STATISTICAL ANALYSIS PLAN

An Evaluation of the Toxicity and Therapeutic Effects of Epstein-Barr Virus-Immune T Lymphocytes Derived from a Normal HLA-Compatible or Haplotype-Matched Donor in the Treatment of EBV-Associated Lymphoproliferative Diseases or Malignancies and Patients with Detectable Circulating Levels of EBV DNA who are at High Risk for EBV-Associated Lymphoproliferative Diseases

Protocol Number:	95-024
Version:	1.0
Date:	07 JULY 2019
Authors:	PPD , PhD

TABLE OF CONTENTS

LIST OF ABBREVIATIONS AND DEFINITIONS OF TERMS

Abbreviation	Definition
AID	acquired immunodeficiency
AIDS	acquired immune deficiency syndrome
AE	adverse event
CI	confidence interval
CR	complete response
CTL	cytotoxic T lymphocyte
CTCAE	Common Toxicity Criteria for Adverse Events
DLT	dose-limiting toxicity
DNA	deoxyribonucleic acid
DOR	duration of response
DRR	durable response rate
EBV	Epstein-Barr virus
EBV-CTLs	EBV cytotoxic T lymphocytes
EBV-PTLD	EBV post-transplant lymphoproliferative disorder
EBV-LPD	EBV-associated lymphoproliferative disease
LMS	leiomyosarcoma
LLOQ	lower limit of quantification
GvHD	graft-versus-host disease
HCT	hematopoietic (stem) cell transplant
HLA	human leukocyte antigen
IPD	important protocol deviations
NE	not evaluable
NPC	nasopharyngeal carcinoma
ORR	objective response rate
OS	overall survival
PD	progressive disease
PFS	progression-free survival
PID	primary immunodeficiency
PR	partial response
PTLD	post-transplant lymphoproliferative disorder
R/R	refractory/relapsed
SAE	serious adverse event
SAP	statistical analysis plan

Abbreviation	Definition
SD	stable disease [used in context of response assessment] standard deviation [used in context of analysis methods]
SOC	system organ class
SOT	solid organ transplant
SOP	standard operating procedure
tab-cel	tabelecleucel
TESAE	treatment-emergent serious adverse event
TTP	time to progression
TTR	time to response
ULOQ	upper limit of quantification

1 INTRODUCTION

The purpose of this statistical analysis plan (SAP) is to provide details of the statistical analyses that have been outlined within version A23 of the tabelecleucel (ATA129; tab-cel®) Study 95-024 protocol, dated 14 March 2018.

2 STUDY OVERVIEW

2.1 Study Objectives

- To evaluate in a phase 1/2 dose-escalating trial both the toxicities and therapeutic potential of adoptive immunotherapy with Epstein-Barr virus (EBV)-specific cytotoxic T lymphocytes (EBV-CTLs) derived from human leukocyte antigen (HLA)-histocompatible or at least HLA-haplotype-matched related donors in the treatment of EBV-induced lymphomas or other EBV-associated malignancies in severely immunocompromised hosts and organ allograft recipients who are at high risk for this complication, and to complete a single selected dose level phase 2 extension of this study to identify the probability of achieving a complete response (CR) of EBV lymphoma with EBV-specific T-cell therapy in allogenic hematopoietic stem cell transplant (HCT) recipients and immunodeficient patients.
- To evaluate the in vivo biodistribution, expansion and duration of engraftment of successive doses of transferred EBV-reactive lymphocytes within immunocompromised histocompatible or HLA-haplotype matched hosts afflicted with EBV-associated lymphoproliferative diseases and to correlate these findings with the diseased hosts' T-cell populations, general immune status and capacity to generate allospecific antidonor response.
- To determine the incidence, kinetics and durability of pathological and/or clinical responses of EBV-induced lymphomas to treatment with infusions of EBV-CTLs derived from histocompatible EBV-seropositive related donors.

2.2 Study Design

This is a dose-escalation, phase 1/2 clinical study of treatment of patients with EBV-associated lymphoproliferative diseases or malignancies. Each subject consenting to participate in this study will be treated with in-vitro expanded third-party-derived at least 2 HLA alleles matched EBV-CTLs, transplant donor-derived HLA-compatible EBV-CTLs, and/or autologous EBV-CTLs. Based on the expected risk of graft-versus-host disease (GvHD) due to CTL infusion, the study population was divided into the following 2 groups:

Group 1: subjects at low risk of GvHD, which included subjects with HCT, severe congenital immunodeficiency, or those with antineoplastic drug-induced immunodeficiency

Group 2: subjects with high risk of GvHD, which included solid organ transplant (SOT) recipients, acquired immunodeficiency syndrome (AIDS) patients, and EBV-associated malignancies

As specified in the protocol, the initial planned dose-escalation scheme was not followed, and the dose limiting toxicity (DLT) data were not collected. Therefore, DLTs will be not be summarized.

