CTOTC-08 STATISTICAL ANALYSIS PLAN
25 September 2019

B Cell Targeted Induction to Improve Outcomes in Pediatric Lung Transplantation

CTOTC-08

VERSION: **1.0**

DATE: **25 September 2019**

SPONSOR: Division of Allergy, Immunology, and Transplantation
National Institute of Allergy and Infectious Diseases – NIH
[redacted]

PREPARED BY: Rho
[redacted]

CTOTC-08 STATISTICAL ANALYSIS PLAN ACKNOWLEDGMENT AND SIGNATURE SHEET

CTOTC-08

Approved: _____ Date: _____
[redacted]
Principal Investigator

Approved: _____ Date: _____
[redacted]
Protocol Chair

Approved: _____ Date: _____
[redacted]
NIAID Medical Monitor

Approved: _____ Date: _____
[redacted]
Rho Scientist

Approved: _____ Date: _____
[redacted]
Rho Statistician

Document History

Version	Date	Change(s)	Author
0.1	31Oct2018	Initial Draft	[redacted]
0.2	25Mar2019	Revised Draft	[redacted]
0.3	16May2019	Revised Draft	[redacted]
0.4	10Sep2019	Revised Draft	[redacted]
1.0	25Sep2019	Final revisions	[redacted]

Table of Contents

1. PROTOCOL SYNOPSIS

Title	B cell Targeted Induction to Improve Outcomes in Pediatric Lung Transplantation
Clinical Phase	Phase 2
Number of Sites	7 US Sites
IND Sponsor/Number	NIAID/IND Number 121403
Primary Study Objective	To determine whether rituximab induction along with standard of care immunosuppression will improve outcomes following pediatric lung transplantation.
Secondary Study Objectives	1. To determine the effects of rituximab induction on post-transplant immunity in pediatric lung transplant recipients. 2. To assess the safety and tolerability of rituximab. 3. To assess the feasibility of a phone based intervention to decrease tacrolimus trough level variability.
Study Design/Treatment Description	Phase 2, prospective, multi-center, double-blind, randomized, placebo-controlled clinical trial in which 50 primary pediatric lung transplant recipients will be randomized (1:1) to receive either induction therapy with anti-CD20 mAb (375 mg/m2) IV or placebo (IV day 0 and day 12 post-transplant) plus standard of care immunosuppression (thymoglobulin induction, tacrolimus or generic equivalent, MMF or generic equivalent, and steroids).
Primary Endpoint	The primary endpoint is a composite of the earliest time to any of the following events during the follow-up period: • Chronic Allograft Dysfunction • Listed for Retransplant • Death
Secondary Clinical Endpoints	1. Post-transplant clinical outcomes including: a. Incidence of chronic allograft dysfunction, listing for retransplantation, and death during the follow-up period, which will be a minimum of 12 months post-transplant b. Incidence of Primary Graft Dysfunction c. Incidence of Grade A Acute Rejection during the follow-up period, which will be a minimum of 12 months post-transplant d. Incidence of Antibody Mediated Rejection during the follow-up period, which will be a minimum of 12 months post-transplant e. Incidence of tacrolimus variability threshold during the follow-up period, which will be a minimum of 12 months f. Percentage of participants meeting tacrolimus variability threshold who complete tacrolimus variability intervention

	g. Magnitude of change in standard deviation of tacrolimus levels following intervention 2. Post-transplant safety outcomes including: a. Incidence and severity of infection episodes b. Serious adverse events related to rituximab
Secondary Mechanistic Endpoints	1. Incidence and kinetics of DSA and autoantibodies, specifically Collagen V (ColV) and k-alpha-1 tubulin (kα1T) 2. Frequency, kinetics, phenotype and function of peripheral B cells 3. Frequency, kinetics and cytokine profiles of allo- and auto reactive T cells 4. Incidence and quantity of B cells and B cell proximity to other cells in the graft tissue
Accrual Objective	50 randomized pediatric lung transplant recipients
Study Duration	4.5 years (3.5 year accrual + 1-4.5 year follow-up period) All participants will be followed for a minimum of 1 year post-transplant. Participants will continue follow-up visits (up to 4.5 years post-transplant) until the last participant completes 1 year of follow-up.
Enrollment Inclusion Criteria	1. Subject and/or parent/guardian must be able to understand and provide informed consent 2. Less than or equal to 21 years of age 3. Candidate for primary lung transplant (listed for lung transplant) 4. Female and male subjects with reproductive potential must agree to use FDA approved methods of birth control for 12 months after completion of treatment 5. Adequate bone marrow function based on the following criteria: a. ANC > 1000mm^3 b. Platelets > 100,000/mm^3 c. Hemoglobin > 7 g/dL d. AST or ALT <2x Upper Limit of Normal unless related to primary disease
Enrollment Exclusion Criteria	1. Inability or unwillingness of a participant to give written informed consent or comply with study protocol 2. Multi-organ transplant 3. Previous treatment with rituximab (Rituxan®) 4. History of severe allergic anaphylactic reactions to humanized or murine monoclonal antibodies 5. History of severe reaction to previous therapy with IVIG 6. History of Burkholderia cenocepacia 7. History of anti-CD20 therapy

