STATISTICAL ANALYSIS PLAN

Study Title:	The ROCKET Study: A Phase 2, Single-arm, Multicenter Trial to Determine the Efficacy and Safety of JCAR015 in Adult Subjects with Relapsed or Refractory B-Cell Acute Lymphoblastic Leukemia
Study Number:	015001
NCT Number:	NCT02535364
Protocol Version (Date):	Amendment 5.1 (14 Jul 2016)
Analysis Plan Version:	4.0
Analysis Plan Date:	09 Nov 2016

DISCLOSURE

REDACTED STATISTICAL ANALYSIS PLAN

JCAR15 Protocol 015001

The ROCKET Study: A Phase 2, Single-arm, Multicenter Trial to Determine the Efficacy and Safety of JCAR015 in Adult Subjects with Relapsed or Refractory B-Cell Acute Lymphoblastic Leukemia

The information contained in the attached report is the property of Celgene and should not be shared or used for any purpose other than that for which it was provided.

Celgene is committed to providing information about its clinical trials to researchers and patients with the goal of furthering science and enhancing healthcare worldwide. Laws and regulations require, however, that Celgene protects patient privacy. The company may further have legal or contractual obligations not to disclose commercial or technical information provided by or related to certain partner companies or vendors.

The attached report is presented in its original format, but certain information has been redacted in order to comply with the aforementioned obligations or to protect Celgene's confidential commercial information. The redactions are based on the following principles:

- Redacted information has been replaced by grey space, maintaining original spacing and pagination.
- Any information that might allow the identification of individuals has been redacted for anonymization.
- Attachments to this report that contain confidential information are not made available. Such attachments include those that contain identifiable patient information, such as subject listings, narratives, and profiles. They also may contain confidential commercial information such as methodologies, and hypothesis generating and exploratory analyses.
- Cross-references to these attachments (such as links to subject listings in Section 16.2) are not redacted from the body of the report. However, the hyperlinks in the electronic document are no longer functional.
- Information about Celgene vendors and their services are redacted because many contracts prohibit disclosure of that information. Further, laws and regulations prevent us from disclosing certain information about our vendors or their services because it is protected by copyright.

Information about Celgene's redaction policies and the availability of additional data from this report may be found at http://www.celgene.com/research-development/clinical-trials/clinical-trials-data-sharing/.

STATISTICAL ANALYSIS PLAN

Study Title:	The ROCKET Study: A Phase 2, Single-arm, Multicenter Trial to Determine the Efficacy and Safety of JCAR015 in Adult Subjects with Relapsed or Refractory B-Cell Acute Lymphoblastic Leukemia
Name of Test Drug:	JCAR015
Study Number:	015001
Protocol Version (Date):	Amendment 5.1 (14 Jul 2016)
Analysis Type:	Interim and Final
Analysis Plan Version:	
Analysis Plan Date:	09 Nov 2016
Analysis Plan Author:	

TABLE OF CONTENTS

LIST OF ABBREVIATIONS

AE	adverse event
AESI	adverse event of special interest
ALL	acute lymphoblastic leukemia
ATC	Anatomical-Therapeutic-Chemical
AUC	area under the concentration vs time curve
BMI	body mass index
BOR	best overall response
CAR	chimeric antigen receptor
CI	confidence interval
CIF	cumulative incidence function
C_{max}	maximum concentration
CNS	central nervous system
CR	complete remission
CRF	case report form
CRi	complete remission with incomplete blood count recovery
CRS	cytokine release syndrome
CSF	cerebrospinal fluid
CTCAE	Common Terminology Criteria for Adverse Events
DOR	duration of remission
DSMB	data safety monitoring board
ECOG	Eastern Cooperative Oncology Group
EFS	event-free survival
EOS	end-of-study
HLGT	high-level group term
HLH	hemophagocytic lymphohistiocytosis
HLT	high-level term
HRQL	health-related quality-of-life
HSCT	hematopoietic stem cell transplant
ICU	intensive care unit
ID	identification
IEC	Independent Ethics Committee
IRB	Institutional Review Board
IRC	independent review committee
KM	Kaplan-Meier
LLT	lower-level term
LTFU	long-term follow-up
MMSE	Mini Mental State Examination
MRD	minimal residual disease
MSKCC	Memorial Sloan Kettering Cancer Center
ORR	overall remission rate
OS	overall survival
Ph-	Philadelphia chromosome-negative
PK	pharmacokinetic(s)
PP	per protocol

PT	preferred term
QC	quality control
qPCR	quantitative polymerase chain reaction
RCR	replication-competent retrovirus
RFS	relapse-free survival
SAE	serious adverse event
SAP	statistical analysis plan
sCRS	severe cytokine release syndrome
SD	standard deviation
SE	standard error
SOC	system organ class
TEAE	treatment-emergent adverse event
TLS	tumor lysis syndrome
$T_{1/2}$	half-life
T_{max}	time to maximum concentration
WHO-DD	World Health Organization Drug Dictionary

1 Introduction

This document details the planned statistical analysis methods for Protocol 015001, a Phase 2, single-arm, multicenter trial to determine the efficacy and safety of JCAR015 in adult subjects with morphologic relapsed or refractory B-cell acute lymphoblastic leukemia (ALL). This statistical analysis plan (SAP) is based on Amendment 5.1 of the study protocol, dated 14 July 2016.

1.1 Primary Study Objective and Study Design

The primary study objective is to evaluate the efficacy of JCAR015 as measured by the overall remission rate (ORR) after the final JCAR015 infusion in subjects with morphologic evidence of disease, based on independent review committee (IRC) assessment. The study is divided into two sequential parts, Part A and Part B; subjects will be screened and will provide informed consent before initiating any study procedures in Part A of the study.

In Part A, subjects will undergo leukapheresis to enable T-cell product generation and will then receive cytoreductive chemotherapy (and/or supportive care, at the Investigator's discretion) while the JCAR015 cell product is being manufactured. Upon recovery from any toxicities or any illnesses arising during Part A and upon availability of JCAR015 cell product, subjects with evidence of hematopoietic recovery as assessed by peripheral counts will undergo a bone marrow examination to assess the burden of disease (complete remission [CR] or morphologic disease). Eligible subjects will proceed to Part B of the study. Subjects without evidence of hematopoietic recovery will continue to be monitored for evidence of marrow recovery or morphologic relapse.

In Part B, eligible subjects will receive two infusions of JCAR015 T cells, administered 14 to 28 days apart. Each infusion is preceded by lymphodepleting chemotherapy with cyclophosphamide alone. Response will be evaluated at Day 28 following the final JCAR015 infusion. An interim analysis will be conducted between subjects who receive cyclophosphamide lymphodepleting chemotherapy followed by either JCAR015 (Protocol 015001) or MSKCC 1928z CAR T cells (Protocol 09-114) to assess clinical comparability of the respective investigational drug products.

Subjects treated in Part B of the study will be analyzed in two distinct groups (Group 1 and Group 2) based on their disease status prior to JCAR015 treatment (morphologic evidence of disease [≥ 5% lymphoblasts in bone marrow] or CR with presumed minimal residual disease [MRD], respectively). The primary efficacy analysis will be conducted in subjects with morphologic evidence of leukemia at the time of the first JCAR015 infusion (Group 1) who receive lymphodepleting chemotherapy with cyclophosphamide alone. Subjects who achieve a morphologic CR or complete remission with incomplete blood count recovery (CRi) after cytoreductive chemotherapy (Group 2) will not be included in the primary analysis but will be included in the secondary efficacy, safety, manufacturing, and exploratory analyses.

Bone marrow examinations will be repeated at Months 3, 6, and 12 after the final JCAR015 infusion or until the subject requires alternative therapy for his or her disease. Response assessment will utilize guidelines based on Cheson et al. (Cheson 2003) and NCCN ALL Guideline (National Comprehensive Cancer Network 2014). Safety, relapse, and survival will be assessed until the end of the treatment and primary follow-up phase (12 months after the final JCAR015 infusion). Post-study follow-up for survival, relapse, long-term toxicity, and viral vector safety and persistence

will continue under a separate long-term follow-up (LTFU) protocol for up to 15 years after the final JCAR015 infusion as per health authority regulatory guidelines.