2.3 Sample Size

The initial rationale for the sample size of this study was the following: For subjects in group 1, the phase 2 portion of the study was to accrue a minimum of 20 subjects and a maximum of 47 subjects, using a Simon two-stage design. Enrollment of 20 subjects is needed in the first stage. If 8 or more subjects achieve CR then we 27 additional subjects will be enrolled, for a total of 47 subjects. If at least 20 subjects achieve CR then this regimen will be recommended for further evaluation. This design has 90% power to distinguish between CR of 35% and 55% with a type I error of 10%.

2.4 Source of Study Treatment

As specified in Section 2.2, study treatment administered to subjects in this study included the following 3 types of EBV-CTL products.

2.4.1 Third-party Cytotoxic T Lymphocytes

Tabelecleucel, or third-party EBV-CTLs, refer to EBV-targeted CTLs generated from high resolution HLA-typed, EBV-seropositive, unrelated third-party donor T cells: ie, CTLs generated from a donor other than the subject's original allogeneic SOT or HCT transplant donor.

In this study, subjects could change to tabelecleucel with a different HLA restriction from a different donor either due to a lack of efficacy from the current cell line or insufficient supply.

Only the change to tabelecleucel with a different HLA restriction due to lack of efficacy is considered as treatment switch ("Restriction Switch")

2.4.2 Transplant donor derived Cytotoxic T Lymphocytes

Transplant donor-derived EBV-CTLs refers to T cells generated from the subject's original allogeneic SOT or HCT donor.

2.4.3 Autologous Cytotoxic T Lymphocytes

Autologous EBV-CTLs refers to T cells generated from the subjects themselves.

3 STUDY ENDPOINTS

Response-related endpoints were not specified in the protocol and are, therefore, defined in this SAP. The efficacy analyses of these response-related endpoints will be based on the disease assessments by the investigator.

3.1 Primary Endpoint

- Objective response rate (ORR)

3.2 Secondary Endpoints

- Overall survival (OS)
- Duration of response (DOR)
- Progression-free survival (PFS)
- Durable response rate (DRR)
- Time to progression (TTP)
- Time to response (TTR)

4 ANALYSIS SETS AND SUBGROUPS

In general, summary tables will include only the subjects in the specified analysis set or subgroup, and cohorts will be presented as columns.

4.1 Full Analysis Sets

The Full analysis set is defined as all subjects who received at least 1 dose of study treatment. All efficacy and safety analyses will use the full analysis set except for the endpoints of DOR and TTR, which will be conducted on the subjects in the full analysis set with an objective response (CR or partial response [PR]).

4.2 Subject Cohorts

Analysis cohorts are defined based on subjects' disease background, prior EBV^+ disease therapy, and source of study treatment received (as defined in Section 2.4). The cohort definitions are specified in Table 1. With respect to prior EBV^+ disease therapy, subjects with prior rituximab therapy data reported will be assigned to the relevant refractory/relapsed (R/R) rituximab cohort; subjects with no prior rituximab therapy data reported will be assigned to the rituximab-naïve cohort.

All analyses will be presented by the cohorts and combined cohorts defined in this SAP, not by the groups defined in the protocol. Cohorts to be used in the analysis depend on the data availability.

Table 1: Cohorts

Cohort Number	Cohort Name	Subjects Enrolled in The Study
1	Tab-cel HCT EBV$^+$ PTLD R/R Rituximab	Y
2	Tab-cel HCT EBV$^+$ PTLD Rituximab Naïve	Y
3	Tab-cel SOT EBV$^+$ PTLD R/R Rituximab	N
4	Tab-cel SOT EBV$^+$ PTLD R/R Rituximab + Chemo	Y
5	Tab-cel AID LPD	Y
6	Tab-cel PID LPD	Y
7	Tab-cel Lymphoma	N
31	Tab-cel LMS	Y
32	Tab-cel NPC	Y
61	Tab-cel Other Solid Tumor	N
71	Tab-cel Viremia	N
101	EBV-CTL (Transplant donor derived) HCT EBV$^+$ PTLD R/R Rituximab	Y
102	EBV-CTL (Transplant donor derived) HCT EBV$^+$ PTLD Rituximab Naïve	Y
103	EBV-CTL (Transplant donor derived) SOT EBV$^+$ PTLD	Y
106	EBV-CTL (Transplant donor derived) PID LPD	Y
107	EBV-CTL (Transplant donor derived) Lymphoma	Y
171	EBV-CTL (Transplant donor derived) Viremia	Y
201	Autologous	Y

