NOVARTIS

Clinical Development

CTL019

CTL019B2205J (14BT022) / NCT02228096

RAP Module 3 – Detailed Statistical Methodology

Author:

Document type: RAP Documentation

Document status: Final 2.0

Release date: 12-Feb-2015

Number of pages: 45

NOVARTIS

Document History

Version	Date	Changes
0.1	7-Nov-2014	Draft for Penn/CHOP review
0.2	21-Nov-2014	Draft incorporating Penn's comments
1.0	28-Jan-2015	Version sent to sDLT for review
2.0	12-Feb-2015	Final version

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
AUMC	Area under the first moment 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
██	██████████
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
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
██	█████████████

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 CTL09B2205J: 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 chemo-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 (12 months), Secondary Follow-up (Relapse Follow-up), and Survival Follow-up. Efficacy will be assessed at months 1, 2, 3, 4, 5, 6, 9, and 12 months post-infusion 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). Safety will be assessed until the end of treatment and primary follow-up phase. 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.

Figure 2-1   Study design
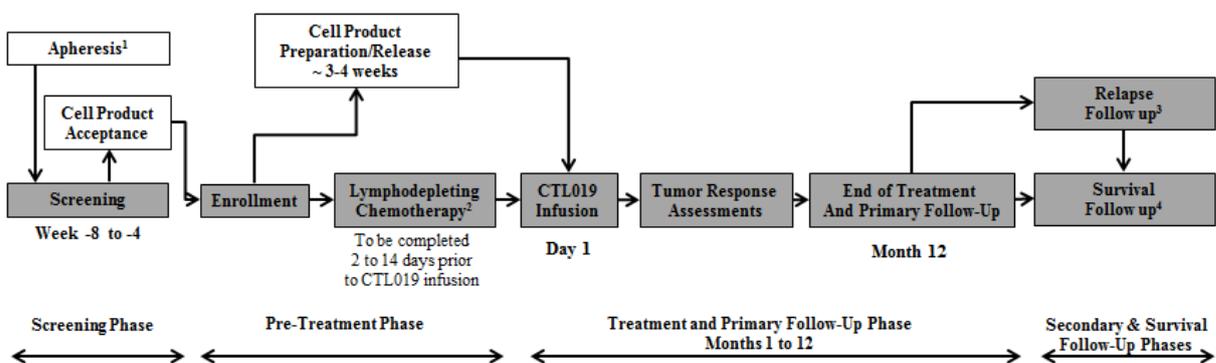
1   Performed prior to Study Entry
2   As indicated per protocol
3   If patients complete or prematurely discontinue from the Treatment and Primary Follow-Up Phase while in remission
4   Survival followed until end of trial. 15-year Long term safety follow-up conducted under separate protocol

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 $\leq 1,000$ 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 of 2 to 5 x 10^6 autologous CTL019 transduced cells per kg body weight, with a maximum dose of 2.5 x 10^8 autologous CTL019 transduced cells will be administered. For patients with manufactured cell numbers falling below the above recommended dose ranges, CTL019 therapy will still be administered, however the minimum number of manufactured cells to be infused is 2 x 10^7.

Disease status, PK, ▮▮▮▮▮▮ and safety will be assessed in the Treatment and Primary Follow-up phase for up to 12 months post CTL019 infusion. For patients who complete or prematurely discontinue from the 12 months primary follow-up phase while in remission, they will enter the secondary follow up phase for relapse status. Patients who relapse will be followed for survival in the survival follow up phase.

The end of study is defined as the last patient's last visit (LPLV), which is the last patient's End of Treatment and Primary Follow-Up (EOT) visit (Month 12 evaluation), or the time of premature withdrawal.

Patients may continue to be followed under the current protocol for relapse and survival until Last Patient Last Visit (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 as measured by overall remission rate (ORR), which includes CR and CR with incomplete blood count recovery (CRi) as determined by independent review committee (IRC)assessment	ORR (= CR + CRi) per IRC assessment; See Protocol Appendix 1 for response definition
Key secondary	
Not applicable	Not applicable
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	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)	DOR, i.e. the time from achievement of CR or CRi whichever occurs first to relapse or death due to ALL
Evaluate the percentage of patients who achieve a CR or CRi associated with minimal residual disease (MRD) negative bone marrow	Percentage of patients with CR or CRi with MRD negative bone marrow, among all patients who achieve a CR or CRi
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 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	- q-PCR detected DNA encoding anti-CD19 chimeric antigen receptor (CTL019) in blood, bone marrow and CSF

