

darTreg Statistical Analysis Plan

Donor-Reactive Regulatory T Cell Therapy in Liver Transplantation

Version: 1.0

Date: 06MAY2020

Study Sponsor: The National Institute of Allergy and Infectious Diseases (NIAID)

NCT02188719

PREPARED BY

Rho, Inc.

A series of five horizontal black bars of varying lengths, with the text 'Rho, Inc.' positioned to the left of the first bar.

Confidential

This document is confidential and proprietary to the National Institute of Allergy and Infectious Diseases. Acceptance of this document constitutes agreement by the recipient that no unpublished information contained herein will be reproduced, published, or otherwise disclosed without the prior written approval of the National Institute of Allergy and Infectious Diseases, except that this document may be disclosed to appropriate Institutional Review Boards under the condition that they keep the information confidential.

APPROVALS

Approved:

Treg Principal Investigator

Date:

NIAID Medical Monitor

Date:

Rho Senior Reviewer

Date:

Rho Senior Statistician

Date:

TABLE OF CONTENTS

1. INTRODUCTION	7
2. STUDY DESIGN	7
2.1 Safety Outcomes	9
2.2 Safety Objective	9
5. ANALYSIS SAMPLES.....	11
5.1 Screen Failure Analysis	11
6. DEMOGRAPHIC AND OTHER BASELINE CHARACTERISTICS	11
7. INTERIM ANALYSES AND DATA MONITORING.....	12
8. STATISTICAL METHODOLOGY AND STATISTICAL ANALYSES.....	12
8.1 Analysis Methods	12
9. ADDITIONAL SAFETY EVALUATIONS.....	13
9.1 Overview of Safety Analysis Methods.....	13
9.2 Adverse Events	13
9.3 Deaths/Graft Losses, Serious Adverse Events, and Other Significant Adverse Events	13
9.4 Clinical Laboratory Evaluation	14
9.5 Vital Signs, Physical Findings, and Other Observations Related to Safety..	14
9.5.1 Vital Signs	14
9.5.2 Physical Examinations	14
9.5.3 Protocol/For-Cause Biopsies	14
9.5.4 Creatinine/GFR	14

List of Abbreviations

AE	Adverse Event
ALT	Alanine Aminotransferase
CFR	Code of Federal Regulations
CI	Confidence Interval
CMV	Cytomegalovirus
CNI	Calcineurin Inhibitor
CRF	Case Report Form
CTCAE	Common Terminology Criteria for Adverse Events
DAIT	Division of Allergy, Immunology, and Transplantation
dr	Donor Reactive
darTregs	Donor-Reactive T Regulatory Cells
DSMB	Data Safety Monitoring Board
EBV	Epstein Barr Virus
EVR	Everolimus
FACS	Fluorescence Activated Cell Sorting
FDA	Food and Drug Administration
GCP	Good Clinical Practice
GFR	Glomerular Filtration Rate
GMP	Good Manufacturing Practice
GVHD	Graft Versus Host Disease
HBV	Hepatitis B Virus
HCC	Hepatocellular Carcinoma
HCV	Hepatitis C Virus
HIV	Human Immunodeficiency Virus
ICH	International Conference on Harmonization
ICU	Intensive Care Unit
IgG	Immunoglobulin G
IND	Investigational New Drug
IRB	Institutional Review Board
IS	Immunosuppression
MELD	Model for End Stage Liver Disease
MFC	Multiparameter Flow Cytometry
MIHC	Multiplex Immunohistochemistry
MLR	Mixed Lymphocyte Reaction
AE	Adverse Event
ALT	Alanine Aminotransferase
CFR	Code of Federal Regulations
CI	Confidence Interval
CMV	Cytomegalovirus
CNI	Calcineurin Inhibitor
CRF	Case Report Form
CTCAE	Common Terminology Criteria for Adverse Events
DAIT	Division of Allergy, Immunology, and Transplantation
dr	Donor Reactive
darTregs	Donor-Reactive T Regulatory Cells