1.2 Secondary and Exploratory Objectives

The secondary objectives of this study are:

1. To evaluate the duration of remission
2. To evaluate the percentage of subjects who achieve a CR or CRi with no evidence of MRD in the bone marrow
3. To evaluate the safety and tolerability of JCAR015 therapy
4. To evaluate disease control (relapse-free survival, event-free survival, and overall survival)
5. To characterize the cellular pharmacokinetic (PK) profile of JCAR015, including the quantity and persistence in the peripheral blood and bone marrow
6. To characterize the prevalence and incidence of humoral immune responses to JCAR015
7. To evaluate the ORR at Month 6 following the final JCAR015 infusion
8. To compare the safety and tolerability of the JCAR015 cell product with the Memorial Sloan Kettering Cancer Center (MSKCC) 1928z CAR T cell product
9. To evaluate the percentage of subjects who achieve a morphologic remission within 6 months after the final JCAR015 infusion and then proceed to hematopoietic stem cell transplant (HSCT)

1.3 Sample Size and Power

In a previous Phase 2 study of vincristine sulfate liposomal injection (HBS407) in adult subjects with Philadelphia chromosome-negative (Ph- ALL) in second or greater relapse, the ORR (CR + CRi as determined by an IRC) was 15.4% (95% CI [7.6%, 24.5%]). Based on the null hypothesis of ORR ≤ 20% and an alternative hypothesis of ORR = 50%, 50 subjects in the primary efficacy analysis set will provide > 95% power to demonstrate statistical significance at one-sided 0.025 significance level, and will provide an ample safety database.

Assuming ORR ≤ 30% under the null hypothesis (Center for Drug Evaluation and Research 2014), and an alternative hypothesis of ORR = 50% or ORR = 60%, 50 subjects in the primary efficacy analysis set will still provide ~76% or > 95% power to demonstrate statistical significance at the one-sided 0.025 significance level, respectively.

An interim analysis is planned based on approximately ten subjects in the primary efficacy analysis set (see definition in Section 3.1.3). This interim analysis will assess analytical and clinical comparability between the JCAR015 cell product used in this study and the 1928z CAR T-cell product used in the ongoing Phase 1 study (MSKCC Protocol 09-114).

Assuming that 15% of subjects enrolled on Part A will not be infused with JCAR015 due to reasons such as manufacturing failure or worsening of subject's medical condition and that 25% of subjects enrolled on Part A will achieve a CR from cytoreductive chemotherapy and will be treated in

Group 2 of Part B, approximately 110 subjects need to be enrolled to ensure that a minimum of 50 subjects are included in the primary efficacy analysis set.

2 Type of Planned Analysis

2.1 Data Safety Monitoring Board

An independent data safety monitoring board (DSMB) will review cumulative study data over the course of the study to evaluate safety, protocol conduct, and scientific validity and integrity of the trial. The detailed role and responsibilities of the DSMB and the scope of analysis to be provided to the DSMB are provided in a mutually agreed upon charter, which defines the DSMB membership, meeting logistics, and meeting frequency.

The DSMB will be convened prior to enrollment of the first subject on the protocol and will meet approximately quarterly throughout the trial and as needed to address any safety issues that may arise. The DSMB will provide advice to the Sponsor as outlined in the DSMB charter.

2.1.1 Safety Stopping Rules for Individual Events

Unexpected serious adverse events (SAEs) that are related to JCAR015 will be the primary criteria for pausing or stopping the study. Review of these SAEs, and any decision to pause enrollment or terminate the study, will be determined by the DSMB, the Sponsor, and the Medical Monitor.

Study enrollment will be paused pending notification of the DSMB and appropriate regulatory authorities if any subject experiences any of the following events within 30 days of a JCAR015 cell product infusion:

- Life-threatening (Grade 4) toxicity attributable to protocol therapy that is unexpected, unmanageable (i.e., does not resolve to Grade 3 or lower within 7 days), and unrelated to chemotherapy

  Expected toxicities include up to Grade 4 cytokine release syndrome (CRS), neurologic toxicity (e.g., confusion, aphasia, seizures, convulsions, lethargy, and/or altered mental status), fever, hypotension, hypoxia, tumor lysis syndrome (TLS), and disseminated intravascular coagulation. In addition, admission to the intensive care unit (ICU), the need for dialysis, and/or the need for mechanical ventilation are also expected. The expected toxicities may also result in secondary toxicities of Grade 4 renal toxicity, hepatic toxicity, or other organ involvement.
- Death related to JCAR015 therapy

The study will be terminated for the following reasons:

- Any subject develops uncontrolled JCAR015 proliferation that is unresponsive to treatment
- Any subject develops detectable replication competent retrovirus (RCR) during the study
- The Sponsor, Institutional Review Board (IRB)/Independent Ethics Committee (IEC), or DSMB decides that subject safety may be compromised by continuing the study
- The Sponsor decides to discontinue the development of JCAR015 in this indication or the development of JCAR015 for all indications

2.1.2 Safety Monitoring Boundaries

Besides the safety stopping rules for an individual safety events specified above, safety monitoring boundaries based on Bayesian framework (Thall 1994, Yao 2013) have also been included to help detect safety signals that may occur during the course of the study. The boundaries are based on the cumulative number of subjects who have received lymphodepleting chemotherapy with cyclophosphamide alone and at least one dose of JCAR015 and who experienced either of the following events within 30 days of a JCAR015 cell product infusion:

- A Grade 3, JCAR015-related, treatment-emergent adverse events of special interest (AESIs) that is unmanageable or fails to resolve to Grade 2 or better after 21 days
- A Grade 4, JCAR015-related, treatment-emergent AESI that is unmanageable or fails to resolve to Grade 3 or better after 7 days

AESIs include CRS, neurological toxicity, macrophage activation syndrome, organ dysfunction, hemophagocytic lymphohistiocytosis (HLH) symptoms, TLS, and allergic reaction (see Section 7.1.4).

As of 03 February 2015, in the MSKCC Study 09-114, the most significant toxicities have been severe cytokine release syndrome (sCRS) and neurotoxicity. In this study, nine out of 17 (53%) subjects with morphologic disease at the time of T-cell infusion developed sCRS and/or ≥ Grade 3 neurotoxicity; and three out of 16 (19%) subjects with MRD at the time of T-cell infusion developed severe CRS and/or Grade ≥ 3 neurotoxicity. Based on the assumed ratio of morphological and MRD subjects to be treated with JCAR015 in this study (i.e., ~25% subjects with MRD only) and the MSKCC study data, we assume a prior distribution for the probability of the occurrence of the AESI, $\pi_1 \sim \text{Beta}(10,12)$. We further assume a background incidence rate, $\pi_0 = 30\%$, i.e., the risk of developing SAEs or dying without receiving further treatment for this study population. For any i^{th} subject, a criterion function is defined as $P(\pi_1 - \pi_0 > \delta \mid i, X_i = x_i) > 0.9$, where $\delta = 15\%$ representing a 50% relative increase in risk over the background incidence rate. This criterion implies that the study will be paused for DSMB review as soon as there is at least 90% probability that JCAR015 therapy causes an increased risk of AESI by 15% or more over the background incidence rate. The calculation of Bayesian safety boundary was performed using the Multc Lean software package (MD Anderson Cancer Center, Houston, TX; available online at https://biostatistics.mdanderson.org/SoftwareDownload). The boundary values are shown in Table 1.

Whenever the safety boundaries are crossed, enrollment will be paused and an ad hoc DSMB meeting will be held to review the data. The data to be reviewed by the DSMB are specified in the DSMB charter and DSMB report template. The DSMB may override the safety stopping boundaries to recommend study continuation based on clinical and overall risk-benefit considerations.

Table 1: Safety Monitoring Boundaries Based on Thall and Simon (Thall 1994)

Number of JCAR015-Treated Subjects*	Incidence	Number of JCAR015-Treated Subjects*	Incidence
1-5	No Stopping	48	27
6	6	50	28
8	7	52	29
9	8	54	30
11	9	56	31
13	10	58	32
15	11	60	33
17	12	62	34
19	13	64	35
21	14	66	36
23	15	68	37
25	16	70	38
27	17	73	39
29	18	75	40
32	19	77	41
34	20	79	42
36	21	81	43
38	22	83	44
40	23	85	45
42	24	87	46
44	25	89	47
46	26	90	48

* Includes subjects who have received lymphodepleting chemotherapy with cyclophosphamide alone and at least one infusion of JCAR015

2.1.3 Efficacy Monitoring Boundaries

Because the initial clinical trials with CD19-targeted CAR T cell therapies have been associated with a high rate of clinical remission induction, Bayesian efficacy monitoring boundaries have been implemented to continuously monitor the efficacy of JCAR015 during the duration of the study. The boundaries are constructed for the purpose of efficacy monitoring and will not be used to stop the trial early for positive outcome.