	8. Persistent hypogammaglobulinemia (IgG < lower level of normal for age based on local laboratory ranges or 400 g/dL for >2 months) and/or IVIG replacement therapy 9. Positive blood culture, sepsis or other disease process with hemodynamic instability at time of enrollment 10. Any history of serologic positivity to HIV, HBsAg, HBcAb and HCV Ab 11. History of malignancy less than 2 years in remission of malignancy (any history of adequately treated in-situ cervical carcinoma, or adequately treated basal or squamous cell carcinoma of the skin will be permitted) 12. Any condition, including psychiatric disorders, that in the opinion of the investigator would interfere with the subject's ability to comply with study requirements 13. Participation in another investigational trial within 4 weeks of enrollment 14. Currently lactating or plans to become pregnant during the timeframe of the study follow-up period 15. Past or current medical problems or findings from physical examination or laboratory testing that are not listed above, which, in the opinion of the investigator, may pose additional risks from participation in the study, may interfere with the participant's ability to comply with study requirements or that may impact the quality or interpretation of the data obtained from the study
Randomization Inclusion Criteria	1. Serum IgG immunoglobulin level greater than lower level of normal for age based on local laboratory ranges or 400 mg/dL within 90 days prior to randomization 2. Female subjects of childbearing potential must have a negative pregnancy test within 48 hours of transplant 3. Negative for Hepatitis B infection (if at time of transplant, subject does not exhibit effective immunization, the subject should be re-tested)
Randomization Exclusion Criteria	1. Use of an induction agent other than Thymoglobulin® 2. Renal insufficiency requiring hemodialysis or ultrafiltration 3. Inability to obtain intravenous access 4. Positive blood culture, sepsis or other disease process with hemodynamic instability at time of transplant 5. Use of investigational agent(s) within 5 half-lives of the investigational drug or 4 weeks, whichever is longer 6. Receipt of a MMR vaccine within 30 days prior to randomization 7. Any condition that, in the opinion of the investigator, would interfere with the subject's ability to comply with study requirements
Study Stopping Rules	Satisfaction of any of the following stopping rules in study subjects at any time of follow-up in the treatment arms will trigger an ad hoc DSMB Safety Review:

	• Any occurrence of confirmed PML • Incidence of death of 30% or more subjects • Incidence of at least mild acute rejection of 35% or more • Incidence of humoral rejection of 25% or more • Incidence of primary graft dysfunction of 50% or more • Incidence of PTLD of 5% or more • Incidence of infections of any type requiring hospitalization of 40% or more
Individual Subject Stopping Rules	Individuals who meet any of the criteria listed below will not receive the second dose of rituximab 1. Serious adverse event casually related to the rituximab infusion 2. Acute pulmonary infectious process with evidence of graft dysfunction 3. Positive blood culture, sepsis or other disease process with hemodynamic instability 4. Renal insufficiency requiring hemodialysis or ultrafiltration 5. Inability to obtain intravenous access 6. Use of an investigational drug after the first dose of placebo or rituximab 7. Any other event which in the opinion of the principal investigator may pose additional risk to the participant

2. INTRODUCTION

This statistical analysis plan includes pre-planned analyses related to the study objectives outlined in the protocol.

