



Published in final edited form as:

Transplantation. 2009 May 27; 87(10): 1520–1529. doi:10.1097/TP.0b013e3181a484d7.

Early outcomes of thymoglobulin and basiliximab induction in kidney transplantation: Application of statistical approaches to reduce bias in observational comparisons¹

Lisa M. Willoughby, PhD², Mark A. Schnitzler, PhD³, Daniel C. Brennan, MD⁴, Brett W. Pinsky, MPH³, Nino Dzebisashvili, MS³, Paula M. Buchanan, MPH³, Luca Neri, MD³, Lisa A. Rocca-Rey, MD³, Kevin C. Abbott, MD,MPH⁵, and Krista L. Lentine, MD,MS^{3,6}

²Department of Research Methodology, Saint Louis University, St. Louis MO 63108

³Center for Outcomes Research, Saint Louis University School of Medicine, 3545 Lafayette Avenue, Salus Center 2nd Floor, St. Louis MO 63104

⁴Division of Nephrology, Washington University School of Medicine, 6107 Queeny Tower, One Barnes-Jewish Hospital Plaza, St. Louis, MO 63110

⁵Nephrology Service, Walter Reed Army Medical Center, Washington, DC, 20307-5001

⁶Division of Nephrology, Saint Louis University School of Medicine, St. Louis, MO, 63104

Abstract

Background—Retrospective comparison of treatment-related kidney transplant outcomes may be facilitated by multivariable statistical adjustments and case-matching.

Methods—We studied Organ Procurement and Transplantation Network registry data for kidney transplants in 2001 to 2005 managed with thymoglobulin, basiliximab or no antibody induction and discharge maintenance immunosuppression regimens of tacrolimus and mycophenolate mofetil. The primary outcome was the six-month, Food and Drug Administration-approved composite endpoint of rejection, graft failure, or death. Outcomes according to induction exposure were compared using logistic regression, exposure likelihood matching, and outcome risk score matching.

Results—All statistical approaches demonstrated lower rates of the six-month triple endpoint with thymoglobulin compared with basiliximab when steroids were present, with approximately 22% adjusted, relative reduction by logistic regression and 3% absolute reductions by matching approaches. When steroids were absent, risk reduction among thymoglobulin versus basiliximab-treated patients was of larger magnitude but borderline statistical significance. Triple endpoint incidence was lower with both induction regimens compared to no induction across methods. Estimated sample sizes necessary to detect the observed differences between induction types in the presence of steroids in a prospective trial ranged from 1600 to nearly 7000 patients.

Conclusions—Consistency across statistical approaches suggests superiority of thymoglobulin compared to basiliximab or no antibody induction therapy for six-month kidney transplant outcomes

¹Funding Sources: Dr. Brennan received support from a grant from the National Institute of Diabetes Digestive and Kidney Diseases (NIDDK), P30DK079333. Dr. Lentine received support from a grant from the NIDDK, K08DK073036. Additional support was provided to Dr. Schnitzler from Genzyme, Inc.

Corresponding author: Krista L. Lentine, MD, MS, Saint Louis University Center for Outcomes Research, Salus Center, 2nd Floor, 3545 Lafayette Avenue, St. Louis, MO 63130, Phone: (314) 977-9477, Fax: (314) 977-1101, E-mail: lentine.krista@stanfordalumni.org.

Institution at which work was performed:

Saint Louis University Center for Outcomes Research, St. Louis, MO

in the modern immunosuppression era. As the sample sizes necessary to power a prospective superiority trial are likely prohibitive, studies such as these provide clinically relevant information that may not be otherwise attainable.

Keywords

Basiliximab; Immunosuppression; Outcome assessment; Registries; Renal transplantation; Thymoglobulin

Introduction

Antibody induction agents in renal transplantation are highly effective in reducing acute rejection (1) and ultimately in preserving allograft function (2). Use of induction agents has increased over recent years and by 2005, nearly 75% of renal transplants in the United States were performed with induction therapy compared with 39% utilization in 1998 (3). Although thymoglobulin was the most commonly used induction agent in 2005 (3), only basiliximab and daclizumab, both antibodies against the inter-leukin-2 receptor (IL2R Abs), are currently approved by the Food and Drug Administration (FDA) for use as induction agents in kidney transplantation.

There is emerging evidence that thymoglobulin may be associated with favorable outcomes compared with IL2R Abs in some populations. A recent randomized control trial of 278 renal organ recipients at high baseline risk for rejection or delayed graft function demonstrated lower incidence of acute rejection at one year post-transplant with thymoglobulin compared to basiliximab in the context of maintenance cyclosporine and mycophenolate mofetil (MMF) (4). A large, observational, registry-based study of transplants in 1998–2003, in which less than half of subjects received tacrolimus (FK), suggested an approximate 10% relative reduction in one-year rejection but similar two-year graft survival with thymoglobulin compared to IL2R Abs (5). In contrast, small studies in low immunologic-risk patients using cyclosporine as the calcineurin inhibitor suggest similar long-term outcomes with thymoglobulin and ILR Abs, but small sample sizes limit study power (6–8).

Tac has replaced cyclosporine as the most commonly used calcineurin inhibitor in recent years (3), a practice supported by findings of superior rejection risk and slightly better one-year graft survival with low-dose FK compared to low-dose cyclosporine in the context of IL2R induction, MMF and steroids in a large multi-national trial (9). A six-month multi-center trial also found lower rates of six-month clinically-apparent, biopsy-proven and steroid-resistant acute rejection in thymoglobulin-induced patients randomized to tacrolimus-based versus cyclosporine-based maintenance immunosuppression (10). The comparative efficacy of thymoglobulin and basiliximab in the context of FK-based maintenance immunosuppression remains controversial.

Sample size requirements for detection of small but important differences in allograft outcomes after kidney transplantation may prohibit examination of questions of interest within the framework of a randomized trial. As an alternative, the large numbers of observations in national registries provide powerful data for examination of a wide range of clinical outcomes associated with treatment regimens in “real-life” outside of clinical trials, which often restrict participation to subjects not representative of the larger population of interest. Challenging the benefit of statistical power, however, is the lack of treatment randomization. In recent years, important progress has been made in statistical methods for reducing bias from non-randomization in observational analyses with techniques that account for baseline clinical characteristics that may influence outcome (11–13).

In this study, we applied three different statistical analytic approaches – multivariate logistic regression, exposure likelihood matching, and outcome risk score matching – to retrospectively compare early graft outcomes according to antibody induction regimens within registry data of the Organ Procurement and Transplantation Network (OPTN). Specifically, we investigated associations of thymoglobulin, basiliximab or no antibody induction with a composite six-month outcome of acute rejection, graft failure and death in patients receiving Tac and MMF maintenance immunosuppression. We also estimated the sample sizes necessary to detect the observed differences in prospective trials.