Abbreviations: AID, acquired immunodeficiency; chemo, chemotherapy; CTL cytotoxic T lymphocyte; EBV, Epstein-Barr virus; HCT, hematopoietic cell transplant; LPD, lymphoproliferative disorder; N, no; NPC, nasopharyngeal carcinoma; PID primary immunodeficiency; post-transplant lymphoproliferative disorder; R/R, refractory/relapsed; SOT, solid organ transplant; tab-cel, tabelecleucel; Y, yes

Disposition, demographics, baseline characteristics, exposure and safety will be summarized as in the following headers:

Table a: HCT Cohorts:

Tab-cel HCT EBV+ PTLD			EBV-CTL (Transplant donor derived) HCT EBV+ PTLD			Total
R/R Rituximab (cohort 1)	Rituximab Naïve (cohort 2)	Total (cohort 1+2)	R/R Rituximab (cohort 101)	Rituximab Naïve (cohort 102)	Total (cohort 101+102)	(cohort 1, 2, 101, 102)

Table b: *SOT Cohorts:*

Tab-cel SOT EBV+ PTLD R/R Rituximab+Chemo (cohort 4)	EBV-CTL (Transplant donor derived) SOT EBV+ PTLD (cohort 103)	Total (cohort 4+103)

Table c: *Other cohorts and All cohorts:*

Other EBV+ Disorders with ID				HCT/SOT EBV+ PTLD	EBV+ Malignancy without ID			
AID LPD (cohort 5)	PID LPD (cohort 6+106)	Viremia (cohort 171)	LMS (cohort 31)	Total (cohort 5, 6+106, 171, 31)	(cohort 1,2,101,102, 4, 103)	EBV-CTL (Transplant donor derived) Lymphoma (Cohort 107)	Tab-cel NPC (cohort 32)	Total (cohort 107+32)	Total (all cohorts)

Table d: *Other cohorts and All cohorts – tabelecleucel only:*

Other EBV+ Disorders with ID				HCT/SOT EBV+ PTLD			EBV+ Malignancy without ID	
AID LPD (cohort 5)	PID LPD (cohort 6)	LMS (cohort 31)	Total (cohort 1,2, 101,102, 4, 103)	HCT EBV+ PTLD (Cohort 1+2)	SOT EBV+ PTLD (Cohort 4)	Total (cohort 1,2,4)	Tab-cel NPC (cohort 32)	Total (all tab-cel cohorts)

Efficacy will be summarized as in the following headers:

Table a: HCT Cohorts:

Tab-cel HCT EBV+ PTLD			EBV-CTL (Transplant donor derived) HCT EBV+ PTLD		
R/R Rituximab (cohort 1)	Rituximab Naïve (cohort 2)	Total (cohort 1+2)	R/R Rituximab (cohort 101)	Rituximab Naïve (cohort 102)	Total (cohort 101+102)

Table b: *SOT Cohorts:*

Tab-cel SOT EBV+ PTLD R/R Rituximab+Chemo (cohort 4)	EBV-CTL (Transplant donor derived) SOT EBV+ PTLD (cohort 103)

Table c: *Other cohorts and All cohorts:*

Other EBV+ Disorders with ID				EBV+ Malignancy without ID	
AID LPD (cohort 5)	PID LPD (cohort 6+106)	Viremia (cohort 171)	LMS (cohort 31)	EBV-CTL (Transplant donor derived) Lymphoma (Cohort 107)	Tab-cel NPC (cohort 32)

Table d: *Other cohorts and All cohorts – tabelecleucel only:*

Other EBV+ Disorders with ID			EBV+ Malignancy without ID
AID LPD (cohort 5)	PID LPD (cohort 6)	LMS (cohort 31)	Tab-cel NPC (cohort 32)

Cohort Assignment for Subjects with > 1 Source of Study Treatment

Some subjects may receive EBV-CTLs from > 1 donor source type (Section 2.4) during the study. As a general rule for efficacy analysis, any transplant donor-derived or autologous EBV-CTL treatment received after the initiation of tabelecleucel will be considered as other anti-underlying EBV+ disease therapy. For such subjects, the efficacy analysis may be censored due to the initiation of this therapy in sensitivity analysis.

The efficacy and safety data for a subject who starts with tabelecleucel and then switches to either transplant donor-derived or autologous EBV-CTLs will be presented under the relevant tabelecleucel cohort.

Each subject who first receives an EBV-CTL product other than tabelecleucel and then switches to a different non-tabelecleucel EBV-CTL product (eg, transplant donor-derived followed by autologous) will be considered as 2 independent subjects for the purposes of analysis. The efficacy and safety data for such a subject will be assigned to the relevant study treatment source cohort based on the initial dose of each product type.

