

**T_{REG}, CD4 T_{EM}, AND CD8 T_{EMRA} FAIL TO PREDICT ACUTE CELLULAR REJECTION
IN LIVING-DONOR RENAL ALLOGRAFT RECIPIENTS**

by

John T. Walters

BS Chemistry, BA History and Philosophy of Science, University of Pittsburgh, 2008, 2008

Submitted to the Graduate School of
the Department of Biostatistics
Graduate School of Public Health in partial fulfillment
of the requirements for the degree of
Master of Science

University of Pittsburgh

2016

UNIVERSITY OF PITTSBURGH
GRADUATE SCHOOL OF PUBLIC HEALTH

This thesis was presented

by

John T. Walters

It was defended on

December 2, 2016

and approved by

Ada O Youk, PhD, Associate Professor, Department of Biostatistics, Graduate School of
Public Health, University of Pittsburgh

Diana Metes, MD, Professor, Department of Surgery and Immunology, School of Medicine,
University of Pittsburgh

Thesis Director: Jeanine M. Buchanich, M.Ed., Ph.D., Research Assistant Professor,
Department of Biostatistics, Graduate School of Public Health, University of Pittsburgh

Copyright © by John T. Walters

2016

Jeanine M. Buchanich, MEd, PhD

T_{REG}, CD4 T_{EM}, AND CD8 T_{EMRA} FAIL TO PREDICT ACUTE CELLULAR REJECTION IN LIVING-DONOR RENAL ALLOGRAFT RECIPIENTS

John T. Walters, MS

University of Pittsburgh, 2016

ABSTRACT

Since it was pioneered successfully, renal transplantation remains the only option for patients with stage-5 chronic, severe, and end-stage renal disease for whom dialysis treatment complications preclude its continued use. Currently, and historically, the supply of suitable living-donor allografts is far less than their clinical need, and this gap cannot be offset by transplanting available cadaveric kidneys. Immuno-suppressive and -induction therapies have been used to impair the recipients' immune response against the allograft. Without such therapies many recipients' immune systems would reject the organ, causing the patient to again experience renal failure. It is therefore incumbent upon the public health community to ascertain the most effective treatment modality for transplanted organs to ensure that these limited resources are utilized in the most efficient manner possible.

Thymoglobulin and Basiliximab induction therapies are two induction treatments available clinically for kidney transplantation. Within this study, patients undergoing treatment using either of these induction-agents had their circulating T cell phenotypes analyzed and compared. The goal of this study was to produce a statistical model, based on Generalized Estimating Equation methodology, that would predict episodes of acute cellular rejection (ACR) between the day of the transplant and up to one year after transplantation. The selection of potential covariates was based upon previously identified T cell markers.

Due to small sample size, missing data, and both left- and right-truncation of the data, the model was not able to discriminate between patients that underwent ACR from those that did not based upon the *a priori* markers of interest. However, the analysis did identify different memory T cell proportions that may be predictive of ACR in patients treated with Thymoglobulin or Basiliximab and warrants further study. The public health impact of a predictive model would be to increase the quality and duration of life in individual patients and reduce the burden of End Stage Renal Disease (ESRD) within the US population as a whole.

TABLE OF CONTENTS

1.0	INTRODUCTION.....	1
2.0	CLINICAL BACKGROUND FOR RENAL TRANSPLANTATION.....	4
2.1	RENAL ALLOGRAFT OUTCOMES.....	5
2.2	LYMPHOCYTE MARKERS.....	6
2.2.1	CD3, CD4, and CD8	6
2.2.2	CD45RO, CD45RA, and CD62L	7
2.2.3	FoxP3 and CD25	7
2.2.4	T Cell Subsets of Interest	8
3.0	ANALYSIS METHODOLOGY	9
3.1	GENERALIZED LINEAR MODELS.....	9
3.1.1	GLM Assumptions.....	10
3.1.2	GLM Limitations.....	11
3.2	GENERALIZED ESTIMATING EQUATIONS.....	11
3.2.1	GLM and GEE Similarities	12
3.2.2	Marginal GEE Models	12
3.2.3	GEE Assumptions.....	13
3.2.4	GEE Correlation Matrices.....	14
3.3	MODEL SELECTION CRITERIA.....	15
3.4	GEE MODEL APPLICATION.....	16
3.5	MISSING DATA AND GEE ANALYSIS	17
3.5.1	MCAR.....	17

3.5.2	MAR.....	18
3.5.3	MNAR.....	18
3.5.4	Study Data and the MCAR Assumption	18
4.0	STUDY DESIGN AND IMPLEMENTATION.....	20
4.1	STUDY HYPOTHESES.....	20
4.2	STUDY IMPLEMENTATION.....	21
4.3	BLOOD SAMPLE ACQUISITION.....	22
4.4	PHENOTYPIC ANALYSIS OF LYMPHOCYTES.....	23
4.4.1	Leukocyte Preparation.....	23
4.4.2	Data Acquisition	24
5.0	RESULTS	25
5.1	DATA DESCRIPTION	25
5.2	MISSING DATA.....	25
5.3	PARAMETER SELECTION	28
5.4	MODEL CONSTRUCTION.....	29
6.0	DISCUSSION	32
6.1	ANALYSIS LIMITATIONS.....	33
7.0	CONCLUSION.....	37
	BIBLIOGRAPHY	38

LIST OF TABLES

Table 2.1: T cell Memory Subsets Defined by CD45RO and CD62L	7
Table 3.1: Examples of GEE Correlation Matrices	15
Table 5.1: Study Participant Demographics	25
Table 5.2: GEE Single-Parameter Model Values	29
Table 5.3: Model A Values	29
Table 5.4: Model B Values	30
Table 5.5: Model C Values	30
Table 5.6: Model D Values	30
Table 5.7: Model E Values	31
Table 5.8: Model F Values.....	31

LIST OF FIGURES

Figure 4.1: Collection Gates	24
Figure 5.1: Collected Data	27

1.0 INTRODUCTION

Chronic stage-5 kidney disease, also called end-stage renal disease (ESRD), is a persistent disease endemic within the United States. From 1996 to 2006 the incidence rate (per million per year) of ESRD increased from 328.1 to 386.4. Since its peak in 2006, the incidence rate fell to 352.2 in 2013; however, the incident cases of ESRD have remained steady over the same time period at roughly 101,000 cases per year. As of 2013, ESRD's four main causes were diabetes, hypertension, glomerulonephritis, and cystic kidney disease comprising 53.4%, 34.9%, 9.2%, and 2.6% of new cases, respectively. As a result of the steady incidence and persistent nature of ESRD, as well as an etiology which includes a factor comorbid with another increasing public health concern, obesity, the prevalence of ESRD has increased steadily each year from 1996 to 2013. The primary means of treating ESRD is via hemodialysis or, more infrequently, peritoneal dialysis. Because dialysis treatments only address the symptoms of ESRD, these modalities do not reduce the prevalence of the disease. Renal transplantation, however, addresses the cause of ESRD by providing the patient with a functional kidney, and eliminates the need for regular dialysis treatments.^{[1], [2]}

The first successful human kidney transplant was performed in 1954 by Dr. Joseph E. Murray in Boston's Peter Bent Brigham Hospital (now Brigham and Women's Hospital).^[3] While it was a remarkable feat of surgical technique, it was only a long-term successful transplant because the patient was given a kidney from his identical twin brother, meaning that the transplanted organ

was genetically identical (syngeneic) to the host.^[4] As a result, the new kidney did not face the typical host vs. allogenic (of different genetic composition) transplanted organ (allograft) immunological response seen in the transplantation of kidneys from one individual to another with non-identical DNA. A common type of such an immunological response towards a solid organ transplant is acute cellular rejection.^{[5], [6]}