Methods

Data and Sample

Data were drawn from the OPTN Standard Transplant Analysis and Research Files for kidney transplants performed between January 2001 and December 2005. All recipients were aged 18 years or older at the time of transplant with no more than one previous transplant, and received kidneys from non-human leukocyte antigen (HLA) identical donors aged 6 years or older. We excluded recipients of multi-organ transplants, and those with peak panel reactive antibody (PRA) levels at discharge greater than 50%, or indicated research study participation within 6 months of transplant. We also excluded transplants in which recipient or the donor was known to be seropositive for hepatitis virus (B or C) or human immunodeficiency virus.

The study sample was limited to transplants managed with thymoglobulin, basiliximab or no induction therapy plus FK and MMF maintenance immunosuppressant agents, with or without corticosteroids, at discharge. The FDA-approved dosing for basiliximab involves one dose before graft reperfusion and a second dose on day four post-transplant. Thymoglobulin induction is most commonly initiated intra-operatively followed by daily doses through day four (4,14), although dosing may be delayed for cytopenias or vary by center-specific protocol. Daclizumab was not considered in the study design because the administration protocol approved by the FDA involves dosing every 14 days to eight weeks post-transplant. This schedule is impractical, and would not be comparable to thymoglobulin or basiliximab in clinical settings or within a blinded trial.

We compared two treatment groups at a time, separating comparisons according to presence or absence of corticosteroids at discharge. Patients receiving no induction therapy were considered the reference group when involved in a comparison. Patients treated with basiliximab constituted the reference group when compared to thymoglobulin-treated subjects.

Outcome measure

The outcome of interest was the FDA-specified composite triple endpoint of allograft rejection, graft failure or patient death used in clinical trials of immunosuppression efficacy. We ascertained occurrence of these events within six months following transplant from OPTN records. Due to the potential for reporting lags to the OPTN, discharge and follow-up records reported within 273 days from the time of transplant were considered. Patient mortality data reported to the OPTN were supplemented with information from the Social Security Death Master File. Acute rejection episodes were defined as biopsy confirmed rejection, the administration of immunosuppressant agents administered for rejection treatment (i.e., antilymphocyte globulin, OKT3, thymoglobulin, and/or steroids), or other reported episodes based on clinical assessment.

Statistical Analyses

Statistical analyses were performed with SAS for windows software, version 9.1 (SAS Institute, Inc., Cary, N.C.). Covariates considered for all statistical approaches included

recipient, donor and transplant factors shown in Table 1. Differences in the frequency distributions of baseline covariates according to induction regimen were compared by the chi-square test. Logistic regression models were designed to compare relative risk of the triple endpoint (odds ratio, OR) between two treatment groups at a time (e.g., thymoglobulin plus corticosteroids versus basiliximab plus corticosteroids). Adjusted models were constructed using stepwise logistic regression with all covariates entered in the first step and treatment status entered in the second, with significance threshold of $P < 0.05$ for variable entry and removal.

In order to reduce selection bias due to non-randomized treatment allocation we performed one-to-one exposure likelihood matching using the greedy matching algorithm of the SAS GMATCH macro (15). This algorithm applies the approach of Rosenbaum which, grounded in research in observational studies, discrete mathematics and computer science research in matching in graphs and networks, identifies matches according to a vector of matching variables (11). Best matches are found by minimizing the weighted sum of the absolute “distances” in clinical covariate values. Logistic regression models predicting the likelihood of “treatment group” exposure status were constructed in cohorts including all subjects eligible for a given comparison based on induction and steroid use, where exposure of interest was defined as thymoglobulin in comparisons to basiliximab, or as the form of antibody induction of interest in comparisons to no antibody induction. Distance is defined as $D_{ij} = \text{SUM} \{ W_k * |X_{ik} - X_{jk}| \}$, where X_{ik} and X_{jk} are the value of variable $X(k)$ for subject i and subject j , respectively, and W_k , the weight assigned to matching predictor variable k , is the parameter estimate for that covariate produced by the logistic regression model predicting exposure status. D_{ij} is thus the weighted sum of the differences over the number of matching predictor variables X (with index k). In the GMATCH macro, the group in each paired comparison with the smallest number of patients forms the starting point for the match process, determining the minimum group size to which a comparator sample will be matched. “Greedy” matching algorithms involve sequential progression through subject lists to find the best available match at each step (16). The subject (j) matched to subject (i) is the one with the smallest D_{ij} within the allowed distance limits. In a one-to-one matching, once a match is formed the pair is removed from the set of subjects available for future matches.

Outcomes risk matching was performed to match subjects in treatment groups of interest with reference patients according to the expected likelihood of the triple-endpoint estimated in each reference group based on observed baseline characteristics. Thus, in contrast to the exposure likelihood matching in which logistic regression models were constructed in all patients eligible for a comparison, here the logistic parameters were estimated only in “reference patients”, defined as basiliximab in comparisons to thymoglobulin or as no antibody induction in comparisons to either form of antibody induction. After distance parameter estimation the matching algorithm proceeded as above.

Observed differences in triple endpoint frequencies between matched groups were compared by the McNemar test for paired proportions. Statistical significance was defined as $P < 0.05$. We estimated the sample sizes necessary to detect observed differences in endpoint proportions at $P < 0.05$ with 80% power in superiority (1-tailed) trial design by SAS Proc Power procedure (17).

The GMATCH procedure also allows specification of the maximal distance allowed for matching. Larger maximal distances facilitate matching greater numbers of patients at the expense of less matching precision. Shorter distances allow greater matching accuracy at the expense of sample size reduction. We conducted sensitivity analyses of the robustness of the matching procedures by adjusting the relative difference in either exposure or outcome-matching between compared groups. The most liberal matching schemes produced matches

that allowed a maximum distance of 2 units between the intervention and reference. The most conservative matching schemes allowed for no distance between reference and comparison patients, such that they had equal weighted sums of predictor values.

Results

Group Difference in Covariate Frequencies

We identified 19,137 eligible patients transplanted in 2001 to 2005 who received maintenance FK and MMF at transplant discharge. As shown in Table 1, distributions of baseline clinical characteristics differed across induction groups. Among patients who received corticosteroids at discharge, those administered thymoglobulin induction therapy compared to basiliximab or no antibody induction were more likely to be black race, non-Hispanic, obese, employed at transplant, and to have angina history, prolonged pretransplant dialysis duration, and renal failure from causes other than diabetes, hypertension or glomerulonephritis. In the context of steroids, thymoglobulin-treated patients were also more likely to be sensitized, to receive DR-mismatched kidney or kidneys from donors who were deceased, older, obese, or with a history of cerebrovascular death, hypertension, or diabetes. Thymoglobulin was used more commonly in patients with delayed graft function. Fewer differences according to antibody induction status were observed among patients without corticosteroids at discharge. Among patients without steroids, those treated with thymoglobulin were more likely to have history of angina, donor-negative recipient-positive cytomegalovirus seropairing, to receive kidneys from donors who were deceased or older than 60 years of age, and to have experienced delayed graft function.