Objective	Endpoint
available)	- Cmax, Tmax, AUCs and other relevant PK parameters of CTL019 in blood, bone-marrow, CSF if available
	- Maximum extent of expansion of CTL019 in blood (Cmax/C0,hr)
	- Persistence of CTL019 in blood and bone marrow (based on Time>LOQ, Time>other threshold CTL019 levels, MRTlast)
Describe the incidence of newly acquired and confirmed immunogenicity to CTL019	- Incidence of newly acquired and confirmed immunogenicity and anti-CTL019 assay titers
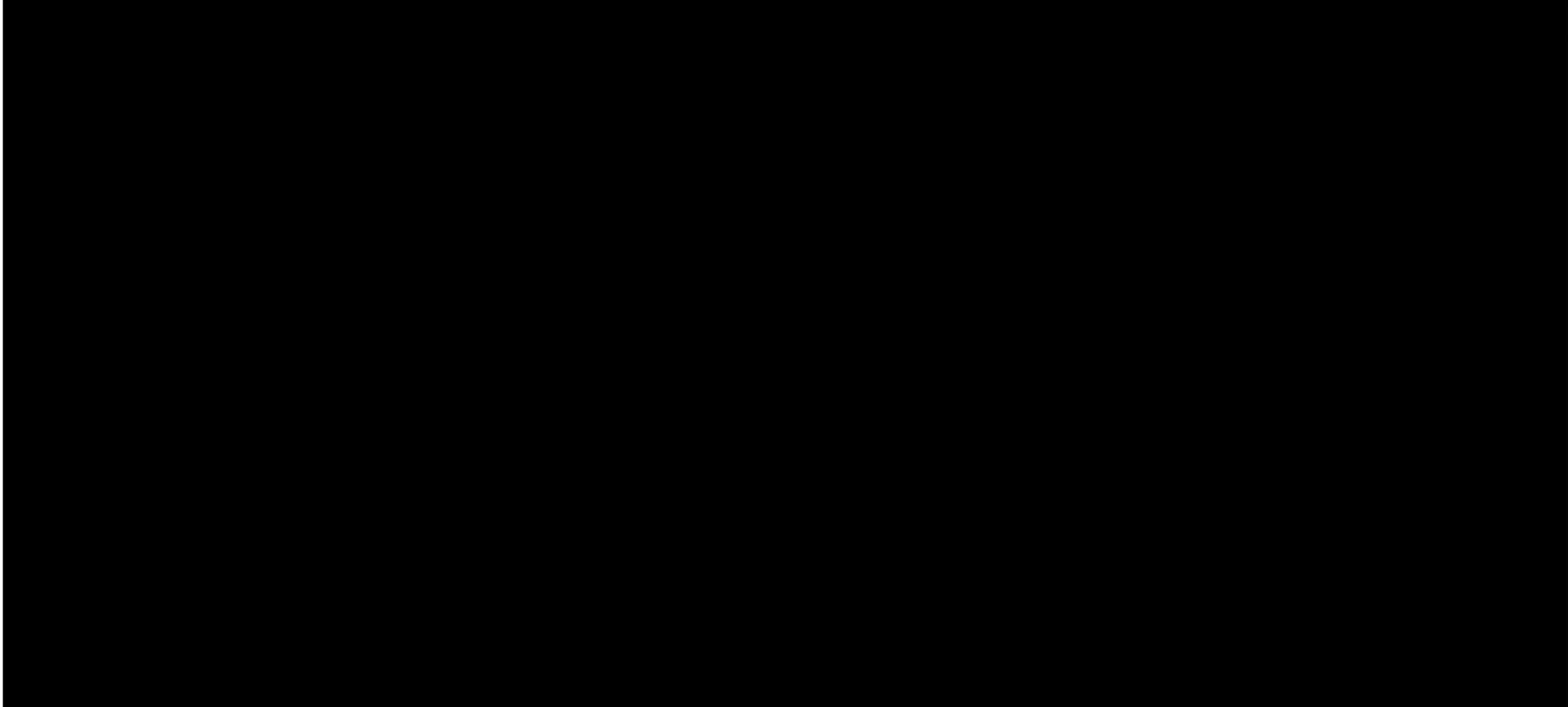
Objective	Endpoint

3   Definitions and general methodology

3.1   Definitions

3.1.1   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.

3.1.2   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" eCRF for the indication "Lymphodepleting".

3.1.3   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.

3.1.4   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 3.1.2); 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 3.1.3).

3.1.5   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.

3.1.6   Baseline

For ***baseline disease evaluations***, the most current bone marrow assessment prior to CTL019 infusion and lymphodepleting chemotherapy will be used as the baseline bone marrow assessment. The blood count and extramedullary disease assessments (physical exam, CSF, imaging if indicated) that are most proximal to and prior to CTL019 infusion and lymphodepleting chemotherapy the above defined baseline bone marrow assessment should

then be used as the baseline blood count and baseline extramedullary disease assessments, respectively.

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.

3.1.7 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 3-1. 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. Doses of 0 are allowed.
Any specific efficacy assessment date if available	Evaluation is marked as 'done'.
Laboratory/PK collection dates	Sample collection marked as 'done'.
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).

3.1.8 Lost to follow-up

Patients will be considered as lost to follow-up for time to event analysis 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).

3.2 Data Included in the analysis

Data from all participating centers will be combined for each study.

The primary analysis for each study will be performed after 50 patients have received CTL019 infusion and completed 6 months from study day 1 infusion or discontinued earlier.

Selected efficacy and safety analysis will be updated annually. 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 Full analysis set (FAS) will be used as the primary efficacy analysis set. 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 is defined as the point at which the patient meets all inclusion/exclusion criteria, and the patients' apheresis product is received and accepted by the manufacturing facility.

Full Analysis Set (FAS)

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

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 (PPS) consists of a subset of the patients in the FAS 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 2×10^6/kg CTL019 transduced cells will also be excluded.

Pharmacokinetic Analysis Set (PAS)

The pharmacokinetic analysis set (PAS) consists of FAS patients who have at least one sample providing evaluable pharmacokinetic (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.

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.

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 and Acute Lymphoblastic Lymphoma (ALL) studies.

The overall disease response is determined at a given evaluation using the criteria described in Table 3-2 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** • Trilineage Hematopoiesis (TLH) and < 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 CNS symptom assessment), and • If additional assessments (e.g. CSF assessment by 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
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, or • Platelets ≤ 100 x 10^9/L, or

Response category	Definition
	• 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, No response 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 can be assessed based on a partial evaluation (e.g. a relapse is assessed from blood alone). The assessment date for relapse is calculated as the earliest date of all assessments that reveal a relapse.

3.5 Time-to-event definitions

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

$$\text{Time to event} = \text{event date} - \text{start date} + 1 \text{ (in days)}$$

When no post-baseline response assessments are available, the date of CTL019 infusion will be used as end date when time is to be censored at last post-baseline response assessment, 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.