3. GENERAL ANALYSIS AND REPORTING CONVENTIONS

The following analyses and reporting conventions will be used:

- Categorical variables will be summarized using counts (n) and percentages (%) and will be presented in the form "n (%)." Percentages will be rounded to one decimal place.
- Numeric variables will be summarized using n, mean, standard deviation (SD), median, minimum (min), maximum (max). The min/max will be reported at the same level of significance as original data. The mean and median will be reported at one more significant digit than the precision of the data, and SD will be reported at two more significant digits than the precision of the data.
- The median will be reported as the average of the two middle numbers if the dataset contains an even number of observations.
- Test statistics including *t* and *z* test statistics will be reported to two decimal places.
- *P*-values will be reported to three decimal places if greater than or equal to 0.001. If less than 0.001, the value will be reported as "<0.001." A *p*-value can be reported as "1.000" only if it is exactly 1.000 without rounding. A *p*-value can be reported as "0.000" only if it is exactly 0.000 without rounding.

If departures from these general conventions are present in the specific evaluations section of this SAP, then those conventions will take precedence over these general conventions.

4. ANALYSIS SAMPLES

Intent-to-Treat (ITT) Sample – The ITT sample is all randomized subjects who receive at least a portion of the initial rituximab/placebo infusion. This sample will be used for evaluation of the primary and secondary clinical endpoints. Subjects will be analyzed in the group to which they were randomized, regardless of compliance with the protocol or withdrawal from the study.

Per-Protocol (PP) Sample – The PP sample is the subset of subjects from the ITT sample who complete the full induction protocol, with complete administration of the prescribed dose at both the initial and subsequent infusions, and do not have any major protocol deviations that impact the subject's evaluability for the primary endpoint. This sample will capture subjects who both tolerate rituximab well and recover as expected from their transplant.

Screening Sample – All consented (i.e., enrolled) subjects will constitute the screening sample. This sample will be used to describe population characteristics.

Safety Sample – All subjects from the screening sample who have the pre-transplant blood draw performed will constitute the safety sample. This sample will be used for all safety summaries. Subjects in this sample who were randomized will be analyzed in the group to which they received, regardless of which group they were randomized to, compliance with the protocol, or withdrawal from the study. If a subject received any rituximab the subject will be analyzed with the rituximab group.

5. STUDY SUBJECTS

5.1. Disposition of Subjects

The disposition of all enrolled subjects will be summarized in tables and presented in listings using the screening sample. A consort diagram will also be constructed to serve as a visual depiction of the flow of subjects through the study.

The numbers and percentages of subjects randomized (and in each analysis sample) will be displayed by randomized treatment group and overall. Reasons for early termination from the study and visit completion statistics will be presented. The reasons for discontinuing study treatment early will also be presented for any subjects who prematurely discontinue study treatment.

5.2. Demographic and Other Baseline Characteristics

Summary descriptive statistics, with no formal group comparisons, for baseline and demographic characteristics will be reported for the ITT sample by treatment group and overall. Additionally, these characteristics will be described for subjects who were screened but were not eligible for treatment. Characteristics to be summarized include age, race, ethnicity, sex, cause of death, and chest trauma for donors, and age, race, ethnicity, sex, pre-operative diagnoses, and medical histories for recipients. Additional relevant characteristics may be included.

6. STUDY OPERATIONS

6.1. Protocol Deviations

Protocol deviations will be listed by site with information such as severity of the deviation (major or non-major), type of deviation, date of occurrence, and the reason for the deviation. Protocol deviations may also be summarized in tabular format by severity and type of deviation.

6.2. Treatment Adherence

Treatment Groups:

- **Rituximab Group** – induction therapy with anti-CD20 mAb (375 mg/m^2) IV on day 0 and day 12 post-transplant plus standard of care immunosuppression (thymoglobulin induction, tacrolimus or equivalent, MMF or equivalent, and steroids).
- **Placebo Group** – induction therapy with placebo IV on day 0 and day 12 post-transplant plus standard of care immunosuppression (thymoglobulin induction, tacrolimus or equivalent, MMF or equivalent, and steroids).

Randomization to one of the two groups will occur once the subject has received the transplant and is deemed hemodynamically stable. The first infusion must be administered with 12 hours of returning to the ICU following transplant. The second infusion should be administered on day 12 (± 2 days) post-transplant.

Adherence will be summarized for the ITT sample. The status of each infusion will be classified as complete, partial, or not done and the duration (or extent) of exposure will be calculated by determining the number of days between first and second infusions (+1) for each subject in the

sample. These variables will then be summarized using appropriate statistics in a table and subject-level details will be presented in a listing.

7. ENDPOINT EVALUATION

7.1. Overview of Efficacy Analysis Methods

7.1.1. Multicenter Studies

Study subjects will be recruited from 7 study sites.