Group Differences on the Rates of Triple Endpoint

Logistic Regression Comparisons—Among patients receiving steroids at discharge, relative risk of triple endpoint by logistic regression was lower in thymoglobulin-induced patients compared to those who received basiliximab (unadjusted OR 0.88, 95% CI 0.79–0.99; covariate-adjusted OR 0.78, 95% CI 0.69–0.87) or no antibody induction (unadjusted OR 0.78, 95% CI 0.71–0.86; adjusted OR 0.64, 95% CI 0.58–0.71). Risk of the triple endpoint in patients treated with basiliximab was lower than those who did not receive antibody induction (unadjusted OR 0.88, 95% CI 0.79–0.97; adjusted OR 0.82, 95% CI 0.74–0.92). When steroids were absent at discharge, the adjusted risk reduction for the triple endpoint in thymoglobulin-treated compared to basiliximab-treated patients was significant and appeared somewhat larger than in steroid-free samples (unadjusted OR 0.75, 95% CI 0.52–1.08; adjusted OR 0.66, 95% CI 0.44–1.00). Triple endpoint risk was markedly lower in thymoglobulin-treated patients than in patients without antibody induction in the absence of steroids (unadjusted OR 0.40, 95% CI 0.29–0.56; adjusted OR 0.36, 95% CI 0.25–0.52). Without steroids, triple endpoint risk with basiliximab was significantly lower than with no antibody induction only in unadjusted comparison (unadjusted OR 0.54, 95% CI 0.36–0.82; adjusted OR 0.69, 95% CI 0.42–1.11).

Acute rejection was the most common component of the triple endpoint across treatment groups (Figure 1). Effect sizes in analyses performed for rejection alone were similar to those obtained for the composite endpoint.

Matching for Exposure Likelihood—After matching for expected likelihood of exposure the “intervention arm” based on baseline characteristics, risk of the six-month triple endpoint was significantly lower in thymoglobulin versus basiliximab-induced groups regardless of steroid status (Figure 2). The largest absolute difference occurred when no steroids were used at discharge, with 4.7% absolute reduction in outcome frequency in the thymoglobulin versus basiliximab group compared with 2.7% absolute reduction in the presence of steroids. Compared to no antibody induction, outcome frequency after exposure likelihood matching

was lower for induction by both thymoglobulin (3.2% and 10.5% reductions with and without steroids, respectively) and basiliximab (2.8% and 7.6% reductions with and without steroids, respectively).

Requiring a closer match in the exposure likelihood distance vector reduced available sample sizes for comparison and led to loss of statistical significance in endpoint differences with thymoglobulin versus basiliximab induction when the maximum allowable difference in exposure likelihood distance was less than approximately 0.4 in the presence of discharge steroids or 0.8 in the absence of steroids (Figure 3). However, a pattern of lower events in the thymoglobulin group per the point estimates was preserved.

Outcome Risk Score Matching—Following matching for expected reference risk of the triple endpoint based on baseline characteristics, six-month triple-endpoint frequency was significantly lower by 3.1% with thymoglobulin versus basiliximab in presence of steroids and trended lower by 3.9% when steroids were absent ($P=0.05$). Compared to no antibody induction, outcome frequency after matching by reference outcome risk was lower for induction by both thymoglobulin (4.6% and 11.3% reductions with and without steroids, respectively) and basiliximab (3.0% and 7.8% reductions with and without steroids, respectively).

Sensitivity analyses varying the stringency of matching in the distance vector revealed that, regardless of steroid status at discharge, the rates of triple endpoint associated with thymoglobulin were generally lower compared with basiliximab (Figure 4). While all comparisons were statistically significant when steroids were absent, statistical significance was lost with conservative matches requiring less than 0.2 distance-unit differences in baseline clinically expected outcome risk.

Sample Size Estimates—Estimated sample sizes that would be required to detect the observed unadjusted, exposure likelihood matched, and outcome risk score matched event proportion differences in a clinical trial are presented in Table 2. Due to the relatively smaller differences in observed event frequencies when steroids were present at discharge, the estimated number of subjects was much higher than when steroids were absent. The smallest overall sample sizes were estimated using event frequencies produced with the outcome risk score matching procedure.

Discussion

We examined the OPTN registry to compare six-month outcomes according to induction protocol (thymoglobulin, basiliximab, or no antibody induction) in kidney transplant recipients who were also treated with maintenance FK and MMF, with or without corticosteroids. To adjust for potential selection biases underlying induction protocol choice and statistically approximate randomized treatment assignment, we applied three analytic approaches in this investigation: logistic regression, matching by clinically predicted likelihood of induction exposure, and matching by clinically predicted outcome risk. We found that thymoglobulin was associated with superior outcomes compared with basiliximab, and that both forms of induction appeared superior to no antibody induction. Results were generally consistent across analytic approaches, and were robust in sensitivity analyses with varied stringencies for matching precision. Sample size simulations indicated that in the presence of steroids, thousands of subjects would be needed to conduct superiority trials for induction type or for either regimen versus no induction.

Patients in the OPTN registry are not randomized to treatments, but rather treatments are determined in the course of real-life clinical practice. We observed that transplants performed

with thymoglobulin induction were more commonly associated with traditional risk factors for graft and patient loss such as recipient sensitization and black race, deceased donors, older donor age, donor stroke, and delayed graft function. While the reasons for differences in clinical profiles across treatment groups cannot be determined from observational data, it is common practice for transplant centers to rely on standard protocols, patient risk profiles, or both to determine which induction agents, if any, to use (18). We therefore applied several statistical methods to adjust for observed baseline clinical differences and found consistent results.

The largest differences in the six-month composite outcome according to induction regimen were observed in recipients who were steroid-free at transplant discharge. Early attempts at steroid avoidance/withdrawal have been associated with unacceptable rates of acute rejection and graft failure (19–21); however, concomitant antibody induction has yielded acceptable results in some samples including low immunologic risk groups, pediatric samples, and patients receiving FK and sirolimus (22–24). Notably, a recent open-label non-inferiority trial comparing 12-month renal function as the primary endpoint among 336 de novo transplant recipients randomized to steroid avoidance, early steroid withdrawal, or “standard-steroids” in the context of basiliximab induction, cyclosporine and MMF found that estimated GFR was highest in the standard-steroids group and did not satisfy the a priori non-inferiority criterion (25). The 12-month incidence of a composite outcome of acute rejection, graft loss or death was also less favorable in the steroid-free and steroid withdrawal compared to standard steroids group. Data on the comparative efficacy of antibody induction agents with steroid-avoidance/withdrawal maintenance immunosuppression are sparse. Consistent with our results, one small study of 89 type 1 diabetic kidney transplant recipients treated with FK, MMF and rapid steroid withdrawal found superior six-month rejection free survival with thymoglobulin versus basiliximab (26).

