

Author's Accepted Manuscript

Development and validation of a major adverse transplant event (MATE) score to predict late graft loss in pediatric heart transplantation

Christopher S. Almond, Helena Hoen, Joseph W. Rossano, Chesney Castleberry, Scott Auerbach, Lingyao Yang, Ashwin K. Lal, Melanie D. Everitt, Matthew Fenton, Seth A. Hollander, Elfriede Pahl, Elizabeth Pruitt, David N. Rosenthal, Doff B. McElhinney, Kevin P Daly, Manisha Desai



PII: S1053-2498(17)31728-X
DOI: <http://dx.doi.org/10.1016/j.healun.2017.03.013>
Reference: HEALUN6486

To appear in: *Journal of Heart and Lung Transplantation*

Cite this article as: Christopher S. Almond, Helena Hoen, Joseph W. Rossano, Chesney Castleberry, Scott Auerbach, Lingyao Yang, Ashwin K. Lal, Melanie D. Everitt, Matthew Fenton, Seth A. Hollander, Elfriede Pahl, Elizabeth Pruitt, David N. Rosenthal, Doff B. McElhinney, Kevin P Daly and Manisha Desai, Development and validation of a major adverse transplant event (MATE) score to predict late graft loss in pediatric heart transplantation, *Journal of Heart and Lung Transplantation*, <http://dx.doi.org/10.1016/j.healun.2017.03.013>

This is a PDF file of an unedited manuscript that has been accepted for publication. As a service to our customers we are providing this early version of the manuscript. The manuscript will undergo copyediting, typesetting, and review of the resulting galley proof before it is published in its final citable form. Please note that during the production process errors may be discovered which could affect the content, and all legal disclaimers that apply to the journal pertain.

**DEVELOPMENT AND VALIDATION OF A MAJOR ADVERSE TRANSPLANT EVENT (MATE) SCORE TO PREDICT
LATE GRAFT LOSS IN PEDIATRIC HEART TRANSPLANTATION**

Christopher S. Almond MD, MPH^{1*}, Helena Hoen MS^{1*}, Joseph W. Rossano MD², Chesney Castleberry MD³,
Scott Auerbach MD⁴, Lingyao Yang MS¹, Ashwin K. Lal MD⁵, Melanie D. Everitt MD⁴, Matthew Fenton MD⁶,
Seth A. Hollander MD¹, Elfriede Pahl MD⁷, Elizabeth Pruitt MPH⁸, David N. Rosenthal MD¹, Doff B.

McElhinney MD MS¹, Kevin P Daly MD^{3*}, and Manisha Desai, PhD^{1*}

On behalf of the Pediatric Heart Transplant Study (PHTS) Group Registry

From ¹Stanford University, the Department of Pediatrics (Cardiology), Palo Alto, CA; ²Children's Hospital of Philadelphia, the Department of Cardiology, Philadelphia, PA; ³St. Louis Children's Hospital, the Division of Pediatric Cardiology, St. Louis, MO; ⁴The Children's Hospital Colorado, Division of Pediatric Cardiology, Aurora, CO; ⁵Primary Children's Hospital, the Division of Pediatric Cardiology, Salt Lake City, UT; ⁶Great Ormond Street Hospital, London, England; the ⁸Pediatric Heart Transplant Study Group, University of Alabama, Birmingham, AL; Lurie Children's hospital, Chicago, IL; and ⁶Boston Children's Hospital, the Department of Cardiology, Boston, MA

*†The authors contributed equally to the report.

Please address correspondence to: Dr. Christopher S. Almond, Stanford University, 750 Welch Road, Suite 325, Palo Alto, CA 94304. Fax (650) 725-8343, Phone (650) 723-7913, christopher.almond@stanford.edu

Subject Codes: [37] CV surgery: transplantation, [110] Congestive heart failure, [41] Pediatric and congenital heart disease, including cardiovascular surgery, ventricular assistance, cardiomyopathy

ABSTRACT

BACKGROUND: There is inadequate power to perform a valid clinical trial in pediatric heart transplantation (HT) using a conventional endpoint because the disease is rare and hard endpoints like death or graft loss infrequent. We sought to develop and validate a surrogate endpoint involving the cumulative burden of post-transplant complications to predict death/graft loss in order to power a randomized clinical trial of maintenance immunosuppression in pediatric HT.

METHODS: Pediatric Heart Transplant Study (PHTS) data were used to identify all children who underwent an isolated orthotopic heart transplant (HT) between 2005 and 2014 who survived to 6 months post-HT. Longitudinal mixed effects modeling was used to develop and evaluate a surrogate endpoint comprised of six major adverse transplant events (MATE) (ACR, AMR, Infection, CAV, PTLD and CKD) occurring between 6 and 36 months where individual events were defined according to international guidelines. Two-thirds of the study cohort was used for score development; one-third to test the score.

RESULTS: Among 2,118 children, 6.4% experienced graft loss between 6 and 36 months post-HT while 39% developed CKD, 34% ACR, 34% infection, 9% AMR, 4% CAV and 2% PTLD. The best predictive score involved a simple MATE score sum yielding a concordance probability estimate (CPE)-statistic of 0.74. Whereas the power to detect non-inferiority (NI) assuming a NI hazard ratio of 1.45 in graft survival was 10% (assuming 200 subjects and 6% graft loss rate), the power to detect NI assuming a 2-point non-inferiority margin was >85% using the MATE score.

CONCLUSION: The MATE score reflects the cumulative burden of major adverse transplant events and has acceptable prediction characteristics for death/graft loss post-heart transplant. The MATE score may be useful as a surrogate endpoint to power a clinical trial in pediatric heart transplantation.

Abstract: 289 words.

KEYWORDS: orphan and rare disease, pediatrics, surrogate endpoints, study design, heart transplantation, heart failure, risk-prediction model, time-varying hazards, serious adverse events, congenital heart disease.

BACKGROUND

At present there are no FDA approved immunosuppressants for pediatric heart transplant. Nor has there ever been a successful randomized clinical trial in pediatric heart transplantation. While many study design challenges exist, perhaps the single greatest challenge is the lack of statistical power stemming from the rareness of the disease (small sample size) and low event rate (e.g., death or graft loss). For example, to detect an inordinately large two-fold difference in graft survival, which might not emerge for 5-10 years given the favorable transplant outcomes in the current era, one would need over 500 subjects to achieve adequate power. To detect a more realistic treatment difference, such as a 10% improvement in graft loss over a 3-year study period, over 6,000 subjects would be required.