A subject who first receives an EBV-CTL product other than tabelecleucel, switches to tabelecleucel, and then switches to a non-tabelecleucel EBV-CTL product, will be considered as 2 independent subjects, based on the first 2 study treatments received, for the purposes of analysis. The efficacy and safety data for such a subject will be assigned to the relevant study treatment source cohort based on the initial dose of the first 2 product types. The third EBV-CTL product will be considered as other anti- post-transplant lymphoproliferative disorder (PTLD) therapy.

Data for subjects assigned to the autologous EBV-CTL cohort will not be included in the summary tables or figures. The efficacy and safety data for these subjects during autologous EBV-CTL treatment will be provided in the listings.

Regardless of the number of products received, all subjects will be counted only once in the 'Overall Total' combined cohort for safety analyses.

4.3 Subgroups

Some of the safety and efficacy analyses may be repeated for the following subgroups within the full analysis set, if plausible considering the number of subjects in the subgroups.

- Age (< 18 vs. ≥ 18, < 16 vs. ≥ 16)
- Gender (male vs. female)
- Race (White, Other races, Unknown or Missing)

Other subgroups may be included in the analysis depending on the availability of the data.

5 INTERIM ANALYSIS AND EARLY STOPPING GUIDELINES

No formal interim analysis is planned.

6 GENERAL PRINCIPLES

6.1 General

The Full Analysis Set will be used for all the safety analyses and efficacy analyses except endpoints of DOR and TTR, which will be conducted on the subjects in the full analysis set with an objective response (CR or PR). Some analyses may be repeated in the selected subgroups specified in Section 4.3 depending on data availability.

Unless otherwise specified, all continuous variables will be summarized using descriptive statistics, which will include the number of subjects with a valid measurement (n), mean, standard deviation (SD), median, 25% quantile (Q1), 75% quantile (Q3), minimum and maximum. All categorical variables will be summarized using frequencies and percentages. The exact binomial 2-sided 95% confidence interval (CI) will be provided for all binary efficacy endpoints, including ORR and DRR. Kaplan-Meier (K-M) estimates along with their corresponding 95% CI will be calculated for time-to-event endpoints.

The following terms may be used for this SAP and tables, figures and listings (TFLs):

- **Study treatment:** Study treatment refers to all the protocol specified treatment administered to subjects.
- **Enrolled subjects:** refers to subjects who receive study treatment.
- **Enrollment date:** refers to the date that the inform consent form was signed.
- **Study day:** The subject's time on study will be determined in study days. Day 1 is defined as the day of the first study treatment administration. Study day is defined as the date of interest minus day 1 plus 1 if the date is on or after day 1. If the date is before day 1, the study day is defined as the date of interest minus day 1.
- **Baseline:** refers to day 1 prior to the first administration of study treatment. The baseline value of a parameter (eg, laboratory tests and efficacy endpoints) is defined as the last value prior to the first study treatment administration.
- **Duration of treatment:** defined as the date of the last dose of study treatment minus the date of the first dose of study treatment + 1
- **Treatment exposure period:** from day 1 through 30 days after the last administration of study treatment.
- **Restriction Switch:** Subjects could change to tabelecleucel with a different HLA restriction from a different donor either due to a lack of efficacy from the current cell line or insufficient supply. "Restriction switch" refers to only the change to tabelecleucel with a different HLA restriction due to lack of efficacy.
- **Year:** A year consists of 365.25 days
- **Month:** A month consists of 365.25/12 days

6.2 Missing Data Handling

In general, missing data will not be imputed. Only year of birth was collected during the study. To calculate age, 01 July was used to impute the birthday and month.

6.3 Visit Window

Depending on the statistical analysis method, single values may be required for each analysis window. For example, change from baseline by visit usually requires a single value per visit. When a single value is needed, the following rule(s) will be used:

- In general, the baseline value will be the last non-missing value on or prior to the date of the first dose of study treatment, unless specified differently.
- For post-baseline visits, if there are multiple assessments in a same visit window, the assessment closest to the corresponding dose date will be used in the analysis. For example, if after cycle 1 day 1 (first infusion) lab samples are collected on day 3 and day 5 before cycle 1 day 8 (second infusion), the lab assessment collected on day 3 will be used for cycle 1 day 1.