The proposed efficacy monitoring rule is: $Pr(ORR>20\%|data) < 0.2$. That is, based on the observed data to determine if the posterior probability of ORR exceeding 20% (the alternative primary hypothesis) is less than 20%. Whenever the boundary is crossed, an ad hoc DSMB meeting will be held to review the study data, and an enrollment pause will be triggered pending the DSMB's recommendations. To calculate the efficacy monitoring boundary values, we assume a non-informative prior Beta(1,1), and the ORR as a binary response (CR/CRi: Yes or No) which follows a binomial distribution; therefore, the posterior distribution is still a beta distribution. The analysis will be based on the primary efficacy analysis set (see Section 3.1.4). The ORR or disease response

evaluation and baseline disease status will be based on Investigators' assessments. The boundary values are shown in Table 2.

Table 2: Efficacy Monitoring Boundaries

Number of Subjects in the Primary Efficacy Analysis Set	Number with CR or CRi per Investigator's Assessment
6	2
10	3
14	4
19	5
23	6
27	7
31	8
36	9
40	10
45	11
49	12
54	13
58	14
60	15

2.2 Interim Analysis

A formal interim analysis is planned after approximately ten subjects in the primary efficacy analysis set are treated and are followed for at least 1 month after the final JCAR015 infusion. For analysis of efficacy, the morphologic disease status at baseline and response will be based on the Investigators' assessments. In addition to safety and efficacy assessments, this interim analysis will assess the comparability between the JCAR015 cell product used in this study and the cell product used in the ongoing Phase 1 study (MSKCC Protocol 09-114) among patients who received cyclophosphamide lymphodepleting chemotherapy. In addition to ten subjects in the primary efficacy analysis set, it is estimated that approximately four additional subjects will achieve a CR from the cytoreductive chemotherapy administered in Part A. These subjects will be assigned to Group 2 and will receive JCAR015. Treatment-emergent adverse events (TEAEs) and AESI will be summarized for all JCAR015-treated subjects at the time the interim analysis is performed. Safety data from a cohort of subjects with similar baseline characteristics and who receive a similar CAR T-cell dose on MSKCC Protocol 09-114 will be summarized. The assessment of comparability will be descriptive and will not involve inferential statistics.

2.3 Primary Analysis

The primary analysis will be carried out after all subjects in the primary efficacy analysis set have been followed for at least 6 months from the start of Part B or have discontinued the study due to any reason.

2.4 Final Analysis

The final analyses will be carried out after all subjects have completed or discontinued the study due to any reason.

3 Analysis Sets and Data Conventions

Analysis sets define the population to be included in an analysis. The type of analysis sets and their definitions are provided in this section. The number and percentage of subjects eligible for each analysis set will be provided as part of the subject disposition summary. A by-subject listing of reasons for exclusion from the analysis sets will be provided.

3.1 Analysis Sets

3.1.1 Screened Set

The screened set will include all subjects who have signed informed consent.

Subjects who are withdrawn from the study because of prolonged illness during Part A or failure to generate a JCAR015 dose that meets the required quality control (QC) and release criteria may re-enroll in the study at a later time provided that the subject meets all eligibility criteria. A new subject ID will be assigned, and the subject will be counted twice in the screened set.

3.1.2 Enrolled Set

The enrolled set will include all subjects who are enrolled into Part A and undergo leukapheresis.

3.1.3 Primary Efficacy Analysis Set

The primary efficacy analysis set will include all subjects with morphologic evidence of leukemia, as determined by the IRC, at the time of the first JCAR015 infusion who receive lymphodepleting chemotherapy with cyclophosphamide alone and at least one infusion of JCAR015. The primary efficacy analysis set will be used for the primary analyses of ORR and duration of response (DOR).

3.1.4 JCAR015-Treated Morphological Disease Analysis Set

The JCAR015-treated morphological disease analysis set will include all subjects with morphologic evidence of leukemia as determined by the IRC at the time of JCAR015 infusion, and who receive at least one infusion of JCAR015 (i.e., Part B, Group 1).

3.1.5 JCAR015-Treated Analysis Set

The JCAR015-treated analysis set is the primary analysis set for safety analyses and will include all subjects who receive at least one infusion of JCAR015 in Part B of the study (i.e., Part B, Group 1 and Group 2).

3.1.6 Per Protocol Analysis Set

The per protocol (PP) analysis set represents a subset of the primary efficacy analysis set and includes subjects who are compliant with the major requirements of Part B of the study protocol. The main reasons for excluding a subject from the PP analysis set may include:

- The subject has a disease other than B-cell ALL at baseline
- The subject's prior therapy does not match the protocol requirements in terms of number and types of prior therapeutic regimens
- The documentation of the subject's disease is missing or incomplete

Subjects who only receive a single dose of JCAR015 in Group 1 will be included in the PP analysis set. The detailed classification of subjects to be excluded from the PP analysis will be made by the study team prior to primary data analysis or database lock.

3.1.7 Pharmacokinetic Analysis Sets

The PK analysis set includes data from subjects in the JCAR015-treated analysis set who have the necessary baseline and on-study cellular PK (i.e., expansion and persistence) measurements to provide interpretable results for the specific parameters of interest.

3.2 Data Handling Conventions

The following data conventions will be applied to all analyses:

- Data from all clinical trial sites will be combined for the final analysis.
- Analysis results will be presented using descriptive statistics. For categorical variables, the number and percentage of each category will be presented; for continuous variables, the number of subjects [n], mean, standard deviation [SD] or standard error [SE], median, first quartile [Q1], third quartile [Q3], minimum, and maximum will be presented. The mean, median, Q1, Q3, minimum, and maximum will be displayed to the reported number of digits; SD and SE to the reported number of digits plus 1.
- All data collected in the study will be presented in by-subject listings for the appropriate analysis set.
- In general, missing data will not be imputed unless methods for handling missing data are specified. By-subject listings will be sorted by subject identification (ID) number, visit date, and time (if applicable). Data collected on log forms, such as adverse events (AEs), will be presented in chronological order within subject.
- Age (in years) at first JCAR015 infusion will be used for analyses and presentation in listings.
- If there is a significant degree of non-normality, analyses may be performed on log-transformed data or non-parametric methods, as appropriate.
- Outliers will be identified during the data management and data analysis processes, but no sensitivity analyses will be conducted to evaluate the impact of outliers on efficacy or safety outcomes, unless specified otherwise. All data will be included in the data analyses.
- For time-to-event efficacy endpoints, the interval is calculated as:

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

  Time to event (in months) = (event date - start date + 1)/30.4375

- Since each AE with a change in toxicity grade will be recorded in the CRF as a separate AE record from the start of lymphodepleting chemotherapy in Part B through 30 days after the final JCAR015 infusion, for duration of adverse event captured in this period, the following calculation will be used:

  If an event resolves completely, i.e. an AE with an outcome of "Recovered/Resolved", the duration is calculated as:

  Duration of AE = AE end date – AE start date + 1

  If an event does not resolve completely, i.e., an AE with an outcome of "Recovering/Resolving", the duration is calculated as:

  Duration of AE = AE end date – AE start date.

- Calculation of follow-up time for time-to-event endpoints (e.g, overall survival [OS], DOR, relapse-free survival [RFS], event-free survival [EFS]) will be performed using the reverse Kaplan-Meier method (Schemper 1996).

3.3 Visit Windows

3.3.1 Definition of Study Day

The study day is the day relative to the first infusion date of JCAR015. It will be calculated from the date of first JCAR015 infusion and derived as follows:

- For post-infusion timepoints, assessment date – first JCAR015 infusion date + 1
- For timepoints prior to the first infusion, assessment date – first JCAR015 infusion date

Study Day 1 is defined as the day of first JCAR015 infusion.

3.3.2 Analysis Windows

Subject visits might not occur on protocol-specified days. Therefore, for the purpose of analysis, observations will be assigned to analysis windows (see Table 3).