4.3 Background and demographic characteristics

The Full Analysis Set (FAS) 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

Medical history and ongoing conditions, including cancer-related conditions and symptoms at the time of informed consent will be summarized and listed. Separate summaries will be presented for ongoing and historical medical conditions. 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.

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 ATC class, and preferred term.

Patients will also be classified and summarized by their response status at study entry:
- Primary refractory: If patient did not ever have a complete remission (CR) prior to the study
- Relapse without SCT: If patient has not had SCT, had a CR from other therapy and relapsed prior to the study
- Relapse after SCT: If patient has had SCT and relapsed after SCT 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 at screening is provided to the manufacturing facility for CTL019 product manufacturing, and hence 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 summarized. 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.

Transfusions during the study will be listed.

Rescue medications are medications given for severe CRS due to CTL019 cells. Number of patients requiring rescue medications (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) rate, which includes CR and CR with incomplete blood count recovery (CRi) as determined by 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 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).

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 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 ORR will be summarized along with the 2-sided 95% exact Clopper-Pearson confidence intervals. 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 or equal to 20% can be rejected. The primary efficacy endpoint, ORR will be analyzed based on the data observed in the FAS.

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 4.6.2.

4.6.1.4 Supportive analyses

The primary analysis will also be performed on the Enrolled Set and PPS using the same methodology. In addition, analysis will also be performed using all patients who satisfy all clinical eligibility criteria.

4.6.2 Secondary efficacy

No formal hypothesis testing is planned other than for the primary objective. The 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.

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

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 response evaluation at the "Month 6" scheduled visit of the protocol will be used to determine CR or CRi. If patient proceeds to SCT before the scheduled Month 6 visit, this endpoint is considered as not met. If the patient continues in the study beyond Month 6 but Month 6 evaluation for response is missing, then it will be imputed as CR or CRi only if both the adequate evaluations immediately prior to and following the Month 6 evaluation indicate CR or CRi. In this case, if patient proceeds to SCT before the latter of the two evaluations, this endpoint is considered as not met.

The evaluation with the overall response assessment date between day 168 and day 228 is deemed as the "month 6" visit. If there are more than one such evaluations, then the one that is closest to day 183 will be regarded as the month 6 evaluation.

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.

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

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.

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 4.6.2.1.

4.6.2.3   Duration of remission (DOR)

Duration of remission (DOR) 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 ALL 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

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 (Method 2) 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 (Method 2) will be performed.

In both above analyses (Methods 1 and 2), death due to reason other than ALL will be considered as a censoring event. However, if there are at least 5 patients who die due to reasons other than ALL, then additional sensitivity analyses (Methods 3 and 4) will be performed in which death due to reason other than ALL will be regarded as the competing risk event to other events of interest (relapse or death due to ALL). Methods 3 and 4 will handle SCT in the same fashion as Methods 1 and 2 respectively.

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

Table 4-1   Analyses of duration of remission (DOR)

	Death due to reason other than ALL	SCT after remission
Method 1	Censor at last adequate disease assessment	Censor at time of SCT
Method 2	Censor at last adequate disease assessment	Ignore SCT

Method 3	Competing risk analysis	Censor at time of SCT
Method 4	Competing risk analysis	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 2, 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.

For Method 3 and Method 4, 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 are at least 5 patients with competing risk event.
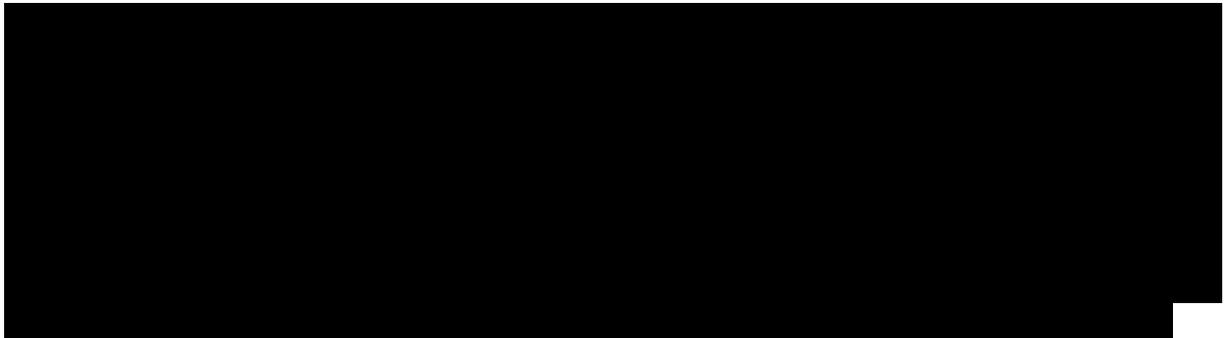
4.6.2.4　CR/CRi with MRD negative bone marrow

The percentage of patients with MRD negative bone marrow at the first time of achieving a CR or CRi, among all patients who achieve CR or CRi, will be summarized along with exact 95% CI.

4.6.2.5　Relapse free survival (RFS)

Relapse free survival (RFS) 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
- 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 4.6.2.3 for the rationale). In addition, a sensitivity analysis of RFS will be performed without censoring SCT.

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.

4.6.2.6   Event free survival (EFS)

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

- Death from any cause
- Relapse
- Treatment failure: Defined as no response in the study and discontinuation from the study due to any of the following reasons:
  - 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 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
- 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 4.6.2.3 for the rationale). In addition, a sensitivity analysis of EFS will be performed without censoring SCT.

EFS will be assessed in all patients (FAS). 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.

4.6.2.7   Overall survival (OS)

Overall survival (OS) 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. No censoring will be done in case of SCT. Thus, patients should be followed-up for survival also in case of SCT.

OS will be assessed in all patients (FAS). 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.

