CTOT-19 STATISTICAL ANALYSIS PLAN

Randomized Controlled Trial of Infliximab (Remicade®) Induction Therapy For Deceased Donor Kidney Transplant Recipients

CTOT-19

VERSION: **1.0**

DATE: **14 December 2021**

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

PREPARED BY: Rho
[redacted]

CTOT-19 STATISTICAL ANALYSIS PLAN ACKNOWLEDGMENT AND SIGNATURE SHEET

CTOT-19

Signature and Date: _____ [redacted] Protocol Chair
Signature and Date: _____ [redacted] Principal Investigator
Signature and Date: _____ [redacted] NIAID Medical Monitor
Signature and Date: _____ [redacted] Rho Scientist
Signature and Date: _____ [redacted] Rho Statistician

Document History

Version	Date	Change(s)	Author
0.1	05 May 2021	Initial Draft	[redacted]
0.2	25 May 2021	Revised Draft	[redacted]
0.3	24 August 2021	Revised Draft	[redacted]
0.4	02 December 2021	Revised Draft	[redacted]
1.0	14 December 2021	Final revisions	[redacted]

Table of Contents

1. PROTOCOL SYNOPSIS

Title	Randomized Controlled Trial of Infliximab Induction Therapy For Deceased Donor Kidney Transplant Recipients
Short Title	Effects of inhibiting early inflammation in kidney transplant patients
Clinical Phase	Phase II
Number of Sites	15
Health Authority Applications/ Sponsor Number	NIAID/US IND 124633 NIAID/HC CTA File # HC6-24-C185918, Control #185918 Clinicaltrials.gov # NCT02495077
Study Objectives	The objective of the study is to determine the efficacy of intravenous infliximab administered at the time of transplantation, prior to reperfusion, on 2-year kidney transplant survival and function.
Study Design	This is a Phase II, multicenter, randomized, double blind, placebo-controlled, 2-arm clinical trial of 300 deceased donor kidney transplant recipients. Subjects will be randomized 1:1 to the experimental or control arms (150 subjects per arm).
Primary Endpoint(s)	The difference between the mean 24-month eGFR (modified MDRD) in the experimental vs. control arms.
Secondary Endpoint(s)	**Efficacy Endpoints** **1.** Proportion of subjects with biopsy proven acute cellular rejection (BPAR) within **a)** 6 months and **b)** 2 years of transplant **2.** BANFF grades of first Acute Cellular Rejections (ACR) within 6 month of transplant **3.** Proportion of subjects with biopsy proven acute cellular rejection (BPAR) or borderline rejection within **a)** 6 months and **b)** 2 years of transplant **4.** Proportion of subjects with biopsy proven acute antibody mediated rejection (AMR) within 6 months and 2 years of transplant **5.** Proportion of subjects with biopsy proven acute antibody mediated rejection AMR or suspicious for AMR within **a)** 6 months and **b)** 2 years of transplant **6.** BANFF grades of first AMR within 6 months of transplant **7.** Proportion of subjects with BANFF chronicity scores ≥ 2 on 24 month biopsy **8.** Change in BANFF chronicity scores between implantation and 24 month biopsies **9.** eGFR (as measured by both MDRD and CKD-EPI) **a)** Change in eGFR between 3 months and 24 months **b)** Change in eGFR between post-transplant nadir and 24 months **c)** eGFR on days 7, 30, 90, and 180 post-transplant **10.** Proportion of subjects with: **a)** Death or graft failure within 2 years **b)** Only graft failure within 2 years

11. Each of the following:
- **a)** Proportion of subjects that required at least one dialysis treatment within the first week after transplantation
- **b)** Number of dialysis sessions in the first 8 weeks post-transplantation.
- **c)** Duration of DGF defined as time from transplantation to the last required dialysis treatment
- **d)** The incidence of primary non-function (PNF), defined as for dialysis-dependency for more than 3 months
- **e)** Change from baseline (immediately after surgery) in serum creatinine and serum creatinine concentration at 24, 48, and 72 hours.

12. Days from transplantation until event (ACR, AMR, or hospitalization for infection and or malignancy)

13. Rate of Slow Graft Function (SGF):
- **a)** The proportion of patients with a serum creatinine of more than 3 mg/dL at day 5 post-transplant,
- **b)** Creatinine reduction ratio (CRR) on day 2 (defined as the first creatinine on day 2 divided by the first creatinine after surgery)
- **c)** Creatinine reduction ratio (CRR) on day 7 (defined as the first creatinine on day 7 divided by the first creatinine after surgery
- **d)** The proportion of patients whose day 5 serum CRR was less than 70%
- **e)** The proportion of patients whose day 2 serum CRR was less than 30%
- **f)** Proportion of subjects who need dialysis after 1 week.

Safety/Complication Endpoints

1. Proportion of subjects with:
- **a)** Any infection requiring hospitalization or resulting in death
- **b)** Mycobacterial or fungal infections

2. Proportion of subjects with CMV viremia that require a change in immunosuppression or anti-viral treatment as per standard of care at the site

3. Proportion of subjects with BK viremia that require a change in immunosuppression or anti-viral treatment as per standard of care at the site

4. Proportion of subjects with malignancy

5. Proportion of subjects with impaired wound healing manifested by wound dehiscence, wound infection, or hernia at the site of the transplant incision.