Table 3: Analysis Windows

Nominal Visit	Nominal Day	Lower Limit	Upper Limit
Baseline (Dose #1)	Day 1 (Dose #1)		
Post Dose #1 Day 2	2	2	2
Post Dose #1 Day 4	4	3	5
Post Dose #1 Day 7	7	6	9
Post Dose #1 Day 11	11	10	12
Post Dose #1 Day 14	14	13	21
Post Dose #1 Day 28	28	22	44

Nominal Visit	Nominal Day	Lower Limit	Upper Limit
Dose #2	Day 1 (Dose #2)		
Post Dose #2 Day 2	2	2	2
Post Dose #2 Day 4	4	3	5
Post Dose #2 Day 7	7	6	9
Post Dose #2 Day 11	11	10	12
Post Dose #2 Day 14	14	13	21
Post Dose #2 Day 28	28	22	44
Post Final Dose Month 2	60	45	75
Post Final Dose Month 3	90	76	105
Post Final Dose Month 4	120	106	135
Post Final Dose Month 5	150	136	165
Post Final Dose Month 6	180	166	225
Post Final Dose Month 9	270	226	315
Post Final Dose Month 12	360	316	≥ 360

Visits according to the above windows will be determined for each variable analyzed unless noted otherwise in the following sections. If multiple non-missing numeric observations exist in a window, records will be chosen based on the following rules:

- For baseline, the last available record on or prior to the date of the first dose of JCAR015 will be selected. If there are multiple records with the same time or no time recorded on the same day, the average (arithmetic or geometric mean, as appropriate) will be used for the baseline value.
- For post-baseline visits:
  - The record closest to the nominal day for that visit will be selected.
  - If there are two records equidistant from the nominal day, the later record will be selected.
  - If there is more than one record on the selected day, the average will be taken unless otherwise specified.

If multiple valid, non-missing categorical observations exist in a window, records will be selected as follows:

- For baseline, the last available record on or prior to the date of the first JCAR015 infusion will be selected. If there are multiple records with the same time or no time recorded on the same day, the value with the lowest severity will be selected.
- For post-baseline visits, the most conservative value (e.g., the highest severity for a lab abnormality) within the window will be selected.

4 Subject Disposition

4.1 Subject Enrollment and Disposition

A summary of subject enrollment will be provided for each clinical trial site and overall. The summary will present the number and percentage of subjects enrolled.

Based on the screened set, the number and percentage of subjects who are eligible for Part B of the study, but who discontinue due to prolonged illness during Part A or failure to successfully generate a JCAR015 cell product, will be summarized. Available information on subjects who were screened but not enrolled into Part B of the study will be listed.

A summary of subject disposition will be provided for the primary efficacy analysis set, JCAR015-treated, and PP analysis sets. This summary will present the number and percentage of subjects in each of the categories listed below:

- Ongoing
- Completed study
- Discontinued the study with reasons for premature discontinuation
- Entered the long-term follow-up study

The denominator for the percentage calculation will be the total number of subjects in the appropriate analysis set. A cumulative frequency table showing subject disposition will also be provided.

The following by-subject listings will be provided in order to support the above summary tables:

- Reasons for premature discontinuation of study drug and/or study
- Reasons for screening failure
- Reasons for exclusion from the PP analysis set

4.2 Exposure to Study Treatment

The total cells infused and total JCAR015-transduced cells infused during each JCAR015 dose, as well the overall total number of cells infused and JCAR015-transduced cells infused, will be listed and summarized using descriptive statistics.

The number (%) of subjects who receive one or two JCAR015 infusions will be summarized. Time to the second JCAR015 infusion (in days) from Day 1 will be provided for subjects who receive two infusions. Reasons for subjects who didn't receive the second JCAR015 infusion will be summarized and listed.

Total on-study follow-up time will be summarized using descriptive statistics. The on-study follow-up time is defined as the end-of-study (EOS) date minus first dose date plus 1, and will be expressed in months using up to one decimal place. If subjects are continuing on study, the last known alive date will be used to impute the EOS date for the purpose of the calculation. The number and percentage of subjects followed for at least 1 day, as well as 2, 4, 6, 8, 10, and 12 months will be summarized. Total follow-up time, including long-term follow-up, will be provided.

4.3 Protocol Deviations

Protocol deviations will be identified by the study team. Significant protocol deviations will be summarized by type of deviation based on the JCAR015-treated and primary efficacy analysis sets. A listing will be provided for all protocol deviations.

5 Baseline Characteristics

5.1 Demographics

Demographics, including sex, race, ethnicity, and age, will be summarized for the primary efficacy analysis and JCAR015-treated analysis sets. Age will be calculated as:

> Age (years) = (date of first JCAR015 infusion – date of birth + 1) / 365.25 (rounded down to an integer)

Number and percentage of subjects with age < 40, ≥40 to < 65, and ≥65 years will also be summarized. A data listing will be presented for date of birth and the above demographic data.

5.2 Baseline and Cytogenetic Characteristics

Other baseline characteristics, including body weight (in kg), height (in cm), body mass index (BMI; in kg/m^2), will be summarized using descriptive statistics for the primary efficacy analysis and JCAR015-treated analysis sets. BMI is calculated as:

> BMI (kg/m^2) = weight / $(height)^2$ (round to nearest 1 decimal point)

Baseline Eastern Cooperative Oncology Group (ECOG) performance status will also be summarized. Subjects' prior response status (i.e., refractory, relapse without transplant, or relapse after transplant), prior HSCT Status (i.e., Yes or No), number of relapses, baseline lymphoblast (%) based on bone marrow per pathologist and per flow cytometry, and estimated marrow cellularity (%) will be summarized. Subjects' cytogenetic characteristics, Philadelphia chromosome status (i.e., positive or negative), Karyotype (i.e., normal or abnormal), florescence in situ hybridization result (i.e., normal or abnormal) will be summarized. To determine a subject's baseline disease status (Group 1 or Group 2), the bone marrow aspirate result will be used by default unless it is not available or is inevaluable (e.g., sample quality is poor due to hemodilution). In such cases, the bone marrow biopsy results will be used.

A by-subject listing of baseline and cytogenetic characteristics will also be provided.

5.3 Medical History

Medical history will be listed for the JCAR015-treated and primary efficacy analysis sets.

6 Efficacy Analysis

6.1 Primary Endpoint and Analysis Methods

6.1.1 Definition of the Primary Efficacy Endpoint

The primary endpoint of the study is the IRC-reviewed ORR after the final JCAR015 infusion. The ORR is defined as the proportion of subjects with a best overall response (BOR) of either CR or CRi. The BOR is the best disease response recorded from the time of the final JCAR015 infusion (i.e., Dose #2 for subjects who complete JCAR015 treatment and Dose #1 for subjects who receive only one dose of JCAR015) until the start of another anticancer therapy. Best response will be assigned according to the following order: CR, CRi, no response or unknown. Subjects with an unknown or missing response will be included in the denominator in calculations of ORR.

In order for the BOR to be categorized as CR or CRi, there must be no clinical evidence of relapse at a minimum of 4 weeks (28 days) after the initial achievement of CR or CRi. Any subject not experiencing a CR or CRi, or having disease recurrence within 4 weeks after achieving a CR or CRi will be treated as a non-responder. Also, any subject with CR or CRi will be treated as a non-responder in the primary analysis of the ORR endpoint if the subject receives another anticancer therapy or HSCT within 4 weeks (28 days) after the initial achievement of CR or CRi (that is, before it is possible to confirm that there is no clinical evidence of recurrence for 4 weeks). Any additional assessments (e.g., bone marrow, cerebrospinal fluid [CSF] assessment by lumbar puncture, central nervous system [CNS] imaging, biopsy, etc.) performed for the purpose of disease response evaluation must also support a response of remission.

6.1.2 Statistical Hypothesis for the Primary Efficacy Endpoint

The study will test the hypothesis that ORR > 20% against the null hypothesis that the ORR ≤ 20% at a 1-sided 2.5% level of significance, powered for ORR = 50%, i.e.,

H_0: ORR ≤ 20% versus H_1: ORR > 20%

6.1.3 Analysis of the Primary Efficacy Endpoint

The ORR will be calculated along with the 2-sided 95% exact Clopper-Pearson confidence intervals. The number and proportion of subjects who were evaluated as CR, CRi, no response, or unknown will also be tabulated.

The primary analysis for ORR will be conducted in the primary efficacy analysis set based on the IRC assessments. The study will be considered successful if the lower bound of the 2-sided 95% confidence interval for ORR is greater than 20% such that the null hypothesis is rejected.

In addition, the following sensitivity analyses will be performed:

- ORR based on the PP analysis set using the same analysis method as described above
- ORR determined by the Investigators based on the primary efficacy analysis set, where the baseline morphological status is determined by the Investigators.

6.2 Analysis Methods for Secondary Endpoint

Secondary endpoint analyses will be performed to further assess the efficacy of JCAR015 by combining data collected in this protocol together with data from the LTFU protocol, where appropriate.