An important aspect of this study is that statistical adjustments considered a wide range of clinical factors known to be associated with transplant outcomes. We employed multivariable logistic modeling, exposure likelihood matching, and outcome risk score matching in order to cross-validate results. The three approaches share a common goal of minimizing bias associated with treatment assignment through statistical adjustment of covariates but are technically different. In multivariate logistic regression, the set of considered covariates is assumed to account for covariate-associated outcome variance and thus improve assessment of the portion of outcome variance related to treatment. However, independence of the errors is also assumed, and violation of this assumption with non-random assignment may produce bias. Exposure likelihood matching classifies patients by characteristics relevant to intervention assignment and does not depend on the relationships between covariates and outcome. We thus also employed outcome risk score matching, a method designed to match subjects based on the predicted likelihood of outcome suggested by baseline characteristics. The rationale for this approach is similar to that of multiple logistic regression analysis, except that maximum likelihood estimates are derived from associations observed in a reference group, thus minimizing biases for intervention subjects with over- or under-represented covariates. While each statistical approach may have idiosyncratic strengths and limitations for controlling bias, consistency of results across methods speaks to the robustness of findings. Notably in our sample size simulations, smallest overall group sizes were estimated using outcome risk score matching, suggesting that controlling patient profiles related to graft failure, acute rejection, and/or death in clinical trial design may improve power.

While our results are suggestive, caution is warranted. Induction regimens may not be the only explanation for group-associated outcome differences as the decision to implement a particular induction agent and to use steroids and other immunosuppressant agents may be influenced by factors not available in the registry, such as post-surgical hypertensive status or center-specific treatment preferences. A major assumption for all three statistical approaches applied in this

study is that all relevant covariates are considered in the analyses. Results may be impacted by factors such as maintenance immunosuppression levels and other center-specific ancillary treatment protocols. We included patients who received FK and MMF at discharge on an intent-to-treat basis and there were likely patients who switched maintenance regimens during the observation period. We studied six-month outcomes as an initial investigation, and longer-term assessments are warranted. Further, although the triple endpoint of rejection, graft failure and death is FDA-approved, it does not encompass other clinical events that are important in comprehensive understanding of post-transplant outcomes such as non-fatal infections and malignancies. Despite these limitations, the OPTN registry contains high-quality baseline and follow-up data that have been used to study many topics in transplantation, and recent studies have shown high level of agreement between OPTN records and other administrative and clinical data sources (27–29).

In conclusion, we explored a national registry of kidney transplant recipients to compare six-month outcomes associated with antibody induction among large patient samples receiving maintenance FK and MMF. We applied three distinct statistical approaches to adjust for possible selection bias. Composite outcomes associated with thymoglobulin were favorable to that of basiliximab, and both induction methods appeared superior to no antibody induction across analytic approaches, suggesting that findings are robust. As the sample size necessary to detect small but significant induction-related differences in six-month outcomes are likely to be prohibitive for execution of clinical trials, observational approaches may provide valuable information that would not be otherwise attainable.

Abbreviations

FDA	Food and Drug Administration
FK	Tacrolimus
IL2R Abs	Inter-leukin-2 receptor
MMF	Mycophenolate mofetil
OPTN	Organ Procurement and Transplantation Network
OR	Odds ratio
PRA	Panel reactive antibody

Acknowledgments

The data reported here have been supplied by United Network for Organ Sharing as the contractor for the OPTN. The interpretation and reporting of these data are the responsibility of the authors and should in no way be seen as representing official policy of or interpretation by the OPTN or the U.S. Government. Dr. Lentine received support from a grant from the National Institute of Diabetes Digestive and Kidney Diseases (NIDDK), K08DK073036. Dr. Brennan received support from a grant from the NIDDK, P30DK079333. Additional support was provided to Dr. Schnitzler from Genzyme.

References

1. Castro MC, Araujo LM, Nahas WC, Arap S, David-Neto E, Ianhez LE. Induction versus noninduction therapy in kidney transplantation: considering different PRA levels and different induction therapies. *Transplant Proc* 2004;36(4):874. [PubMed: 15194299]
2. Cherikh WS, Kauffman HM, McBride MA, Maghirang J, Swinnen LJ, Hanto DW. Association of the type of induction immunosuppression with posttransplant lymphoproliferative disorder, graft survival, and patient survival after primary kidney transplantation.[see comment]. *Transplantation* 2003;76(9):1289. [PubMed: 14627905]
3. OPTN/SRTR Annual Report. 2007. http://www.ustransplant.org/annual_reports/current/

4. Brennan DC, Daller JA, Lake KD, Cibrik D, Del Castillo D, Thymoglobulin Induction Study G. Rabbit antithymocyte globulin versus basiliximab in renal transplantation.[see comment]. *New England Journal of Medicine* 2006;355(19):1967. [PubMed: 17093248]
5. Patlolla V, Zhong X, Reed GW, Mandelbrot DA. Efficacy of anti-IL-2 receptor antibodies compared to no induction and to antilymphocyte antibodies in renal transplantation. *Am J Transplant* 2007;7(7):1832. [PubMed: 17564638]
6. Lebranchu Y, Bridoux F, Buchler M, et al. Ithat thymoglobulin yielded *Am J Transplant* 2002;2(1):48.
7. Mourad G, Rostaing L, Legendre C, Garrigue V, Thervet E, Durand D. Sequential protocols using basiliximab versus antithymocyte globulins in renal-transplant patients receiving mycophenolate mofetil and steroids. *Transplantation* 2004;78(4):584. [PubMed: 15446319]
8. Al Najjar A, Etienne I, Le Pogamp P, et al. Long-term results of monoclonal anti-IL2-receptor antibody versus polyclonal antilymphocyte antibodies as induction therapy in renal transplantation. *Transplantation Proceedings* 2006;38(7):2298. [PubMed: 16980070]
9. Ekberg H, Tedesco-Silva H, Demirbas A, et al. Reduced exposure to calcineurin inhibitors in renal transplantation. *N Engl J Med* 2007;357(25):2562. [PubMed: 18094377]
10. Charpentier B, Rostaing L, Berthoux F, et al. A three-arm study comparing immediate tacrolimus therapy with antithymocyte globulin induction therapy followed by tacrolimus or cyclosporine A in adult renal transplant recipients. *Transplantation* 2003;75(6):844. [PubMed: 12660513]
11. Rosenbaum PL. Optimal Matching for observational studies. *J Am Statistical Association* 1989;84(408):1024.
12. D'Agostino RB Jr. Propensity score methods for bias reduction in the comparison of a treatment to a non-randomized control group. *Stat Med* 1998;17(19):2265. [PubMed: 9802183]
13. Haro JM, Kontodimas S, Negrin MA, Ratcliffe M, Suarez D, Windmeijer F. Methodological aspects in the assessment of treatment effects in observational health outcomes studies. *Appl Health Econ Health Policy* 2006;5(1):11. [PubMed: 16774289]
14. Gaber OA, Schnitzler MA, Willoughby LM, Stirnemann PM. Thymoglobulin induction in living donor renal transplant recipients: Report from TAILOR registry. *Am J Transplant* 2006;6:293.
15. Mandrekar, JN.; Mandrekar, SJ. An introduction to matching and its applications using SAS. *Proceedings of the 29th Annual SAS Users' Group International (SUGI) Conference*; 2004. p. 208
16. Bergstralh EJ, Kosanke JL, Jacobsen SJ. Software for optimal matching in observational studies. *Epidemiology* 1996;7(3):331. [PubMed: 8728456]
17. Bauer, D. Proc Power in SAS 9.1. *Proceedings of the 29th Annual SAS Users' Group International (SUGI) Conference paper*; p. 195
18. Knight RJ, Kerman RH, Schoenberg L, et al. The selective use of basiliximab versus thymoglobulin in combination with sirolimus for cadaveric renal transplant recipients at low risk versus high risk for delayed graft function. *Transplantation* 2004;78(6):904. [PubMed: 15385812]
19. Ahsan N, Hricik D, Matas A, et al. Prednisone withdrawal in kidney transplant recipients on cyclosporine and mycophenolate mofetil--a prospective randomized study. *Steroid Withdrawal Study Group. Transplantation* 1999;68(12):1865. [PubMed: 10628766]
20. Vanrenterghem Y, Lebranchu Y, Hene R, Oppenheimer F, Ekberg H. Double-blind comparison of two corticosteroid regimens plus mycophenolate mofetil and cyclosporine for prevention of acute renal allograft rejection. *Transplantation* 2000;70(9):1352. [PubMed: 11087152]
21. Kasiske BL, Chakkera HA, Louis TA, Ma JZ. A meta-analysis of immunosuppression withdrawal trials in renal transplantation. *J Am Soc Nephrol* 2000;11(10):1910. [PubMed: 11004223]
22. Laouad I, Halimi JM, Buchler M, et al. Recipient age and mycophenolate mofetil as the main determinants of outcome after steroid withdrawal: analysis of long-term follow-up in renal transplantation. *Transplantation* 2005;80(6):872. [PubMed: 16210979]
23. Oberholzer J, John E, Lumpaopong A, et al. Early discontinuation of steroids is safe and effective in pediatric kidney transplant recipients. *Pediatr Transplant* 2005;9(4):456. [PubMed: 16048597]
24. Woodle ES, Vincenti F, Lorber MI, et al. A multicenter pilot study of early (4-day) steroid cessation in renal transplant recipients under simulect, tacrolimus and sirolimus. *Am J Transplant* 2005;5(1):157. [PubMed: 15636625]