Mechanistic Endpoints

1. Sensitivity, specificity, PPV, and NPV of biomarkers, including PRT, urinary CXCL9, blood genomic profile, and 3 month allograft genomic profile (alone and/or in combination) to predict:
- **a)** Incident biopsy-proven acute rejection.
- **b)** Graft loss
- **c)** Chronic graft injury, as measured by 2-year eGFR

2. Each of the following:
- **a)** Inflammatory gene expression profiles
- **b)** Frequency of donor reactive T-cells
- **c)** Frequency and function of Treg
- **d)** Proportion of subjects with de novo DSA within 24 months
- **e)** Fibrogenic gene expression profiles
- **f)** Amount of peritubular capillary loss by histology

Ancillary Endpoints

	1. Percentage of predicted prednisone bottle openings as measured by a medication event monitoring system (MEMS®) in the first 3 months post-transplantation **2.** Standard deviation of the monthly tacrolimus trough levels from 6 months post-transplantation to 2 years post-transplantation
Accrual Objective	300 patients will be enrolled and randomized into the two treatment groups (150 Experimental arm, 150 Control Arm) across approximately 10 participating centers.
Study Duration	5 years (3 year accrual period + 2 year follow up period) • 24 months for primary and majority of secondary endpoints • 5 years with Reduced follow-up to assess patient and graft survival and serum creatinine
Treatment Description	There are two arms/groups in this study, the Experimental Group and the Control Group. Each group will receive the following: • **Experimental Group**: rabbit anti-thymocyte globulin (rATG, Thymoglobulin) is co-administered with anti-TNFa (infliximab/Remicade®) followed by maintenance therapy with tacrolimus, either Mycophenolate Mofetil/MMF or Mycophenolate Acid/MPA (or their generic equivalents) and prednisone • **Control group:** rabbit anti-thymocyte globulin (rATG, Thymoglobulin) plus placebo (Sterile normal saline) induction followed by maintenance therapy with tacrolimus, either Mycophenolate Mofetil/MMF or Mycophenolate Acid/MPA (or their generic equivalents) and prednisone
Inclusion Criteria	**1.** Adult (>18 years of age) male and female recipients (all races and ethnicities) **2.** Subject must be able to understand and provide consent **3.** Recipients of deceased donor kidney transplants (including re-transplants) **4.** Negative crossmatch, actual or virtual, or a PRA of 0% on historic and current sera as determined by each participating study center. **5.** Donor kidneys from deceased donors and donors after cardiac death (DCD) with Kidney Donor Profile Indices (KDPI) ranging from >20 to <95 **6.** Female participants of childbearing potential must have a negative pregnancy test upon study entry **7.** Subjects must have a negative test result for latent tuberculosis (TB) infection (PPD, QuantiFERON, ELISPOT). Subjects who have a negative test result for latent TB infection within 1 year of transplant date are eligible for enrollment and no further action is required. Subjects who have a negative test for latent TB infection that is greater than 1 year old are eligible for enrollment but are required to have a repeat test prior to transplantation. Samples for testing latent TB infection can be obtained during the hospital admission for the transplant but must be collected prior to the initiation of immunosuppression and prior to transplant. The results of this repeat test will determine the next step. • If the test is **negative**, no further action is required • If the test is **positive**, **lost**, **in determinant** or **unavailable the subject must be treated for latent TB infection** (if the subject is enrolled, received study drug/placebo and transplanted).

	○ Treatment for latent TB infection is required for all in this group. ○ The treatment regimen (including the specific drug, dose and duration of treatment) will be determined by the local Infectious Disease team. • If the subject has no latent TB testing available, then the subject cannot be enrolled Subjects may be enrolled if they have not had a test for latent TB infection within one year of the transplant if samples for the test are collected prior to the administration of immunosuppressive drugs and prior to transplant.
Exclusion Criteria	**1**. Inability or unwillingness of a participant to give written informed consent or comply with study protocol **2**. Recipients of living donor transplants **3**. Presence of other transplanted solid organs (heart, lung, liver, pancreas, small intestines) or co-transplanted organ **4**. HIV+ recipients **6.** Hepatitis B surface antigen positive kidney transplant recipients **7**. Hepatitis B core antibody positive kidney transplant recipients **8.** Hepatitis B negative kidney transplant recipients that receive transplant from Hepatitis B core antibody positive donor **9.** Hepatitis C virus positive (HCV+) patients who are either untreated or have failed to demonstrate sustained viral remission for more than 12 months after anti-viral treatment **10.** Recipients with a previous history of invasive fungal infection **11.** Recipients with a previous history of active TB disease **12.** Recipients with a positive test for latent TB infection (PPD, QuantiFERON, ELISPOT), regardless of previous therapy. **13.** Any severe infection at the time of transplantation. Severe infection determination will be made by the local site investigator. **14.** Severe congestive heart failure (NYHA functional class III or higher) **15.** Subjects with a known hypersensitivity to any murine/ mouse proteins **16.** Subjects with any history of receiving any anti-TNF products **17.** Subjects in whom rATG or infliximab might not be tolerated **18.** Subjects with less than 3000/mm3 WBC **19.** Subjects with less than 100,000/mm3 platelets counts **20.** Subjects with systolic blood pressure <100 mm/Hg