DSMB	Data Safety Monitoring Board
EBV	Epstein Barr Virus
EVR	Everolimus
FACS	Fluorescence Activated Cell Sorting
FDA	Food and Drug Administration
GCP	Good Clinical Practice
GFR	Glomerular Filtration Rate
GMP	Good Manufacturing Practice
GVHD	Graft Versus Host Disease
HBV	Hepatitis B Virus
HCC	Hepatocellular Carcinoma
HCV	Hepatitis C Virus
HIV	Human Immunodeficiency Virus
ICH	International Conference on Harmonization
ICU	Intensive Care Unit
IgG	Immunoglobulin G
IND	Investigational New Drug
IRB	Institutional Review Board
IS	Immunosuppression
CTCAE	Common Terminology Criteria for Adverse Events
DAIT	Division of Allergy, Immunology, and Transplantation
dr	Donor Reactive
darTregs	Donor-Reactive T Regulatory Cells
DSMB	Data Safety Monitoring Board
EBV	Epstein Barr Virus
EVR	Everolimus
FACS	Fluorescence Activated Cell Sorting
FDA	Food and Drug Administration
GCP	Good Clinical Practice
GFR	Glomerular Filtration Rate
GMP	Good Manufacturing Practice
GVHD	Graft Versus Host Disease
HBV	Hepatitis B Virus
HCC	Hepatocellular Carcinoma
HCV	Hepatitis C Virus
HIV	Human Immunodeficiency Virus
ICH	International Conference on Harmonization
ICU	Intensive Care Unit
IgG	Immunoglobulin G
IND	Investigational New Drug
IRB	Institutional Review Board
IS	Immunosuppression
MELD	Model for End Stage Liver Disease
MFC	Multiparameter Flow Cytometry
MIHC	Multiplex Immunohistochemistry
MLR	Mixed Lymphocyte Reaction
MMF	Mycophenolate Mofetil
mSAP	Mechanistic Statistical Analysis Plan

mToR	Mammalian Target of Rapamycin
NIAID	National Institute of Allergy and Infectious Disease
PI	Principal Investigator
PTLD	Post -Transplant Lymphoproliferative Disorder
PBMC	Peripheral Blood Mononuclear Cell
SACCC	Statistical and Clinical Coordinating Center
SAE	Serious Adverse Event
SAP	Statistical Analysis Plan
SAR	Suspected Adverse Reaction
sBc	Stimulated B Cell
SOC	Standard of Care
SOP	Standard Operating Procedure
SWFI	Sterile Water for Injection
TAC	Tacrolimus
Tconv	Conventional T Cell
Treg	Regulatory T Cell
TSDR	Treg-Specific Demethylation Region
ULN	Upper Limit of Normal
UNOS	United Network for Organ Sharing

1. INTRODUCTION

This statistical analysis plan (SAP) only includes analyses related to the clinical endpoints outlined in the protocol. Mechanistic analyses will be performed at the Immune Tolerance Network (ITN), and a separate mechanistic statistical analysis plan (MSAP) will be created to detail the planned analyses. Relevant clinical data from the study will be submitted to the ITN Biomarker and Discovery Research (BDR) and ITN Bioinformatics Groups (BiG) to permit the mechanistic data analyses.