For summary of laboratory data by visit window, it is possible that large time gaps exist between dosing cycles. To account for this, a cycle is defined to start at the first dose of the cycle and end at the first dose of the next cycle or first dose of the cycle + 50 days, whichever is earlier. For the last cycle, the cycle ends at first dose of the cycle + 50 days or last dose of the cycle + 30 days, whichever is earlier. Within a cycle, additional visits may be defined based on the actual infusion date of each dose. The window of an infusion visit starts at the infusion date and ends at the day before the next infusion date. If an infusion is the last infusion for a cycle, the infusion visit window ends at the day before the start of next cycle or the infusion date + 27 days, whichever is earlier.

Disease assessments performed by investigator will be grouped into one visit if the gap between two consecutive assessments is less than or equal to 14 days.

6.4 Data Handling and Electronic Transfer of Data

Analysis Data Model (ADaM) data sets will be prepared for statistical analysis. ADaM data are derived from the Study Data Tabulation Model (SDTM) data. SDTM data are derived from raw data entered into the clinical trial database.

Data files were provided from the site via secured transfer. The data was mapped from the raw files to the fields available in the clinical trial database. The formatting of the data into data entry packets was done in accordance with Atara Work Instruction. The data was then transferred via secured server to the vendor for entry into the trial database using the vendor's standard operating procedures and data entry conventions.

6.5 Validation of Statistical Analysis

The statistical analysis validation procedure will be conducted in accordance with Atara's standard operating procedures (SOPs).

7 STUDY SUBJECTS

7.1 Study Disposition

Subject disposition, including the number of subjects screened, the number of subjects deemed eligible and the number of subjects enrolled/treated, will be summarized. Treatment status (completed or discontinued) with reasons for treatment discontinuation and end of study status (completed or discontinued) with reasons for not completing the study will also be summarized. The analysis will be repeated for the subgroups defined in Section 4.3.

7.2 Important Protocol Deviations

Important protocol deviations (IPDs) will be reviewed prior to the final database lock. A by-subject listing of IPDs will be provided, which will include IPD category, subcategory.

7.3 Demographic and Baseline Characteristics

The following demographic and baseline characteristics for full analysis set and the subgroups defined in Section 4.3 will be summarized using descriptive statistics. Baseline weight will be defined as the last screening weight.

- Age at baseline: descriptive statistics and n (%) for the following:
  - < 2, 2 to < 12, 12 to < 16, 16 to < 65, and ≥ 65 years
  - < 16 and ≥ 16 years
  - < 2, 2 to < 12, 12 to < 18, 18 to < 65, and ≥ 65 years
  - < 18 and ≥ 18 years
- Sex (male, female)
- Ethnicity (Hispanic, non-Hispanic)
- Race (White, Black or African American, Asian, Unknown)
- Time from initial diagnosis of EBV^+ disorder to first dose of study treatment (months)
- Disease risk subgroups (for subjects ≥ 16 years and for subjects ≥ 18 years)
  a) Age < 60 years (low risk), age ≥ 60 years
  b) Serum lactate dehydrogenase (LDH) concentration: normal (low risk), elevated (> upper limit of normal [ULN])
  Time from initial diagnosis to first dose of study treatment (months)

The following disease characteristics will be summarized for HCT and SOT subjects only:

- Central nervous system (CNS) disease
- Extranodal disease at baseline (including bone marrow [BM])
- Lymph node disease
- > = 3 sites of disease
- Time from transplant to diagnosis of EBV^+ PTLD

7.4 Extent of Exposure

Descriptive statistics will be provided for average number of weight-adjusted cells per dose (10^6 cells/kg), number of lots subjects received, number of treatment cycles, number of doses, and duration of treatment (date of last dose of tabelecleucel – date of first dose of tabelecleucel + 1). Summary and listing of the subjects who undergo Restriction Switch will also be provided.

The analysis will be repeated for the subgroups defined in Section 4.3

8 EFFICACY ANALYSES

For the primary efficacy analysis, all response assessment data will be included regardless of how many Restriction Switches the subjects had.

Efficacy analyses will be conducted by cohorts and combined cohorts, as specified in Section 4.2.

Disease assessments performed by investigators regardless of modalities will be grouped into one timepoint assessment if the gap between two consecutive disease assessments is less than or equal to 14 days. This grouping will continue for the same timepoint assessment until the gap between two consecutive disease assessments exceeds 14 days.

The ranking for assessment modalities is specified in Table 2. If there are multiple assessments with the same modality in the same timepoint assessment, the worst/worse assessment (from the best to the worst in the order of not evaluable [NE], CR, PR, stable disease [SD] and progressive disease [PD]) will be selected for that modality. If there are assessments based on multiple modalities for a timepoint assessment, then the overall response for the timepoint assessment will be decided by the rank of the modalities specified in Table 2. For example, for a timepoint assessment, a PR was assessed per biopsy and a CR was assessed per physical examination. The assessment per biopsy will override that of physical examination for this timepoint assessment. If PD is determined as the response for a timepoint assessment, the earliest assessment date among all the disease assessments for this timepoint assessment will be considered as the PD date. If SD, PR or CR is determined as the response for a time assessment, the latest assessment date among all the disease assessments for this timepoint assessment will be considered as the response date.