One established strategy to decrease sample size is to increase the event rate by using a composite endpoint^{1,2}. In adult cardiovascular trials, MACE (major adverse cardiac events) is perhaps the most widely used composite endpoint, and includes mortality, non-fatal MI and non-fatal stroke. A second strategy to decrease sample size is to use an ordinal rather than a binary endpoint³⁻⁵. The Modified Rankin Scale (MRS), a 6-point stroke severity scale, has been used successfully to reduce sample size in stroke trials³. We hypothesized that by combining both of these strategies; we could develop a novel surrogate endpoint for a pediatric heart transplant trial comprised of major adverse transplant events (MATE) using a sample size that is sufficient yet feasible to conduct a randomized trial in pediatric HT medicine.

Therefore, the specific aims of this study are (1) to develop and evaluate a surrogate endpoint comprised of major adverse transplant events (MATEs) that predicts graft loss⁶; and (2) to estimate the sample size

necessary for a randomized clinical trial. The broader purpose of this study is to find better ways to measure the safety and efficacy of competing immunosuppressant regimens to improve the overall safety and survival of children undergoing heart transplantation.

METHODS

Study Population and Data Source

Pediatric Heart Transplant Study Group (PHTS) Registry: PHTS Registry data were used to identify all children <18 years of age who underwent primary orthotopic heart transplant between January 1, 2005 and December 31, 2014 and survived to 6 months post-HT. Subjects undergoing heart re-transplant, multi-organ transplant and heterotopic HT were excluded. The PHTS database is an event-driven pediatric heart transplant registry that collects information from 54 North American centers, 2 centers in England and one in Brazil (see **SUPPLEMENTAL TABLE 1** in the Appendix for a complete list of participating centers). The details of the registry have been reported previously⁷. The PHTS registry was selected because it captures the vast majority of pediatric heart transplants in the US while also providing the most detailed information available regarding individual adverse events, severity level, and outcome. Each center maintains Institutional Review Board approval or obtains exemption according to institutional guidelines. The data were obtained under a limited data use agreement data set provided to Stanford University for this analysis.

Study Definitions and Endpoints

The primary aim of this study was to develop and validate a surrogate endpoint comprised of major adverse transplant events occurring between 6 and 36 months post-transplant that can predict late graft loss after pediatric heart transplantation. The 30-month window from 6 to 36 months post-HT was chosen because it simulated the projected time frame of a randomized clinical trial being planned by the authors and FDA to evaluate the safety and efficacy of everolimus conversion at 6 months after pediatric heart transplant. The baseline characteristics of the study cohort were defined at the time of transplant. The level of

hemodynamic support was categorized into five mutually exclusive categories: extracorporeal membrane oxygenation (ECMO), ventilator, VAD, inotropes and oral therapy as reported previously.⁸ Cardiac diagnosis was categorized as congenital heart disease (CHD) or other. Creatinine clearance (CrCl) was estimated using the Schwartz Formula.⁹ The primary endpoint of the study was death or graft loss leading to re-transplantation.

A priori MATE definitions and scoring

SUPPLEMENTAL TABLE 2 summarizes the *a priori* definitions of MATE; their severity levels and scores (weights) based on clinical experience, which served as an initial framework for model development. The six MATEs were defined according to published guidelines from the International Society of Heart and Lung Transplantation (ISHLT)¹⁰⁻¹³, the National Kidney Foundation (NKF)^{14, 15}, and the World Health Organization (WHO)¹⁶. Levels of illness severity and/or disease progression were also defined using the same guidelines^{10, 12-15} and assigned an initial weight of 0 for no adverse event, 1 for a mild event, 2 for a moderate event, 3 for a severe event and 4 for an event resulting in death or graft loss. For infection and post-transplant lymphoproliferative disease (PTLD), we used risk factors for mortality identified in the ISHLT and WHO guidelines to generate low-, medium- and high-risk categories that were subsequently evaluated in the model development phase. Patients treated for rejection without a biopsy were categorized as severe rejection if they had evidence of hemodynamic compromise treated with inotropes (a PHTS definition analyzed previously), moderate rejection if they had mild hemodynamic compromise; all other patients with treated rejection were categorized as mild rejection.¹⁷ Adverse event definitions were then mapped to PHTS data fields. In most cases adverse event definitions could be mapped with reasonable precision because of the detailed information available in PHTS on adverse events. In cases where several possibilities existed, we analyzed variables based on the closest clinical definition and then according to what optimized model performance as outlined below.

MATE score development and evaluation

Two-thirds of study cohort was randomly assigned to the training set, with the remaining one-third assigned to the test set. Summary statistics are presented as median (interquartile range [IQR]) or number (percent). A wide variety of MATE score strategies were considered covering a range of (1) coding strategies for adverse events, (2) weighting strategies to score the severity of adverse events assuming a 1-unit worsening did not carry the same risk across all adverse events or severity levels, and (3) imputation strategies where data were missing. Because the data were longitudinal and event-driven, missing values for adverse events not reported when another event was reported (e.g. when clinical rejection was reported, a value of PTLD was not usually reported) were imputed using a multiple imputation method¹⁸.

Statistical Analysis

A longitudinal mixed effects model was used to predict the risk of graft loss based on the cumulative burden of adverse events between 6 and 36 months post-HT as determined by the MATE score using a forward stepwise selection technique. Model performance was evaluated based on the CPE-statistic, or concordance probability estimate-statistic, a C-statistic suitable for analyzing longitudinal data¹⁹. The fully conditional specification was used for multiple imputation of missing values assuming a non-decreasing adverse event severity function.^{18, 20, 21}

RESULTS

Patient Cohort

Of 2,671 patients who underwent HT during the study period, 2,118 met the study inclusion criteria, (494 died or experienced graft loss in the first 6 months following HT, 48 were excluded for multi-organ transplant and 11 were excluded for heterotopic transplants. Of these, 1412 subjects (two-thirds) were

randomly assigned to the training cohort and 706 to the test cohort. Overall, there were no significant differences in the baseline characteristics of the two cohorts, which are summarized in **TABLE 1**.