25. Vincenti F, Schena FP, Paraskevas S, Hauser IA, Walker RG, Grinyo J. A randomized, multicenter study of steroid avoidance, early steroid withdrawal or standard steroid therapy in kidney transplant recipients. *Am J Transplant* 2008;8(2):307. [PubMed: 18211506]
26. Heilman RL, Reddy KS, Mazur MJ, et al. Acute rejection risk in kidney transplant recipients on steroid-avoidance immunosuppression receiving induction with either antithymocyte globulin or basiliximab. *Transplant Proc* 2006;38(5):1307. [PubMed: 16797289]
27. Stirnemann PM, Takemoto SK, Schnitzler MA, et al. Agreement of immunosuppression regimens described in Medicare pharmacy claims with the Organ Procurement and Transplantation Network survey. *Journal of the American Society of Nephrology* 2006;17(8):2299. [PubMed: 16825328]
28. Gilmore AS, Helderman JH, Ricci JF, et al. Linking the US transplant registry to administrative claims data: expanding the potential of transplant research. *Med Care* 2007;45(6):529. [PubMed: 17515780]
29. Buchanan PM, Schnitzler MA, Brennan DC, et al. Novel methods for tracking long-term maintenance immunosuppression regimens. *Clin J Am Soc Nephrol* 2008;3(1):117. [PubMed: 18077785]

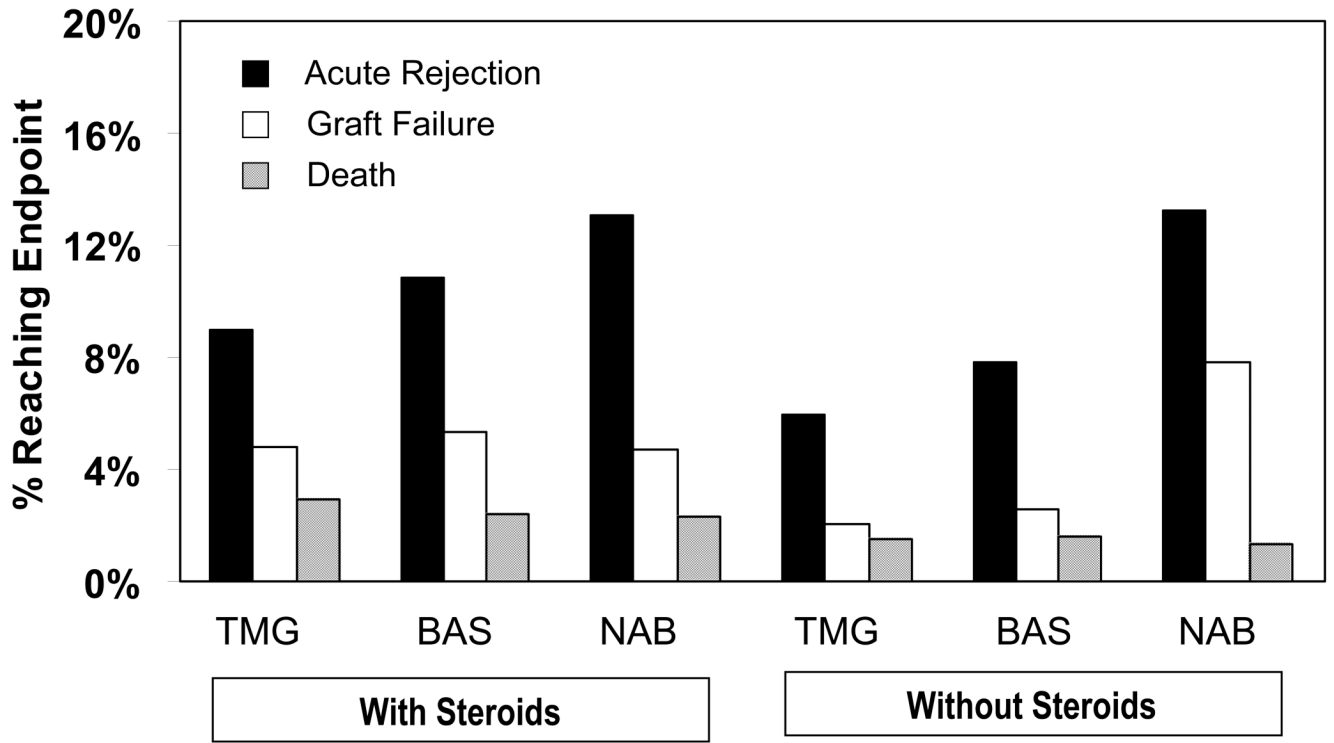
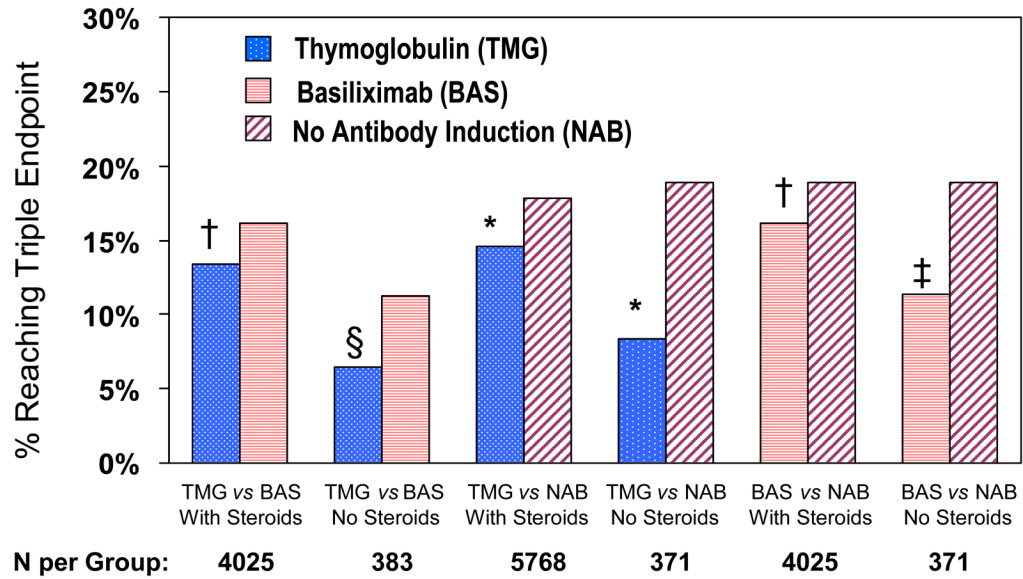