Acute cellular rejection (ACR) episodes are characterized by large numbers of infiltrating host immune cells. Those cells can form dense clusters within the allograft tissue at particular locations, and they may also be profuse throughout the allograft. The presence of such immunocellular infiltrate creates two functional threats to the longevity of the allograft: first, high numbers of activated, allograft-specific adaptive immune cells (such as T cells) directly damage the allograft; second, clusters of these cells can serve to facilitate proliferation of allograft-specific T cells that will further damage the allograft.^[7] Of particular interest are CD4 and CD8 memory cells, which are recognized as critical mediators of ACR, and T regulatory cells, which serve to blunt the activity of T cells that might cause ACR.^[6] For the duration of an ACR episode, the allograft's ability to function is reduced, and even after the episode is resolved clinically, damage to the tissue persists. As such, ACR episodes in kidney transplant recipients present a significant threat to the continued function of the kidney allograft, and may ultimately lead to the complete loss of the transplanted organ.^[5]

Due to the functional impact of episodes of ACR, their discovery after a patient becomes symptomatic is not ideal. Currently they can only be defined pathologically via needle biopsy, a procedure which removes a small cylindrical tissue sample from the kidney allograft for histological analysis. By the time they have been discovered, host immune cells have most likely

done permanent damage to the allograft, limiting its ability to function within its recipient and increasing the risk for additional ACR episodes.^[6]

While the benefits of preventing the loss of a kidney are dramatic for any particular recipient, from an epidemiological perspective, diminishing the rate of renal allograft loss from ACR is imperative. The pool of potential transplant recipients continues to grow at a faster rate than the availability of donor allografts, and this has caused more patients to wait longer for transplants. In 2002, 2012, and 2016, the number of patients awaiting kidney transplant (and the percentage waiting longer than 5 years for a transplants) was 51,004 (9.7%), 92,885 (13.9%), and 99,333 (15.8%), respectively.^{[2], [8]} The loss of function of an allograft exacerbates the increasing disparity between donor organ availability and need. In 2014 and 2015, the number of allografts transplanted into patients receiving their second transplant was 2,003 and 2,319, respectively, which accounted for 11.7% and 13.0% of all kidney transplants performed in those years.^[2] Once the primary allograft is lost, patients awaiting subsequent transplants experience longer wait-times, with 23.8% waiting longer than 5 years for another allograft compared with 14.6% waiting as long for their first allograft.^[1] Therefore, in order to reduce the burden of medical care required for patients after the failure of their primary kidney allograft, and reduce the waiting period for all patients receiving kidney transplants, a means of predicting ACR events before irreparable damage, or loss, of a kidney allograft occurs is needed.

2.0 CLINICAL BACKGROUND FOR RENAL TRANSPLANTATION

The immune response to an allogenic transplant is for host's immune cells to attack allogeneic cells and tissues that express foreign MHC molecules. Acute and chronic T cell mediated rejection (TCMR) are two means by which a functioning allograft can be harmed so severely that it will no longer function.^{[6], [9]} Great strides have been made in reducing the host response against allografts, and as a result, kidney transplantation is the best available treatment for patients experiencing end-stage renal failure.^[10] In addition to the application of corticosteroids^[11] and calcineurin inhibitors^[12] used to suppress the recipient's immune system, a method for reducing the risk of TCMR is the use of induction therapy at the time of the organ implantation.^{[13], [14], [15], [16]}

Induction therapy artificially induces immunological non-response to the allograft by means of manipulation of the hosts' adaptive and/or innate immune systems. Depleting (Thymoglobulin) and non-depleting (Basiliximab) therapies are currently FDA approved and available in the clinic. While tolerance has been achieved in animal models with different induction therapy regimens, complete and lasting tolerance induction remains elusive in the field of human solid organ transplant.^{[13], [17], [18]} Allograft tolerance is the ultimate goal of transplantation research, and for the purposes here, it will be defined as the absence of host immunological attack of, or damage to, the allograft. This definition includes insults to the allograft caused by both the innate and adaptive immune systems and all sub-categories thereof, i.e., TCMR, ACR, chronic humoral rejection, dendritic cell mediated rejection, etc.^{[7], [9]}

2.1 RENAL ALLOGRAFT OUTCOMES

The rates for long-term renal allograft survival have improved markedly since the advent of new pharmacological modalities. The half-life of deceased donor renal allografts increased from 6.6 years in 1989 to 8.0 years in 1995 then 8.8 years in 2005.^[19] Additionally, survival rates of transplanted kidneys have reached 92.0% and 96.5% for patients receiving deceased- and living-donor allografts, respectively, one year after transplantation.^[19] While strides have been made in prolonging long-term graft survival in kidney transplant patients, there is still a significant, and persistent, proportion of recipients who experience ACR fewer than 90 days after transplantation, despite increasingly potent inductive and immunosuppressive regimens.^[5] Thymoglobulin® and Basiliximab have been shown to lower the incidence of ACR episodes versus placebo and previous treatment modalities, and are considered to be part of standard-of-care drug protocols for patients receiving renal allografts.^{[20], [21]}

Despite the known benefits of using induction protocols, their use also has drawbacks. When T cell depleting agents, such as Thymoglobulin, are used, the patients' immune cells gradually become reconstituted. The new populations of T cells quickly fill the place of T cell populations which developed over the patients' lifetimes. Therefore, the reconstituted T cell memory subsets differ from those that were depleted, and the new subsets that proliferate in the patient are associated with increased instances of delayed ACR.^{[15], [19]}

In addition, the non-depleting agents, such as Basiliximab, lower regulatory T cells (T_{REG}), which are associated with long-term graft survival and graft tolerance, by blocking the IL-2R α ability to signal.^{[22], [23]} Thus, non-depleting induction agents also have the potential to interfere with natural regulatory mechanisms and potential to skew the T cell population to favor alloreactive T cells.^[24] This means that, while such therapies have been able to lengthen mean

graft survival of allografts, they have been shown to be imperfect treatments, as they defer attack on the allograft (albeit for years) rather than prevent the host's rejection of the allograft. Only through further exploration can it be known if the T cell memory profiles created by induction therapies can be further modulated to entirely avoid TCMR or ACR rather than simply postponing these allograft-damaging episodes.

2.2 LYMPHOCYTE MARKERS

Throughout this study phenotypic analysis and flow cytometry, are used to identify subsets of lymphocytes. are used to identify subsets of lymphocytes. Receptors on the surface of the cells or transcription factors in the nucleus, were chosen to identify cells that play important roles during ACR or allograft acceptance.

2.2.1 CD3, CD4, and CD8

Clusters of differentiation (CD) -3, -4, and -8 are proteins expressed on the cellular surface of T cells and are markers of determining cellular type. CD3 is the signaling component of the T cell receptor (TCR), and is used throughout this study to differentiate T cells from non-T cells. CD4 and CD8 are two different co-receptors for the TCR and are indicative of the particular functionality of T cells. T cells have either CD4 or CD8 expression once they leave the thymus and circulate throughout the body, but not both. CD4+ T cells are often called helper T cells as their release of cytokines upon antigen recognition influences the activation and proliferation of

cytotoxic T cells. Cytotoxic T cells are CD8+, and they are able to attack and kill cells with the use of perforin and granzymes proteins, which they produce upon activation.^[25]

2.2.2 CD45RO, CD45RA, and CD62L

The transmembrane protein CD45 has multiple isoforms, but the two of interest here are CD45RO and CD45RA, of which only one is expressed on the surface of a T cell at a time. CD62L, or L-selectin, is a protein which drives cells into secondary lymphoid tissues such as lymph nodes, tonsils, and the spleen. Using a matrix of CD45RO versus CD62L expression, T cells can be broken down into the four functional categories seen Table 2.1:^[25]

Table 2.1: T cell Memory Subsets Defined by CD45RO and CD62L

	CD45RO⁻	CD45RO⁺
CD62L⁺	Naïve (T _N)	Central Memory (T _{CM})
CD62L⁻	Effector Memory, RA ⁺ (T _{EMRA})	Effector memory (T _{EM})

2.2.3 FoxP3 and CD25

Forkhead box protein P3 (FoxP3) is located in the cellular nucleus. This protein is an essential transcription regulator as it suppresses transcription of genes encoding for pro-inflammatory cytokines like interleukin-2 and interferon-gamma. FoxP3 is used for the identification of T_{REG} in CD4+ T cells along with a high (as opposed to low and medium) expression of CD25, a marker of activation of T cells, and a lack of CD127 (IL-7R α) expression.^[25]

2.2.4 T Cell Subsets of Interest

The identification of subsets of immune cells depends upon the markers expressed by those cells. The subsets of interest identified from previous studies with similar cohorts, but a different induction agent, are T_{REG} defined as being CD3⁺/CD4⁺/CD25^{hi}/FoxP3⁺, CD4 T_{EM} defined as being CD3⁺/CD4⁺/CD62L⁻/CD45RO⁺, and CD8 T_{EM}RA defined as being CD3⁺/CD8⁺/CD62L⁻/CD45RO⁻.^{[17], [23], [26]} These cells types were examined in this study to determine if changes in their frequency could predict ACR episodes.