1.1 Background and Scientific Rationale

Although ongoing refinement of immunosuppression (IS) regimens has substantially reduced the incidence of acute rejection after solid organ transplant, long-term outcomes have stagnated partly due to morbidity and mortality associated with generalized, lifelong IS. The traditional approach to IS has emphasized non-specific suppression of T cell responses. The more recent elucidation of regulatory T cells (Tregs) and their importance in regulating immune responses has encouraged the reconfiguration of IS regimens to favor Treg development and function with the ultimate goal of inducing graft tolerance (Waldmann, 2008) (Kang, 2007) (Walsh, 2004) (Yeung, 2009) (Sanchez-Fueyo, 2006) (Sagoo, 2008) (Long, 2009). Multiple animal models have shown that adoptive transfer of Tregs can mitigate graft rejection and, in combination with “Treg-supportive” IS regimens, can induce long-term tolerance (Kang, 2007) (Riley, 2009) (Issa, 2010) (Nadig, 2010). Treg-supportive IS regimens have included the initial de-bulking of donor-specific T cells. Thymoglobulin®, a commonly used T-cell depleting agent in transplantation, appears to relatively spare Tregs (Sewgobind, 2009), thereby increasing Treg: T conventional cell (Tconv) ratio. Additionally, mammalian Target of Rapamycin (mTOR) inhibitors, a class of drug to which everolimus (EVR) belongs, suppress effector T-cells while fostering Treg development (Demirkiran A. T., 2008) (Demirkiran A. V., 2009). We aim to translate these basic and clinical findings into a viable clinical protocol. We propose to test the use of donor reactive Tregs (darTregs) in the context of a Treg-supportive IS regimen as an approach to induce liver transplant tolerance. For several reasons, the liver transplant setting is ideal to evaluate the safety of Treg therapy as a strategy to either increase the likelihood of and/or accelerate the development of tolerance. First, liver allografts appear to be more tolerogenic than other allografts. Compared to recipients of other organs, liver transplant recipients require less IS and are relatively spared from both humoral and chronic rejection. Second, acute rejection, albeit common, is readily treated without long-term sequelae, a key advantage for testing novel immunomodulatory agents and tolerance induction strategies. Third, emerging data from IS withdrawal trials in liver transplant recipients indicate that the rate of spontaneous tolerance increases over time after transplantation, from less than 10% within 3 years to as high as 80% after 10 years (Bohne, 2012) (Sanchez-Fueyo A. , 2011) (Feng S. U., 2012). As the first step toward a long-term goal of Treg immunotherapy, we propose to determine the safety of darTregs in combination with a Treg-supportive IS regimen in adult, de novo liver transplant recipients. We hypothesize that ex vivo-expanded darTregs administered to adult, de novo liver transplant recipients in combination with a Treg-supportive IS regimen will be safe. Our study also aims to describe for the first time the persistence of administered darTreg in human, as well as to describe the effects of Treg-supportive IS on darTreg and drTconv after transplantation. If successful, our study will define an approach for the therapeutic administration of Tregs, establish a new paradigm for the design of IS regimens, and set the stage for subsequent efficacy trials.

2. STUDY DESIGN

This is a three center, open-label, dose escalation, pilot study in which adult subjects undergoing primary solitary liver transplant will receive a Treg-supportive IS regimen alone (Cohorts 1a and 1b) or a Treg-supportive IS regimen followed by a single infusion of autologous, darTregs (Cohort 2: 50 million

darTregs; Cohort 3: 200 million darTregs; or Cohort 4: 800 million darTregs). There are three successive stages in the study with specific eligibility criteria that participants must meet prior to proceeding to the next stage. As shown in *Figure 1*, the three stages are:

- Screening and Enrollment
- Treg Supportive IS Regimen and,
- darTreg Infusion

Subjects will be followed for 40 weeks after transplant, during which clinical data along with peripheral blood (PBMCs and serum) and liver biopsy samples will be collected and analyzed.