Overall graft survival and association with adverse events

Overall, the median follow-up time after HT was 3.7 years (range 0.5 to 10 years post-transplant), during which 230 patients experienced a graft loss event (11%). **FIGURE 1** depicts the Kaplan-Meier graft survival by individual adverse event using the highest illness severity value observed in the 6-36 month post-HT window. Overall, the association between the highest illness severity and graft loss was reasonably strong for cardiac allograft vasculopathy (CAV), acute cellular rejection (ACR), antibody mediated rejection (AMR), PTLD and infection but weaker for chronic kidney disease (CKD). When the window of time was restricted to the 30-month window between 6 and 36 months, 6.4% of subjects had a graft loss event.

Final MATE Score

The best fitting or final MATE score is summarized in **TABLE 2** after considering a wide range of possible variable coding, weighting and imputation strategies (**SUPPLEMENTAL TABLE 3**). Overall, in the test cohort, the concordance probability estimate (CPE) statistic for the final MATE score was 0.74 (**FIGURE 2**) and corresponded to the simple MATE sum of scores accumulated between 6 and 36 months with a modest modification to the *a priori* CKD categories. Whereas in the *a priori* model we divided CKD risk categories based on CKD stage (**SUPPLEMENTAL TABLE 2**), we found the best fitting model assigned CKD stage 1 and 2 to the lowest risk group, stage 3A to the mild risk group, 3B to moderate risk, and stages 4 and 5 to the highest risk group. Lastly, to preclude multiple non-fatal low-risk adverse events summing to a total MATE score greater than death or graft loss (score of 4 in the *a priori* model), we assigned a maximum score of 24 to death or graft loss – equivalent to a score of 4 (graft failure) for all 6 major adverse transplant events. **Figure 2** depicts the relative contributions of each adverse event to the CPE statistic, illustrating how even low

frequency events such as PTLD may still make an important contribution to the ability of the model to predict death/graft loss.

FIGURE 3 summarizes the observed distribution of imputed MATE scores using linear regression exclusive of graft loss between 6 and 36 months post-HT. The imputed scores (totaling 19,360 observations in 706 patients) appears normally distributed with a mean of 10 and standard deviation of 4 units. Using the adverse event classifications in the final MATE model, the frequency of adverse events between 6 and 36 months included CAV in 4% of visits, PTLD in 5%, CKD 9%, ACR 36%, AMR 4%, and infection 45%. The association between naïve MATE score categories and long-term graft survival is depicted in **FIGURE 4A**. The association between naïve MATE scores categories for 3 adverse events (CKD, CAV, and ACR) of special interest to the FDA for a trial involving a proliferation signal inhibitor is depicted in **FIGURE 4B**.

TABLE 3 illustrates the clinical application of the MATE score in a hypothetical patient using the non-decreasing criterion described above. At 6 months post-HT, a 10 year-old girl is free of adverse events. At 7 months post-HT she develops 2R ACR for which she is assigned a score of “2” for ACR, and “0” for the remaining adverse events because none is reported. At 14 months post-HT, she developed CMV disease treated with ganciclovir, for which she is assigned an infection score of “3”, retained a score of “2” for ACR and assigned “0” for the remaining adverse events. On her annual evaluation at 24 months post-HT, she was found to have moderate CAV and a GFR of 58 ml/min/1.73m² (down from 68 ml/min/m²), for which she is assigned a “2” for CAV and a “1” for CKD. At 30 months post HT, her CAV is found to be severe for which she is assigned a “3” for CAV and is listed for HT. At 34 months post-HT receives a new heart (graft loss) for which she is assigned a score of “24”, corresponding to a “4” across all six MATE categories. This example illustrates how recurring events and events of worsening severity are factored into the overall MATE score. Although we considered the effect of multiple adverse events (e.g. rejection) in the model development

phase, the model performed best when the highest severity level alone was included in the model.

Sample Size Considerations

To illustrate the utility of the MATE score to reduce sample size, we estimated the sample sizes necessary for a randomized clinical trial of two immunosuppressant regimens between 6 and 36 months post-HT. Assuming an accrual period of 24 months, a minimum follow-up time of 30 months, a normal distribution of MATE scores with standard deviation of 4 units (**FIGURE 3**) and a within-subject correlation ranging from 0.2 to 0.5, a cohort of 200 subjects (100 in each arm) would have >80% power to detect non-inferiority assuming a non-inferiority margin of 2-MATE units using a one-sided two-sample t-test at the 0.025 level of significance. With this sample size of 200 subjects, there is only 12% power to detect a relatively large 45% difference in graft survival assuming a baseline graft loss rate of 6% between 6 and 36 months.

DISCUSSION

In this study, we developed and evaluated a MATE risk-prediction score that predicts death or graft loss in pediatric heart transplant patients. The risk-prediction score is comprised of 6 MATEs: CKD, CAV, ACR, AMR, infection and PTLD and reflects the cumulative burden of major adverse transplant events. The MATE score performs well in predicting graft loss in children followed from 6 months to 36 months post-HT and therefore may be useful as a surrogate endpoint for clinical trials in pediatric heart transplantation.

Our findings are consistent with previous studies that have shown CAV^{22, 23}, rejection^{24, 25}, PTLD^{26, 27} and infection^{28, 29} to be associated with graft loss and death. To date, however, no studies have incorporated these traditional transplant complications into a risk-prediction tool that has been evaluated in its ability to predict graft loss. Previous heart transplant risk prediction studies have been validated; however, they have been limited to covariates present at the time of heart transplantation^{8, 30, 31}. Thus these models have

excluded transplant complications such as rejection and infection that occur after transplant and may be time-varying. To our knowledge, this is the first report to develop and evaluate a risk-prediction model involving multiple time-varying adverse events post-HT and demonstrate that the model can predict graft loss based on the cumulative progression of events through 36 months post-HT with strong concordance.

We chose a 30-month time window for the MATE analysis beginning at 6 months post-HT to simulate the projected timeframe of a randomized clinical trial involving conversion of maintenance immunosuppression at 6 months. Because the density of adverse events after 6 months post-HT is considerably lower than the first 6 months, we felt that a timeframe shorter than 30 months may be insufficient time for a critical number of adverse events to occur. On the other hand, a longer follow-up timeframe would be less feasible to conduct because of increases in trial costs and greater patient attrition.