Following database lock, if potentially impactful site-to-site variation is noted, basic descriptive analyses of baseline demographics, medical history, and key study endpoints will be repeated for each site individually in order to allow qualitative exploration of site-to-site variability.

7.1.2. Assessment Time Windows

Allowable visit windows are detailed in Section 8 of the protocol.

Unscheduled visits may also occur throughout the study if a subject develops symptoms of a respiratory viral infection (RVI) or a bronchoscopy is performed for suspicion of rejection or infection.

Generally, all data will be included in analyses, regardless of time of assessment.

7.2. Primary Endpoint

The primary endpoint is a composite of the earliest time from transplant to any of the following events during the follow-up period:

- Chronic allograft dysfunction (defined as the occurrence of confirmed BOS grade 0-p or higher or a diagnosis of Obliterative Bronchiolitis (OB)),
- Listed for re-transplant,
- or Death.

7.2.1. Computation of the Primary Endpoint

First, each component will be derived separately as described below.

- **Chronic allograft dysfunction**
  - BOS component – Subjects with at least one confirmed diagnosis of BOS grade 0-p or higher, as defined in Section 7.5.3 of the protocol, will be considered to have met the BOS component of chronic allograft dysfunction. Rho will programmatically monitor the PFT data and notify sites of suspected BOS. Confirmation (or rule out) of BOS will then be made by the site PI and documented in the BOS PFT Decline CRF.

    Subjects who do not meet this component must have at least one set of post-transplant PFTs reported in order to be considered evaluable and, as such, classified as not meeting the endpoint.

    The date/day of the first occurrence of confirmed BOS will be used for analyses.
  - OB component – Subjects with at least one episode of OB reported will be considered to have met the OB component of chronic allograft dysfunction. OB will be identified either as a reason for hospitalization or as a specific histopathological finding noted on a biopsy.

Subjects who do not meet this component must have at least one post-transplant biopsy reported in order to be considered evaluable and, as such, classified as not meeting this component.

For patients unable to perform PFTs (and will therefore have an unknown status for the BOS component), chest imaging data will be reviewed by the study Principal Investigator and Protocol Chair for any subjects designated as not meeting the OB component. The purpose of this review will be to identify any imaging changes consistent with airway obstruction that were not identified in the transbronchial biopsy tissue. If this comprehensive review of the imaging data yields a diagnosis of OB, the subject's OB classification will be changed from No to Yes, and the date of the imaging exam will be used as the date of OB diagnosis.

The date/day of the first occurrence of OB will be used for analyses.

- Once the status of the two components above is known, the chronic allograft dysfunction status for each subject will be classified per the rule below.
  - Chronic allograft dysfunction= Yes if BOS= Yes or OB= Yes
    - For these subjects, if more than one component is met then the earliest date/day will be used for analysis. Otherwise, the date/day of the only component met will be used for analysis.
  - Chronic allograft dysfunction= No if BOS= No and OB= No
    - For subjects who do not meet either component, data will be censored at the last time point where a subject's status is known.
  - Chronic allograft dysfunction= Unknown and considered not evaluable if
    - BOS= No and OB= Unknown *or*
    - BOS= Unknown and OB= No *or*
    - BOS= Unknown and OB= Unknown

- **Listed for re-transplant** – Subjects with a reported date of listing for re-transplant will be considered to have met this component of the primary endpoint. The date/day of re-listing will be used for analysis.

  Subjects otherwise reported as not having been re-listed will be classified as not meeting this component. Data will be censored at the last time point where a subject's status is known.

  Any subject not meeting the criteria to be classified as Yes or No for re-listings will be considered not evaluable for this component and classified as Unknown.

- **Death** – Subjects with a reported date of death will be considered to have met this component of the primary endpoint. All other subjects will be classified as not meeting this component. The date/day of death will be used for analysis; otherwise, data will be censored at the last time point where a subject's status is known. Once the status of each of the above components is determined, the primary endpoint status for each subject will be classified per the rule below.

- Primary Endpoint= Yes if

  Chronic allograft dysfunction= Yes or Listed for re-transplant= Yes or Death= Yes

  - For these subjects, if more than one component is met then the earliest date/day will be used for analysis. Otherwise, the date/day of the only component met will be used for analysis.

- Primary Endpoint= No if
  Chronic allograft dysfunction= No and Listed for re-transplant= No and Death= No
  - For subjects who do not meet any of the components, data will be censored at the last time point where a subject's status is known.