While in general if any therapy directed towards the disease being treated, other than protocol-specified treatment, is initiated, response data after the initiation date of such therapy may be censored. However, given the fact that some anti-underlying disease therapies were actually allowed per the protocol, in the primary analysis of disease assessment related endpoints, none of such therapies were considered for censoring rule. As a sensitivity analysis, response data after the initiation of such therapies are censored, and such analysis is conducted for the endpoints including ORR, DOR, and PFS based on the Full Analysis Set.

Table 2: Rank of the Modalities (1 indicates highest rank)

Modality	Priority for non-Viremia	Priority for Viremia
BIOPSY	1	1
PATHOLOGY	1	1
BONE MARROW	1	1
PET	2	2
PET SCAN	2	2
PET/CT SCAN	2	2
CT SCAN	2	2
MRI	2	2
ULTRASOUND	3	3
THALLIUM SCAN	4	4
PHYSICAL EXAM	4	4
BLOOD	4	1
CLINICAL LABS	4	4

Abbreviations: CT, computed tomography; MRI, magnetic resonance imaging; PET, positron emission tomography

8.1 Analysis of Primary Endpoints

The primary endpoint ORR and its corresponding 95% CI will be reported by cohorts and combined cohorts specified in Section 4.3. The primary endpoint ORR is the proportion of subjects who have achieved a CR or PR during the study.

Sensitivity analysis will be conducted for ORR excluding response data after the initiation of any non-protocol anti-underlying disease therapies.

CCI [redacted]

8.2 Analysis of Secondary Endpoints

8.2.1 Overall Survival

Overall survival is defined as the time from the first dose of tabelecleucel to the date of death for any cause. Subjects who are lost to follow-up or still alive will be censored on the last known-to-be-alive date. All data including those collected during the follow-up period, if any, will be summarized.

8.2.2 Duration of Response

Duration of response is calculated based on subjects who achieve CR or PR. DOR is defined as the time from the date of initial response until (1) progression after the last response or (2) death due to any cause.

Only deaths within 90 days after the last valid disease evaluation (ie, not NE) will be counted as events in the DOR definition above.

For subjects without an event of death or disease progression, DOR is censored at the last valid (not NE) disease evaluation date.

In a sensitivity analysis, if any anti-underlying disease therapy is initialized before the event date/censoring date, DOR is censored at the last assessment date prior to the therapy initialization.

8.2.3 Progression-free Survival

Progression-free survival is defined as the time from the first dose of study treatment to either of following events whichever occurs first: (1) progression after the last response or (2) death due to any cause.

Death within 90 days after last valid disease evaluation (ie, not NE), or first study treatment dose date if there is no valid post-baseline disease evaluation, will be counted as events in the PFS definition

For subjects without an event as defined above, PFS will be censored at

- Last valid (not NE) post -baseline disease evaluation date
- First dose date if there is no valid post-baseline disease evaluation.

As a sensitivity analysis, PFS will be defined as the time from the first dose of tabelecleucel to either of following events, whichever occurs first, (1) the first progression or (2) death due to any cause. All other details and censoring rules are the same as above for the primary PFS analysis.

In another sensitivity analysis, if any anti-underlying disease therapy is initialized before the event date/censoring date, PFS is censored at the last assessment date prior to the therapy initialization.

8.2.4 Durable Response Rate

A response with a duration > 6 months is considered a durable response, and the DRR is defined as the proportion of subjects in the Full Analysis Set with a durable response.

Summary will also be provided for clinical benefit rate which is defined as the proportion of subjects who have achieved a CR, PR or SD assessed at least 28 days after the first dose date of tabelecleucel.

8.2.5 Time to Progression

Time to progression (TTP) is defined as the time from the date of the first dose of study treatment to progression after the last response. Death within 90 days after the last valid post-baseline disease evaluation, or after first dose date if there is no valid post-baseline disease evaluation, will be counted as an event only if the death reason is due to disease progression. Data will be censored on the death date if such deaths are not due to disease progression. The other censoring rules are the same as the ones used for the primary analysis of PFS in Section 8.2.3.

8.2.6 Time to Response

Time to response (TTR) is calculated only for subjects who achieve CR or PR on study. TTR is defined as the time from the date of the first dose of tabelecleucel to the date of the response (PR or CR whichever occurs first). Descriptive statistics will be provided. Similarly, time to best response will also be summarized for subjects who achieve CR or PR.