We were somewhat surprised to discover that a naïve MATE sum performed better than a variety of more complex weighting strategies. We did not want to make the assumption that the risk attributable to a one-unit increase was uniform within or across major adverse events. Ultimately, however, it turned out that the simplest weight strategy (e.g. 0,1,2,3 corresponding to none, mild, moderate, and severe, respectively) performed the best. We also found it interesting that no one adverse event made an outsized contribution—as reflected in the CPE statistic corresponding to each MATE—suggesting that including each of the major adverse events or their synergy was important to the performance of the overall score in predicting graft failure.

This study has several important implications. First, the MATE score may be a useful surrogate endpoint for designing an immunosuppression clinical trial in pediatric cardiac transplantation. In contrast to the fields of adult heart transplant³², pediatric kidney^{33, 34} and liver³⁵ transplantation, there has yet to be a randomized

trial of immunosuppression therapy in pediatric heart transplant patients. Most believe the greatest challenge to designing a clinical trial is the lack of statistical power stemming from the rareness of the disease and the low endpoint frequency (e.g. death and/or graft loss). This fundamental problem (type II statistical error) has beleaguered a number of well-organized but negative pediatric cardiovascular trials³⁶⁻³⁸ where, in retrospect, the trial's primary endpoint has drawn scrutiny³⁹⁻⁴¹. The MATE score seeks to fill this gap by creating a variable that has leverage to power a trial but also is associated with outcome. The association between adverse events and graft survival appears to hold true for the overall MATE score comprised of all 6 adverse events as well as subset of 3 adverse events of interest to the FDA (CKD, CAV and ACR) in evaluating the risk-benefit profile of a proliferation signal inhibitor for pediatric heart transplantation.

The use of surrogate endpoints to design clinical trials has gained widespread support in recent years from academics, industry and the US Food and Drug Administration⁴¹⁻⁴³. In recent years, the FDA has hosted a series of workshops on developing valid surrogate endpoints for clinical trials in solid-organ transplantation^{42, 44, 45}. Workshops have emphasized that not all surrogate endpoints are equally valid. To serve as a valid surrogate endpoint for a regulatory trial, the FDA requires the surrogate endpoint satisfy the Prentice Criteria^{6, 42, 44, 46}. These include: (1) that the surrogate endpoint is correlated with a clinical outcome, (2) the surrogate endpoint captures the net effect of treatment on the clinical outcome, (3) the pattern characterizing the effects of the surrogate endpoint on the clinical outcome must be determined, and (4) the degree of deviation in that pattern is known. Here, we found (1) the MATE score is correlated with graft loss, (2) the MATE score captures a broad range of ways in which immunosuppressant may be linked to graft loss (rejection, infection, CKD, CAV, AMR and PTLD), (3) the pattern characterizing the effect of the MATE score on graft loss is known (**FIGURES 2 and 3**), and (4) the degree in deviation of the pattern is known (**FIGURES 3 and 4**). These characteristics suggest the MATE score may satisfy the Prentice Criteria.

Prospective data collection in the form of a clinical trial would provide the best opportunity to confirm that the MATE score satisfies the Prentice Criteria.

This study has several limitations. First, we identified patients retrospectively through a national registry of transplant recipients, which creates the opportunity for selection bias. However, PHTS currently captures nearly 80% of transplant centers in the US, and the characteristics of patients reported to PHTS and not reported to PHTS appear relatively similar based on internal comparisons with UNOS data, suggesting that selection bias did not play a significant role in the study's findings. Second, PHTS data fields could not be mapped exactly to consensus definitions of adverse events in all cases, creating the potential for misclassification bias. However, in the large majority of cases, the definitions could be mapped precisely to PHTS data fields because of the PHTS's level of coding detail. Where there was uncertainty, we created several coding options that were tested in the score development stage, and performed well in the testing cohort regardless of the coding strategy used suggesting the overall effect on the scoring system was small. Misclassification bias of CKD stage could be caused by choice of formula to estimate the GFR. We sought to minimize this effect by using the most updated version, the modified Schwartz formula, which is also the preferred formula for children with existing renal dysfunction. Lastly, because PHTS is an event-driven database, missing data could threaten the validity of the findings. To address this, we evaluated a wide variety of imputation techniques in the training cohort as well as the sensitivity of the findings to the imputation technique and found that the results were robust. While this limitation remains relevant to a retrospective analysis of historical registry data, it is worth noting it would be less relevant in a prospective clinical trial where data on all adverse events can be incorporated into the data collection to minimize missing data.

CONCLUSION

In summary, we have developed and evaluated a surrogate endpoint for graft loss in pediatric heart transplant recipients that is comprised of 6 major adverse transplant events (MATEs). The surrogate endpoint, known as the MATE score, predicts graft loss well (C-statistic 0.74) and may be useful as a primary endpoint for FDA clinical trials in pediatric heart transplantation given it appears to satisfy the Prentice Criteria for surrogate endpoints. This is important because the lack of a surrogate endpoint has been one of the greatest challenges to designing clinical trials which evaluate promising new therapies in pediatric heart transplantation, and perhaps throughout solid organ transplantation. While survival after pediatric heart transplant has improved considerably in the current era, many children still fail to survive into adulthood because of a predictable set of transplant complications. Novel clinical trial designs and endpoints affords the pediatric community the best prospect for evaluating promising medical therapies involving rare diseases.

Acknowledgements

Dr. Desai and Ms. Hoen and Yang had full access to the data in the study and take responsibility for the integrity of the data and the accuracy of analysis. The data were supplied by the PHTS to Stanford University under a limited data use agreement. The interpretation and reporting of these data are the responsibility of the authors and not an official policy of or interpretation by the PHTS. We would like to thank Ms. FeiFei Qin of the Stanford University Quantitative Sciences Unit for validating statistical coding; and Dr. Charlie Canter at Washington University in St. Louis, and Ms. Miranda Zinsman of Cardiovascular Surgery for their thoughtful review of the manuscript. Most importantly, we would like to thank the centers participating in the PHTS, especially the heart transplant clinical and data coordinators who assume primary responsibility for submitting PHTS data.

Funding Sources

This study was funded in part by generous support of the Kate Marra family to the Pediatric Advanced Cardiac Therapies (PACT) Team Research Fund at Stanford University.