	21. Subjects with symptomatic orthostatic hypotension or currently requiring Midodrine for blood pressure support. **22.** Subjects from or who have traveled to endemic areas with a history of active histoplasmosis or with a chest x-ray consistent with previous active histoplasmosis (no serological testing required). Endemic regions determined by site based on local standard of care **23.** Subjects currently or formerly residing in regions of the US that are highly endemic for coccidioidomycosis, and who have a positive serologic test for coccidioidomycosis. Endemic regions determined by site based on local standard of care. * * (Subjects currently or formerly residing in regions of the US that are highly endemic for coccidioidomycosis, and who have a negative pre-transplant serologic test for coccidioidomycosis are eligible for enrollment only if they receive fluconazole 200 mg/day for the duration of the study. Serologic testing for coccidioidomycosis is required for subjects that currently or formerly resided in regions that are highly endemic for coccidioidomycosis only, per the site's standard of care). **24.** Recipients are excluded if the local site decides to treat the recipient with fluconazole because of diagnosis or suspicion of fungal infection in the donor. **25.** Subjects that receive IVIG treatment within 3 months of transplant or planned IVIG treatment peri-transplant. **26.** Use of an investigational agent within 4-weeks prior to study entry
Study Stopping Rules	Satisfaction of any of the following stopping rules at any time during the post-transplant (treatment) follow-up will trigger an *ad hoc* DSMB Safety Review. • Any single occurrence of a life-threatening or fatal AE that is possibly, probably, or definitely related to either the investigational agent (infliximab/infliximab placebo) or a study mandated procedure. **Across both treatment arms:** • Incidence of PTLD of 1% or more subjects • Incidence of Adverse Events for tuberculosis active disease of 1% or more subjects • Incidence of invasive fungal infection of 3% or more subjects • Incidence of coccidioidomycosis of 1% or more subjects • Incidence of histoplasmosis of 1% or more subjects • Incidence of death of 10% or more subjects **Within either treatment arm:** • Incidence of infection of any type requiring hospitalization of 40% or more subjects • Incidence of graft loss of 20% or more subjects • Incidence of BPAR (Banff Grade 1 or higher) or AMR based on local read of 25% or more subjects

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 transplanted and randomized subjects who receive the infliximab/placebo infusion. This sample will be used for efficacy and some safety analyses. Transplanted subjects who receive an infusion, but have no additional follow up will be used in this population. Subjects will be assigned to the treatment they were randomized to, regardless of which treatment they received.

Per-Protocol (PP) Sample – The PP sample is the subset of subjects from the ITT sample who receive the entire infliximab/placebo infusion, complete the full induction protocol, are compliant with their medications, and do not have any major protocol deviations that preclude analysis of the primary endpoint.

Safety Sample - All subjects who receive any infliximab/placebo infusion. This sample will be used for most safety summaries/analyses. Subjects will be assigned to the treatment they received, regardless of what they were randomized to.

Delayed Graft Function (DGF) Sample –All subjects with delayed graft function. This is defined as any dialysis in the first 7 days. This sample will be used for all DGF related analyses.

Screening Sample – All consented subjects will constitute the screening sample. This sample will be used for disposition summaries.

5. STUDY SUBJECTS

5.1. Disposition of Subjects

The disposition of all subjects will be summarized in tables and presented in listings, as needed, 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 enrolled, randomized, and transplanted will be displayed by treatment group and overall. The number of subjects in each analysis sample will be presented. Additionally, the numbers and percentages of subjects who complete the study or terminate early, along with reasons for early termination, will be summarized. Other information relevant to disposition will be summarized by treatment group and study visit as needed.

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. These characteristics, to the extent known, will also be summarized for subjects who were screened but were not eligible for study participation. Characteristics to be summarized are identified below. Additional relevant characteristics may be included.

- Recipient: age, race, ethnicity, sex, primary reason for transplant, cPRA%, dialysis pre-transplant, length and modality of pre-transplant dialysis, EBV and CMV serology status, pump perfusion status, resistive index
- Donor: age, race, ethnicity, sex, cause of death, donation after cardiac death, CMV serology status, KDPI, cold ischemia time

6. STUDY OPERATIONS

6.1. Protocol Deviations

Only major protocol deviations were collected for this study. Protocol deviations will be listed by site with information such as type of deviation, date of occurrence, and the reason for the deviation. Protocol deviations may also be summarized in tabular format by type of deviation.