Figure 1. Observed frequencies of individual components of the triple endpoint by induction regimen, with and without discharge corticosteroids.

Exposure-Likelihood Matched Comparisons



Outcome-Risk Score Matched Comparisons

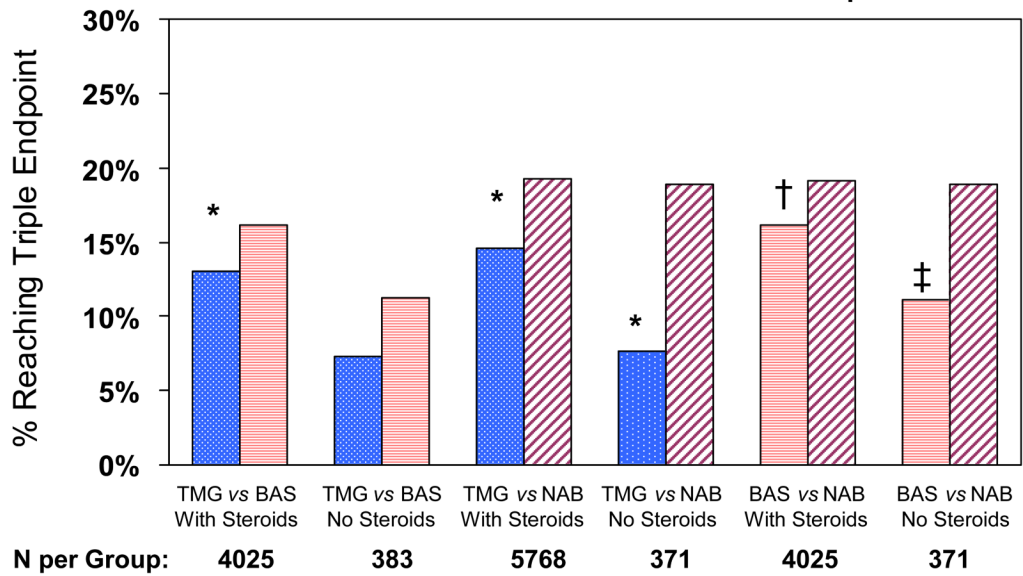
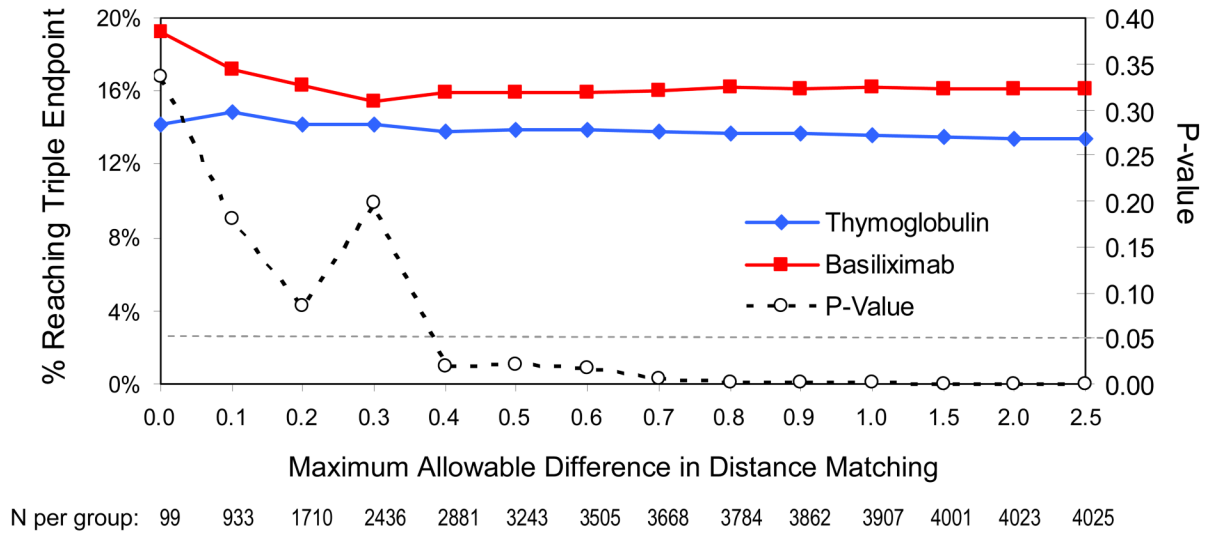


Figure 2. Exposure likelihood and outcome risk-matched comparisons of six-month triple endpoint by induction regimen and discharge steroid use.