Disclosures

None

REFERENCES

1. Gomez G, Gomez-Mateu M and Dafni U. Informed choice of composite end points in cardiovascular trials. *Circ Cardiovasc Qual Outcomes*. 2014;7:170-8.
2. Sankoh AJ, Li H and D'Agostino RB, Sr. Use of composite endpoints in clinical trials. *Stat Med*. 2014;33:4709-14.
3. Bath PM, Geeganage C, Gray LJ, Collier T and Pocock S. Use of ordinal outcomes in vascular prevention trials: comparison with binary outcomes in published trials. *Stroke*. 2008;39:2817-23.
4. Roozenbeek B, Lingsma HF, Perel P, Edwards P, Roberts I, Murray GD, Maas AI, Steyerberg EW, Group IS and Collaborators CT. The added value of ordinal analysis in clinical trials: an example in traumatic brain injury. *Crit Care*. 2011;15:R127.
5. Whitehead J. Sample size calculations for ordered categorical data. *Stat Med*. 1993;12:2257-71.
6. Prentice RL. Surrogate endpoints in clinical trials: definition and operational criteria. *Stat Med*. 1989;8:431-40.
7. Shaddy RE, Naftel DC, Kirklin JK, Boyle G, McGiffin DC, Towbin JA, Ring WS, Pearce B, Addonizio L and Morrow WR. Outcome of cardiac transplantation in children. Survival in a contemporary multi-institutional experience. Pediatric Heart Transplant Study. *Circulation*. 1996;94:II69-73.
8. Almond CS, Gauvreau K, Canter CE, Rajagopal SK, Piercey GE and Singh TP. A risk-prediction model for in-hospital mortality after heart transplantation in US children. *Am J Transplant*. 2012;12:1240-8.
9. Schwartz GJ, Munoz A, Schneider MF, Mak RH, Kaskel F, Warady BA and Furth SL. New equations to estimate GFR in children with CKD. *J Am Soc Nephrol*. 2009;20:629-37.
10. Berry GJ, Burke MM, Andersen C, Bruneval P, Fedrigo M, Fishbein MC, Goddard M, Hammond EH, Leone O, Marboe C, Miller D, Neil D, Rassl D, Revelo MP, Rice A, Rene Rodriguez E, Stewart S, Tan CD, Winters GL, West L, Mehra MR and Angelini A. The 2013 International Society for Heart and Lung Transplantation Working Formulation for the standardization of nomenclature in the pathologic diagnosis of antibody-mediated rejection in heart transplantation. *J Heart Lung Transplant*. 2013;32:1147-62.
11. Husain S, Mooney ML, Danziger-Isakov L, Mattner F, Singh N, Avery R, Ison M, Humar A, Padera RF, Lawler LP, Fisher A, Drew RJ, Gould KF, Sole A, Studer S, Munoz P, Singer LG, Hannan M and Definitions IIDCWGo. A 2010 working formulation for the standardization of definitions of infections in cardiothoracic transplant recipients. *J Heart Lung Transplant*. 2011;30:361-74.
12. Mehra MR, Crespo-Leiro MG, Dipchand A, Ensminger SM, Hiemann NE, Kobashigawa JA, Madsen J, Parameshwar J, Starling RC and Uber PA. International Society for Heart and Lung Transplantation working formulation of a standardized nomenclature for cardiac allograft vasculopathy-2010. *J Heart Lung Transplant*. 2010;29:717-27.
13. Stewart S, Winters GL, Fishbein MC, Tazelaar HD, Kobashigawa J, Abrams J, Andersen CB, Angelini A, Berry GJ, Burke MM, Demetris AJ, Hammond E, Itescu S, Marboe CC, McManus B, Reed EF, Reinsmoen NL,

- Rodriguez ER, Rose AG, Rose M, Suciu-Focia N, Zeevi A and Billingham ME. Revision of the 1990 working formulation for the standardization of nomenclature in the diagnosis of heart rejection. *J Heart Lung Transplant*. 2005;24:1710-20.
14. Hogg RJ, Furth S, Lemley KV, Portman R, Schwartz GJ, Coresh J, Balk E, Lau J, Levin A, Kausz AT, Eknoyan G, Levey AS and National Kidney Foundation's Kidney Disease Outcomes Quality I. National Kidney Foundation's Kidney Disease Outcomes Quality Initiative clinical practice guidelines for chronic kidney disease in children and adolescents: evaluation, classification, and stratification. *Pediatrics*. 2003;111:1416-21.
 15. Levey AS, Eckardt KU, Tsukamoto Y, Levin A, Coresh J, Rossert J, De Zeeuw D, Hostetter TH, Lameire N and Eknoyan G. Definition and classification of chronic kidney disease: a position statement from Kidney Disease: Improving Global Outcomes (KDIGO). *Kidney Int*. 2005;67:2089-100.
 16. Swerdlow SH WS, Chadburn A, et al . WHO Classification of Tumours of Haematopoietic and Lymphoid Tissues. Lyon: IARC Press; 2008. Post-transplant lymphoproliferative disorders.
 17. Everitt MD, Pahl E, Schechtman KB, Zheng J, Ringewald JM, L'Ecuyer T, Naftel DC, Kirklin JK, Blume ED, Bullock EA, Canter CE and Pediatric Heart Transplant Study I. Rejection with hemodynamic compromise in the current era of pediatric heart transplantation: a multi-institutional study. *J Heart Lung Transplant*. 2011;30:282-8.
 18. Little RJ, and Donald B. Rubin. Statistical analysis with missing data. In: John Wiley & Sons, ed.; 2014.
 19. Gonen M and Heller G. Concordance probability and discriminatory power in proportional hazards regression. . *Biometrika*. 2005;92:965-970.
 20. Allison PD. Missing data. In: S. Publications, ed.; 2001(136).
 21. van Buuren S. Multiple imputation of discrete and continuous data by fully conditional specification. *Stat Methods Med Res*. 2007;16:219-42.
 22. Kindel SJ, Law YM, Chin C, Burch M, Kirklin JK, Naftel DC, Pruitt E, Carboni MP, Arens A, Atz AM, Dreyer WJ, Mahle WT and Pahl E. Improved Detection of Cardiac Allograft Vasculopathy: A Multi-Institutional Analysis of Functional Parameters in Pediatric Heart Transplant Recipients. *J Am Coll Cardiol*. 2015;66:547-57.
 23. Pahl E, Naftel DC, Kuhn MA, Shaddy RE, Morrow WR, Canter CE, Kirklin J and Pediatric Heart Transplant S. The impact and outcome of transplant coronary artery disease in a pediatric population: a 9-year multi-institutional study. *J Heart Lung Transplant*. 2005;24:645-51.
 24. Chin C, Naftel DC, Singh TP, Blume ED, Luikart H, Bernstein D, Gamberg P, Kirklin JK, Morrow WR and Pediatric Heart Transplant Study G. Risk factors for recurrent rejection in pediatric heart transplantation: a multicenter experience. *J Heart Lung Transplant*. 2004;23:178-85.
 25. Dipchand AI, Kirk R, Edwards LB, Kucheryavaya AY, Benden C, Christie JD, Dobbels F, Lund LH, Rahmel AO, Yusef RD, Stehlik J, International Society for H and Lung T. The Registry of the International Society for Heart and Lung Transplantation: Sixteenth Official Pediatric Heart Transplantation Report--2013; focus theme: age. *J Heart Lung Transplant*. 2013;32:979-88.
 26. Webber SA, Naftel DC, Fricker FJ, Olesnevich P, Blume ED, Addonizio L, Kirklin JK, Canter CE and Pediatric Heart Transplant S. Lymphoproliferative disorders after paediatric heart transplantation: a multi-institutional study. *Lancet*. 2006;367:233-9.
 27. Manliot C, Pollock-Barziv SM, Holmes C, Weitzman S, Allen U, Clarizia NA, Ngan BY, McCrindle BW and Dipchand AI. Post-transplant lymphoproliferative disorder in pediatric heart transplant recipients. *J Heart Lung Transplant*. 2010;29:648-57.
 28. Canter C, Naftel D, Caldwell R, Chinnock R, Pahl E, Frazier E, Kirklin J, Boucek M and Morrow R. Survival and risk factors for death after cardiac transplantation in infants. A multi-institutional study. The Pediatric Heart Transplant Study. *Circulation*. 1997;96:227-31.