6.2. Treatment Adherence

Treatment Groups:

Experimental Group: rabbit anti-thymocyte globulin (rATG, Thymoglobulin) is co-administered with anti-TNFa (infliximab/Remicade®) followed by maintenance therapy with tacrolimus, either Mycophenolate Mofetil/MMF or Mycophenolate Acid/MPA (or their generic equivalents) and prednisone

Control Group: rabbit anti-thymocyte globulin (rATG, Thymoglobulin) plus placebo (Sterile normal saline) induction followed by maintenance therapy with tacrolimus, either Mycophenolate Mofetil/MMF or Mycophenolate Acid/MPA (or their generic equivalents) and prednisone

Randomization occurred after enrollment and prior to transplant surgery. Participants were

randomized to either the Experimental group (receive Remicade®) or the Control Group (receive placebo). This was a blinded study so participants and the study team did not know the treatment assignment of the participant during the course of a subject's participation.

Adherence will be summarized for the ITT sample. Infusion status will be classified as complete or partial and also reported as the percentage of expected dose received, if needed. These variables will then be summarized in a table using appropriate statistics 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 were recruited from 15 study sites. Since randomization was carried out using fixed block sizes and stratified by enrolling site, study data will be analyzed as a whole, and no formal accommodation for site-to-site variation will be made.

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 at the time of an unscheduled (i.e., for cause) biopsy. A follow-up visit will be conducted 2-6 weeks after an unscheduled biopsy visit.

Additionally, subjects may have semi-annual (± 2 months) visits between months 24 and 60 post-transplant, up until the last transplanted subject completes 24 months of follow-up. These visits are only intended to collect a core lab specimen for eGFR, graft status, and, in some cases, a local creatinine result.

Generally, all data will be included in analyses, regardless of time of assessment. Exceptions to this include when the endpoint has a built in time component (e.g., within 24 months), in which case only data collected up to and including the protocol-specified upper window of time for the corresponding visit may be included. Endpoint-specific windows are specified in the subsections below.

7.2. Primary Endpoint

The primary endpoint is the difference between the mean 24-month eGFR in the experimental vs. control arms.

7.2.1. Computation of the Primary Endpoint

All post-transplant serum creatinine results received from the central laboratory and collected at protocol visits at month 1, 3, 6, 12, 18, and 24 will be used to estimate GFR via the Modification of Diet in Renal Diseases (MDRD) equation. Any creatinine sample collected on or after the date of graft failure will be excluded from the eGFR calculation.

MDRD equation for eGFR (mL/min/1.73m^2):

$175 \times (SCr)^{-1.154} \times (Age)^{-0.203} \times (0.742 \text{ if female}) \times (1.212 \text{ if African American})$

SCr=serum creatinine measured in mg/dL from the central laboratory

Age=age in years, rounded down to the nearest integer, at time of the SCr collection

Only data for subjects with known values of sex and race will be included in the evaluation of the primary endpoint. Subjects who reported race as 'Unknown or Not Reported' will be considered non-African American when calculating eGFR.

7.2.2. Primary Analysis of the Primary Endpoint

All subjects in the ITT sample will be used for the primary analysis. Mean eGFR of each of the two treatment arms will be compared within a restricted maximum likelihood repeated measures model (MMRM) framework, wherein repeated measures of the available eGFRs (as outlined in section 7.2.1) will be modeled using fixed effects of treatment group, collection time point, and the interaction between time point and treatment group. A random subject effect will also be included in the model to appropriately adjust for the correlation of having multiple eGFR values for each subject.

Let τ denote the "treatment effect at 24 months", which will be defined as the difference in 24-month population mean eGFR for the experimental treatment minus the corresponding placebo mean. τ is a parameter of primary interest; the MMRM will be used to estimate τ.

The primary null and alternative hypotheses are: $H_0: \tau = 0$ vs. $H_A: \tau \neq 0$. The MMRM will be used to test these hypotheses at the α= 0.05 Type I error rate.

SAS PROC MIXED (ver. 9.4 or higher) will be used to fit the MMRM to the data, estimate the parameters, and test the primary hypothesis along with others.

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

```
Proc mixed;
   class trtnm(ref='Placebo') subjid;
   model gfrmdrd=trtnm colldy trtnm*colldy / ddfm=kr;
   random intercept colldy / subject=subjid type=<xx>;
   lsmeans trtnm / at colldy=730 cl pdiff;
   estimate 'Experimental vs. Control at M24' trtnm 1 -1
      trtnm*colldy 730 -730 / cl;
run;
```

The `type=<xx>` covariance structure used will be determined following an evaluation of the AIC, AICC, and BIC model fit statistics for several candidate covariance structures (e.g., Unstructured, Toeplitz, and Variance components). The best fitting covariance structure will be used in the final model.

The estimated mean eGFR at 24 months, treatment group difference, and two-sided 95% confidence intervals will be summarized. Results associated with the F-test for the contrast/estimate will be reported. Additional graphical summaries of the eGFR over time will be used to further illustrate any differences between the two treatment groups. MMRM assumptions will be evaluated graphically via standard residual analyses. If the assumptions fail, a suitable transformation (e.g., natural logarithms) will be utilized and the data will be reanalyzed on the transformed scale. Missing data will not be imputed for the primary analysis; however, a sensitivity analysis will be employed to assess the impact of missing data due to subjects who terminate the study early or experience graft loss.