6.2.1 Duration of Remission

Duration of remission (DOR) is defined as the interval from the first documentation of CR or CRi to the earlier date of relapse or death due to ALL. The first documentation of CR or CRi is defined as the latest of all dates of required measurements to establish the response. DOR will be evaluated based on the IRC evaluations for the primary efficacy analysis set who achieve a CR or CRi.

In case a subject 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. A response is considered to be adequate if the assessment was performed and the outcome of assessment was other than "unknown" or "not done". The censoring reason could be:

- Ongoing
- Completed the study
- Discontinued the study
- Received a new anticancer therapy (see below for handling HSCT)
- Experienced an event after missing at least two consecutive scheduled disease assessments

Death due to a reason other than ALL will be considered as a competing risk event to other events of interest (relapse or death due to ALL).

Because HSCT is an important treatment option in responding subjects, it is appropriate to consider the date of HSCT as a censoring date, instead of censoring at the last disease assessment date. However, censoring due to HSCT will overestimate the rate of relapse and therefore may be considered inappropriate for the main analysis when a substantial number of subjects choose to receive HSCT (European Medicines Agency 2010). If a subject undergoes HSCT after achieving a CR or CRi in response to JCAR015, relapse or survival status after HSCT will be collected. In such cases, the date of relapse or death due to ALL after HSCT will be used for the calculation of DOR as a sensitivity analysis.

Another sensitivity analysis will be performed by censoring death due to reasons other than ALL instead of considering it as the competing risk event to other events of interest (relapse or death due to ALL). The proposed analyses for DOR are:

- Competing risk analysis and censoring at time of HSCT
- Competing risk analysis and ignoring HSCT as a censoring event
- Censoring at the last adequate disease assessment and censoring at the time of HSCT
- Censoring at the last adequate disease assessment and ignoring HSCT as a censoring event

The first method (i.e., competing risk analysis and censoring at time of HSCT) will be considered as the primary analysis for DOR.

For the competing risk analyses, the cumulative incidence function (CIF) method will be used to estimate the probability of the event of interest (Cheson 2003, Kim 2007). For the other analyses, the Kaplan-Meier (KM) method will be used to estimate the median DOR along with 95% CI. The estimated percentage of subjects with continued response at 6 months and 12 months will be presented with 95% confidence intervals using the CIF or the KM method, as appropriate.

Duration of molecular remission (per IgH gene sequencing assay) is defined as the interval from the first documentation of an MRD-negative CR or CRi to the earlier date of molecular relapse (appearance of MRD), morphologic relapse, or death due to ALL. The duration of molecular remission will be analyzed using the same methods as DOR described above.

6.2.2 Percentage of Subjects Who Achieve a CR or CRi with No Evidence of MRD in the Bone Marrow

The percentage of subjects with an MRD-negative bone marrow (i.e., below the level of detection for the IgH gene sequencing assay) at the time of achieving CR or CRi will be summarized along with the exact 95% confidence interval.

This analysis will be conducted in the JCAR015-treated and primary efficacy analysis sets.

6.2.3 Relapse-free Survival

Relapse-free survival (RFS) is defined as the interval from the first documentation of CR or CRi to the earlier date of relapse or death due to any cause. The relapse date is defined as the earliest date of all assessments that lead to a relapse. RFS will be evaluated based on the primary efficacy analysis set who achieve a CR or CRi.

If a subject does not experience a relapse or death 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. The censoring reason could be:

- Ongoing
- Completed the study
- Discontinued the study
- Received a new anticancer therapy (also see below for handling HSCT)

- Experienced an event after missing at least two consecutive scheduled disease assessments

Subjects who proceed to HSCT after JCAR015 infusion will be censored at the time of HSCT. A sensitivity analysis will be performed without censoring HSCT. The KM method will be used to estimate 6-months RFS rate, 12-month RFS rate, and median RFS along with the 95% CI.

6.2.4 Event-free Survival

Event-free survival (EFS) is defined as the interval from the date of the first JCAR015 infusion to the earliest of the following events: death from any cause, relapse, or treatment failure (defined as no response and subsequent discontinuation from the study for adverse event, lack of efficacy or progressive disease, or new anticancer therapy). EFS will be evaluated based on the JCAR015-treated and primary efficacy analysis sets.

In case a subject does not have an EFS event prior to data cutoff, EFS will be censored at the date of the last adequate disease assessment on or prior to the earliest censoring event. The censoring reason could be:

- Ongoing
- Lost to follow-up
- Withdrawal of consent
- Received a new anticancer therapy (also see below for handling HSCT)
- Experienced an event after missing at least two consecutive scheduled disease assessments

In the case of treatment failure, the event date will be set to Study Day 1 (European Medicines Agency 2010).

Subjects who proceed to HSCT after JCAR015 infusion will be censored at the time of HSCT. A sensitivity analysis will be performed without censoring for HSCT. The KM method will be used to estimate the 6-month EFS rate, 12-month EFS rate, and median EFS along with the 95% CI.

6.2.5 Overall Survival

Overall survival (OS) is defined as the interval from the date of the first JCAR015 infusion to the date of death due to any reason. The OS analysis will be performed using both the JCAR015-treated and primary efficacy analysis sets and will include all available survival information with long-term follow-up data. Data from surviving subjects will be censored at the last time that the subject is known to be alive. No censoring will be done in the case of HSCT. OS will be evaluated based on the JCAR015-treated and primary efficacy analysis sets.

The KM method will be used to estimate the 6-month OS rate, 12-month OS rate, and median OS along with the 95% CI.

6.2.6 Percentage of Subjects Who Achieve CR or CRi at Month 6 from the Time of the Final JCAR015 Infusion to the Month 6 Post-infusion Assessment

The percentage of subjects who achieve a CR or CRi at Month 6 without HSCT between the final JCAR015 infusion and the Month 6 response assessment will be summarized along with the exact

95% confidence interval. The time of proceeding to HSCT is defined as the time of commencing the conditioning regimen as required for HSCT.

This analysis will be conducted in the primary efficacy analysis set based on the IRC assessments.

6.2.7 Percentage of Subjects Who Achieve a CR or CRi and Then Proceed to HSCT While in Remission within 6 Months after the Final JCAR015 Infusion

The percentage of subjects who achieve a CR or CRi following the final JCAR015 infusion and then proceed to HSCT while in remission by the Month 6 response assessment will be summarized along with the exact 95% confidence interval. The time of proceeding to HSCT is defined as the time of commencing the conditioning regimen as required for HSCT.

This analysis will be conducted in the primary efficacy analysis set based on the IRC assessments.

6.3 Efficacy Subgroup Analysis

Subgroup analysis will be performed on the following variables based on the subject's baseline status:

- Age: < 40, ≥ 40 to < 65, ≥ 65 years
- Prior response status: refractory, relapse without transplant, relapse after transplant
- Philadelphia chromosome status: positive versus negative
- Prior HSCT Status: Yes versus No
- Number of relapses prior to first JCAR015 infusion: One relapse versus two or more relapses
- Sex: Male versus Female
- Race: White vs Other Races
- Ethnicity: Hispanic or Latino versus non-Hispanic or Latino

Subgroup analyses will be performed for all the primary and secondary efficacy endpoints listed in Section 6.1 and Section 6.2, and will only be performed if there are at least five subjects in each subgroup. Some grouping of classes will be considered if there are too few subjects in some subgroups.

7 Safety Analysis

Safety data will be summarized for the JCAR015-treated analysis set.

7.1 Adverse Events

Adverse events will be coded using MedDRA. System organ class (SOC), high-level group term (HLGT), high-level term (HLT), preferred term (PT), and lower-level term (LLT) will be provided in the AE dataset.

The severity of each adverse event will be graded by the Investigator using Common Terminology Criteria for Adverse Events (CTCAE) version 4.03, unless otherwise specified in the protocol. If CTCAE criteria do not exist for a given event, the Investigator should use one of the following: Grade 1 (mild), Grade 2 (moderate), Grade 3 (severe), Grade 4 (life-threatening), or Grade 5 (fatal) to describe the maximum intensity of the adverse event. CRS toxicity will be assessed by the protocol-specified grading scale. The severity of events for which the Investigator does not record severity will be categorized as "missing" for tabular summaries and data listings.

Related AEs are those for which the Investigator selects "Related" to JCAR015 on the AE case report form (CRF). Relatedness will always default to the Investigator's choice, not that of the medical monitor. Events for which the Investigator does not record relationships to JCAR015 will be considered related to JCAR015 for summary purposes. However, by-subject data listings will show the relationship as missing.