29. Schowengerdt KO, Naftel DC, Seib PM, Pearce FB, Addonizio LJ, Kirklin JK and Morrow WR. Infection after pediatric heart transplantation: results of a multiinstitutional study. The Pediatric Heart Transplant Study Group. *J Heart Lung Transplant*. 1997;16:1207-16.
30. Davies RR, Russo MJ, Mital S, Martens TM, Sorabella RS, Hong KN, Gelijns AC, Moskowitz AJ, Quaegebeur JM, Mosca RS and Chen JM. Predicting survival among high-risk pediatric cardiac transplant recipients: an analysis of the United Network for Organ Sharing database. *J Thorac Cardiovasc Surg*. 2008;135:147-55, 155 e1-2.
31. Schumacher KR, Almond C, Singh TP, Kirk R, Spicer R, Hoffman TM, Hsu D, Naftel DC, Pruitt E, Zamberlan M, Canter CE, Gajarski RJ and Investigators PSG. Predicting graft loss by 1 year in pediatric heart transplantation candidates: an analysis of the Pediatric Heart Transplant Study database. *Circulation*. 2015;131:890-8.
32. Eisen HJ, Tuzcu EM, Dorent R, Kobashigawa J, Mancini D, Valantine-von Kaeppler HA, Starling RC, Sorensen K, Hummel M, Lind JM, Abeywickrama KH, Bernhardt P and Group RBS. Everolimus for the prevention of allograft rejection and vasculopathy in cardiac-transplant recipients. *N Engl J Med*. 2003;349:847-58.
33. Benfield MR, Tejani A, Harmon WE, McDonald R, Stablein DM, McIntosh M, Rose S and Group CS. A randomized multicenter trial of OKT3 mAbs induction compared with intravenous cyclosporine in pediatric renal transplantation. *Pediatr Transplant*. 2005;9:282-92.
34. Sarwal MM, Ettenger RB, Dharnidharka V, Benfield M, Mathias R, Portale A, McDonald R, Harmon W, Kershaw D, Vehaskari VM, Kamil E, Baluarte HJ, Warady B, Tang L, Liu J, Li L, Naesens M, Sigdel T, Waskerwitz J and Salvatierra O. Complete steroid avoidance is effective and safe in children with renal transplants: a multicenter randomized trial with three-year follow-up. *Am J Transplant*. 2012;12:2719-29.
35. Alvarez F, Atkison PR, Grant DR, Guilbault N, Jones AB, Kim PS, Kneteman NM, Laurin L, Martin SR, Murphy GF, Paradis K, Shapiro J, Smith LJ and Superina RA. NOF-11: a one-year pediatric randomized double-blind comparison of neoral versus sandimmune in orthotopic liver transplantation. *Transplantation*. 2000;69:87-92.
36. Hsu DT, Zak V, Mahony L, Sleeper LA, Atz AM, Levine JC, Barker PC, Ravishankar C, McCrindle BW, Williams RV, Altmann K, Ghanayem NS, Margossian R, Chung WK, Border WL, Pearson GD, Stylianou MP, Mital S and Pediatric Heart Network I. Enalapril in infants with single ventricle: results of a multicenter randomized trial. *Circulation*. 2010;122:333-40.
37. Monagle P, Cochrane A, Roberts R, Manlihot C, Weintraub R, Szechtman B, Hughes M, Andrew M, McCrindle BW and Fontan Anticoagulation Study G. A multicenter, randomized trial comparing heparin/warfarin and acetylsalicylic acid as primary thromboprophylaxis for 2 years after the Fontan procedure in children. *J Am Coll Cardiol*. 2011;58:645-51.
38. Shaddy RE, Boucek MM, Hsu DT, Boucek RJ, Canter CE, Mahony L, Ross RD, Pahl E, Blume ED, Dodd DA, Rosenthal DN, Burr J, LaSalle B, Holubkov R, Lukas MA, Tani LY and Pediatric Carvedilol Study G. Carvedilol for children and adolescents with heart failure: a randomized controlled trial. *JAMA*. 2007;298:1171-9.
39. Canter CE. Preventing thrombosis after the Fontan procedure not there yet. *J Am Coll Cardiol*. 2011;58:652-3.
40. Gidding SS. The importance of randomized controlled trials in pediatric cardiology. *JAMA*. 2007;298:1214-6.
41. Singh TP. Letter by Singh regarding article, "Enalapril in infants with single ventricle: results of a multicenter randomized trial". *Circulation*. 2011;123:e374.
42. Food and Drug Administration. Biomarkers and Surrogate Endpoints. July 2015. Accessed November 28, 2015 at <http://www.fda.gov/downloads/NewsEvents/Newsroom/FactSheets/UCM455106.pdf>.