7.2.2. Primary Analysis of the Primary Endpoint

As the data allow, a Cox proportional hazards model will be used to compare treatment groups on the time to endpoint, allowing for right-censoring at the end of the follow-up period or at time of loss to follow-up. Estimates of the hazard ratios (HR) and their two-sided 90% confidence intervals will be produced. HR estimates <1 will indicate rituximab efficacy. For example, if an HR<0.85 is considered as indicating clear efficacy, then an HR estimate and 90% confidence interval of 0.33 (0.16, 0.68) will provide excellent support for clear efficacy; whereas an HR= 0.55 (0.28, 1.07) would provide fair to good support for efficacy, even though it would fail to reach significance for the usual test of HR= 1.

The model will be run on both the ITT and PP samples, as events allow.

Model diagnostics will include an evaluation of the proportional hazards assumption and of each observation for influence on model parameters. The proportional hazards assumption will be assessed by testing for a significant interaction between treatment group and the time-to-event variable.

An example of SAS code to be used for the primary model:

```
Proc phreg;
   class treatment_group(ref='Placebo');
   model event_day * endpoint_status(0) = treatment_group intterm /
risklimits ties=efron alpha=0.1;
   intterm= treatment_group_01 * event_day;
run;
```

Note: *intterm* represents the interaction term that can be used to test the proportional hazards assumption and *treatment_group_01* represents *treatment_group* coded as a 0, 1 variable where 0=placebo and 1=rituximab. This term will be used for testing the assumption only and will not be included in the final model.

To estimate the hazard function of each of the components, separately, the following models will be used.

An example of SAS code to be used for the analysis of competing risks:

```
*This is the model for Death;
Proc phreg;
   class treatment_group(ref='Placebo');
   model event_day * endpoint_status(0,1,2) = treatment_group /
risklimits ties=efron alpha=0.1;
run;
```

```
*This is the model for Listed for Re-transplant;
Proc phreg;
   class treatment_group(ref='Placebo');
   model event_day * endpoint_status(0,1,3) = treatment_group /
risklimits ties=efron alpha=0.1;
run;

*This is the model for Chronic Allograft Dysfunction;
Proc phreg;
   class treatment_group(ref='Placebo');
   model event_day * endpoint_status(0,2,3) = treatment_group /
risklimits ties=efron alpha=0.1;
run;
```

7.2.3. Sensitivity Analyses of the Primary Endpoint

Given the likelihood of little to no events for one or more of the 3 components of the primary endpoint, a reduced competing risks model will be used excluding any component not observed. Additionally, a time-to-first-event model utilizing the Kaplan-Meier framework will be employed to produce estimates of the event-free rate of the composite endpoint and its individual components.

Furthermore, demographic and other clinical covariates (e.g., the presence or absence of pre-transplant auto or allo-immunity) may be included in these secondary analyses. The decision of what to include may be based on clinical experience, literature review, or exploratory data analysis done following database lock but prior to unblinding the treatment assignments.

An example of SAS code to be used:

```
Proc lifetest plots=(survival(atrisk));
   time event_day * endpoint_status(0);
   strata treatment_group;
run;
```

7.3. Secondary Clinical Endpoints

All secondary endpoint analyses will use the ITT sample, unless otherwise specified, and will be assessed during the post-transplant follow-up period, which will be a minimum of 12 months.

7.3.1. Post-transplant Clinical Endpoints

Variable: Secondary incidence endpoints including:

- Chronic allograft dysfunction, defined as indicated for the primary endpoint
- Listing for re-transplant, defined as indicated for the primary endpoint
- Death, defined as indicated for the primary endpoint
- Primary Graft Dysfunction (PGD), defined at 0, 24, 48, and 72 hours post-transplant per Section 7.5.4 of the protocol
- Grade A Acute Rejection, reported locally as defined in Section 7.5.1 of the protocol

- Antibody Mediated Rejection of Grade II or III, reported locally as defined in Section 7.5.2 of the protocol
    - Additionally, if the data allow, AMR will be evaluated according to the 2016 ISHLT consensus guidelines published in the Journal of Heart and Lung Transplantation.

Analysis: For each of the above incidence endpoints, the proportion of subjects with events will be reported by treatment group with a corresponding 90% confidence interval, and for PGD and the rejection endpoints the treatment groups will be compared using a Chi-square or Fisher's Exact test. The incidence of grade 2 or higher PGD will be summarized overall and at each of the 4 time points separately. Additionally, for acute rejection and antibody meditated rejection, number of events per person time (e.g., year) will be estimated.