4.7 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 (12 months from CTL019 infusion)	Safety Set

4.7.2 Adverse events (AEs)

Reporting of AEs (except for CRS and GVHD) will be based on MedDRA (latest version per database lock) and Common Terminology Criteria for Adverse Events (CTCAE) version 4.03. 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.

Adverse events of special interest (AESI) are described in Table 4-3 below (the terms could be further updated and finalized before database lock). The current search criteria of AESI are based on limited experience from ongoing clinical studies without an accurate assessment of causality. The search criteria of the AESI may be updated prior to reporting. AESI that occur within 8 weeks of the last CTL019 infusion will be summarized by group term and preferred term.

Table 4-3   **Adverse events of special interest (AESI) search criteria based on MedDRA v17.0**

AESI group term	MedDRA term	Type
Syndromes	Cytokine Release Syndrome	PT
	Histiocytosis haematophagic	PT
	Tumor Lysis Syndrome	PT
Cytokine Release Syndrome Symptoms	Pyrexia	PT
	Myalgia	PT
	Hypotension	PT
	Dyspnoea	PT
	Tachypnoea	PT
	Capillary Leak Syndrome	PT
	Hypoxia	PT
	Organ Failure	PT
	Acute Respiratory Distress Syndrome	PT
Tumor Lysis Syndrome Symptoms	Hyperkalaemia	PT
	Hyperphosphataemia	PT
	Hyperuricaemia	PT
	Hypocalcaemia	PT
Histiocytosis Haematophagic Symptoms	Splenomegaly	PT
	Marrow depression and hypoplastic anaemias	HLT
	Haemolysis	PT
	Disseminated intravascular coagulation	PT
	Blood triglycerides increased	PT
	██████████████	█
Organ Dysfunction	Hepatic enzymes and function abnormalities	HLT
	Renal Failure and Impairment	HLT
	Confusion and disorientation	HLT
	Encephalopathy	PT
Allergic Reaction	Infusion related reaction	PT

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 during study follow-up, 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 of special interest (AESI), regardless of study drug relationship, by group term, preferred term and maximum grade
- Adverse events of special interest (AESI), suspected to be study drug related, by group term, 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

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

4.7.3 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 laboratory tests for Safety Set:

- Shift tables using CTCAE grades to compare baseline to the worst post-infusion value within 12 months from infusion.

    o 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 within 12 months from infusion, with descriptive statistics of baseline value, worst post-infusion value and the change.

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

Immunogenicity assessment will include prevalence of immunogenicity against CTL019 (pre-existing), incidence of immunogenicity (sum of treatment-induced and treatment-boosted pre-existing assay titers), proportion of patients making transient anti-CTL019 assay titers, proportion of patients making sustained anti-CTL019 assay titers. Descriptive summaries will be provided.

4.7.6 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 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). The age category immediately above the patient's exact age should be used. SDS is actually a Z score that measures the distance from the population mean in units of standard deviations. That is, SDS < 0 refers to values lower than the population mean, and for example SDS ≤ -1.645 refers to values in the lowest 5%. (The 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 3 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 5th/95th 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.

4.7.7   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.

4.7.8 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 summarized descriptively by time windows defined in Section 5.4 and will be listed by subject.

4.8 Pharmacokinetic analysis

PAS will be used for all PK analysis.

The following parameters will be displayed graphically:

Q-PCR generated CTL019 transduced T-cell concentration versus time and mean CTL019 concentration versus time in peripheral blood, bone-marrow and CSF, and ▮▮▮▮▮▮▮▮▮▮▮▮▮▮▮▮▮▮▮▮▮▮▮▮▮▮▮▮▮▮▮

The PK parameters listed in Table 4-4 will be estimated from the individual concentration versus time profiles using a non-compartmental approach within Phoenix$^®$ (Pharsight, Mountain View, CA). All concentrations below the limit of quantification are set to zero in the source dataset and will be labeled as below the limit of quantification in the concentration data listings. As a result concentrations below the limit of quantification will be treated as zero in summary statistics and treated as missing for the calculation of geometric mean and geometric CV%. For the calculation of the PK parameters, a zero value will be imputed for the values below the limit of quantification after the administration and during the expansion phase. The values below the limit of quantification during the elimination phase used for $T_{1/2}$ estimation will not be imputed and will be considered as missing.

Table 4-4   Noncompartmental pharmacokinetic parameters

AUC 0 - T_{max}	The AUC from time zero to T_{max} in peripheral blood (mass x time x volume-1)
AUC T_{max} - 28d or 84d	The AUC from time T_{max} to day 28 or 84 or other disease assessment days, in peripheral blood, (mass x time x volume-1)
AUC 0 - 28d or 84d	The AUC from time zero to day 28 or 84 or other disease assessment days, in peripheral blood (mass x time x volume-1)

C_{max}	The maximum (peak) observed in peripheral blood drug concentration after single dose administration (mass x volume - 1)
T_{max}	The time to reach maximum (peak) peripheral blood drug concentration after single dose administration (time)
$T_{1/2}$	The half-life associated with the disposition phase slopes (alpha, beta, gamma etc.) of a semi logarithmic concentration-time curve (time) in peripheral blood
$(C_{max} / C_{0,hr})$	In vivo CTL019 T-cell expansion in peripheral blood
T_{LOQ}	Duration of time at or above LOQ (or other threshold CTL019 concentrations if applicable): Persistence of CTL019 T-cell based on longest consecutive time above the LOQ of CTL019 in peripheral blood after CTL019 administration; other clinically relevant threshold CTL019 concentrations may be used to determine persistence (time)
MRTlast	Mean residence time from the time of dosing to the time of the last measurable concentration in peripheral blood.
* For flow cytometry kinetic values, units of % may be used in place of mass*volumn-1 for above parameters.	