43. FDA Guidance for Industry: Providing Clinical Evidence of Effectiveness for Human Drug and Biological Products. U.S. Department of Health and Human Services Food and Drug Administration Center for Drug Evaluation and Research (CDER) Center for Biologics Evaluation and Research (CBER). <http://www.fda.gov/downloads/Drugs/.../Guidances/ucm078749.pdf>. Accessed September 15, 2015. 1998.
 44. US Food and Drug Administration Workshop: Surrogate Endpoints for Clinical Trials in Kidney Transplantation; Public Workshop. September 28, 2015, Rockville, MD. Accessed November 28, 2015 at <http://www.fda.gov/Drugs/NewsEvents/ucm449248.htm>.
 45. Archdeacon P, Chan M, Neuland C, Velidedeoglu E, Meyer J, Tracy L, Cavaille-Coll M, Bala S, Hernandez A and Albrecht R. Summary of FDA antibody-mediated rejection workshop. *Am J Transplant*. 2011;11:896-906.
 46. Guidance F. FDA Guidance for Industry: Clinical Trial Endpoints for the Approval of Cancer Drugs and Biologics. U.S. Department of Health and Human Services Food and Drug Administration Center for Drug Evaluation and Research (CDER) Center for Biologics Evaluation and Research (CBER). 2007.
- Figure Legends**

FIGURE 1.

Kaplan-Meier freedom from death or graft loss by major adverse transplant event. Categories for individual event types are defined in **TABLE 2**. **A.** Chronic kidney disease (CKD) Major Adverse Transplant Event, by CKD Stage ($P=0.04$). **B.** Coronary artery vasculopathy Major Adverse Transplant Event, by ISHLT severity grade ($P<0.01$). **C.** Acute cellular rejection Major Adverse Transplant Event, by ISHLT severity grade ($P<0.01$). **D.** Antibody-mediated rejection Major Adverse Transplant Event, by ISHLT severity grade ($P<0.01$). **E.** Infection Major Adverse Transplant Event ($P=0.01$). **F.** Post-transplant lymphoproliferative disorder Major Adverse Transplant Event ($P<0.01$).

FIGURE 2.

CPE statistics for the best-fitting model.

FIGURE 3.

Observed distribution of naïve sum of imputed MATE scores using linear regression in the test cohort.

FIGURE 4.

A. Association between naïve MATE score categories and graft survival between 6 and 36 months post-heart transplant for all major adverse transplant events. **B.** Association between naïve MATE score categories and graft survival between 6 and 36 months post-heart transplant for three adverse events: renal dysfunction, coronary artery vasculopathy and cellular rejection.

TABLE 1. Baseline characteristics of study cohort conditional upon survival to 6 months post HT.

Characteristic	All patients (N=2,118)	Model Development Cohort (N=1,412)	Model Validation Cohort (N=706)
Age at transplant (y)	4.1 (0.7, 2.7)	4.1 (0.7, 12.9)	4.1 (0.7, 12.4)
<1 y	660 (31%)	446 (32%)	214 (30%)
1-5 y	451 (21%)	290 (21%)	161 (23%)
6-12 y	416 (20%)	279 (20%)	137 (19%)
>12 y	591 (28%)	397 (28%)	194 (27%)
Female	987 (47%)	649 (46%)	338 (48%)
Weight at transplant (kg)	15.0 (6.9, 41.4)	15.0 (6.9, 41.0)	15.0 (7.0, 42.2)
<10 kg	784 (37%)	530 (38%)	254 (36%)
10-25 kg	550 (26%)	357 (25%)	193 (27%)
26-50 kg	419 (20%)	278 (20%)	141 (20%)
>50 kg	365 (17%)	247 (17%)	118 (17%)
Status 1 at transplant	1963 (94%)	1299 (93%)	664 (95%)
Blood type O	959 (45%)	628 (45%)	331 (47%)
African American	454 (21%)	311 (22%)	143 (20%)
Congenital Heart Disease	925 (44%)	618 (44%)	307 (43%)
Support at transplant ¹			
Medical	1349 (64%)	895 (63%)	454 (64%)
VAD	375 (18%)	246 (17%)	129 (19%)
Ventilator	376 (18%)	255 (18%)	121 (17%)
ECMO	63 (3%)	38 (3%)	25 (4%)
Total bilirubin (mg/dL)	0.7 (0.4, 1.4)	0.70 (0.4, 1.4)	0.70 (0.4, 1.4)
Serum creatinine (mg/dL)	0.4 (0.3, 0.7)	0.5 (0.3, 0.7)	0.4 (0.3, 0.6)
GFR (ml/min/1.73m ²) ²	90.6 (68.8, 117.2)	90.0 (67.7, 115.6)	92.2 (71.6, 119.8)
Chronic Kidney Disease Stage 3-5	24 (1%)	18 (1%)	6 (1%)
Dialysis	24 (2%)	18 (2%)	6 (2%)
Location at transplant			
Not hospitalized	505 (24%)	332 (24%)	173 (25%)
ICU	972 (46%)	636 (45%)	336 (48%)
Maximum PRA			
PRA ≥10%	434 (23%)	291 (24%)	143 (23%)
PRA ≥25%	291 (16%)	192 (16%)	99 (16%)
Ischemic time (hours)	3.6 (3.0, 4.3)	3.6 (3.0, 4.3)	3.6 (2.9, 4.2)
Public health insurance	1114 (53%)	734 (52%)	380 (54%)
Listing year			
2005-2007	478 (23%)	313 (23%)	165 (24%)
2008-2010	691 (33%)	460 (33%)	231 (33%)
2011-2014	905 (44%)	609 (44%)	296 (43%)
Primary immunosuppression at hospital discharge ³			
Tacrolimus/MMF	389 (18%)	246 (17%)	143 (20%)

TAC/MMF/steroids	1135 (54%)	759 (54%)	376 (53%)
Sirolimus	100 (5%)	68 (5%)	32 (5%)
Positive crossmatch (any)	201 (10%)	144 (10%)	57 (8%)

Data are presented as number (percent) or median (interquartile range). VAD, ventricular assist device; ECMO, extracorporeal membrane oxygenation; ICU, intensive care unit; PRA, panel reactive antibody; GFR, glomerular filtration rate; TAC, tacrolimus; MMF, mycophenolate mofetil; CDC, cytotoxic-dependent crossmatch,

¹Mutually exclusive, hierarchical categories, where each patient is assigned to one category based on their highest level of support (ECMO>Ventilator>VAD>inotropes>medical).