An example of SAS code to be used:

```
Proc freq;
   tables treatment_group * endpoint / chisq measures cl alpha=0.1;
run;
```

Variable: Incidence of meeting the tacrolimus variability threshold

Analysis: Among subjects who have at least 3 outpatient tacrolimus trough levels collected that are at least 3 months (i.e., greater than or equal to 3 months) post-transplant, an estimate of the standard deviation of tacrolimus levels will be produced for each subject. This estimate will be a rolling estimate (continually re-assessed) and may use up to 1 year worth of trough levels at any given time. The proportion of subjects with an estimated standard deviation at any time of 2.0 ng/mL or greater will be reported with a corresponding 90% confidence interval.

Variable: Incidence of completion of the Tacrolimus Variability Intervention (TVI)

Analysis: Among subjects who qualified for and agreed to participate in the TVI, the proportion of subjects who complete the intervention will be reported with a corresponding 90% confidence interval.

Variable: Magnitude of change in standard deviation of tacrolimus levels following intervention

Analysis: Among subjects who qualified for and agreed to participate in the TVI, the change in standard deviation of tacrolimus levels following the TVI will be examined. More specifically, a paired t-test will be carried out using the pre-intervention and post-intervention estimates of the standard deviation. An estimate of the mean difference between pre- and post-intervention standard deviation with a corresponding 90% confidence interval and p-value will be produced.

The value used for pre-intervention may be the same estimated standard deviation which qualified the subject for the TVI; however, if additional tacrolimus levels are recorded for a subject prior to their re-affirmation of consent for TVI participation, all outpatient tacrolimus levels in the year prior to commencement of the TVI will be used for estimating the pre-intervention standard deviation.

Likewise, the standard deviation value used for post-intervention will be estimated using all outpatient tacrolimus levels collected in the 180 days following completion of the TVI.

All estimates of the standard deviation pre- or post-intervention will adhere to the same conditions described earlier, which mandate at least 3 outpatient tacrolimus levels collected at least 3 months post-transplant.

Additionally, the change in standard deviation will be analyzed using the same paired t-test approach for pre-intervention vs. 180 days post-enrollment into the TVI, providing an assessment of change during the intervention.

Furthermore, attempts will be made to identify a set of "control" subjects for descriptive comparisons to subjects who enrolled and completed the TVI. These "control" subjects will at a minimum consist of subjects who qualified for but did not enroll in the TVI, and may be supplemented with subjects whose standard deviation exceeded 2.0 ng/mL but were not flagged due to enrollment in the TVI having been terminated. No formal statistical comparison of these two sets of subjects is planned.

7.3.2. Post-transplant Safety Endpoints

The safety sample will be used for analysis of these post-transplant safety endpoints.

Variable: Incidence and severity of infection episodes

Analysis: We will analyze each type of infection separately (bacterial, fungal, or viral, including RVI, CMV, and EBV), as well as for all types of infections combined. We will first determine the incidence of bacterial, fungal and viral infections in the rituximab and control groups using a combination of information from both the local, clinical site and the viral detection core laboratory. The presence of an infection at a given time point will be defined as follows:

- A local report of infection through either the organism specific CRF page or adverse event reporting;
- Presence of symptoms at the time of a study visit without a local report of infection but with a corresponding core lab identified infection.
    - All symptoms reported at the time of a study visit will be considered for any EBV and CMV infections identified by the core lab.
    - For any other infections identified in core lab NP and BAL samples, shortness of breath, new x-ray or other imaging finding, new supplemental oxygen requirement, cough with sputum, and decreased spirometry will be the only symptoms considered.

'Incidence rate' will be the number of infections per patient-month, and we will focus on the estimates and associated 90% confidence intervals for the ratio of such rates (RR), where RR > 1.0 indicates greater infection rates for rituximab. Of course, infection events are biologically related within patient, thus, relative to perfect independence of events, greater proportions of patients will experience more episodes of the same type of infection and greater proportions of patients will experience no episodes. Accordingly, Poisson regression models offset for months of follow-up and adjusted for overdispersion will be used to compare groups on the rate of each type of infection and any infection.