7.2.3. Sensitivity Analyses of the Primary Endpoint

The impact of early terminations will be assessed by looking for differences between the treatment groups in the proportion of subjects who terminate the study before completing the

month 24 visit, the timing of early terminations, and the reasons for early termination. These comparisons will allow for determination of whether or not the missing at random (MAR) assumption holds. Additionally, we will utilize the methods of Ma et al (2) to provide an objective index of the level of non-ignorable missingness. If the MAR assumptions fails, various sensitivity analyses will be employed to examine the impact of missing data on the primary analysis.

The first sensitivity analysis will use a pattern mixture model fitted via multiple imputation (3), and is an example of a "controlled imputation" method (1, 4) and will implement a jump to control (4) approach. For subjects who drop out of the study, this method imputes missing values for both treatment groups from the placebo population, which reduces any bias away from the null hypothesis of no difference between treatment groups.

A second sensitivity analysis will impute a single point for subjects (in either treatment group) that have had a graft loss or have died. Specifically, subjects who have a graft loss prior to 24 months will have an eGFR of 10 imputed for their remaining eGFR collection time points. Subjects who die with a functioning kidney will have a "last observation carried forward" imputation, which means that we would impute their last known eGFR prior to death for all remaining eGFR collection time points. This final eGFR will be based on the central evaluation of eGFR if available and the collection time was within 1 month of death; otherwise, a local assessment of eGFR will be used, if available.

A third sensitivity analysis will compare the eGFR at 24 months between the two treatment groups using a Wilcoxon test. Subjects who experience death or graft loss will be given the worst ranks. Subjects who are missing their 24-month eGFR due to other causes will receive a rank based on an imputed eGFR from the model used in the primary analysis.

7.2.4. Additional Analyses of the Primary Endpoint

If alternative eGFR equations are published in a peer-reviewed journal following conclusion of this trial and the required data inputs were collected for this trial, the primary model will be rerun using eGFRs calculated via one or more of these alternative equations.

Furthermore, the following additional analyses of the primary endpoint are planned.

- A two-sample t-test will be used to compare group means of eGFR at 24 months post-transplant. This analysis will include only the month 24 (± 1 month) eGFR values.
- The two-sample t-test will be repeated using local creatinines, in lieu of central creatinines, when calculating eGFR. The same ± 1 month window will be applied.
- The following analyses will use the same model as the primary analysis with data substitutions as follows:
    - Include central creatinines from the post-month 24 visits;
    - Substitute available local creatinines into the eGFR calculation at time points where a central creatinine result was not available via separate models with and without post-month 24 visits. All local creatinine values collected on or after the month 1 visit that are not within +/- 7 days of an available central creatinine result will be included in these models.
    - Use only local creatinines, in lieu of central creatinines, and separate models with and without post-month 24 visits.
- All analyses noted above may be repeated using eGFRs calculated via one or more of the alternative equations, if published and available.

7.3. Secondary Endpoints

7.3.1. Secondary Clinical Endpoints

The secondary endpoints will be summarized by treatment and, where appropriate, by time point and treatment. Treatment groups will be compared through the use of appropriate statistical methods. The particular method used will depend on the scale of measurement of the endpoint. The accompanying table lists each endpoint, the applicable analysis sample, measurement scale (continuous, dichotomous, or ordinal), and the summary statistics and statistical model to be used for treatment group comparisons. Rejection and renal function endpoints will be based on central laboratory results; graft failure, delayed graft function, and slow graft function endpoints will be based on data recorded in the clinical database. Endpoints denoted with a '#' were not part of the protocol-specified endpoints and were added during finalization of this document.

Secondary analyses will be considered descriptive.

Summary of Proposed Analyses for Secondary Clinical Endpoints

Response Type (Sample)	Response	Measurement Scale	Summary Statistics	Models to test for treatment effects
Incidence and Severity of Rejection Endpoints (ITT, PP)	Note: All biopsies (protocol, clinically indicated, and standard of care) will be utilized for rejection endpoints.			
	Proportion of subjects with biopsy proven acute cellular rejection (BPAR) within a) 6 months and b) 2 years of transplant	dichotomous	proportion + 95% CI	Chi-square test using a window of +21 days at month 6. Month 24 will use a window of + 1 month if the biopsy occurred before 01MAR2020. Otherwise, the +6 month window used during the COVID-19 pandemic will be used for month 24.
	BANFF grades of first Acute Cellular Rejections (ACR) within 6 months of transplant, among subjects who have ACR	categorical	Table	Chi-square test using a window of +21 days at month 6
	Proportion of subjects with BPAR or borderline rejection within a) 6 months and b) 2 years of transplant	dichotomous	proportion + 95% CI	Chi-square test using the same windows as outlined for BPAR
	Proportion of subjects with biopsy proven acute antibody mediated rejection (AMR) within a) 6 months and b) 2 years of transplant	dichotomous	proportion + 95% CI	Chi-square test using the same windows as outlined for BPAR