3.0 ANALYSIS METHODOLOGY

General estimating equations (GEE), are a form of generalized linear models (GLM) that take into account correlation introduced from multiple measurements of an individual.^[27]

3.1 GENERALIZED LINEAR MODELS

Generalized Linear Models (GLM) refers to a large class of statistical models employed to relate outcomes to linear combinations of predictor variables, which can be adapted for many different types of data. These data types include, continuous, ordinal, count, rate, and binary.

In its most basic form, the generalized linear model may be written as seen in Equation 3.1.^[28]

Equation 3.1
$$y_i = X_i\beta + \varepsilon_i \sim \text{iid } N(0, \sigma^2)$$

Equation 3.1 states that the expected value for the random variable, y , for a given case, i , may be predicted by the observed values of covariates for that case, X_i , multiplied by the coefficients for those covariates, β . The error between the predicted value for a given case, y_i , is accounted for by the error term, ε_i . Sample error, ε , for this model is assumed to follow a normal distribution with a mean 0 and standard deviation, σ . This assumption is a consequence of the foundational assumption that the sampled random variables in the model are independent and identically distributed (iid).^[28]

The distribution of expected values, μ_i , of Y_i given observed values of X_i (shown in Equation 3.2) need not be normally distributed.

Equation 3.2
$$\mu_i = E(Y_i|X_i)$$

When fitting a GLM, the most appropriate distribution should be chosen in accordance with the observed data being fit to the model. This is accomplished by utilizing a link function, defined in Equation 3.3.^[28]

Equation 3.3
$$g(\mu_i) = \eta_i$$

In the special case of $\mu = \eta$, the GLM would function the same as classical (Gaussian) linear models. This is known as the identity link. However, GLM can be easily fit for any type of statistical distribution from the exponential family (the pdf of whose “natural form” was described by McCullagh and Nelder (1989), shown in Equation 3.4) to allow proper specification of the outcome variable structure. For a function that is a member of the exponential family, θ is the scalar parameter, φ is the dispersion parameter, and $b(\theta)$, $a(\varphi)$, and $c(y, \varphi)$ are known functions.^[28]

Equation 3.4
$$f_Y(y|\theta) = \exp\left(\frac{y\theta - b(\theta)}{a(\varphi)} + c(y, \varphi)\right)$$

If the proper distribution of the data is identified, so too is the proper variance structure from which it comes, as the first derivative of $b(\theta)$ is the distribution’s mean, and its second derivative is the distribution’s variance. It is important to note that θ in the equations above stand for any number of parameters, so GLM are able to account for an equally large number of variable coefficients, β .^[28]

3.1.1 GLM Assumptions

Applying GLM requires that a set of assumptions be met in order for valid statistical inferences to be made. The assumptions placed upon GLM are as follows:

- 1) The data Y_1, Y_2, \dots, Y_n are independently distributed, i.e., the cases are independent;

- 2) Y_i , the dependent variable, need not be normally distributed, but it will typically take the distribution of one of the members of the exponential family;
- 3) There is a linear relationship between the transformed dependent variable, via the link function, and the independent variables;
- 4) Missing data are missing completely at random (MCAR);
- 5) And, large sample approximations are valid for the data, as they are required for maximum likelihood estimation (MLE) used to estimate the parameter.

Neither normal distribution of errors, nor homogeneity of variance need to be assumed in order to properly employ GLM.^[28]

3.1.2 GLM Limitations

Generalized linear models have two main limitations. First, only linear predictors may be used in the predictor components of the model. Second, requiring independent responses explicitly precludes analysis of the type of individual clustering expected in longitudinal studies.^[28]

3.2 GENERALIZED ESTIMATING EQUATIONS

The Generalized Estimating Equation (GEE) approach is a natural extension to GLM that relaxes the restriction of independent response. This difference is key in using GEE for longitudinal data analysis.^[27]

3.2.1 GLM and GEE Similarities

The most important conceptual similarity between GLM and GEE is the types of models, and the types of hypotheses, they can address. Both methods are designed to explore marginal models, i.e., show the risk of a population given a particular set of treatments or exposures, while only GLM can address conditional models, those designed to address the risk of an individual given certain parameter values.^[27]

3.2.2 Marginal GEE Models

Marginal GEE models are designed and employed to explain population- or cohort-wide probability averages for an outcome based upon covariate values over time. These probabilities average over the random-effects distribution of the population. As such, the primary use of marginal models is to describe both the positive and negative effects of each covariate as it relates to an outcome of interest and also the magnitude of its influence over the outcome while accounting for within group correlation over time.^[27]

Marginal models can give probability for an outcome given treatment or exposures (in the binary case) or weighted risk against continuous variables that applies to the whole cohort. The probability for a given case, p_i , in a binary GLM model is shown in Equation 3.5 (given that $Y_i = 1$ indicates that the presence of a treatment or exposure), where X_i and β are the values for the case's observed variables and model's coefficients for the covariates, respectively.^[27]

Equation 3.5
$$p_i = E(Y_i|X_i) = P(Y_i = 1|X_i) = \frac{\exp(X_i\beta)}{1 + \exp(X_i\beta)}$$

The GLM equation above can be adapted to determine probabilities of outcomes in GEE (see Equation 3.6), and it can also be easily converted to determine odds ratios, OR, between

treatments or exposures (see Equation 3.8). In Equation 3.6 the probability of an outcome at a given time-point, p_{ij} , is determined by the observed values of the covariates at all observed time-points, X_{ij} , and the values of the coefficients of the model, β , given that Y_{ij} equal to one indicates the outcome of interest. In Equation 3.6, 3.7, and 3.8 the subscript i indicates a given subject and subscript j indicates a given time-point of observation.^[27]

Equation 3.6
$$\text{logit}(p_{ij}) = E(Y_{ij}|X_{ij}) = P(Y_{ij} = 1|X_{ij}) = \beta X_{ij}$$

Obtaining an OR from a GEE model for a given exposure or treatment can be done utilizing the coefficients of the model. Equation 3.7 is a GEE model of the simplest case in which only a single, binary treatment or exposure variable, X_{ij} , is observed. The coefficients of the model shown in Equation 3.7 β_0 and β_1 are the baseline coefficient for all subjects observed and the coefficient associated with receiving treatment, respectively.^[27]

Equation 3.7
$$\text{logit}(p_{ij}) = P(Y_{ij} = 1|X_{ij}) = \beta_0 + \beta_1 X_{ij}$$

The OR of an outcome based on receiving a treatment or exposure given the model from Equation 3.7 can be found in Equation 3.8. The OR is defined by ratios of probabilities of outcomes at observed time-points, Y_{ij} , conditioned on the treatment received at those times, X_{ij} .^[27]

Equation 3.8
$$OR = \frac{P(Y_{ij}=1|X_{ij}=1)/P(Y_{ij}=0|X_{ij}=1)}{P(Y_{ij}=1|X_{ij}=0)/P(Y_{ij}=0|X_{ij}=0)} = \frac{\exp(\beta_0 + \beta_1)}{\exp(\beta_0)} = \exp(\beta_1)$$

3.2.3 GEE Assumptions

As is the case with GLM, GEE analysis has a number of assumptions which must be met in order for the model to be employed correctly. They are:^[27]

- 1) The data Y_1, Y_2, \dots, Y_n are not independently distributed, i.e., the cases are correlated or clustered;

- 2) Covariates can have interaction terms and may be non-linear transforms of independent variables;
- 3) Variance does not need to be homogenous between estimated parameters;
- 4) Errors are correlated; and
- 5) Missing data are MCAR.