The severity of infection episodes will be determined by reconciling all infections, as defined above, against those infections which qualified for adverse event (AE) reporting. Since the protocol restricts AE reporting to only those events which meet NCI-CTCAE grading criteria of Grade 3 or higher, any infections not reported as an AE will be assigned a severity of Grade '<3' (effectively combining Grades 1 and 2); all other infections will utilize the assigned CTCAE severity grade of the corresponding AE. The proportion of subjects within each severity grade will be reported by treatment group with a corresponding 90% confidence interval; treatment groups will be compared using a Chi-square of Fisher's Exact test. This presentation will be repeated for the number of events divided by person time, where time is the same across severity groups.

Variable: Serious adverse events (SAEs) related to rituximab

Analysis: For all SAEs, assessment of the relationship to rituximab will be made by the DAIT medical monitor. The proportion of subjects with SAEs determined to be related to rituximab will be reported by treatment group with 90% confidence intervals, and treatment groups will be compared using Chi-square or Fisher's Exact test.

7.4. Secondary Mechanistic Endpoints

The following mechanistic endpoints will be assessed using the ITT sample:

- Incidence and kinetics of DSA and autoantibodies, specifically Collagen V (ColV) and k-alpha-1 tubulin (kα1T)
- Frequency, kinetics, phenotype and function of peripheral B cells
- Frequency, kinetics and cytokine profiles of allo and auto-reactive T cells
- Incidence and quantity of B cells and B cell proximity to other cells in the graft tissue

We will correlate pre and/or post-transplant serum reactivity to each autoantigen and to donor HLA with our primary endpoint to prospectively and independently test the hypothesis that autoantibodies can be used as predictors of incipient injury. We will also correlate the presence of serum antibodies with positive C4d staining of the graft tissue (to be collected as part of the clinical component of the study).

To test the hypothesis that rituximab prevents induction of de novo autoantibodies and or DSA, we will compare titers of autoantibodies (or DSA) between the rituximab and control groups. Development of antibodies to either DSA or self-antigens or both and the kinetics of antibody development (time course following transplantation) will be correlated with clinical evidence of AMR and BOS as determined in the clinical part of study described above and the ISHLT criterion. While the absolute numbers of patients in each group with preexisting autoantibodies or DSA is anticipated to be small (thus limiting power) we will compare and report pre vs post-transplant titers in each case to determine if rituximab impacts preexisting antibody.

The results will be analyzed for development of PGD, development of DSA, diagnosis of AMR and BOS. In addition we will also determine any correlation which may exist between the strength of antibodies detected for DSA, HLA and each antibody to self-antigens either alone or cumulatively and clinical outcomes such as PGD, AMR and BOS.

We will analyze DSA and each type of antibody separately. The DSA and autoantibody data will consist of binary (positive/negative) or ordinal (negative/weak positive/strong positive) responses measured repeatedly at scheduled and unequally spaced intervals. With these features of the data and at this sample size, it may not be feasible to obtain estimates of treatment effect at each time-point using a single statistical model with rituximab and time main effects and their interaction as covariates. Therefore, we will describe the treatment effect over the course of the follow-up period.

The hazard ratio for the association between development of antibodies and clinical outcomes (the endpoint and its components, and the secondary incidence endpoints) will be estimated with its 90% confidence interval using a Cox proportional hazard model with presence or absence of antibodies as time-dependent repeated measures covariates, stratified by treatment group. We will model the association of each antibody on each outcome separately and in combination.

As permitted by the data, we will estimate odds ratios with 90% confidence intervals for the associations between each of the various serum antibodies and C4d staining of the graft tissue. We will estimate these odds ratios at each time point between each pair of antibodies and each antibody and C4d result, both cross-sectional, and cumulative, in which case we would treat any subject with a current or prior positive result as positive for that particular antibody.

Immunophenotyping of B cell and T cell subsets from flow cytometry will be examined. B cell and T cell fractions will be compared between treatment groups within a repeated measures mixed model framework while utilizing samples from the pre-transplant, weeks 4-6, and months

3, 6, 12 and 18 time points. P-values from Tukey's HSD test will be reported to account for multiple comparisons of the mean cell fractions between groups at the various time points.

7.5. Other Endpoints

No other endpoints are specified in the protocol. However, additional exploratory analyses suggested by the primary or secondary analysis results may be performed.

8. SAFETY EVALUATION

8.1. Overview of Safety Analysis Methods

All safety analyses will be carried out using the safety sample and percentages will be calculated based on the number of subjects in the safety sample, unless otherwise noted. Missing safety information will not be imputed. These analyses will not be stratified by site.

Safety will be analyzed in each treatment group through the reporting of AEs, vital signs, physical and neurological examination findings, and changes in routine laboratory values.