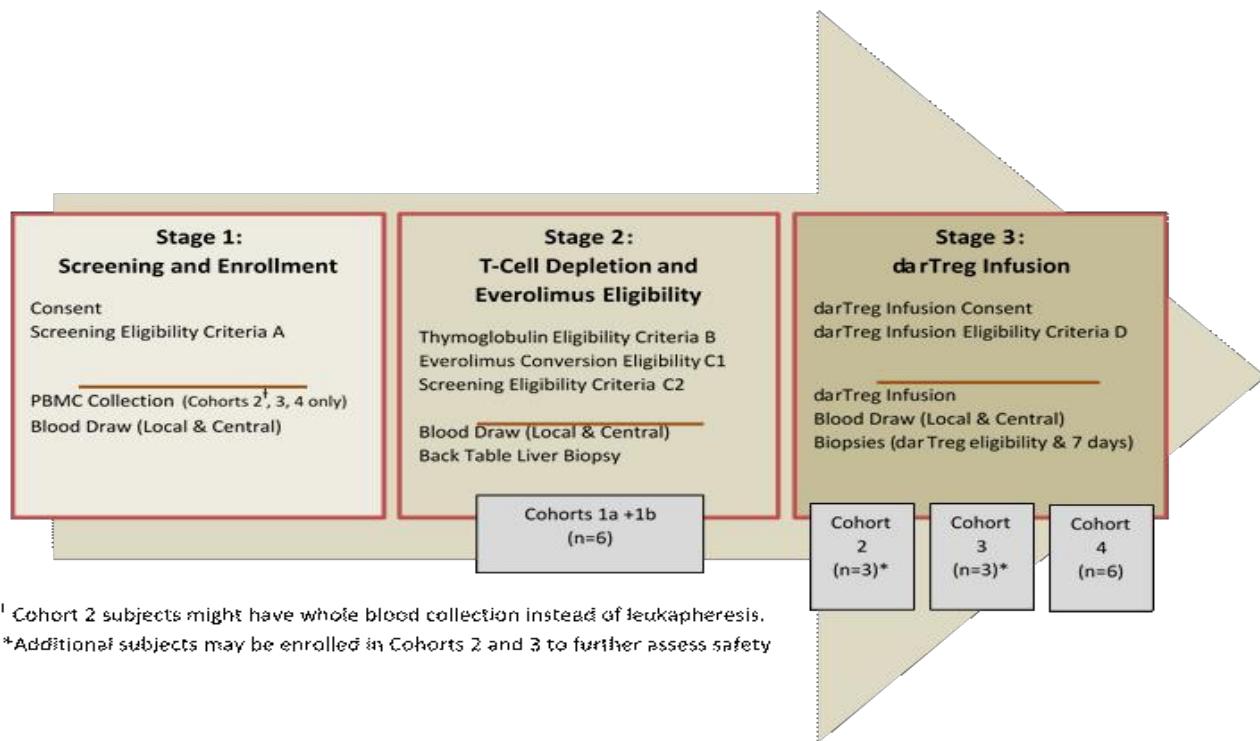


Figure 1. Three Stages of Trial

The four cohorts are summarized below:

- Cohort 1: 3 subjects from Center 1 (Cohort 1a) + 3 subjects from Center 2 (Cohort 1b); these subjects will only receive Treg supportive IS study therapy (Stage 1 and 2; they will not receive darTreg infusion (Stage 3)).
- Cohort 2: 3 subjects meeting eligibility criteria for all 3 Stages will receive 50 million darTregs (Dose A).
- Cohort 3: 3 subjects meeting eligibility criteria for all 3 Stages will receive 200 million darTregs (Dose B).
- Cohort 4: 6 subjects meeting eligibility criteria for all 3 Stages will receive 800 million darTregs (Dose C).

2.1 Safety Outcomes

The safety, tolerability, and dose limiting toxicities of darTreg therapy given within the context of a Treg supportive IS regimen will be evaluated with the rate of the following events within 40 weeks of transplantation:

Clinical outcomes that will be described are:

1. Incidence and severity of biopsy-proven acute and/or chronic rejection
2. Incidence of \geq Grade 3 infections as defined in Section 13.3.1
3. Incidence of wound complications (\geq CTCAE Grade 3)
4. Incidence of anemia, neutropenia, and/or thrombocytopenia (\geq CTCAE Grade 3)
5. Incidence of adverse events attributable to the darTreg infusion including infusion reaction/cytokine releaser syndrome (\geq CTCAE Grade 3), and malignant cellular transformation.