3.2.4 GEE Correlation Matrices

In GEE, correlation matrices are used to account for the lack of independence in clustered observations. Values for the matrix range from zero to one where zero indicates no correlation between observations in a cluster, and values approaching one indicate very high correlation between observations. The matrix is arranged such that the first value in each column and row is for the first observation and each subsequent column and row is for each subsequent observation. Correlation matrices are always square with its main diagonal values equal to one.^[27]

GEE is robust relative to the incorrect specification of correlation structure, however, a model fit with the correct correlation structure will produce more accurate estimates of variable coefficients and standard errors for the coefficients. For a GEE model fit to longitudinal data collected from patients it would be expected that the most appropriate correlation matrix would show that the closer observations are in time, the more highly they are correlated. Stationary and Autoregressive correlation matrices are two examples with this type of time-dependence. Illustration of different correlation matrices is depicted in Table 3.1 where ρ is a given correlation value exclusively between zero and one.^[27]

Table 3.1: Examples of GEE Correlation Matrices

Independence	Exchangeable	Stationary (m=1)	AR1	Unstructured
$\begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix}$	$\begin{bmatrix} 1 & \rho & \rho \\ \rho & 1 & \rho \\ \rho & \rho & 1 \end{bmatrix}$	$\begin{bmatrix} 1 & \rho & 0 \\ \rho & 1 & \rho \\ 0 & \rho & 1 \end{bmatrix}$	$\begin{bmatrix} 1 & \rho & \rho^2 \\ \rho & 1 & \rho \\ \rho^2 & \rho & 1 \end{bmatrix}$	$\begin{bmatrix} 1 & \rho_{12} & \rho_{13} \\ \rho_{21} & 1 & \rho_{23} \\ \rho_{31} & \rho_{32} & 1 \end{bmatrix}$

3.3 MODEL SELECTION CRITERIA

As opposed to GLM, GEE relies on quasi-likelihood estimation rather than maximum likelihood estimation. Thus, methods used for model selection in GLM, such as Akaike information criterion (AIC) and Bayesian information criterion (BIC) are inappropriate for GEE. Quasi-likelihood under the independence model criterion (QIC) is an adaptation of AIC that can be used in its place when evaluating the value of one model over another within GEE. The similarities between AIC and QIC can be seen in Equation 3.9 and 3.10, respectively.^{[29], [30]}

Equation 3.9
$$AIC = -2LL + 2p$$

Equation 3.10
$$QIC = -2Q(\hat{\mu}; I) + 2trace(\hat{\Omega}_I^{-1}\hat{V}_R)$$

In the Equations 3.9 and 3.10 above, LL stands for log likelihood, p indicates the number of parameters, $Q(\hat{\mu}; I)$ is the quasi-likelihood function with an estimated mean of $\hat{\mu}$ given I , the independent covariance structure, $\hat{\Omega}_I$ is the variance estimator under the independence correlation structure, and \hat{V}_R is the robust variance estimator. Both AIC and QIC, therefore, give lower values for models which have a higher (quasi-)likelihood and fewer parameters, all else being equal. QIC can also be used to compare GEE models with different assigned covariance structures if they have the same parameters, as the values in the \hat{V}_R matrix are dependent upon the chosen covariance

structure. By selecting GEE models with the lowest QIC values, the most parsimonious model with the best parameters and goodness-of-fit may be chosen.^{[29], [30]}

The last aspect of model selection resides in the inclusion, or not, of interaction terms. If an interaction between covariates within a model is found to be significant, the parent model and the model constructed by adding the interaction term to it may be compared using a modification of QIC termed QICu. QICu adds to the value of QIC in proportion to the number of parameters within the model. As a result, when two models differ in only one parameter a QICu comparison will weigh the information added by the additional parameter against the preference for as parsimonious a model as possible. Equation 3.11 retains the variable definitions from Equations 3.9 and 3.10 and describes QICu.^[30]

Equation 3.11
$$QICu = -2Q(\hat{\mu}; I) + 2p$$

3.4 GEE MODEL APPLICATION

In order to test the hypotheses that T_{REG}, CD4 T_{EM}, and/or CD8 T_{EMRA} distributions were important predictors of ACR events, GEE models were built. The process of building these models began with description and cleaning the data, determining which parameters were statistically significant, construction and comparison of models containing statistically significant parameters, and ultimately testing the best model against the data to determine its utility in discerning between patients at higher risk of ACR episodes from those that were not.

The strategy for parameter selection was to test parameters of interest individually against the binary outcome variable of a patient experiencing ACR on their subsequent or not, with the inclusion criterion being a p-value of 0.15 or less using the same GEE analysis intended for the

built model. The *a priori* assumption of covariance was that the data would be auto-regressive due to intra-patient correlation.

3.5 MISSING DATA AND GEE ANALYSIS

Missing data are a problem for accurate analysis of data in most statistical setting. When applying GEE to longitudinal data, thorough and accurate characterization of the type the missing data is paramount because incorrect specification of the type of missing data during analysis can lead to biased or inefficient estimators.^[31] Three types of missing data will be discussed: data missing completely at random (MCAR), data missing at random (MAR), and data missing not at random (MNAR).^[32]

3.5.1 MCAR

Data are considered to be MCAR if failure to observe the data does not depend on the outcome of interest, observed covariate values, or unobserved data. Simply put, if data are MCAR, the missing data constitutes a random sample of all data that could have been observed. As a result, no bias is introduced into the inferences drawn from analyzing the data, and this type of missing data can be appropriately analyzed using GEE.^[32]

3.5.2 MAR

Data are MAR when the missingness of the observed data does not depend on the outcome of interest. However, the MAR assumption is less restrictive than the MCAR assumption because it allows for the failure to observe data to depend upon unobserved data. Because MCAR is a special case of MAR, all data that are MCAR are also MAR.^[32] Data that are MAR but not MCAR can introduce bias into GEE analysis. Other methods that appropriately account for data being MAR, such as mixed models, need to be utilized for such data.^[33]

3.5.3 MNAR

Data are MNAR when the failure to observe the data depends upon the data that could have been observed. In this case, missing data mechanisms must often be explicitly modeled, and ultimately, statistical inferences found using this type of methodology will be sensitive to the accuracy with which the missing data are modeled. Adding to the complexity MNAR data is that typically the assumptions made in developing these models may not be verifiable.^[32]

3.5.4 Study Data and the MCAR Assumption

In order to analyze data using GEE the MCAR assumption must be met. With data that are multivariate normal, Little's MCAR test can be used to determine if the MCAR assumption is met.^{[34], [35]} However, this test is not appropriate for application to these data. While the original test has been adapted to accommodate arbitrary missingness, the non-binary covariates come from binomial distributions, some of which have small probabilities. This drove the data to depart from

univariate normality (at least one covariate had $p < 0.05$ for the Shapiro-Wilk test at each time-point) as well as multivariate normality (Doornik-Hansen test values of $p < 0.0001$ at all time-points).

Because no broadly applicable test was available to ascertain if the MCAR assumption was violated, indicator variables for missing data were generated to determine if any were statistically significant in predicting ACR episodes. These variables were assigned a value of one if the datum was missing and zero if it was observed.