Descriptive statistics of PK parameters (mean, standard deviation, coefficient of variation, geometric mean, CV% geometric mean, median, min and max) will be summarized by best overall response. When a geometric mean will be presented, it will be stated as such. A range of values will be presented for selected variables. For T_{max} median values and ranges only will be given.
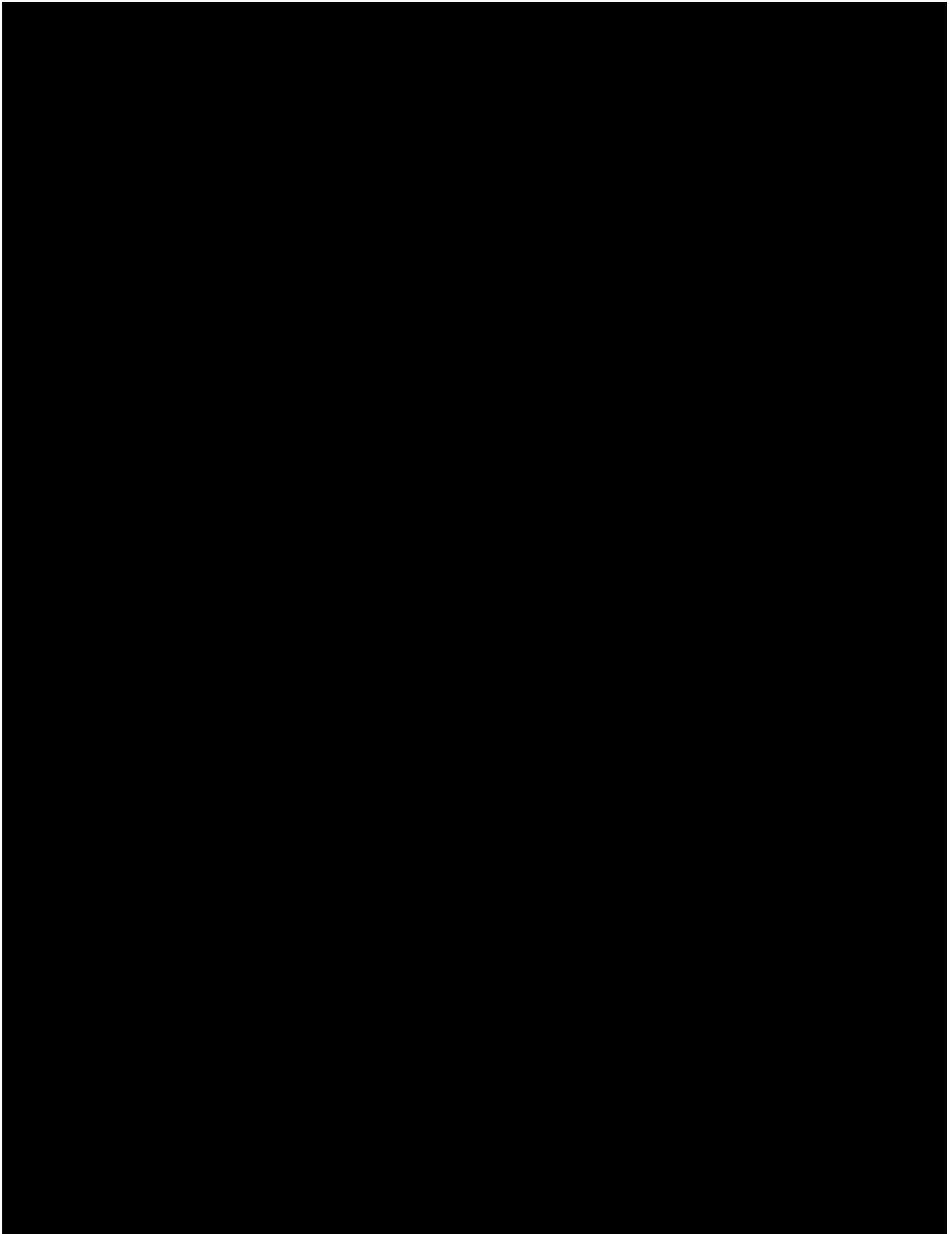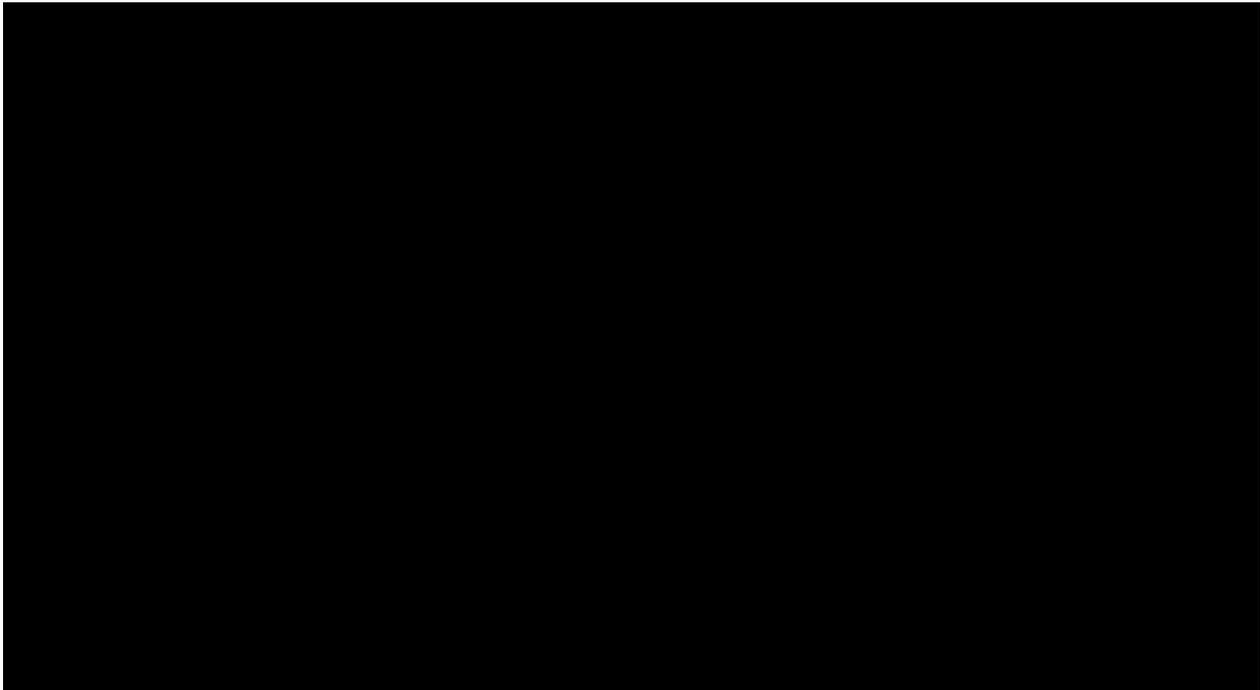
4.10 Apheresis product processing and CTL019 product characteristics