2.2 Safety Objective

This study will evaluate the safety, tolerability, and dose limiting toxicities (DLT's) of a Treg-supportive IS regimen and darTreg infusion for adult, de novo, liver transplant recipients.

3. Sample Size Considerations

This pilot study is designed to evaluate 18 subjects through Stage 3 in a modified 3+3 dose escalation design. Therefore, no formal power and sample size analysis have been performed. However, with 18 evaluable subjects undergoing darTreg supportive IS for 12 weeks of follow-up (6 subjects for 40 weeks of follow-up) and up to 12 evaluable subjects undergoing darTreg infusions for 28 weeks of follow-up, exact 95% confidence intervals on person-week incidence rates will vary in width by analysis period. For example, the following table (**Table 1**) shows incidence rates and confidence intervals for selected numbers of events and person-weeks of 216 (18 subjects x 12 weeks), 336 (12 subjects x 28 weeks), and 384 (6 subjects X 40 weeks and 12 subjects X 12 weeks) in the two analysis periods. It is not known at this time how many subjects may be expected to comprise the entire Safety Sample.

Table 1. Incidence Rates and Confidence Intervals for Selected Numbers of Events and Person-Weeks

Number of Events	Total Person-Weeks	Person-Weeks Incidence Rate	Lower 95% Confidence Limit	Upper 95% Confidence Limit
2	216	0.0093	0.0011	0.0334
4	216	0.0185	0.0050	0.0474
6	216	0.0278	0.0102	0.0605
8	216	0.0370	0.0160	0.0730
10	216	0.0463	0.0222	0.0851
12	216	0.0556	0.0287	0.0970
14	216	0.0648	0.0354	0.1087
16	216	0.0741	0.0423	0.1203
18	216	0.0833	0.0494	0.1317
20	216	0.0926	0.0566	0.1430
2	336	0.0060	0.0007	0.0215
4	336	0.0119	0.0032	0.0305
6	336	0.0179	0.0066	0.0389
8	336	0.0238	0.0103	0.0469

10	336	0.0298	0.0143	0.0547
12	336	0.0357	0.0185	0.0624
14	336	0.0417	0.0228	0.0699
16	336	0.0476	0.0272	0.0773
18	336	0.0536	0.0317	0.0847
20	336	0.0595	0.0364	0.0919
<hr/>				
2	384	0.0052	0.0006	0.0188
4	384	0.0104	0.0028	0.0267
6	384	0.0156	0.0057	0.0340
8	384	0.0208	0.0090	0.0410
10	384	0.0260	0.0125	0.0479
12	384	0.0313	0.0161	0.0546
14	384	0.0365	0.0199	0.0612
16	384	0.0417	0.0238	0.0677
18	384	0.0469	0.0278	0.0741
20	384	0.0521	0.0318	0.0804

4. 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.
- In general, in frequency tables of variables with different categories (e.g., for discontinuation reason, race): if no subjects belong to a certain category, including “Missing”, across all treatment groups, then the printing of this category should be suppressed. Otherwise, if the number of subjects in a category for a treatment group is zero, then a zero should be displayed for the number, and the percentage should be left blank.
- Numeric variables will be summarized using n, mean, standard deviation (SD), median, min, max. The min/max will be reported at 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. Descriptive statistics will be displayed in the order: n, mean, SD, median, min, max
- The median will be reported as the average of the two middle numbers if the dataset contains even numbers.
- 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, then report “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.

- In general, columns with character values will have the header and column values left aligned. Numeric columns will be centered on their decimal place with headers also centered.
- In the first column, if text wraps onto another line indent one additional space. For subgroups, indent two spaces.
- Units of measurement - International units SI will be used for clinical laboratory data as a standard presentation. The metric system will be used whenever possible. Thus, weight will be in kilograms and height in centimeters. Temperature will be presented in Celsius degrees.
- For general footnotes, 'Note:' will come before any bracketed footnotes.
- All listings will be sorted in order of treatment, subject, and time of assessment (e.g., visit, time, and/or event).