4.0 STUDY DESIGN AND IMPLEMENTATION

4.1 STUDY HYPOTHESES

The hypotheses examined within this study stemmed from previous a cross-sectional study conducted within the University of Pittsburgh,^[23] and was supported by observations from other findings in the literature.^{[15], [26], [36], [37]} The aim of the study was to create a predictive statistical model that would predict ACR episodes. Within this study a predictive model is defined as one which would indicate changes in the allograft recipients' immune system during the visit prior to the occurrence of ACR (i.e., if a patient experienced ACR at 90 days post-Tx, a predictive model should indicate a change based on that patient's immunological profile at 30 days post-Tx).

The study was designed to explore whether or not the proportion of T_{REG} relative to the rest of the CD4+ T cell population could be used as a predictor of rejection episodes. T_{REG} have been shown to suppress T cell reactivity in a dose-dependent fashion and correlate with ACR episodes.^[18] As such, the first hypothesis for the study was whether the proportion of T_{REG} from CD4+ T cells could be used to predict patient ACR episodes.

Decreased prevalence of terminally-differentiated effector memory T cells that were also CD45RA+ (CD8 T_{EMRA}), was previously shown to be cross-sectionally correlated with patients undergoing ACR.^[23] It was hypothesized that there may be a longitudinal correlation between the proportion of CD8 T_{EMRA} versus the remainder of patients' CD8+ T cell populations. As a result, the second hypothesis tested in this study was that the proportion of CD8 T_{EMRA} relative to all CD8+ T cells within patients would predict ACR episodes.

The third, and final, hypothesis to be tested in this study resulted from a finding that patients characterized as “reactive” (neither tolerant nor experiencing ACR) had elevated levels of CD4+ CD45RO+ (CD4 T_{EM}) T cells compared with kidney transplant recipients who appeared tolerant.^[23] Resultantly, the third hypothesis to be tested in this study was that the proportion of CD4 T_{EM} out of all CD4+ T cells within patients would predict ACR episodes.

4.2 STUDY IMPLEMENTATION

The data used for this thesis were obtained from blood samples collected as part of a parent clinical study at the University Hospital and Christ Hospital in Cincinnati, Ohio. The parent study was designed as a 12-month, prospective, randomized, dual center, open label pilot study to evaluate the safety and efficacy of Myfortic® (mycophenolic acid) loading regimens in combination with Simulect® (Basiliximab) induction vs. Thymoglobulin induction alone and with Prograf® (tacrolimus) in early corticosteroid withdrawal (hereafter referred to as CWT). This clinical study collected data relevant to standard-of-care treatment for renal transplant recipients; however, only demographic data, date of transplant, and rejection episode dates and severity, were made available to the University of Pittsburgh’s research group.

From the 61 patients enrolled in the CWT, this study’s cohort was selected from any patient who agreed to volunteer additional blood samples for our subordinate lymphocyte characterization and analysis. All of the CWT’s participants had the option to enroll or refuse enrollment in our characterization study while remaining part of the CWT. Samples for this current study were collected from 2011 to 2013. Forty patients donated blood samples this study. Of those, five were

excluded due to pretreatment with an additional drug, not in the original study design. Samples for this current study were collected from 2011 to 2013 (as discussed in Section 5.1.2.).

Demographic comparisons of the participants enrolled in the current study were made. The two-tailed Pearson χ^2 test was used for testing sex and race. The two-sided T-test was used for patient age at the time of transplant, and a comparison of proportion test was used for number of patients. No test was performed on ethnicity due to complete homogeneity within the study population. Fisher's exact test was used to assess statistical significance of patients experiencing ACR episodes between groups.

A panel of fluorochrome-conjugated antibodies was developed to characterize the makeup of peripheral blood leukocyte. The panel examined the proportion of three cell types known to be correlated in a cross-sectional manner with ACR: T_{REG}, CD4 T_{EM}, and CD8 T_{EMRA}.

4.3 BLOOD SAMPLE ACQUISITION

Participants were asked to give up to 75 mL of blood (70 mL in sodium-heparin test tubes and 5 mL in a serum-separating test tube) at each of the following time-points: before transplantation (pre-Tx), 30-; 90-; 180-; 270-; and 360-days after transplantation (post-Tx). The acceptable windows were ± 3 days for the 30-day sample, and ± 7 days for all other post-Tx samples. The pre-Tx samples could be obtained any time prior to the administration of preconditioning or induction drugs. These samples were then shipped overnight to the University of Pittsburgh's Starzl Transplant Institute. Samples that took greater than 24 hours to arrive were discarded.

4.4 PHENOTYPIC ANALYSIS OF LYMPHOCYTES

Lymphocyte counts were obtained from whole blood samples using the automated cell counter. Approximately 1.2 mL of whole blood was retained from the sample for use in staining and flow cytometry analysis. The lymphocytes in the remainder of the samples were then separated via Ficoll-Paque suspension and centrifugation at 1,200 rpm in a centrifuge. The suspended lymphocyte layer was combined and washed within 30 minutes of centrifugation using 1640 RPMI and bovine serum albumin (BSA) solution under sterile conditions. The cells from the sample were counted visually using a hemocytometer and they were preserved cryogenically in a solution of 10% dimethyl sulfoxide (DMSO) in bovine serum albumen (BSA) for follow-on studies.

4.4.1 Leukocyte Preparation

The cells intended for analysis via flow cytometry were stained with fluorochrome-conjugated monoclonal antibodies for CD3, CD4, CD8, CD45RO, CD45RA, CD62L, CD127, and CD25 purchased from eBioscience and Beckton Dickinson. Upon staining, the cells were kept at 4°C in the dark for 30 minutes, and then washed and fixed using 1% paraformaldehyde in DPBS. For the analysis of T_{REG}, eBioscience fixation/permeabilization solutions were utilized to allow intracellular staining of FoxP3. Samples were acquired by flow cytometry within 2 hours of staining, and the raw data output was analyzed with BD's FACSDiva software. All statistical analysis of FACSDiva output was conducted using Stata SE 14 software.

4.4.2 Data Acquisition

Proportional data was used throughout this analysis, and shown in percentages. The number of cells from which the percentages are derived are always larger than 100, often greater than 1,000 or 10,000, so the large number approximation is appropriate for any given observed datum. The system by which the data is obtained is called a gating strategy, because the data are derived by drawing a polygon around areas in scatter plots that correspond to cells that correlate to markers either being present or absent in the cells. Figure 4.1 shows the gates for selecting T_{REG} , $CD4$ T_{EMRA} , and $CD8$ T_{EMRA} populations from their parent populations, namely T cells, $CD4^+$ T cells and $CD8^+$ T cells, respectively.

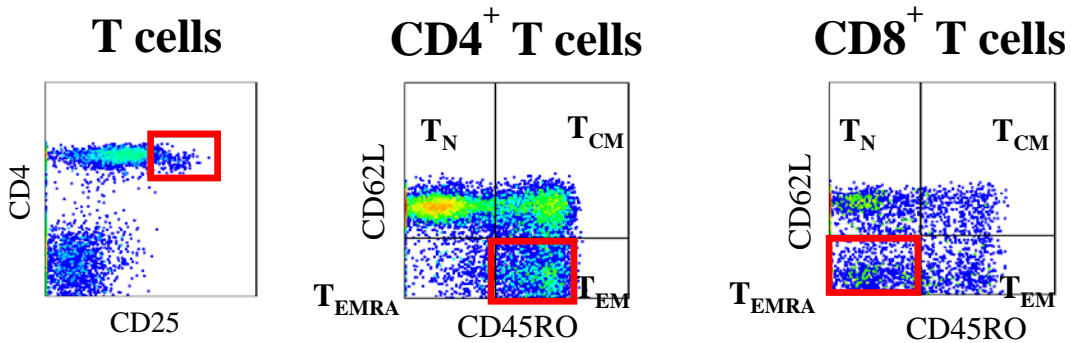


Figure 4.1: Collection Gates

5.0 RESULTS

5.1 DATA DESCRIPTION

The CWT randomized subjects to two arms. Those that volunteered for this study were similar across both treatment groups, with the exception of sex (Table 5.1). Although sex is close to being statistically significant, none of the demographic differences reach statistical significance at the $p = 0.05$ level. There was no statistical difference between the number of patients that experienced at least one ACR episode.