	BANFF grades of first AMR within 6 months of transplant, among subjects who have acute AMR	categorical	Table	Chi-square test using a window of +21 days at month 6
	Proportion of subjects with biopsy proven AMR or suspicious for AMR within a) 6 months and b) 2 years of transplant	dichotomous	proportion + 95% CI	Chi-square test using the same windows as outlined for BPAR
	# Proportion of subjects with locally treated rejection, defined as treatment administered for rejection based on clinical signs or biopsy findings, within a) 6 months and b) 2 years of transplant	dichotomous	proportion + 95% CI	Chi-square test using the same windows as outlined for BPAR
	Proportion of subjects with BANFF chronicity scores ≥ 2 on 24 month biopsy	dichotomous	proportion + 95% CI	Chi-square test using the biopsy reported from Visit 14/Month 24
	Change in BANFF chronicity scores between implantation and 24 month biopsies	categorical	Table	CMH test using the difference of M24 minus BL categorized as 0, 1, 2, or 3+.
	Days from transplantation until event (ACR, AMR, or hospitalization for infection and or malignancy)	time to event	Kaplan-Meier estimates	Kaplan-Meier model with log-rank test
Renal Function Endpoints (ITT, PP)	Note: Any creatinine sample collected on or after the date of graft failure will be excluded from the eGFR calculation for renal function endpoints.			
	Change in eGFR (as measured by both MDRD and CKD-EPI[1]) between 3 months and 24 months	continuous	mean + 95% CI	T-test using the 3 and 24 month scheduled collections. Change calculated two ways using M24 – M3: (1) using the M3 (+/- 14 days) and M24 (+/- 1 month) collections and (2) using all M3 and M24 collections, regardless of timing post-transplant.

	# Change in eGFR (as measured by both MDRD and CKD-EPI[1]) between 6 months and 24 months	continuous	mean + 95% CI	T-test using the 6 and 24 month scheduled collections. Change calculated two ways using M24 – M6: (1) using the M6 (+/- 21 days) and M24 (+/- 1 month) collections and (2) using all M6 and M24 collections, regardless of timing post-transplant.
	Change in eGFR (as measured by both MDRD and CKD-EPI[1]) between post-transplant nadir (lowest eGFR in first 6 months) and 24 months	continuous	mean + 95% CI	T-test using scheduled collections from the first 6 months and month 24. Nadir will be calculated two ways using the values from month 1 to month 6: (1) using only collections from M1 (-7 days) to M6 (+21 days) and (2) using any collections from the M1 visit through the M6 visit, regardless of timing post-transplant. The M1 eGFR value will be excluded from the nadir determination for any subject with DGF. Then change will be calculated two ways using M24 – nadir: (1) using nadir #1 and M24 (+/- 1 month) collections and (2) using nadir #2 and all M24 collections, regardless of timing post-transplant.
	In addition to the planned analysis as indicated above for each of the endpoints (i.e., change in eGFR from month 3 to month 24, month 6 to month 24, and the 6 month nadir to month 24), an exploratory analysis will report the changes estimated using the same MMRM used for analysis of the primary endpoint.			
	Estimated GFR (as measured by both MDRD and CKD-EPI[1]) on days 7, 30, 90, and 180 post-transplant	continuous	mean + 95% CI	MMRM, same as used for analysis of the primary endpoint to generate estimates at days 30, 90, and 180. Day 7 eGFR value will be added to the model to generate an estimate at day 7.

Graft Failure Endpoints (ITT, PP)	Proportion of subjects with death or graft failure within 2 years	dichotomous	proportion + 95% CI	Chi-square test using a window of +1 month
	Proportion of subjects with only graft failure within 2 years	dichotomous	proportion + 95% CI	Chi-square test using a window of +1 month
Delayed Graft Function Endpoints (ITT, PP)	Proportion of subjects that required at least one dialysis treatment within the first week after transplantation	dichotomous	proportion + 95% CI	Chi-square test
(ITT)	Number of dialysis sessions in the first 8 weeks post-transplantation	continuous	mean and/or geometric mean + 95% CI	Poisson regression
(DGF)	Duration of DGF defined as time from transplantation to the last required dialysis treatment	continuous	mean and/or geometric mean + 95% CI	T-test
(ITT, PP)	The incidence of primary non-function (PNF), defined as dialysis-dependency for more than 3 months	dichotomous	proportion + 95% CI	Chi-square test using only subjects who are still on study at 3 months post-transplant (i.e., day 90)
(ITT, PP)	Change from baseline (immediately after surgery) in serum creatinine at 24, 48, and 72 hours	continuous	mean + 95% CI	MMRM with treatment group, visit number, and their interaction term as independent variables
Slow Graft Function (ITT, PP)	Note: Any subject who experienced DGF will be excluded from assessment of SGF endpoints.			
	The proportion of subjects with a serum creatinine of more than 3 mg/dL at day 5 post-transplant	dichotomous	proportion + 95% CI	Chi-square test using creatinine from the discharge visit, if subject discharged prior to day 5
	Creatinine reduction ratio (CRR) on day 2 (defined as the first creatinine on day 2 divided by the first creatinine after surgery)	ratio	mean + 95% CI	T-test
	Creatinine reduction ratio (CRR) on day 5 (defined as the first creatinine on day 5 divided by the first creatinine after surgery)	ratio	mean + 95% CI	T-test using creatinine from the discharge visit, if subject discharged prior to day 5
	The proportion of subjects whose day 5 serum CRR was less than 70%	dichotomous	proportion + 95% CI	Chi-square test using creatinine from the discharge visit, if subject discharged prior to day 5