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.

5. ANALYSIS SAMPLES

The Safety Sample will be comprised of all study subjects who are consented and receive any study therapy, including leukapheresis, Treg-supportive IS and darTreg infusion. This excludes consented subjects who were terminated before receiving any study treatment.

Safety and clinical outcomes will be analyzed within the Safety Sample as a whole and within the following two treatment-related analysis periods during which:

- Subjects receive any darTreg supportive IS
- Subjects receive darTreg infusions at any dose.

Subjects will contribute adverse events and weeks at risk to each analysis period for as long as they are on the corresponding study treatment in that analysis period. Thus, if they progress from darTreg supportive IS to darTreg infusions at any dose, they will contribute subsequent adverse events and weeks at risk to the darTreg infusion analysis period when they initiate any darTreg infusion and that will continue until the end of their follow-up time on study.

5.1 Screen Failure Analysis

At each stage of the study, subject's eligibility will be re-evaluated. Reasons for study discontinuation for screen failures will be summarized at each stage, both overall and within each cohort.

6. DEMOGRAPHIC AND OTHER BASELINE CHARACTERISTICS

Summary descriptive statistics for baseline and demographic characteristics will be provided for all enrolled participants. Demographic data will include age, race, sex, body weight, and height; these data will be presented in the following manner:

- Continuous data (i.e., age, body weight, and height) will be summarized descriptively by mean standard deviation, median, and range.
- Categorical data (i.e., sex and race) will be presented as enumerations and percentages.

Both Recipient and Donor demographic and baseline characteristic data will also be presented in data listings by subject.

7. INTERIM ANALYSES AND DATA MONITORING

There are no interim analyses planned for this study.

The protocol chair, the ITN clinical trial physician, the NIAID medical monitor, and the NIAID Transplant Data and Safety Monitoring Board (DSMB) will periodically review safety data. Enrollment of subjects in the trial and/or potentially progression from one stage of the trial to the subsequent stage for current trial subjects may be suspended at any time if any of these reviews concludes that there are significant safety concerns. Stopping rules are described in Section 18.3.1 of the protocol.

The progress of the study will be monitored by the NIAID Transplant DSMB. The DSMB will be chartered to review safety data and to make recommendations regarding continuation, termination, or modification of the study. The DSMB will formally review the safety data at least yearly. The discontinuation of study treatment will also be periodically reported to the DSMB.

In addition, safety data will be reviewed by the DSMB when an event occurs that is of sufficient concern to the NIAID medical monitor or protocol co-chairs to warrant review, or when an event occurs that could contribute to a pre-defined stopping rule as specified in the Protocol.

Findings will be reported to IRBs and health authorities.

8. STATISTICAL METHODOLOGY AND STATISTICAL ANALYSES

8.1 Analysis Methods

All analyses will be run on the safety population with no sensitivity analyses, sub-group analyses, or covariate adjustments planned.

8.1.1 Safety Outcome Analyses

The safety outcomes will be considered in two ways: the absolute number and the person-week incidence rates (*Section 3.2*) of selected adverse events following transplantation. The person-week incidence rate will be estimated using proportions and exact binomial two-sided 95% confidence intervals for the Safety Sample as a whole and within the two analysis periods defined above. The absolute number and the estimated incidence proportions of the following will be reported:

- Biopsy-proven acute and/or chronic rejection
- Grade ≥ 3 infections as defined in Section 13.3.1 of the Protocol
- Wound complications (\geq CTCAE Grade 3)
- Anemia, neutropenia, and/or thrombocytopenia (\geq CTCAE Grade 3)
- Adverse events attributable to the darTreg infusion including infusion reaction/cytokine releaser syndrome (\geq CTCAE Grade 3), and malignant cellular transformation.