8.3 Subgroup Analysis

The analysis of ORR and OS will be repeated for the subgroups defined in Section 4.3.

9 PHARMACOKINETIC/PHARMACODYNAMIC ANALYSIS

No pharmacokinetic/pharmacodynamic analysis is planned.

10 SAFETY ANALYSIS

For subjects who were assigned to 2 cohorts, safety summaries for the first cohort will include data up to the day prior to the first dose date of study treatment in the next cohort. For summaries

of laboratory parameters, the baseline data for the second cohort will be the last laboratory parameter prior to the date of the first dose of study treatment in the next cohort.

10.1 Summary of Serious Adverse Events

Only serious adverse events (SAEs) were systematically collected in this study, hence all adverse event (AE) summaries will be based on SAEs only. Non-serious AEs will be listed if any.

The Medical Dictionary for Regulatory Activities version 22.0 will be used to code the SAEs to a system organ class (SOC) and a preferred term (PT) within the SOC. The Common Toxicity Criteria for Adverse Events (CTCAE) version 4.0 was used to grade the severity of SAEs.

Treatment-emergent serious adverse events (TESAEs) are defined as follows:

- any SAE that occurred after initiation of the first dose of study treatment through 30 days after the last dose of study treatment or
- any SAE that occurred prior to the first dose but worsened after the first dose or
- any related SAE with a date of onset on or after the first dose of study treatment.

SAEs occurring on the same day as the first dose of study treatment will be counted as treatment-emergent.

All the SAEs collected in this study will be listed.

All TESAEs will be summarized as follows.

- Overall summary of subject incidence for the following TESAE categories (repeat for the subgroups listed in Section 4.3)
    - Any TESAEs
    - TESAEs with worst grade ≥ 3
    - Fatal TESAEs
    - Any Treatment related TESAEs
    - Treatment related TESAEs with worst grade ≥ 3
    - Treatment related Fatal TESAEs
- Subject incidence for the following categories of TESAEs will be summarized by PT in descending order of frequency
    - TESAEs (repeat for the subgroups listed in Section 4.3)
    - Treatment-related TESAEs (repeated for < 16 vs. ≥ 16 and < 18 vs. ≥ 18 years age subgroups)
    - Fatal TESAEs (repeated for < 16 vs. ≥ 16 and < 18 vs. ≥ 18 years age subgroups)
    - Fatal TESAEs

- Subject incidence for the following categories of TESAEs will be summarized in descending order of frequency by PT and worst grade
    - TESAEs
    - TESAEs with worst grade > = 3 (repeat for the subgroups listed in Section 4.3)
    - Treatment-related TESAEs
- Subject incidence of the following categories of TESAEs will be tabulated SOC and PT in descending order of frequency by SOC and then PT within an SOC:
    - TESAEs
    - Treatment-related TESAEs

10.2 Summary of Laboratory Results

Routine clinical laboratory data (ie, hematology and serum chemistry) will be processed by local laboratories. Atara Biotherapeutics will convert the original test results/units to standard and conventional results/units and grade lab test results based on the CTCAE version 4.03 grading and corresponding normal ranges from University of California San Francisco (UCSF)[1] for the lab parameters.

In case the lab test result is less than the lower limit of quantification (LLOQ), a value of half the LLOQ will be used as the numerical test result. If the lab test result is greater than the upper limit of quantification (ULOQ), the ULOQ will be used as the numerical test result.

10.2.1 General Presentation of Laboratory Data

Summary of selected laboratory tests (Appendix 13.1 Selected Laboratory Tests) will include only data collected up to the last dose of study treatment plus 30 days.

For selected laboratory parameters, descriptive statistics of actual values and changes from baseline for the post-baseline laboratory data will be summarized for each visit window. Descriptive statistics of the minimum and maximum post baseline values for each subject will also be summarized. In the case of multiple values in an analysis window, data will be selected for analysis as described in Section 6.3. If either the baseline or post-baseline value is missing, the observation will not be included in the change from baseline summary.

The percentage of change from baseline (mean ± SD) for selected lab parameters vs. visit window, will be plotted.

In addition, grade shift tables from baseline to worst on-treatment value will be provided for selected laboratory parameters.

For laboratory tests with criteria for both increased and decreased levels, analyses for each direction (ie, increased, decreased) will be presented separately.

[1] Available at: http://labmed.ucsf.edu/sfghlab/test/ReferenceRanges.html

In addition, data for sirolimus, tacrolimus and interleukin 6, highly sensitive will be provided in listings.