Listings will be prepared, as needed, for all safety measurements. All listings will be sorted in order of treatment, subject identifier (ID), and time of assessment (e.g., visit, time, and/or event).

8.2. Adverse Events

Per the protocol, only AEs with a severity grade of 3 or higher will be collected. All AEs will be classified by system organ class (SOC) and preferred term (PT), according to the Medical Dictionary for Regulatory Activities (MedDRA) version 17.0. The severity of AEs will be classified using the National Cancer Institute's (NCI's) Common Toxicity Criteria for Adverse Events (CTCAE) version 4. Each AE will be recorded once at the highest severity.

AEs will be collected from the time of enrollment through study termination. Treatment-emergent AEs will be identified as those with an onset date on or after the initial infusion as well as those with onset before the initial infusion but that continued and worsened in severity after the initial infusion. If the start of the AE in relation to the start of study medication cannot be established (e.g., the start date for the AE is missing), then the AE will be considered as having an unknown timing relative to treatment and will be presented in a separate column in the same table as treatment-emergent AEs. All data tabulations will be of only treatment-emergent events and stratified by treatment group while non-treatment-emergent AEs will be listed separately.

An overall summary table will be developed to report the number of events and the number and percentage of subjects having at least one event in the following categories:

- AEs
- AEs indicated as serious
- AEs that lead to study drug discontinuation
- AEs with an outcome of death
- AEs that were reported as being related to rituximab/placebo, blood draw, or bronchoscopy
- AEs reported by severity and maximum severity

In general, when reporting on the relationship of an adverse event to study drug and procedures, the DAIT MM assessment of relationship will be used for all serious AEs and the site investigator assessment of relationship will be used for all non-serious AEs. This convention will apply even when AEs are summarized in aggregate (i.e., serious and non-serious AEs pooled together).

Additionally, the classification of AEs by MedDRA SOC and PT will be summarized for each treatment group and overall. Summary tables will present the total number of events as well as the number and percentage of subjects experiencing at least one event in each SOC/PT combination.

Separate data listings may be generated for treatment-related AEs and AEs leading to study drug discontinuation.

8.3. Deaths and Serious Adverse Events

Serious adverse events (SAEs) will be summarized in a manner consistent with that described in Section 8.2, and when appropriate SAEs may be summarized in the same table as all AEs.

Separate listings detailing each death, including time to and cause of death, will also be created.

8.4. Clinical Laboratory Evaluation

Clinical laboratory measurements include serum chemistry and hematology. Results will be converted to standardized units where possible. For numeric laboratory results, descriptive statistics of laboratory values and the change from baseline of laboratory values will be summarized for each treatment group and overall. For any categorical laboratory results, the number and percentage of subjects reporting each result will be presented for each treatment group and overall.

In addition, or in lieu of summary statistics, clinical laboratory data may be plotted to show patterns over time. For each test with a numeric result, data will be plotted as a spaghetti plot where each subject's values will be plotted and connected by line segments, forming one line per subject. Tests with qualitative results will not be plotted.

Separate listings of laboratory data over time may also be created.

8.5. Vital Signs, Physical Findings, and Other Observations Related to Safety

8.5.1. Vital Signs

Descriptive statistics of vital sign measures will be summarized by time point for each treatment group and overall. Since pre-infusion (i.e., baseline) vitals are not collected, no summary of change from baseline will be presented. Additionally, or in lieu of summary statistics, vital sign measures may be plotted over time by treatment group with individual lines for each subject.

Separate listings of vital sign data over time may also be created.

8.5.2. Physical Examinations

If needed to fulfill regulatory reporting obligations, listings of physical examination data will be produced. These listings may include, but are not limited to, the use (and amount) of supplemental oxygen, supportive ventilation, auscultation examination findings, and neurological findings.

8.5.3. Other Safety Measures

If present, a listing will be produced detailing any neurological examinations done for PML or any reported instances of PTLD.

9. INTERIM ANALYSES AND DATA MONITORING

No interim statistical analyses are planned. However, the progress of the study will be monitored by the Data and Safety Monitoring Board (DSMB). The DSMB will formally review the safety data at least yearly in open and/or closed sessions.

In addition, safety data will be reviewed by the DSMB when an event occurs that is of sufficient concern to the DAIT medical monitor or protocol chair to warrant review, or when an event occurs that could contribute to a protocol-specified stopping rule.