Table 5.1: Study Participant Demographics

	Thymoglobulin	Basiliximab	P-value
Number of Patients	20	15	0.552
Age at Tx (years)	51.1 ± 14.4 (Range 28 to 71)	56.7 ± 12.2 (Range 37 to 79)	0.231
Sex	10 male, 10 female	12 male, 3 female	0.069
Race	16 Caucasian, 4 Black	12 Caucasian, 3 Black	1.000
Ethnicity	20 Non-Hispanic/Latino	15 Non-Hispanic/Latino	N/A
ACR	7	2	0.244

5.2 MISSING DATA

Eleven of 35 patients had data collected for every planned collection time-point. Nine patients' data were left-truncated, seven patients' data were right-truncated, and an additional eight patients had one or more time-point of missed data collection. In total, approximately 26.7% of the data was missing from the 35 patients. Further examination of the MCAR assumption was required before application of GEE. A depiction of the data collected can be seen in Figure 5.1 in which

collected data for a patient a given time-point is shown in green (if there was no ACR) or red (if there was a biopsy-proven ACR episode after the sample collection but before the next sample collection). Missing data are shown by a blank cell.

Patient	Pre-Tx	30 Days	90 Days	180 Days	270 Days	360 Days
01						
03						
06						
12						
17						
19						
23						
26						
28						
30						
31						
32						
35						
36						
37						
47						
49						
54						
56						
57						
59						
61						
76						
78						
79						
82						
83						
84						
85						
89						
90						
92						
93						
96						
97						

Figure 5.1: Collected Data

Although the level of data missed from collection is high the number of samples received for each time-point is consistent: for each of the six time-points from pre-Tx to 360 days post-Tx, the number of collected data points was 26, 26, 27, 25, 24, and 28, respectively. On average 4.6 samples were collect for each subject in the study. Thirteen patients had missing data only as the result of truncation due to late entry into the study or premature termination of the study (accounting for 76% of the missing data). Eleven of the 35 patients had missing data not accounted for by study implementation, which equates to 10.8% of the total time-points which may have been collected from the 35 enrollees.

None of these indicator variables was found to be statistically significant in GEE analysis where ACR episodes were held as the independent variable and each missing indicator was tested as a dependent variable, in turn. Because that missing data due to study design or implementation does not necessarily violate the MCAR assumption and no missing data were statistically significant in predicting the outcome of interest, the MCAR assumption for these data was deemed appropriate.^[38]

5.3 PARAMETER SELECTION

Parameters, and their corresponding coefficients' p-values, may be seen in Table 5.2 ($p < 0.15$ are italicized). Parameters other than T_{REG} , $CD4 T_{EMRA}$, and $CD8 T_{EMRA}$ were included due to their biological relevance in T cell immunity with the possibility that they may help to refine any predictive relationships that might exist between the two variables of interest and ACR events.

Table 5.2: GEE Single-Parameter Model Values

Parameter	Coefficient	p-value		Parameter	Coefficient	p-value
<i>Age, centered</i>	-.04536	<i>0.109</i>		CD25 ^{hi}	.0480918	0.416
<i>Days post-Tx</i>	-.0039507	<i>0.117</i>		T _{REG}	-.0007007	0.469
<i>Treatment</i>	1.141206	<i>0.149</i>		CD4 T _{EM}	-.0048596	0.760
CD4	.0113613	0.241		CD4 T _{EMRA}	.0717864	<i>0.051</i>
CD8	-.0164905	<i>0.118</i>		CD8 T _{EM}	.0011368	0.953
CD4/CD8	.0630022	0.671		CD8 T _{EMRA}	.014741	0.435

5.4 MODEL CONSTRUCTION

For the purpose of preliminary GEE model construction an AR1 correlation structure was assumed. In addition to the five parameters that met the inclusion criterion, the variables T_{REG}, CD4 T_{EM}, and CD8 T_{EMRA} were added to Models A, C, and E, respectively.

Model A was built using the parameters Age (centered), Days post-Tx, Treatment, CD8, CD4 T_{EMRA}, and T_{REG}. GEE analysis of the model showed that it was significant (Wald χ^2 p = 0.0095) but it contained on covariate that was not statistically significant, Days post-Tx. The details for Model A can be seen in Table 5.3.

Table 5.3: Model A Values

Parameter	Coefficient	p-value		Parameter	Coefficient	p-value
Age, centered	0.035529	0.012		CD8	0.0216521	0.002
Thymoglobulin	1.239464	0.026		CD4 T _{EMRA}	0.0738930	0.062
Days post-Tx	0.0029239	0.900		T _{REG}	0.0145422	0.267

Days post-Tx was removed from the covariates in Model A because it had the highest p-value, and Model B was built from the five remaining covariates. Model B was statistically significant (Wald χ^2 p = 0.0059), but T_{REG} was not a statistically significant covariate (p = 0.245). Details for Model B can be found in Table 5.4.

Table 5.4: Model B Values

Parameter	Coefficient	p-value	Parameter	Coefficient	p-value
Age, centered	0.0350208	0.010	CD8	0.0211845	0.001
Thymoglobulin	1.120552	0.012	CD4 T _{EMRA}	0.0736068	0.061
			T _{REG}	0.0140173	0.245

Model C contained the five covariates found to be statistically significant in univariate analysis as well as CD4 T_{EM}. Model C achieved statistical significance (Wald χ^2 p = 0.0061), but it contained covariates that failed to reach statistical significance. Details for Model C can be seen in Table 5.5.

Table 5.5: Model C Values

Parameter	Coefficient	p-value	Parameter	Coefficient	p-value
Age, centered	-0.098	0.007	CD8	-0.077	0.013
Thymoglobulin	2.419	0.026	CD4 T _{EMRA}	0.152	0.058
Days post-Tx	-0.0006	0.834	CD4 T _{EM}	0.028	0.405

Days post-Tx was removed from the pool of covariates in Model C to produce Model D's covariate pool because it had the highest p-value. Model D was statistically significant (Wald χ^2 p = 0.0121), but it contained one covariate that was not, CD4 T_{EM}. Values for Model D can be seen in Table 5.6.

Table 5.6: Model D Values

Parameter	Coefficient	p-value	Parameter	Coefficient	p-value
Age, centered	-0.099	0.006	CD8	-0.079	0.015
Thymoglobulin	2.503	0.010	CD4 T _{EMRA}	0.152	0.057
			CD4 T _{EM}	0.027	0.407

Model E was constructed using the third covariate of interest, CD8 T_{EMRA}, in addition to the five covariates found to be statistically significant in univariate analysis. Model E was statistically significant (Wald χ^2 p = 0.0198), but it contained three covariates that were not statistically significant at the p < 0.05 threshold. See Table 5.7 for Model E details.

Table 5.7: Model E Values

Parameter	Coefficient	p-value		Parameter	Coefficient	p-value
Age, centered	0.0335643	0.006		CD8	0.0247909	0.005
Thymoglobulin	1.123874	0.020		CD4 T _{EMRA}	0.0777823	0.142
Days post-Tx	0.0029216	0.858		CD8 T _{EMRA}	0.0329762	0.249

The covariate with the highest p-value was removed from the covariates in Model E to produce Model F. Model F was statistically significant (Wald χ^2 p = 0.0032), but it contained two covariates that failed to reach statistical significance, CD4 T_{EMRA} and CD8 T_{EMRA}.

Table 5.8: Model F Values

Parameter	Coefficient	p-value		Parameter	Coefficient	p-value
Age, centered	0.032653	0.005		CD8	0.0252462	0.004
Thymoglobulin	1.033318	0.010		CD4 T _{EMRA}	0.0789165	0.149
				CD8 T _{EMRA}	0.0334378	0.253

6.0 DISCUSSION

The analysis of the data showed that neither of T_{REG} , $CD4 T_{EM}$, nor $CD8 T_{EMRA}$ proportions in peripheral blood were associated with ACR events. None of the variables of interest was independently statistically significant in this regard nor were they statistically significant when forced into models containing other statistically significant covariates.