	The proportion of subjects whose day 2 serum CRR was less than 30%	dichotomous	proportion + 95% CI	Chi-square test
	Proportion of subjects who need dialysis after 1 week	dichotomous	proportion + 95% CI	Chi-square test
Alloantibody Endpoints (ITT, PP)	# Proportion of subjects with de novo DSA within 24 months	dichotomous	proportion + 95% CI	Chi-square test applying the same window as outlined for the BPAR endpoint to the alloantibody samples

[1] CKD-EPI equation: 141 x min(Scr/k,1)^{a} x max(Scr/k,1)$^{-1.209}$ x 0.993^{Age} x 1.018 [if female] x 1.159 [if black], where Scr is serum creatinine (mg/dL), k is 0.7 for females and 0.9 for males, a is -0.329 for females and -0.411 for males, min indicates the minimum of Scr/k or 1, and max indicates the maximum of Scr/k or 1.

7.3.2. Safety and Complication Endpoints

The safety sample will be used for analysis of all post-transplant safety endpoints. The specific clinical safety events of interest are infections, hospitalizations, malignancies, and wound healing. The table below lists each safety endpoint, the analysis sample, measurement scale (continuous, dichotomous, or ordinal), and the summary statistics and statistical model that will be used for treatment group comparisons. All safety endpoints will be assessed during the first 24 months post-transplant, using a window of +1 month.

Summary of Proposed Analyses for Safety Endpoints

Response Type (Sample)	Response	Measurement Scale	Summary Statistics	Models to test for treatment effects
Safety Endpoints (Safety)	Proportion of subjects with any infection requiring hospitalization or resulting in death	dichotomous	proportion + 95% CI	Chi-square test
	Proportion of subjects with mycobacterial or fungal infections	dichotomous	proportion + 95% CI	Chi-square test
	Proportion of subjects with CMV viremia that require a change in immunosuppression or anti-viral treatment as per standard of care at the site	dichotomous	proportion + 95% CI	Chi-square test
	Proportion of subjects with BK viremia that require a change in immunosuppression or anti-viral treatment as per standard of care at the site	dichotomous	proportion + 95% CI	Chi-square test
	Proportion of subjects with malignancy	dichotomous	proportion + 95% CI	Chi-square test

	Proportion of subjects with impaired wound healing manifested by wound dehiscence, wound infection, or hernia at the site of the transplant incision	dichotomous	proportion + 95% CI	Chi-square test

7.3.3. Secondary Mechanistic Endpoints

Appendix 1 of the protocol contains the mechanistic assay collection schedule.

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

- Sensitivity, specificity, PPV, and NPV of biomarkers, including PRT, urinary CXCL9, blood genomic profile, and 3-month allograft genomic profile (alone and/or in combination) to predict:
  - Biopsy-proven acute rejection
  - Graft loss
  - Chronic graft injury, as measured by 2-year eGFR
- Each of the following:
  - Inflammatory gene expression profiles
  - Frequency of donor reactive T-cells
  - Frequency and function of Treg
  - Fibrogenic gene expression profiles
  - Amount of peritubular capillary loss by histology

We will test a biomarker panel to predict and diagnose incipient ACR and to determine the accuracy of early post-transplant markers to define individuals most at risk for late graft dysfunction and chronic graft injury. This panel will include reactive T cell (PRT) assays, urinary chemokine, urinary PCR assays, gene expression profiles in peripheral blood, and gene expression profiles in graft tissue. If the comparison of the biomarkers between the two treatment groups (see previous paragraph) yields no differences, we will utilize the entire study population. If, however, we find an influence of the infliximab on the biomarkers, we will limit this investigation to the control subjects. This evaluation will be based primarily on examining the Receiver Operating Characteristic (ROC) Curves generated for each candidate biomarker and the associated classification accuracy in terms of sensitivity and specificity as well as positive and negative predictive value. We will also explore combining the candidate biomarkers to determine whether there are distinct combinations of biomarkers that may improve performance relative to individual biomarkers.