The total cell count, CD3+CD45+ (% and absolute count), CD3+CD4+CD45+ (% and absolute count), CD3+CD8+CD45+ (% and absolute count), total malignant cell count, elutriation (yes/no) and positive selection (yes/no) will be listed and summarized for apheresis product processing(from the "Apheresis product processing" eCRF page).

The following product characteristics for the manufactured product (from the "Certificate of Analysis – Manufactured Product" eCRF pages) will be listed and summarized: Cell Viability on Sentinel Vial (%), CD3+CD45+ (%), Transduction efficiency (%), Vector DNA sequence: CART-19 PCR (copies/cell), VSV-G DNA: RCL, value (copies/ug DNA), as well as the proportion (and reason) of products not meeting safety or other criteria.

The cryopreservation of cell doses for the primary product for each patient will be summarized: the total cell count (cells and cells/kg), CTL019 cell count (cells and cells/kg) and total product volume (mL). If a patient had multiple bags, then the values of the above parameter for each bag will be added up for the patient.

███

4.11   Subgroup analyses

4.11.1   Efficacy subgroup analyses

Subgroup analyses for the primary endpoint will be performed on the following based on the patient's baseline status:

- Age: ≤10 years, >10 years to <15 years, ≥15 years
- Gender: Male, Female
- Race: Asian, Black, Caucasian, Native American, Other, Pacific Islander, Unknown
- Ethnicity: Hispanic or Latino, Chinese, Indian, Japanese, Mixed ethnicity, Other
- Response status at study entry:
    - Primary refractory: If patient did not ever have a complete remission (CR) prior to the study
    - Relapse without SCT: If patient has not had SCT, had a CR from other therapy and relapsed prior to the study
    - Relapse after SCT: If patient has had SCT and relapsed after SCT prior to the study
- 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

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.
- 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.

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.

4.11.2   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 <15 years, ≥15 years
- Gender: Male, Female
- Race: Asian, Black, Caucasian, Native American, Other, Pacific Islander, Unknown
- Ethnicity: Hispanic or Latino, Chinese, Indian, Japanese, Mixed ethnicity, Other
- Prior SCT therapy and residual donor engraftment status: Prior SCT with any degree of residual donor engraftment, SCT without residual donor engraftment, or no prior SCT.

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

4.12   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.

Based on the null hypothesis of ORR ≤ 20% and alternative hypothesis of ORR >20%, 50 patients in the FAS will provide 95% power to demonstrate statistical significance at one-sided cumulative 0.025 level of significance, if the underlying ORR is 45%. In this setting, an ORR of 17/50=34% will be needed to claim success.

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., 63 to 67 patients are estimated to be enrolled to ensure 50 patients are treated.

4.13   Interim analyses

No interim analysis is planned for the study. The DMC will review safety data periodically.

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;

TABLEES outcome/binomial(CL=exact);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

For time to event endpoints (DOR, EFS and OS), the follow up time (in months) is calculated as:

Follow-up time = (Date of event or censoring – Date of first CTL019 infusion + 1)/30.4375.

The study follow up duration (in months) will be calculated as (Analysis cut-off date – Date of first CTL019 infusion + 1)/30.4375).