Five covariates were identified that were potentially predictive of ACR episodes: Age (centered), Days post-Tx, Treatment, $CD4 T_{EMRA}$, and $CD8$. These five predictors have been previously described in the literature as being correlated with ACR episodes. The positive or negative influence of the coefficients for Days post-Tx, Treatment, $CD4 T_{EMRA}$, and $CD8$ are in line with previous findings.^{[1], [16], [37]} Of note, the literature is mixed in regard to whether Thymoglobulin or Basiliximab therapies are more likely to lead to ACR episodes, with some studies concluding that one treatment is more beneficial while others conclude the opposite.^[16] Also in line with previous studied, patients in this study experienced a preponderance of ACR episodes between transplant and 90 days post-Tx; a common problem for kidney tolerance in the clinic.^[19]

Recipient age is known to be a predictor of renal transplant rejection episodes, however increasing age has been shown to be associated with increased risk for ACR,^[39] while in this study the opposite was observed. It has also been noted that increased deceased donor age increases the risk of ACR episodes.^[39] Due to this study not being provided with donor organ age data, the correlation between Days post-Tx and ACR episodes may be the product of a confounding variable.

There are two very important caveats to the lack of statistical significance found in our subsets of interest as compared with previous literature: first, two of the markers were chosen based upon a study using a different induction agent,^[23] and second due to the small number of patients enrolled in the study, the analysis may have not had enough power to detect a true predictor of ACR.

Different induction modalities have varying means of action, and effects, upon recipients' immune system. The differences noted in populations of T_{REG}, CD4 T_{EM}, and CD8 T_{EMRA} came from studying patients that had undergone induction therapy using Alemtuzumab. Thymoglobulin is a T cell depleting agent, whereas Alemtuzumab depletes not only lymphocytes, but also B cells. The presence or absence of these cells has been shown to effect graft tolerance.^[7]

6.1 ANALYSIS LIMITATIONS

These data did not appear to violate the MCAR assumption. However, due to the high amount of missing data and no established test for the assumption, discussion of addressing the missing data is warranted. Should the collected data not reflect the population-at-large, the results of the analysis could contain bias of the estimators or their statistical significance.

Two common methods for dealing with missing data, line (subject) deletion and multiple imputation (MI), were both inappropriate for application here. If line deletion were used, approximately 69% of the subjects from the study would be eliminated from analysis, leaving a sample too small for any meaningful discussion. MI methodology would not aid analysis because MI addresses missing values for a given time-point based upon other variable values for that same time-point. Given that all lymphocyte population data is derived from the presence or absence of

a blood sample, there is almost no missing data for any given patient-time-point unless all of the data for that time-point is missing.

Beyond the common methods mentioned above, there exist additional methodology for analyzing missing data, but these methods are computationally intensive, lack explicit algorithms for general use, and/or rely upon untestable hypotheses whose violation results in highly biased outcomes given non-normal distributions.^{[40], [41]} Three of these methods are the selection model, pattern mixture model, and shared parameter model. The selection model is inappropriate as it relies on the assumption that data measures can predict missingness, and the left- and right-truncation subjects are known to have missingness not related to their observed values. The pattern mixture model was not used because it relies upon there being relatively few patterns of missingness within the data relative to its clusters, and this assumption was violated by there being 16 patterns of missingness, eight of which were unique, among the 35 subjects. The shared parameter model could not be used on this data because it relies on the assumption that within a given subject the chance of a binary outcome increases or decreases in a linear fashion over time, which is known to be false in the observed population for the parameters of interest. Lastly, these three models are well-suited for dropout (right truncation) of data in which data for a given subject exists until the point that the data is censored. Right truncation was only observed in seven of 35 patients, lending further merit to these models not being applied to these data. It has been shown that mixed-effect models can be tailored to produce more efficient estimators than GEE methodology in some cases when truncation and random missing data are present, however since the assumptions of GEE were met, no mixed-effect models were fit to these data.

Exploration of the data's missingness showed that the amount of data missing or present relative to the other time-points was similar across all time-points. There were no outliers in this

regard. Given that all available methods of analysis would have a violated assumption if applied to these data, it was determined that GEE methodology was best suited to analysis in this study.

Although GEE analysis was deemed an appropriate method for use with these data, one of the assumptions, namely MCAR, might be called into question since Little's MCAR test was not applicable here. Although much of the data was missing, its missingness did not depend upon the outcome of interest, as assessed by the missing data's lack of correlation to ACR episodes. Also of note is that 76% of the missing data was due to a cause independent of ACR episodes, due to left- and right-truncation late implementation of this study and termination of the parent study, rather than lack of patient follow-up (which might be negatively correlated with patient self-care and thus with ACR episodes).

While no power calculations were performed for this study's analysis, other work suggests that the sample size achieved in this study may have been too small to detect statistically significant predictors of ACR episodes. Even with a low level of intra-cluster correlation and a relative risk greater than three, single binary exposure/outcome correlations require more than 70 clusters to achieve 90% power at the desire significance level of 0.05.^[35] Although most of the covariates in this study were not binary, the power simulations serve as an example that many more participants may be necessary in a study of this type in order to detect statistically significant covariates, especially those that have been shown to be linked in cross-sectional studies by many previous investigators.^{[6], [15], [17], [18], [22], [24], [26], [36], [37]}

The coefficients for the parameters of interest mirrored the positive or negative effects expected of T_{REG}, CD4 T_{EM}, nor CD8 T_{EMRA} on future ACR episodes. As such, their failure to reach statistical significance of $p < 0.05$ does not refute previous findings, but rather highlights the need for more information to be gathered in similar cohorts to elucidate the predictive strength of

these markers. Such further studies should also explore the ratio of T_{REG} to the two memory subsets of interest, since they were shown to be correlated with patients undergoing ACR as compared with quiescent kidney allograft recipients.^[23] Unfortunately, due to sparse data in this study, modeling of these ratios was too noisy for meaningful analysis.

7.0 CONCLUSION

The analysis of these data showed that T_{REG} , $CD4 T_{EM}$, and $CD8 T_{EMRA}$ cell populations did not statistically significantly predict ACR episodes, possibly due to lack of analytic power stemming from missing data and small sample size.

While none of the hypotheses were confirmed, if the results of the single-variable models are reflective of the kidney transplant population, clinically useful inference can be made. Of interest, $CD4 T_{EMRA}$ has been shown to be predictively associated with liver transplant ACR,^[37] and $CD8$ proportions have also shown association with kidney ACR episodes.^[26] With this in mind, further exploration of these data and studies with more robust data collection, could reveal that either/both $CD4 T_{EMRA}$ and $CD8$ are predictive of ACR.

From a public health perspective, the ongoing challenge of discovering a minimally-invasive technique for predicting ACR remains. This analysis has shown that although the hypothesized parameters did not predict rejection, other markers may be able to do so. Development of such a technique is of paramount importance as it greatly benefits transplant recipients, reduces the burden of health care provision, and will help to alleviate the persistent problem of lower than necessary donor organ supply.