Our hypothesis that anti-TNFa limits activation of donor reactive immunity and prevents initiation and progression of graft fibrosis by dampening early inflammation within the allograft will be assessed by a series of linear mixed models. The assays to assess these effects will include T cell (PRT) assays, urinary chemokine, urinary PCR assays, gene expression profiles in peripheral blood, gene-expression profiles in graft tissue, donor-reactive effector T cells, markers of inflammation, regulatory T cells, donor specific antibody, fibrogenic gene profiles, and peritubular capillary loss. Models will be fit sequentially, with specific models fit to address specific elements of the hypothesized mechanism. Initially we will use regression models to quantify the extent to which infliximab, when added to therapy with ATG, reduces “early” inflammation at the time of transplantation, measured both by cellular reactivity in blood and from post-implantation biopsies. Corresponding linear regression models will assess differences between treatment groups with respect to inflammation measures at 24 mos., and mixed linear models will be used to assess the inflammation measures over the entire time course

subsequent to the "early" evaluation. Additionally, proportional odds models will be used to assess treatment differences in ordinal grade of graft fibrosis at 24 mos. and, through a generalized estimating equation (GEE) approach, over time.

7.4. Exploratory Endpoints

The following exploratory endpoints will be assessed using the ITT sample.

Variable: Standard deviation of the monthly tacrolimus trough levels from 6 months post-transplant to 2 years post-transplant

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 beginning at month 6 post-transplant. This estimate will be a rolling estimate (continually re-assessed with each subsequent outpatient trough level) and may use up to 1 year worth of trough levels at any given time. The estimated standard deviations will be analyzed within the same MMRM framework as the primary endpoint. The estimated mean of the standard deviation at 2 years post-transplant, treatment group difference, and two-sided 95% confidence intervals will be summarized. Additionally, a cross-sectional analysis will be performed using only the month 24 estimated standard deviation for each subject with available data and will estimate the mean of the standard deviation, treatment group difference, and two-sided 95% confidence intervals for comparison to the MMRM results.

Variable: Percentage of predicted prednisone bottle openings as measured by a medication event monitoring system (MEMS®) in the first 3 months post-transplant

Analysis: The number and percentage of subjects who participated in the MEMS study will be summarized at day 14 and months 1 and 3 post-transplant by treatment group and overall. Additionally, the number of pauses in the use of the MEMS cap at months 1 and 3 will be determined for each subject participating in the MEMS study and summarized by treatment group and overall.

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, and changes in routine laboratory measures.

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

8.2. Adverse Events

Per the protocol, only non-serious AEs with a severity grade of 3 or higher were collected, and all serious AEs were collected. All AEs were classified by system organ class (SOC) and preferred term (PT), according to the Medical Dictionary for Regulatory Activities (MedDRA) version 18.1. The severity of AEs was classified using the National Cancer Institute's (NCI's)

Common Toxicity Criteria for Adverse Events (CTCAE) version 4. Each AE was recorded once at the highest severity.

AEs were collected for each subject from the initiation of investigational treatment through study completion or early termination. Based on this reporting period, all AEs will be considered treatment-emergent. All data tabulations will be stratified by treatment group.

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 with an outcome of death
- AEs that were assessed as being related to infliximab/placebo, thymoglobulin, tacrolimus, MMF, prednisone, blood draw, or biopsy
- AEs by severity

In general, when reporting on the relationship of an adverse event to study drug and procedures, the DAIT medical monitor 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, hematology, and urine. 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 either as a spaghetti plot where each subject's values will be plotted and connected by line segments,

forming one line per subject, or as box plots over time and stratified by treatment group. Tests with qualitative results will not be plotted.

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

8.5. Vital Signs 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 were 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. Other Safety Measures

If present, a listing will be produced detailing any reported instances of PTLD or other malignancies.

9. OTHER ANALYSES

Non-immunosuppressive medications reported on the CRF will be categorized for analysis as prior, concomitant, or after study treatment by comparing the medication start and stop dates with the study drug infusion start and stop date/time. Prior medications will have both the medication start and stop dates prior to the study drug infusion date. After medications will have both the medication start and stop dates after the study drug infusion date. All other medications will be classified as concomitant, indicating that use of the medication overlapped with use of the study drug infusion.

The number and percentage of subjects receiving prior, concomitant, and after medications will be presented overall and by medication class within each treatment group for the ITT sample.

Medications will be captured in the following categories: beta blockers, diuretics, ACE/ARB inhibitors, anti-adrenergics, calcium channel blockers, other antihypertensives, statins, non-statins, insulin (rapid-acting, short-acting, intermediate-acting, long-acting), oral hypoglycemia agents, hemapoietic support, and aspirin.

10. INTERIM ANALYSES AND DATA MONITORING

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

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

11. REFERENCES

1. Little R, Yau L. Intent-to-treat analysis for longitudinal studies with drop-outs. Biometrics 1996;52:1324-1333.
2. Ma G, Troxel AB, Heitjan DF. An index of local sensitivity to non-ignorable drop-out in longitudinal modelling. Stat Med. 2005 Jul 30;24(14):2129-50.

3. Kenward M, Multiple Imputation and its Application James Carpenter, ISBN: 978-0-470-74052-1 February 2013.
4. Carpenter, J. R., and Kenward, M. G. , Multiple Imputation and its Application (Statistics in Practice), Wiley. 2010.