5.4 Time windows

In order to summarize the patient reported outcome (PRO), growth data, PK and ████████ 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, ██████████ 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.
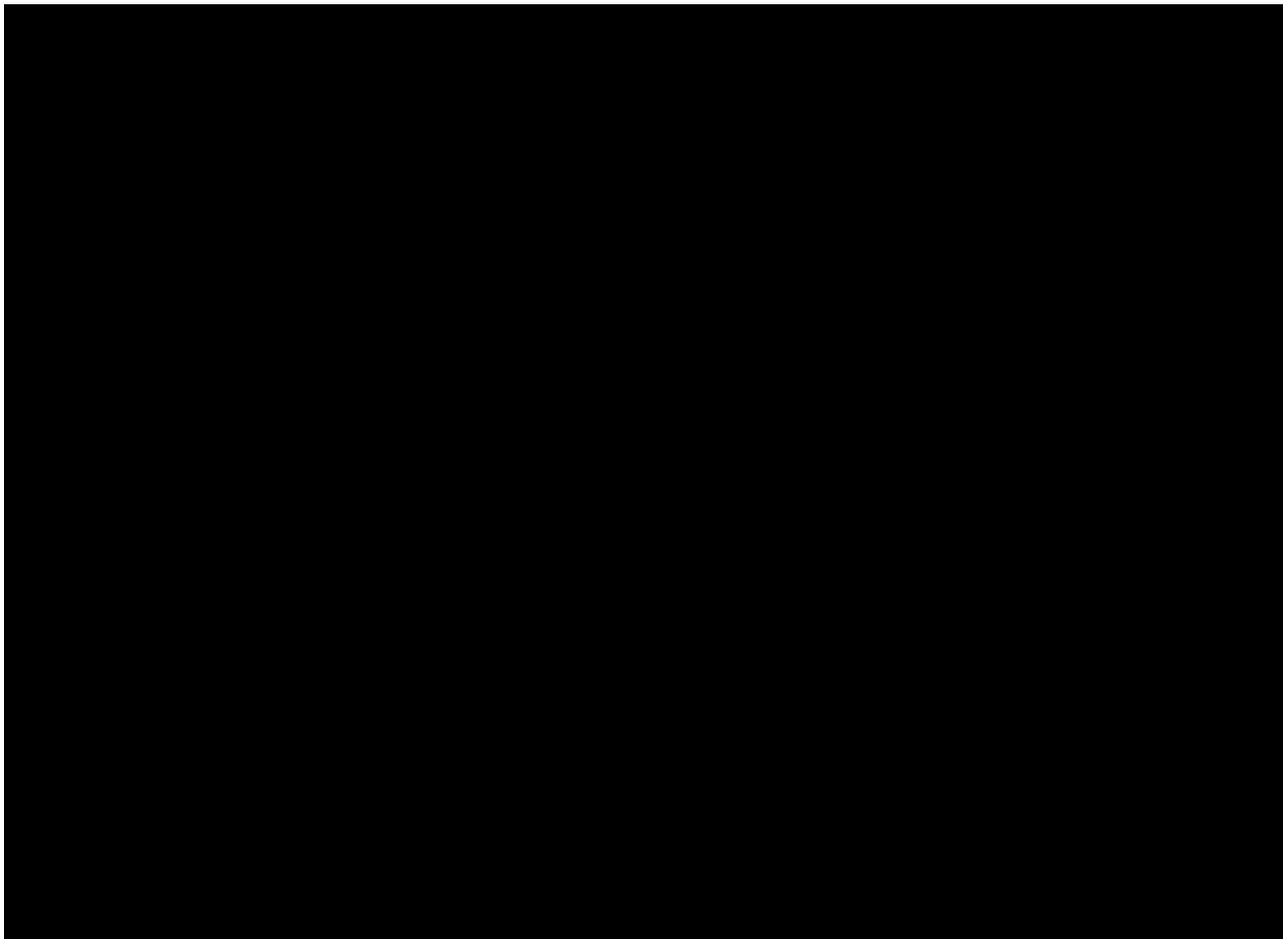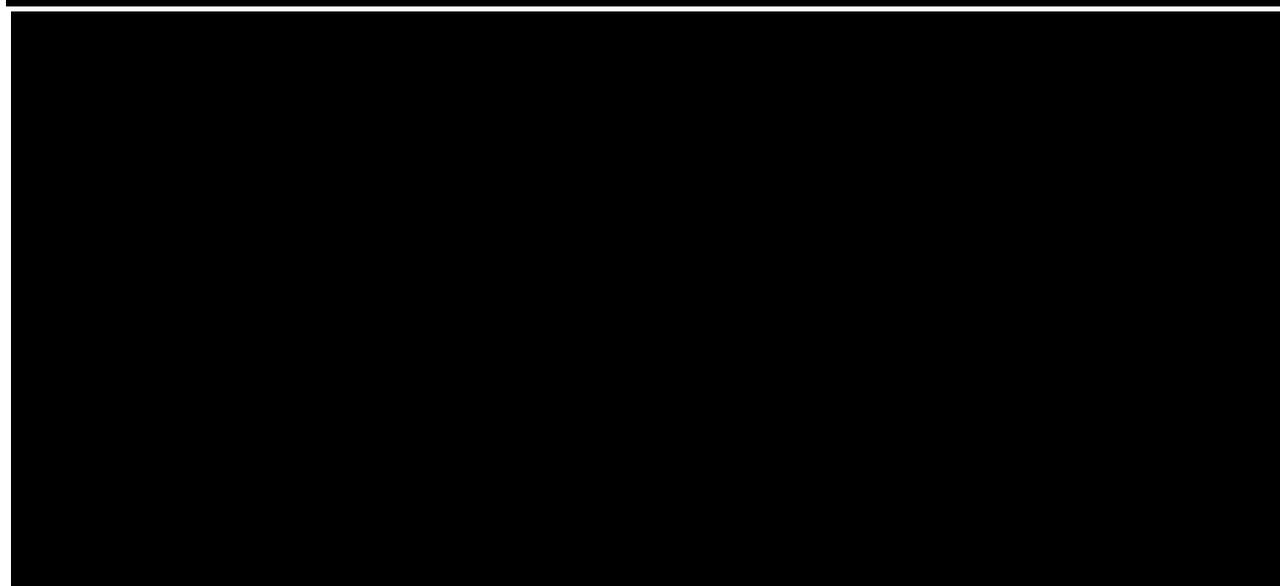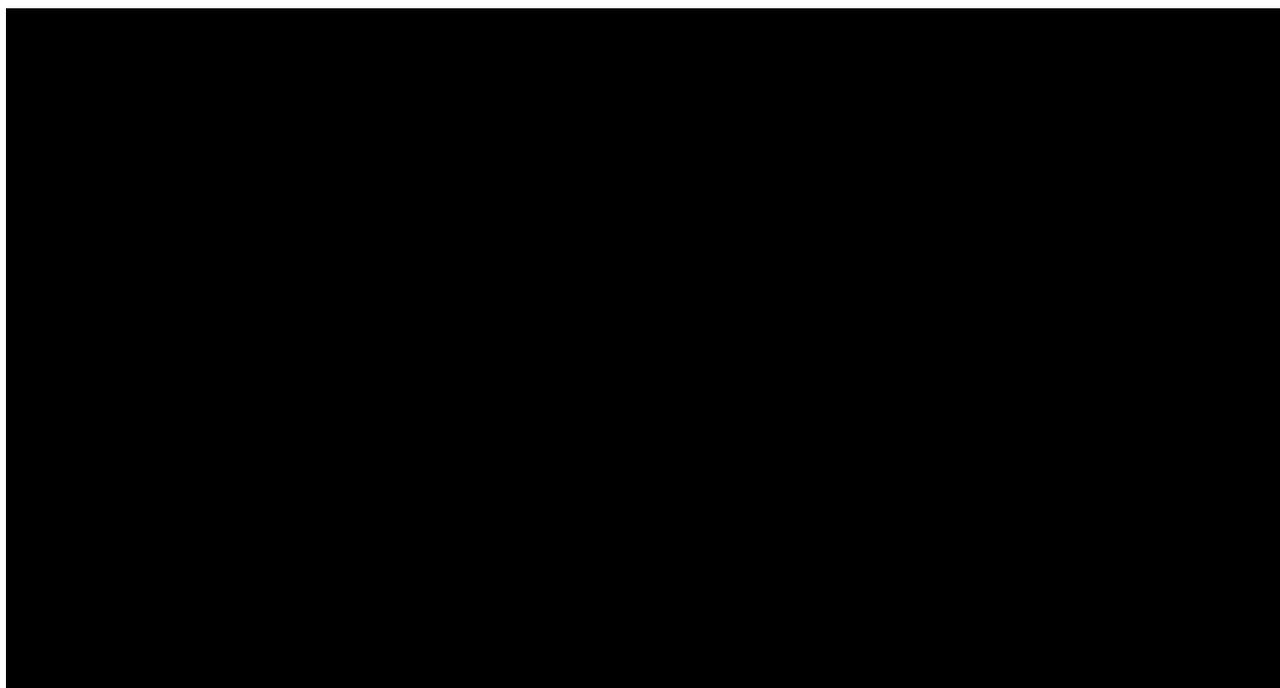
Table 5-3 shows the defined time windows for PK sample.