BIBLIOGRAPHY

- [1] U.S. Health Resources & Services Administration, "OPTN/SRTR 2012 Annual Data Report: Kidney," pp. 18-32, 2012.
- [2] U.S. Health Resources & Services Administration, "OPTN: National Data," 30 Sep 2016. [Online]. Available: <https://optn.transplant.hrsa.gov/data/view-data-reports/national-data/#>. [Accessed 04 Oct 2016].
- [3] J. E. Murray, "Renal Homotransplantation in Identical Twins," *Surg Forum*, vol. 6, p. 432, 1955.
- [4] T. E. Starzl, "The Landmark Identical Twin Case," *Journal of the American Medical Association*, vol. 251, p. 2872, 1984.
- [5] P. J. Morris and S. J. Knechtle, *Kidney Transplantation: Principles and Practice*, Seventh ed., Oxford: Saunders, 2013.
- [6] M. A. Lusco, A. B. Fogo, B. Najafian and C. E. Alpers, "AJKD Atlas of Renal Pathology: Acute T-Cell-Mediated Rejection," *American Journal of Kidney Disease*, vol. 67, no. 5, pp. 29-30, 2016.
- [7] K. Y. Shiu, et al, "B-Lymphocytes Support and Regulate Indirect T-Cell Alloreactivity in Individual Patients with Chronic Antibody-Mediated Rejection," *Kidney International*, vol. 88, no. 3, p. 560, 2015.
- [8] A. K. Khalil, et al, "Re-transplants compared to primary kidney transplants recipients: a mate kidney paired analysis of the OPTN/UNOS database.," *Clinical Transplantation*, vol. 30, no. 5, pp. 566-578, 2016.
- [9] D. Zecher, et al, "Innate Immunity Alone Is Not Sufficient For Chronic Rejection But Predisposes Healed Allografts To T Cell-Mediated Pathology," *Transplant Immunology*, vol. 26, no. 2/3, pp. 113-118, 2012.
- [10] R. A. Wolfe, et al, "Comparison of Mortality in all Patients on Dialysis, Patients on Dialysis Awaiting Transplantation, and Recipients of a First Cadaveric Transplant," *The New England Journal of Medicine*, vol. 341, no. 23, pp. 1725-30, 1999.
- [11] J. D. Pirsch, et al, "A Comparison of Tacrolimus (FK506) and Cyclosporine for Immunosuppression After Cadaveric Renal Transplantation. FK506 Kidney Transplant Study Group," *Transplantation*, vol. 63, no. 7, p. 997, 1997.

- [12] R. Y. Calne, et al, "Cyclosporin A in Patients Receiving Renal Allografts from Cadaver Donors. 1978," *Journal of the American Society of Nephrology*, vol. 9, no. 9, p. 1751, 1998.
- [13] U. M. Thiyagarajan, A. Ponnuswamy and A. Bagul, "Thymoglobulin and its use in Renal Transplantation: A Review," *American Journal of Nephrology*, vol. 37, no. 6, p. 586, 2013.
- [14] K. McKeage and P. L. McCormack, "Basiliximab: A Review of its use as Induction Therapy in Renal Transplantation," vol. 24, no. 1, pp. 55-76, 2010.
- [15] E. Krystufkova, et al, "Regulatory T Cells in Kidney Transplant Recipients: The Effect of Induction Immunosuppression Therapy," *Nephrology, Dialysis, Transplantation*, vol. 27, no. 6, p. 2576, 2012; 2011.
- [16] K. L. Hardinger, D. C. Brennan and C. L. Klein, "Selection of induction therapy in kidney transplantation," *Transplant International*, pp. 662-672, 2012.
- [17] F. Braza, et al, "Regulatory T Cells in Kidney Transplantation: New Directions?," *American Journal of Transplantation*, vol. 15, no. 9, pp. 2288-300, 2015.
- [18] I. R. Ferrer, et al, "Induction of Transplantation Tolerance through Regulatory Cells: From Mice to Men," *Immunological Reviews*, vol. 258, no. 1, pp. 102-16, 2014.
- [19] K. E. Lamb, S. Lodhi and H. U. Meier-Kriesche, "Long the United States: A Critical Reappraisal," *American Journal of Transplantation*, vol. 11, no. 3, pp. 450-62, 2011.
- [20] B. Nashan, et al, "Randomised Trial of Basiliximab Versus Placebo for Control of Acute Cellular Rejection in Renal Allograft Recipients," *The Lancet*, vol. 350, no. 9086, pp. 1192-8, 1997.
- [21] E. D. Deeks and G. M. Keating, "Rabbit Antithymocyte Globulin (Thymoglobulin): A Review of its use in the Prevention and Treatment of Acute Renal Allograft Rejection," *Drugs*, vol. 69, no. 11, pp. 1483-512, 2009.
- [22] H. Waldmann, et al, "Harnessing FOXP3+ Regulatory T Cells for Transplantation Tolerance," *The Journal of Clinical Investigation*, vol. 124, no. 4, pp. 1436-45, 2014.
- [23] C. Macedo, et al, "Long-Term Effects of Alemtuzumab on Regulatory and Memory T-Cell Subsets in Kidney Transplantation," *Transplantation*, vol. 93, no. 8, pp. 813-21, 2012.

- [24] F. Braza, et al, "The Plasticity and Stability of Regulatory T Cells," *Nature Reviews*, vol. 13, no. 6, p. 461, 2013.
- [25] C. Janeway, P. Travers, M. Walport and M. Shlomchik, *Immunobiology*, 6th ed., Garland Science, 2005.
- [26] M. G. Betjes, et al, "Terminally Differentiated CD8+ Temra Cells Are Associated With the Risk for Acute Kidney Allograft Rejection," *Transplantation*, vol. 94, no. 1, pp. 63-69, 15 July 2012.
- [27] K.-Y. Laing and S. L. Zeger, "Longitudinal Data Analysis Using Generalized Linear Models," *Biometrika*, vol. 73, no. 1, pp. 13-22, 1986.
- [28] P. McCullagh and J. A. Nelder, *Generalized Linear Models*, 2nd ed., New York: Chapman & Hall, 1989.
- [29] W. Pan, "Akaike's Information Criterion in Generalized Estimating Equations," *Biometrics*, vol. 57, no. 1, pp. 120-25, 2001.
- [30] J. Cui, "QIC Program and Model Selection in GEE Analyses," *The Stata Journal*, vol. 7, no. 2, pp. 209-220, 2007.
- [31] M. Soley-Bori, "Dealing with Missing Data: Key assumptions and Methods for Applied Analysis," Boston University School of Public Health, Boston, 2013.
- [32] J. G. Ibrahim and G. Molenberghs, "Missing Data Methods in Longitudinal Studies: A Review," *Test*, vol. 18, pp. 1-43, 2009.
- [33] S. Gao and R. Thiebaut, "Mixed-effect Models for Truncated Longitudinal Outcomes with," *Journal of Data Science*, vol. 7, pp. 13-25, 2009.
- [34] R. J. A. Little, "A Test of Missing Completely at Random for Multivariate Data with Missing Values," *JASA*, vol. 83, no. 404, pp. 1198-1202, 1988.
- [35] Z. Li and I. W. McKeague, "Power and Sample Size Calculations for Generalized Estimating Equations via Local Asymptotics," *Statistica Sinica*, vol. 23, pp. 231-250, 2013.
- [36] F. Veronese, et al, "Pathological and Clinical Correlates of FOXP3 Cells in Renal Allografts during Acute Rejection," *American Journal of Transplantation*, vol. 7, no. 4, pp. 914-22, 2007.

- [37] U. A. Gerlach, et al, "Elevation of CD4+ Differentiated Memory T Cells Is Associated With Acute Cellular and Antibody-Mediated Rejection After Liver Transplantation," *Transplantation*, vol. 95, no. 12, pp. 1512-1520, 27 June 2013.
- [38] R. J. A. Little, "Modeling the Drop-Out Mechanism in Repeated-Measures Studies," *Journal of the American Statistical Association*, vol. 90, no. 431, pp. 1112-1121, 1995.
- [39] M. Veroux, et al., "Age is an Important Predictor of Kidney Transplantation Outcome," *Nephrology, Dialysis, Transplantation*, vol. 27, no. 4, pp. 1663-1672, 2012.
- [40] C. K. Enders, "Analyzing Longitudinal Data With Missing Values," *Rehabilitation Psychology*, vol. 56, no. 4, pp. 267-288, Nov 2011.
- [41] Z. Wang, et al, "Statistical resolution of missing longitudinal data in clinical pharmacogenomics," *Advanced Drug Delivery Reviews*, vol. 65, pp. 980-986, 22 Mar 2013.