Adverse events will be identified and captured as SAEs if they meet the definition for SAE specified in the study protocol.

7.1.1 Definition of Treatment Emergent Adverse Event

A JCAR015 treatment-emergent AE (TEAE) is defined as one or both of the following:

- Any AE that occurs or worsens after the first JCAR015 infusion and up to 30 days after the final JCAR015 infusion
- Any AE leading to JCAR015 discontinuation

Any AE occurring after the initiation of another anticancer therapy will not be considered as a JCAR015 TEAE.

7.1.2 Incomplete Dates

If the onset date of the AE is incomplete, then the month and year (or year alone if the month is not recorded) of onset determines whether an AE is treatment-emergent, as long as the AE stop date is not prior to the first JCAR015 infusion.

All AEs with partial onset or stop dates will be identified and the partial dates will be imputed as follows:

- For AE onset date: If the day and month are missing but the year is available, then the imputed day and month will be 01 Jan or the first dosing date if they have the same year, whichever is later. If the day is missing but the month and year are available, then the

imputed day will be the first day of the month or the first dosing date if they have the same month and year, whichever is later.

- For AE stop date: If the day and month are missing but the year is available, then the imputed day and month will be 31 Dec or the data cut-off date if they have the same year, whichever is earlier. If the day is missing but the month and year are available, then the imputed day will be the last day of the month.

An AE with completely missing onset and stop dates, or with the onset date missing and a stop date later than the first JCAR015 infusion date, will be considered to be treatment-emergent.

The incomplete dates will be included in the raw AE datasets without imputation. The imputed dates will be included in the AE analysis dataset for analysis purposes. A flag variable in the AE analysis dataset will be created to identify the imputed dates.

7.1.3 Summary of Adverse Events

All AEs will be listed. The focus of AE summarization will be on JCAR015 TEAEs.

A brief high-level summary of TEAEs will be provided by the number and percentage of subjects who had the following: any AE, any Grade ≥ 3 AE, any JCAR015-related AE, any JCAR015-related Grade ≥ 3 AE, any SAE, any JCAR015-related SAE, any AE leading to discontinuation of JCAR015, and any AE leading to death.

Adverse event summaries will provide the number and percentage of subjects with TEAEs by SOC and PT, based on the JCAR015-treated analysis set as follows:

- All AEs
- All AEs by severity grade
- Grade ≥ 3 AEs
- All JCAR015-related AEs
- JCAR015-related Grade ≥ 3 AEs
- All SAEs
- All JCAR015-related SAEs
- All AEs leading to JCAR015 discontinuation
- All AEs leading to death
- All AEs occurring after 30 days following the final infusion of JCAR015 (i.e., non-treatment-emergent AEs)
- All SAEs recorded between screening and the first infusion of JCAR015 (i.e., non-treatment-emergent SAEs)

Multiple occurrence of the same events will be counted only once per subject in each summary. Adverse events will be summarized and listed first by SOC in descending order of incidence and then by PT (within each SOC), also in descending order of incidence. In summaries by severity grade, the most severe grade will be used for those AEs that occur more than once in an individual subject during the study.

In addition to the above summary tables, all TEAEs and SAEs will be summarized by SOC only and by PT only in order of descending incidence.

TEAEs with onset date between the first JCAR015 infusion and the day prior to second infusion, or EOS visit if a subject does not receive a second treatment, will be summarized.

TEAEs with onset date between the second JCAR015 infusion and EOS will be summarized separately. Unless stated otherwise, the number of subjects who receive two infusions will be used as the N or denominator as applicable.

Furthermore, all AEs reported on or after the date of initiation of the lymphodepleting chemotherapy regimen, up to 30 days after the final JCAR015 infusion, will be summarized. AEs and SAEs related to lymphodepleting chemotherapy will be flagged.

In addition to the summaries described above, data listings will be provided for the following:

- All AEs (with a variable indicating whether the event is treatment-emergent)
- SAEs
- AEs leading to death
- AEs leading to discontinuation of JCAR015

7.1.4 Adverse Events of Special Interest

The following TEAEs are considered to be of special interest:

- **Syndromes**, including the following MedDRA PT:
  - cytokine release syndrome
  - histiocytosis haematophagic
  - tumour lysis syndrome
- **CRS symptoms**, including the following MedDRA PT:
  - pyrexia
  - myalgia
  - hypotension
  - dyspnoea
  - tachypnoea
  - capillary leak syndrome
  - hypoxia
  - organ failure
  - acute respiratory distress syndrome
- **Neurological toxicity**, including the following MedDRA PT:
  - mental status changes
  - delirium
  - confusional state
  - disorientation
  - aphasia
  - seizure
  - seizure-like phenomena
  - encephalopathy
  - tremor
  - myoclonus

  - lethargy
  - depressed level of consciousness
- **Macrophage activation syndrome**, including the following MedDRA PT:
  - hepatosplenomegaly
  - lymphadenopathy
  - pancytopenia
  - disseminated intravascular coagulation
  - hypertriglyceridaemia
  - hyperferritinaemia
  - hypofibrinogenaemia
  - pyrexia
  - hepatic enzyme increased
  - hyponatraemia
  - prothrombin time prolonged
  - activated partial thromboplastin time prolonged
  - fibrin degradation products increased
- **Tumor lysis syndrome**, including the following MedDRA PT:
  - hyperkalaemia
  - hyperphosphataemia
  - hyperuricaemia
  - hypocalcaemia
- **Hemophagocytic lymphohistiocytosis (HLH) symptoms**, including the following MedDRA PT and HLT:
  - splenomegaly (PT)
  - haemolysis (PT)
  - disseminated intravascular coagulation (PT)
  - blood triglycerides increased (PT)
  - serum ferritin increased (PT)
  - Marrow depression and hypoplastic anaemias (HLT)
- **Organ dysfunction**, including the following MedDRA HLT:
  - hepatic enzymes and function abnormalities
  - renal failure and impairment
  - confusion and disorientation
- **Allergic reaction** using the MedDRA preferred term “infusion related reaction”

The search terms of the AESI may be updated prior to reporting. The AESI will be summarized by group term and search terms.

Time to onset and resolution of first event will be summarized. Both KM estimates and descriptive summary statistics will be provided. Time to onset of first event is defined as the time from Day 1 to the start date of first occurrence of the event, i.e., time in days is calculated as (start date of first occurrence of the event) – Day 1 +1. Time to resolution of first event is defined as the time from the start date of the first occurrence of the event to its resolution. In the absence of an event, subjects will be censored at the earliest of the following dates: date of EOS, date of death, or date last known alive on study.

7.2 Laboratory Evaluations

All laboratory data will be listed with a variable indicating whether the result is treatment-emergent (i.e., measured after the first JCAR015 infusion and up to 30 days following the final JCAR015 infusion). The focus of the laboratory data summarization will be on the JCAR015 treatment-emergent laboratory abnormalities (see Section 7.2.2.1) using the JCAR015-treated analysis set.

7.2.1 Numeric Laboratory Results

Summaries of laboratory data will be based on observed data and will be reported using conventional units. Baseline, raw values, and changes from baseline will be summarized using descriptive statistics for each laboratory test specified in the study protocol.

Median (Q1, Q3) values for each laboratory parameter will be plotted over time using a box plot. Mean ± SE of the observed values and changes from baseline for each laboratory parameter will be plotted over time using line plots.

In the case of multiple values in an analysis window, data will be selected for analysis as described in Section 3.3.2.

Laboratory results with a collection date between the first JCAR015 infusion and the day prior to second infusion, or the EOS visit if a subject does not receive a second treatment, will be summarized.

Laboratory results with a collection date between the second JCAR015 infusion and the EOS visit will be summarized separately. Unless stated otherwise, the number of subjects who receive two infusions will be used as the N or denominator as applicable.

Furthermore, all above analyses will be repeated by using the last observation collected prior to or on the date of initiation of the lymphodepleting chemotherapy regimen as baseline.

7.2.2 Graded Laboratory Values

Applicable hematological and serum biochemistry laboratory data will be programmatically graded according to CTCAE, Version 4.03 severity grade [grade laboratory results as Grade 0, mild (Grade 1), moderate (Grade 2), severe (Grade 3), or life threatening (Grade 4)]. Grade 0 includes all non-missing values that do not meet criteria for an abnormality of at least Grade 1. Some laboratory tests have criteria for both increased and decreased levels; analyses for each direction (i.e., increased, decreased) will be presented separately.