* $p < 0.0001$; † $0.0001 \geq p < 0.001$; ‡ $0.001 > p \geq 0.01$; § $0.01 > p \geq 0.05$

A. With Discharge Steroids



B. Without Discharge Steroids

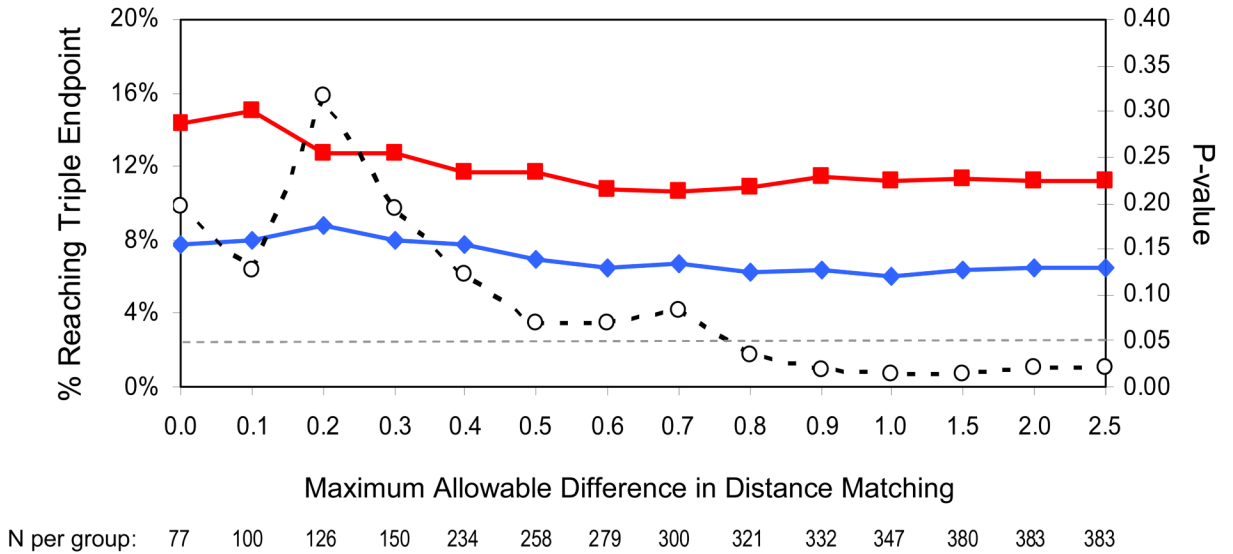
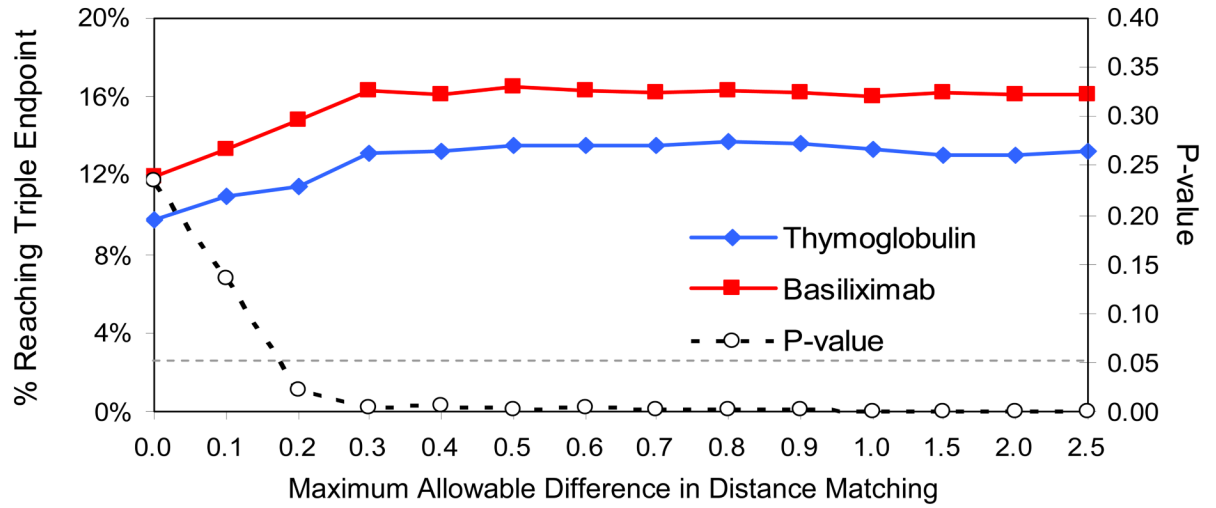


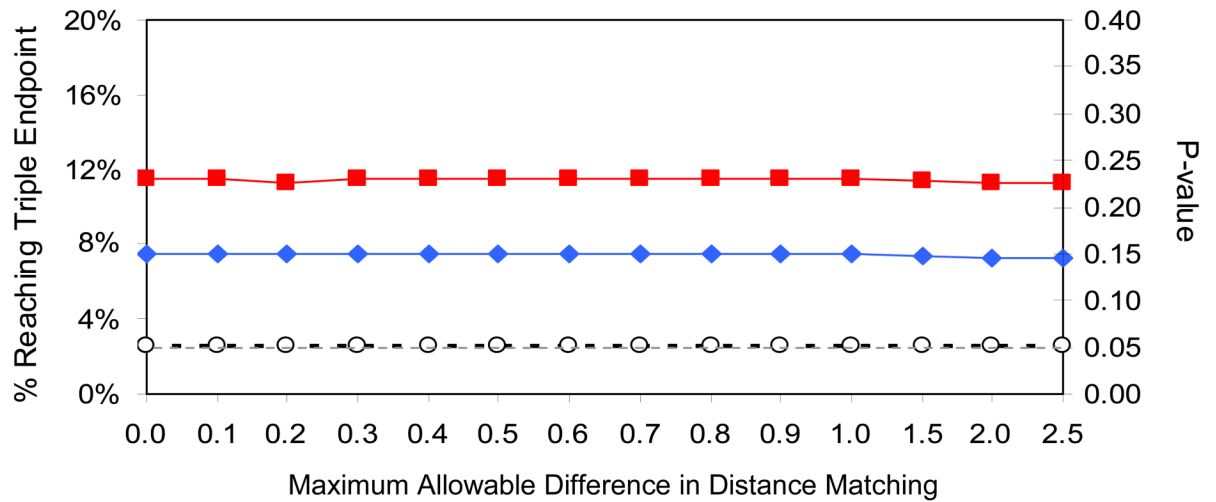
Figure 3. Sensitivity analysis of exposure likelihood-matched comparisons of thymoglobulin versus basiliximab across maximum allowable matching differences.

A. With Discharge Steroids



N per group: 543 855 1024 1945 2216 2697 2781 3151 3259 3354 3624 3946 4020 4025

B. Without Discharge Steroids



N per group: 373 373 373 374 374 374 374 374 374 374 374 379 383 383

Figure 4. Sensitivity analysis of outcome risk score-matched comparisons of thymoglobulin versus basiliximab across maximum allowable matching differences.

Table 1

Patient characteristics by induction therapy and steroid use at discharge.

