute Lymphoblastic Leukemia

O'Brien S, Schiller G, Lister J, et al. (2013) High-Dose Vincristine Sulfate Liposome Injection for Advanced, Relapsed, and Refractory Adult Philadelphia Chromosome-Negative Acute Lymphoblastic Leukemia 34, 676-83

Tanner JM, Goldstein H, Whitehouse RH (1970). Standards for children's height at ages 2-9 years allowing for height of parents. Archives of Disease in Childhood, 45, 755-762.

Topp M, Gokbuget N, Stein A, et al. (2015) Safety and activity of blinatumomab for adult patients with relapsed or refractory B-precursor acute lymphoblastic leukaemia: a multicentre, single-arm, phase 2 study. Lancet Oncol; 16: 57-66.

van der Veer A, Waanders E, Pieters R, et al. (2013) Independent prognostic value of *BCR-ABL1-like* signature and *IKZF1* deletion, but not high *CRLF2* expression, in children with B-cell precursor ALL. Blood; 122(15):2622-2629.

7 Appendix

CDC Growth Charts (http://www.cdc.gov/growthcharts/percentile_data_files.htm) for height (STATAGE) and weight (WTAGE) for children older than 2 years.


statatage.csv


wtage.csv