Table 5-3   **Time windows for PK**

Time Window	Planned visit timing (Study day)	Time Window Definition (Study day)
CTL019 pharmacokinetics by q-PCR in peripheral blood		
W-3 to D-8 Enrollment/Pre-Chemotherapy	Before Study Day -8	≤-1
D1 10 min ± 5 min post-infusion	1	1 to 2
D4±1d	4	3 to 5
D7±1d	7	6 to 10
D14±3d	14	11 to 17

D21±3d	21	18 to 24
D28±7d	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
CTL019 pharmacokinetics by q-PCR in bone marrow aspirate		
W-8 to W-4 Screening	Before Study Day -28	≤-1
D28±7d	28	1 to 59
M3±14d	91	60 to 136
M6±14d	183	≥137
CTL019 pharmacokinetics by q-PCR in CSF		
W-8 to W-4 Screening	Before Study Day -28	≤-1
D28±7d	28	≥ 1
Immunogenicity serum sample;		
Immunogenicity peripheral blood sample		
W-3 to D-8 Enrollment/Pre-Chemotherapy	Before Study Day -8	≤-1
D14±3d	14	1 to 21
D28±7d	28	22 to 59
M3±14d	91	60 to 136
M6±14d	183	137 to 273
M12±14d	365	≥274
Study Day 1 = start date of CTL019		

Table 5-4 shows the defined time windows for growth data and Tanner staging.

Table 5-4   Time windows for growth data and Tanner staging

Time Window	Planned visit timing (Study day)	Time Window Definition (Study day)
Height and Tanner Stage		
W-8 to W-4	Before Study Day -8	≤-1
M6±14d	183	1 to 273
M12±14d	365	≥274
Weight		
W-8 to W-4	Before Study Day -8	≤-1
D28±7d	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
Study Day 1 = start date of CTL019		

5.5   Handling of missing or partial dates

For patients not known to have died prior to the cut-off date:

- All events with start date before or on the cut-off date, and with end date missing or after the cut-off date will be reported as "continuing".

- This approach applies, in particular, to AEs and concomitant medication reports. For these events, the end date will not be imputed and therefore will not appear in the listings.

For patients known to have died prior to or on the cut-off date:

- 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 to the death date.

- This approach applies, in particular, to AEs and concomitant medication reports. For these events, the imputed end date will not appear in the listings.

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.

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/missing AE **end dates**

The following Table 5-5 explains the abbreviations used.

Table 5-5   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 5-6 describes the possible combinations and their associated imputations. In the light grey boxes the upper text indicates the imputation and the lower text the relationship of the AE start date to the treatment start date (TRTSTD).

Table 5-6　　　　**AE partial date imputation algorithm**

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

The following Table 5-7 is the legend to the above table.

Table 5-7　　　　**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 5-8 gives a few examples.

Table 5-8　　　　**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

Partial AE start date	Treatment start date	Relationship	Imputation Calculation	Imputed Date
ddNOV2001	20OCT2001	After	(A)	01NOV2001

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/No Response/UNK. Otherwise, if overall lesion response is relapsed disease, the assessment date

is calculated as the earliest date of all investigation dates at that evaluation number. 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 3.1.7) 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 imputed to the maximum of the full (non-imputed) last contact date and 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

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 and every 3 months thereafter. 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.

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. 3 months), if the last adequate non-relapse assessment occurs on or before Day 136 (i.e. 4.5 months)
- >152 days (i.e. 5 months), if the last adequate non-relapse assessment occurs after Day 136 and on or before Day 167 (i.e. 5.5 months)
- >213 days (i.e. 7 months), if the last adequate non-relapse assessment occurs after Day 167 (i.e. 5.5 months)

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.

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.

statage.csv

wtage.csv