	Steroids at Discharge			No Steroids at Discharge		
	TMG	BAS	NAB	TMG	BAS	NAB
Total N=19,137	5768	4025	7139	1451	383	371
Recipient characteristics						
Age	<i>I</i>					
18–30 years	11.1	12.5	12.4	12.2	11.5	12.1
31–44 years	24.9	25.1	26.2	23.4	21.7	23.7
45–59 years	39.9	38.9	39.8	38.5	40.0	42.3
60+ years	24.2	23.5	21.7	26.0	26.9	21.8
Gender	<i>I, 3</i>					
Female	39.9	37.1	37.7	36.7	33.2	36.7
Male	60.1	62.9	62.3	63.3	66.8	63.3
Race	<i>I, 3</i>					
White	58.0	59.8	56.7	63.4	71.8	63.3
Black	26.0	21.9	23.5	16.8	8.4	18.1
Other	16.0	18.3	19.9	19.8	19.8	18.6
Hispanic ethnicity	10.7	13.5	14.0	11.7	13.8	13.2
ESRD cause	<i>I, 3</i>					
Diabetes	23.4	24.6	22.6	24.2	21.9	24.8
Hypertension	19.8	19.9	22.2	19.1	19.1	20.0
Glomerulonephritis	20.4	22.3	22.1	24.8	23.5	25.6
Other	36.5	33.3	33.1	31.9	35.5	29.7
Pretransplant dialysis duration (months)	<i>I, 3</i>					
No dialysis indicated	17.2	16.3	20	19.2	24.3	18.8
1–12	15.6	17.3	16.4	18.8	18.0	24.3
13–24	14.3	17.2	15.7	16.6	20.6	15.4
25–60	31.0	31.0	26.4	27.8	22.7	27.8
≥61	16.1	12.4	10.9	12.3	8.6	10.0
Dialysis indicated, without reported date	5.9	5.8	10.6	5.3	5.7	3.8

	Steroids at Discharge			No Steroids at Discharge		
	TMG	BAS	NAB	TMG	BAS	NAB
Body mass index	2					
Normal (<25)	1.3	29.1	31.2	29.7	29.0	31.8
Overweight (25 ≤ BMI <30)		33.4	35.9	35.4	34.7	35
Obese (≥30)		34.8	32.2	30.2	35.4	31.5
Not reported		2.6	0.8	4.7	1	1.6
Comorbidities						
Hypertension		81.3 ¹	81.6 ²	73.4	85.2 ³	84.1
Diabetes		30.6 ¹	30.3	28.8	30.1	28.7
Angina		2.1 ^{1,3}	0.7	0.8	4.0 ^{1,3}	1.8
Peripheral vascular disease		4.1 ¹	3.5 ²	2.4	3.8	5.7
Cerebrovascular disease		2.5	2.6 ²	2.0	2.1	1.6
Chronic obstructive pulmonary disease		1.0 ¹	0.9	0.6	0.6	0.3
PRA >10%		17.3 ^{1,3}	14.6	13.4	13.2	12.8
Normal functional status		85.8 ^{1,3}	89.2 ²	79.4	79.9 ^{1,3}	86.7
Employed at transplant		13.4 ^{1,3}	9.0	8.5	21.5 ¹	17.0 ²
Donor characteristics						
Deceased		61.2 ^{1,3}	57.4 ²	51.5	49.4 ^{1,3}	40.0 ²
Age group	1.3 2 1					
6–18 years		7.7	9.8	7.6	7.2	6.3
19–3 years		19.7	22.8	22.2	21.6	22.7
31–44 years		30.5	31.1	32.3	31.4	36.3
45–59 years		32.5	30.3	30.9	32.0	30.6
60+ years		9.6	6.0	7.1	7.9	4.2
Gender						
Female		48.0	47.4	49.2	51.7	53.6
Male		52.0	52.6	50.8	48.3	47.5
Race	1.3 2 3 2					
White		73.4	73.1	70.0	73.1	75.7
Black		12.7	11.5	12.8	11.2	6.5

	Steroids at Discharge			No Steroids at Discharge		
	TMG	BAS	NAB	TMG	BAS	NAB
Other	13.9	15.4	17.2	15.8	17.8	17.0
Hispanic ethnicity	10.8 ^{1,3}	12.2 ²	13.9	11.7	12.8	12.7
Body Mass Index	1.3	2		1		
Normal (<25)	30.5	34.4	29.4	29.4	33.4	29.1
Overweight (25 ≤ BMI <30)	33.1	33.0	31.4	33.2	33.2	34.8
Obese (≥30)	25.9	23.7	23.9	21.6	21.9	27.2
Not reported	10.6	8.9	15.4	15.7	11.5	8.9
Comorbidities						
Cerebrovascular death	27.2 ^{1,3}	20.8	21.4	19.7 ¹	15.4	12.4
Hypertension	16.3 ^{1,3}	11.0	11.7	11.1 ¹	9.7	5.9
Diabetes	3.6 ^{1,3}	2.2	2.7	2.1	1.6	1.1
Transplant Factors						
Delayed Graft Function	20.9 ^{1,3}	12.2 ²	10.3	11.2 ^{1,3}	3.9	5.9
HLA Mismatch	1.3			3		
DR mismatch alone	75.7	73.1	72.3	74.0	68.7	73.9
AB mismatch alone	18.0	20.7	20.4	19.2	19.1	16.7
DR and AB mismatch	0.9	0.9	1.2	1.5	3.9	1.1
No mismatch	5.4	5.3	6.2	5.4	8.4	8.4
CMV sero-pairing	1.3	2		1.3		
Donor-/Recipient+	48.3	47.9	40.9	44.5	37.3	32.1
Donor-/Recipient-	32.4	30.5	38.7	33.3	32.6	32.1
Donor+/Recipient+	13.3	14.4	13.7	15.8	20.4	25.1
Donor+/Recipient-	6.0	7.2	6.6	6.5	9.7	10.8
Cold ischemia time	1.3	2		1.3	2	
0-12 hours	37.5	41.8	37.8	39.0	45.4	61.2
13-24 hours	29.8	27.6	23.0	26.8	18.5	15.1
25-36 hours	9.5	9.8	7.8	8.0	5.2	7.6
37+ hours	1.8	0.9	1.8	0.9	0.5	0.8
Not reported	21.5	20	29.5	25.3	30.3	15.4

BAS, basiliximab; HLA, human leukocyte antigen; ESRD, end-stage renal disease; NAB, no antibody induction; PRA, panel reactive antibody; TMG, thymoglobulin

Chi-square test for difference in frequency distributions:

¹ P<0.05 comparing TMG and NAB

² P<0.05 comparing BAS and NAB

³ P<0.05 comparing TMG and BAS

Table 2

Minimum sample sizes required per group to detect differences in the triple endpoint between treatment groups in a superiority trial (1-tailed test) with 80% power and alpha <0.05.

Comparison	Unadjusted proportions		Propensity Matched		Outcome Risk Matched	
	+ steroids	- steroids	+ steroids	- steroids	+ steroids	- steroids
TMG < BAS	6,876	1,609	2,090	452	1,601	676
TMG < NAB	1,511	139	1,602	131	818	110
BAS < NAB	5,339	270	2,266	277	2,046	257

BAS, basiliximab; NAB, no antibody induction; TMG, thymoglobulin