7.2.2.1 Definition of Treatment-Emergent Laboratory Abnormalities

A JCAR015 treatment-emergent laboratory abnormality is defined as an abnormality that, compared to baseline, worsens by ≥ 1 grade after the first JCAR015 infusion and up to 30 days after the final JCAR015 infusion. Any laboratory abnormality occurring after the initiation of another anticancer therapy will not be considered as JCAR015 treatment-emergent. If baseline data are missing, then any graded abnormality (i.e., an abnormality that is Grade ≥1 in severity) will be considered treatment-emergent.

7.2.2.2 Summaries of Laboratory Abnormalities

All laboratory data will be listed. The focus of laboratory data summarization will be on JCAR015 treatment-emergent laboratory abnormalities.

Summaries (number and percentage of subjects) of baseline, post-baseline at each visit, and worst post-baseline treatment-emergent laboratory abnormalities will be provided. Subjects will be categorized according to the most severe abnormality grade. All summaries of laboratory abnormalities will be based on the JCAR015-treated analysis set.

Shift tables will be presented by showing the change in CTCAE severity grade from baseline to each post-baseline visit and to the maximum post-baseline grade. For parameters for which a CTCAE severity scale does not exist, shift tables will be presented showing change in results from the baseline value (low, normal, and high) to each post-baseline visit and to the worst post-baseline value (low, normal, and high).

Furthermore, all above analyses will be repeated by using the last observation collected prior to or on the date of initiation of the lymphodepleting chemotherapy regimen as baseline.

7.3 Leukemia History

Time since initial B-call ALL diagnosis (in years) to first JCAR015 infusion will be calculated.

Prior ALL therapies include all therapies received prior to the first JCAR015 infusion (including cytoreductive therapies in Part A of this study). Number and type of prior therapies will be summarized using descriptive statistics based on the JCAR015-treated and primary efficacy analysis sets.

Number (%) of subjects who received 1, 2, 3 ... prior therapies (i.e., prior lines of therapy), and number (%) of subjects who received prior allogeneic HSCT and their donor type will be provided. Descriptive statistics will be provided for the last therapy subjects received prior to the first JCAR015 infusion.

7.4 Prior and Concomitant Medications

Concomitant medications will be coded using the World Health Organization Drug Dictionary (WHO-DD) into Anatomical-Therapeutic-Chemical classification (ATC) codes.

Concomitant medications are defined as any medications meeting the following criteria:

- Starting within 90 days after the last infusion of JCAR015
- Starting before and continuing after the first JCAR015 infusion

The handling method for incomplete dates used for AE summaries will be used for concomitant medication summaries (see Section 7.1.2).

Prior medications are defined as any medications that were stopped before the first dose of lymphodepleting chemotherapy.

Number (%) of subjects who used prior and concomitant medications will be presented in tabular form by preferred drug name based on the JCAR015-treated analysis set. The summary tables will

be sorted by descending frequency. Subjects will only be counted once for multiple drug use (by preferred drug name) per subject.

All prior and concomitant medications (other than per-protocol study drugs) will be provided in a by-subject listing sorted by subject number and administration date in chronological order. A variable will be included to flag the prior and concomitant medication. A separate listing will be presented for transfusions during the study.

7.5 Immunogenicity Data

Humoral immunogenicity assessment will include prevalence of immunogenicity (subjects with pre-existing antibodies that bind to JCAR015), incidence of immunogenicity (subjects with treatment-induced or treatment-boosted antibodies that bind to JCAR015), together with antibody titers. Data will be further fractionated to determine the proportion of subjects who make transient versus sustained antibody responses. The assay for humoral immunogenicity will be an ELISA-based assay, detecting antibodies that bind to any potential epitope on the extracellular domain of the JCAR015 protein. T cell-mediated responses to JCAR015 will be evaluated by stimulating PBMC with overlapping peptides from JCAR015 and measuring the cytotoxic T lymphocyte response using an ELISpot assay.

7.6 Physical Exam (Including Vital Signs)

A physical examination will be conducted at specified timepoints indicated in the study protocol. The exam will also include an assessment of the presence of extramedullary disease (i.e., hepatomegaly, splenomegaly, skin/gum infiltration, testicular masses). Extramedullary involvement will be assessed at screening and at each visit for response assessment. For visits where disease response is assessed, the assessment results will be recorded on the applicable CRF. Significant physical exam findings identified during screening for Part A should be recorded in the medical history CRF. Physical examination results will be summarized at each visit.

Vital signs (blood pressure [mm Hg], body temperature (ºC), respirations [breaths/min], heart rate [beats/min], and oxygen saturation via pulse oximetry), and weight (kg) will be collected at specified visits according to protocol. Vital signs at each visit and change from baseline at each visit will be summarized for the JCAR015-treated analysis set using descriptive statistics.

In the case of multiple values in an analysis window, data will be selected for analysis as described in Section 3.3.2.

7.7 Safety Subgroup Analysis

Subgroup analysis will be performed for key safety summaries based on the following variables:

- Disease burden: morphologic disease at the time of JCAR015 treatment (i.e., Group 1) versus no morphologic disease at the time of JCAR015 treatment (i.e., Group 2)
- Age: < 40, ≥ 40 to < 65, ≥ 65 years
- Prior response status: refractory, relapse without transplant, relapse after transplant
- Philadelphia chromosome status: positive versus negative
- Prior HSCT Status: Yes versus No
- Number of relapses prior to the first JCAR015 infusion: one relapse versus two or more relapses
- Sex: Male versus Female
- Race: White versus Other Races
- Ethnicity: Hispanic or Latino versus non-Hispanic or Latino
- Lymphodepleting chemotherapy regimen: cyclophosphamide versus fludarabine plus cyclophosphamide

Subgroup analyses will be based on the JCAR015-treated analysis set and will only be performed if there are at least five subjects in each subgroup. Some grouping of classes will be considered if there are too few subjects in some subgroups.

8 Pharmacokinetic Analysis

Assays for the analysis of various cellular PK parameters include a quantitative polymerase chain reaction (qPCR) assay to detect JCAR015 cells (transgene copies/µg of DNA) in peripheral blood, and bone marrow, as well as flow cytometry analysis to detect JCAR015 cells in peripheral blood. Other tissues (e.g., CSF) may be analyzed for the presence of JCAR015 cells using qPCR. Details of the sample collection procedures for these assays are provided in the 015001 Central Laboratory Manual.

The assessment of JCAR015 cellular PK will measure the relative percent of peripheral white blood cells and bone marrow mononuclear cells expressing the JCAR015 transgene, and, using a CAR detection antibody, the number of CD4 and CD8 positive T cells that express the JCAR015 transgene protein per microliter of blood.

For all subjects treated with JCAR015, the in vivo PK profile of JCAR015 cells in target tissues (blood, bone marrow, and CSF) will be characterized, including C_{max}, T_{max}, AUC, and other relevant PK parameters. The maximum extent of expansion of JCAR015 in the blood will be determined ($C_{max}/C_{post\text{-}dose\ Day\ 2}$), along with the persistence of JCAR015 in the blood and bone marrow, based on flow cytometry assay (time above threshold JCAR015 level) and the qPCR assay (time above lower limit of quantification).

The following PK parameters will be displayed graphically where possible: qPCR-based JCAR015 concentration versus time in peripheral blood, bone marrow, and CSF; flow cytometry-based JCAR015 concentration versus time in peripheral blood.

The following PK parameters, along with other relevant PK parameters, will be estimated from the individual concentration-time profiles using a non-compartmental analysis approach: $AUC_{0\text{-}Tmax}$, $AUC_{Tmax\text{-}56d}$, C_{max}, T_{max}, $T_{1/2}$, $C_{max}/C_{post\text{-}dose\ Day\ 2}$, and time above lower limit of quantification. All concentrations below the limit of detection or quantitation, or missing data, will be labeled as such in the concentration data listings. Concentrations below the limit of detection will be treated as zero in summary statistics.

Descriptive statistics for PK parameters will be categorized by clinical response and will include mean, standard deviation, coefficient of variation, minimum, and maximum. Ranges of values may be presented for selected variables. Median values and ranges will be used for T_{max} as it is typically evaluated by a non-parametric method.

Finally, the relationship between the use of steroids, occurrence of immunogenicity, and other relevant covariates and cellular PK will be explored.