Descriptive statistics will also be presented for:

- The severity of biopsy-proven acute and/or chronic rejection
- Reasons for termination of subjects who did not receive leukapheresis
- Use of concomitant medications

9. ADDITIONAL SAFETY EVALUATIONS

9.1 Overview of Safety Analysis Methods

All safety analysis will be carried out using the safety sample defined in Section 5 above unless otherwise noted. Missing safety information will not be imputed. Safety will be analyzed through the reporting of adverse events, vital signs, physical examinations, and changes in routine laboratory values.

9.2 Adverse Events

All adverse events will be classified by system organ class and preferred term, according to a standardized thesaurus (MedDRA version 17.0). The severity of AEs will be classified using the National Cancer Institute's Common Toxicity Criteria for Adverse Events, with the exception of infections which will be graded according to protocol definitions.

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 (SAEs)
- AEs that lead to study discontinuation
- AEs with an outcome of death
- AEs reported by maximum severity

In addition, adverse events classified by MedDRA system organ class and preferred term will be summarized for each study cohort and overall for each of the following:

- All AEs
- AEs by maximum severity
- AEs by study stage

The summary tables will present the total number of events as well as the number and percentage of subjects experiencing the events. When reporting the number of AEs, if the same AE occurs for a subject on multiple occasions the event will be counted once for each occurrence. When reporting the number of subjects experiencing the events, a subject will only be counted once if they ever experience an event within the particular system organ class or preferred term. Percentages will be based on the number of subjects in the analysis population.

A summary table of incidence rates adverse events classified by system organ class and MedDRA preferred term will be provided. Incidence rate for a system organ class or preferred term will be defined as the number of adverse events divided by the person-time since the subject enrolled.

Separate data listings will be provided for AEs leading to study discontinuation.

9.3 Deaths/Graft Losses, Serious Adverse Events, and Other Significant Adverse Events

SAEs will be listed and summarized in the same manner described in section 9.2. Separate displays listing and summarizing death, graft loss, grade 3 or higher infections (per protocol definitions), severe acute cellular rejection (histological or clinical), chronic rejection, PTLD, and malignancies (excluding recurrent HCC and skin cancer), will also be created.

9.4 Clinical Laboratory Evaluation

Data listings will be provided for clinical laboratory measurements including serum chemistry, hematology, and liver tests. They will be sorted by subject ID, laboratory parameter and time of assessment. Results will be standardized to the international system of units (SI), where possible. Laboratory normal ranges for serum chemistry and hematology results will be included and out-of-range flags (H or L) will be used to denote abnormal values. Standard reference ranges will be used instead of site specific normal ranges. Liver tests will be compared to the subject's baseline test values to determine abnormal values.

Selected serum chemistry and hematology tests with numeric results will be plotted to show patterns over time. Tests with qualitative results (such as 'present' or 'positive') will not be plotted.

9.5 Vital Signs, Physical Findings, and Other Observations Related to Safety

9.5.1 Vital Signs

Descriptive statistics of vital signs results and change from baseline of vital signs will be summarized for each study group and overall. Data listings will be provided for vital signs measurements. They will be sorted by treatment group, subject, vital sign parameter and time of assessment.

9.5.2 Physical Examinations

Physical examination results of normal, abnormal, and not done will be summarized as frequencies and percentages by body system and visit. Data listings will be provided for physical examination results and sorted by treatment group, subject, body system and time of assessment.

9.5.3 Protocol/For-Cause Biopsies

Protocol and for-cause biopsies will be summarized as frequencies for each cohort and overall at each timepoint. Data listings for each individual subject will be provided with detailed biopsy information (i.e.- fibrosis, inflammation, diagnosis, etc.).

9.5.4 Creatinine/GFR

Creatinine/GFR values will be plotted over time for each subject to assess kidney function. Data listings of GFR values will be provided for all subjects with at least one abnormal GFR result.