10.2.2 Treatment-Emergent Laboratory Abnormalities

Treatment-emergent laboratory abnormalities are defined as values that increase by at least 1 toxicity grade from baseline at any post baseline time point, up to and including the date of last dose of tabelecleucel plus 30 days. If the relevant baseline laboratory value is missing, any abnormality of at least grade 1 observed at any post baseline time point up to and including the date of last dose of tabelecleucel plus 30 days will be considered treatment emergent.

Number of subjects (%) with treatment emergent laboratory abnormalities will be summarized for selected laboratory parameters (Appendix 13.1 Summery of Selected Laboratory Tests). The number of subjects with any post baseline value for each laboratory test will be the denominator for the percentage of subjects with the corresponding treatment-emergent laboratory abnormality.

10.3 Exposure to Prior Therapy/Concomitant Medication/Subsequent Therapy

Number of lines of prior systemic EBV associated disease therapies and numbers of subjects who received different prior therapies will be summarized as below:

- Rituximab monotherapy
- Chemotherapy
    - Anthracycline-based therapy (CHOP/R-CHOP)
    - Brentuximab vedotin
    - Plantinum-based therapy
    - Gemicitabine-based therapy
    - Other
- Radiotherapy
- Other therapy
    - Surgery/procedure
    - T-cell therapy
    - Stem cell transplant
    - Checkpoint inhibitor
    - Other

The concomitant and post treatment therapies will be provided in a listing. Therapies which overlapped with the study treatment duration will be considered as concomitant therapies. Therapies which started after last dose of study treatment or started during study treatment and continued after last dose will be considered as post treatment therapies. If a therapy cannot be

categorized due to missing dates, the during or post study treatment flags provided in the EDC will be used to identify concomitant and post-treatment therapies, respectively.

Number (%) of subjects with concomitant immunosuppressive medications will be provided for subjects with HCT and SOT EBV^+ PTLD or leiomyosarcoma (LMS).

10.4 Pregnancies

No data for pregnancies are available in this study.

11 CHANGES FROM PROTOCOL-SPECIFIED ANALYSIS

This study was originally designed and performed at a single academic center. Other than the primary endpoint of CR, specific endpoints and analyses of efficacy and safety data were not defined in the protocol. Therefore, Atara has defined these statistical components in this SAP, including the following:

- No formal hypothesis test will be conducted on CR, although the sample size calculation is a two-stage design based on CR. Point estimates and CIs will be provided for the rates of CR/PR for the cohorts and combined cohorts defined in Section 4.2 whenever feasible. In addition, statistics for the DOR and durable response are also provided.
- In the protocol, the primary objective included assessment of CR. To be consistent with Atara's other tabelecleucel studies, ORR is considered as the primary endpoint.
- All the analyses will be presented by the cohorts and cohorts combined defined in this SAP (Section 4.2), not by the groups defined in the protocol.

12 REFERENCES

1. Choquet S, Oertel S, LeBlond V, et al. Rituximab in the management of post-transplantation lymphoproliferative disorder after solid organ transplantation: Proceed with caution. Ann Hematol 2007; 86: 599–607

13 APPENDIX

13.1 Summary of Lab Tests

APPENDIX: SUMMARY OF LAB TESTS

Laboratory Test	Direction of Abnormality	Summary ofTreatment-emergent Abnormality and Grade Shift Table	Summary of Laboratory Test Value by Visit window
WBC COUNT	decreased	Y	Y
PLATELETS	decreased	Y	Y
ABSOLUTE NEUTROPHIL	decreased	Y	Y
ABSOLUTE LYMPHOCYTE	increased and decreased	Y	Y
LACTATE DEHYDROGENASE (LDH)	Not applicable	Not applicable	Y
HEMOGLOBIN	increased and decreased	Y	Y
CREATININE	increased	Y	Y
TOTAL BILIRUBIN	increased	Y	Y
ALBUMIN	decreased	Y	Y
ASPARTATE AMINOTRANSFERASE (AST)	increased	Y	Y
ALANINE AMINOTRANSFERASE (ALT)	increased	Y	Y
ALKALINE PHOSPHATASE (ALK)	increased	Y	Y
SIROLIMUS	Not applicable	Not applicable	Y
TACROLIMUS	Not applicable	Not applicable	Y

Signature Page for CCI [redacted]

Reason for signing: Approved	Name: PPD Role: Biostatistics Date of signature: 20-Jul-2019 03:50:30 GMT+0000

Reason for signing: Approved	Name: PPD Role: Biostatistics Date of signature: 20-Jul-2019 14:39:50 GMT+0000

Reason for signing: Approved	Name: PPD Role: Clinical Science Date of signature: 23-Jul-2019 01:07:56 GMT+0000

Signature Page for CCI [redacted]