8.1 Pharmacokinetic Subgroup Analysis

Subgroup analysis will be performed for key PK summaries based on the following variables:

- Disease burden: morphologic disease at the time of JCAR015 treatment (i.e., Group 1) versus no morphologic disease at the time of JCAR015 treatment (i.e., Group 2)
- Sex: male versus female
- Cellular and/or humoral immunogenicity: present (positive) or absent (negative)
- Race: White vs Other Races
- Ethnicity: Hispanic or Latino versus non-Hispanic or Latino
- Lymphodepleting chemotherapy regimen: cyclophosphamide versus fludarabine plus cyclophosphamide

Subgroup analyses will be based on the PK analysis set and will only be performed if there are at least five subjects in each subgroup. Some grouping of classes will be considered if there are too few subjects in some subgroups.

9 Biomarker Analysis

A separate biomarker analysis plan will be prepared to detail pharmacodynamics and biomarker analyses.

10 Exploratory Analysis

10.1 Peripheral Blood B-cell Count

Peripheral blood will be analyzed at study assessment timepoints for the number of non-malignant B cells, leukemic blasts, CAR+ T-cell number and T-cell subset distribution, and recovery of normal B cells, or disappearance of CAR+ T cells, whichever occurs later. Baseline, raw values at each visit, and change from baseline will be summarized.

10.2 Severe Cytokine Release Syndrome

Severe CRS is defined as Grade ≥ 3 CRS based on the protocol-specified grading scale. Time to the occurrence of first sCRS and duration of sCRS will be summarized as described in Section 7.1.4.

The time to resolution from the initiation of anti-cytokine therapy will be analyzed using descriptive statistics and the KM method.

Changes in inflammatory markers (including C-reactive protein and serum ferritin), other cytokines, and the relationship with sCRS will be explored using multivariate analysis. Peak levels of interleukin-6, ferritin levels, and C-reactive protein levels will be compared between subjects who develop and do not develop sCRS using Wilcoxon rank sum test.

The relationship of baseline disease burden, clinical response, and PK/PD parameters may be explored using multivariate and univariate analysis method. Wilcoxon rank sum test will be used to detect associations between biomarkers and CRS severity. Median (Q1, Q3) values of blasts in the bone marrow at baseline will be plotted using box plots for subjects who develop and do not develop sCRS.

In vivo JCAR015 cell expansion occurring after both the first and the second JCAR015 infusions will be plotted over time for subjects with sCRS.

10.3 Severe Neurological Toxicity

Severe neurological toxicity is defined as Grade ≥ 3 neurotoxicity. Time to the occurrence of first symptoms consistent with severe neurotoxicity, and duration of neurotoxicity symptoms, will be summarized as described in Section 7.1.4. The time to resolution from the initiation of treatment will be analyzed using descriptive statistics and KM method.

Neurocognitive assessments will be performed using the Mini Mental State Examination (MMSE). This is a multi-item instrument that examines orientation, registration, attention, calculation, recall, visuospatial abilities, and language. The maximum score is 30, with higher scores indicating better cognitive function. MMSE total score change from baseline at each timepoint collected will be summarized for all subjects, separately for subjects who develop severe neurotoxicity and those who do not.

Neuropsychological assessment results and change from baseline at each timepoint collected will be summarized for all subjects and separately for subjects who develop severe neurotoxicity and those who do not.

10.4 Health-Related Quality of Life

The health-related quality-of-life (HRQL) assessments are based on the FACT-Leu questionnaire (Cella 2012). The FACT-Leu questionnaire includes subscales for physical well-being (PWB, 7 items); social/family well-being (SWB, 7 items); emotional well-being (EWB, 6 items); functional well-being (FWB, 7 items); and additional concerns (Leukemia-Specific Subscale, LeuS, 17 items). The FACT-Leu scoring guide identifies those negatively stated items that must be reversed before being added to obtain subscale totals. Negatively stated items are reversed by subtracting the response from "4". After reversing proper items, all subscale items are summed to a total, which is the subscale score. For all FACT-Leu scales and symptom indices, a higher score is associated with a better quality of life. The scores in the following items need to be reversed:

- Physical well-being: all individual items
- Social/family well-being: none
- Emotional well-being: five individual items (except for the GE2)
- Functional well-being: none
- Additional concerns: all individual items (except C6 and An7)

The subscale scores will be a summation of all individual item scores within each subscale. If ≤ 50% of item scores are missing, the subscale score will be calculated by multiplying the sum of the item scores by the number of items in the subscale, then dividing by the number of non-missing item scores.

Prorated subscale score = [sum of item scores] × [N of items in subscale] / [N of items answered]

The following composite scores will be derived from the above subscale total scores:

- Trial Outcome Index (TOI, score range: 0-124) = PWB + FWB + LeuS
- Functional Assessment of Cancer Therapy-General (FACT-G total score, score range: 0-108) = PWB + SFWB + EWB + FWB
- FACT-Leu Total Score (score range: 0-176) = PWB + SFWB + EWB + FWB + LeuS

The total scores will be set to missing if 20% or more of the included items are missing or any of the component subscales are missing. TOI scores are set to missing if any of the component subscales are missing.

The HRQL analyses will be based on the primary efficacy analysis and JCAR015-treated analysis sets. The mean and change from baseline for each subsequent assessment will be summarized for the subscale and composite scores. The best change from baseline during the study, defined as the highest positive value among all post-baseline visits minus the baseline value, will also be summarized. The mean change from baseline ± SE over time will be plotted. The minimally important differences (MID) have been identified for the different subscales (Trask 2012, Trask 2013): PWB 2-3 points; SFWB, not available; EWB, 2 points; FWB, 2-3 points; FACT-G, 3-7 points; FACT-LeuS, 4-7 points; TOI, 5-6 points; and FACT-Leu Total, 6-12 points.

FACT-Leu compliance rates at each time point will be calculated. Compliance is defined as having answered at least one question at an assessment timepoint. Compliance rate at each assessment timepoint for each questionnaire will be calculated as the number of subjects who completed at least one question divided by the total number of subjects available at that assessment timepoint.

A data listing for each individual item, the subscale scores, and the composite scores will be presented for each subject at each visit.

10.5 ECOG Performance Status

The mean and change from baseline for each subsequent assessment will be summarized for the ECOG performance status based on the primary efficacy and JCAR015-treated analysis sets.

10.6 Time to Response

Time to response is defined as the interval from the final JCAR015 infusion to the first documentation of CR or CRi. Time to response will be summarized using IRC assessments based on the primary efficacy analysis set who achieve a CR or CRi.

11 Changes from Protocol-Specified Analyses

There are no deviations from the protocol-specified analyses.

12 Revisions to the SAP

This SAP was the final version and the only one used for study analysis.

13 References

1. Cella D, Jensen SE, Webster K, et al. Measuring health-related quality of life in leukemia: the Functional Assessment of Cancer Therapy--Leukemia (FACT-Leu) questionnaire. Value Health 2012;15(8):1051-1058.
2. Center for Drug Evaluation and Research. Statistical Review and Evaluation for Blincyto (blinatumomab). Biologics License Application #BLA 125557/00. 2014.
3. Cheson BD, Bennett JM, Kopecky KJ, et al. Revised recommendations of the International Working Group for Diagnosis, Standardization of Response Criteria, Treatment Outcomes, and Reporting Standards for Therapeutic Trials in Acute Myeloid Leukemia. J Clin Oncol 2003;21(24):4642-4649.
4. European Medicines Agency. Appendix 2 to the Guideline on the evaluation of Anticancer Medicinal Products in Man (CPMP/EWP/205/95 Rev. 3) on Confirmatory studies in Haematological Malignancies. 2010.
5. Kim HT. Cumulative incidence in competing risks data and competing risks regression analysis. Clin Cancer Res 2007;13(2 Pt 1):559-565.
6. National Comprehensive Cancer Network. NCCN Clinical Practice Guidelines in Oncology. Acute Lymphoblastic Leukemia. v1. 2014.
7. Schemper M and Smith TL. A note on quantifying follow-up in studies of failure time. Control Clin Trials 1996;17(4):343-346.
8. Thall PF and Simon R. Practical Bayesian guidelines for phase IIB clinical trials. Biometrics 1994;50(2):337-349.
9. Trask PC, Cella D, Besson N, et al. Health-related quality of life of bosutinib (SKI-606) in imatinib-resistant or imatinib-intolerant chronic phase chronic myeloid leukemia. Leuk Res 2012;36(4):438-442.
10. Trask PC, Cella D, Powell C, et al. Health-related quality of life in chronic myeloid leukemia. Leuk Res 2013;37(1):9-13.
11. Yao B, Zhu L, Jiang Q, et al. Safety monitoring in clinical trials. Pharmaceutics 2013;5(1):94-106.