²GFR is estimate using the modified Schwartz formula ($GFR \text{ in ml/min/1.73m}^2 = (ht \text{ in cm}) \times (0.413)/\text{serum creatinine in mg/dl, at transplant}$).

³Immunosuppression regimen at the time of hospital discharge from the hospitalization where heart transplantation was performed.

There were no significant differences in the baseline characteristics between the two cohorts.

TABLE 2. Final Major Adverse Transplant Events Score

	0 (No Disease)	1 (Mild Disease)	2 (Moderate Disease)	3 (Severe or High-Risk Disease)	24 PATIENT DEATH GRAFT LOSS
Chronic Kidney Disease (CKD)^{14, 15}	Stage 1-2 >60 ml/min/1.73 m ²	Stage 3A 45-59 ml/min/1.73 m ²	Stage 3B B: 30-44 ml/min/1.73 m ²	Stage 4-5 <30 ml/min/1.73 m ² , dialysis or renal transplant	Death/ Graft Loss
Coronary Allograft Vasculopathy (CAV)¹²	ISHLT CAV₀/None	ISHLT CAV₁	ISHLT CAV₂	ISHLT CAV₃	Death/ Graft Loss
Acute Cellular Rejection (ACR)¹³	Grade 0/None	ISHLT Grade 1R	ISHLT Grade 2R	ISHLT Grade 3R	Death/ Graft Loss
Antibody-mediated Rejection (AMR)¹⁰	pAMR grade 0/None	ISHLT pAMR1	ISHLT pAMR2	ISHLT pAMR3	Death/ Graft Loss
Infection¹¹	None	Low-risk All recorded infections treated with PO or IM antibiotics and not meeting “moderate- risk” or “high- risk” criteria	Moderate-Risk All infections treated with a full course of IV antibiotics but not meeting “high-risk” criteria	High-Risk Life-threatening infection treated with IV anti- microbials with evidence of one of the following: (1) fungal disease, (2) bacterial infection with enterococcus, pseudomonas or staph aureus, or (3) CMV or EBV infection, (4) blood stream infection or endocarditis, or (5) hemodynamic compromise requiring ECMO support.	Death/ Graft Loss
Post-transplant Lymphoproliferative Disease (PTLD)¹⁶	None	Low-risk Polymorphic PLTD without “high-risk” or intermediate risk features present	Moderate risk Non-polymorphic PTLD without “high-risk” features present	High-risk CNS involvement Bone Marrow Involvement Or ≥3 sites involved regardless of cell type	Death/ Graft Loss

TABLE 3. Illustration of the MATE score in a hypothetical child who develops acute rejection and CMV infection complicated by coronary artery disease and ultimately graft loss. This illustration demonstrates a non-decreasing MATE score value where the risk of earlier adverse events is carried forward contributing to the overall risk.

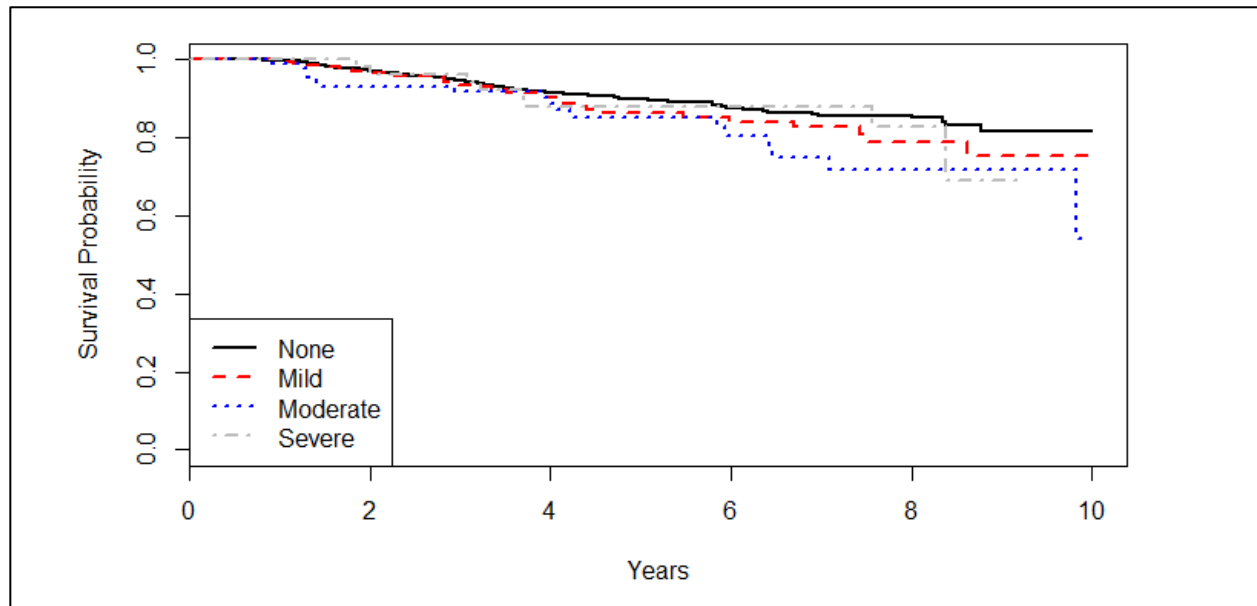
Subject ID	Time Since Transplant	Chronic Kidney Disease	Coronary Artery Vasculopathy	Acute Cellular Rejection	Antibody-Mediated Rejection	Infection	PTLD	Graft Loss*	MATE Sum with Single Imputation
1	7 mos.	0	0	2	0	0	0	0	2
1	14 mos.	0	0	2	0	3	0	0	5
1	24 mos.	1	2	2	0	3	0	0	8
1	30 mos.	1	3	2	0	3	0	0	9
1	34 mos.	4	4	4	4	4	4	1	24

Abbreviations: ID, Identification; Mos., Months; PTLD, post-transplant lymphoproliferative disease.

*To avoid the scenario where multiple minor adverse events could sum to a score in a living patient that is higher than death/graft failure (4), by convention patients with graft loss are assigned a score of 24, equivalent to a graft loss score of 4 across all 6 domains ($4 \times 6 = 24$).

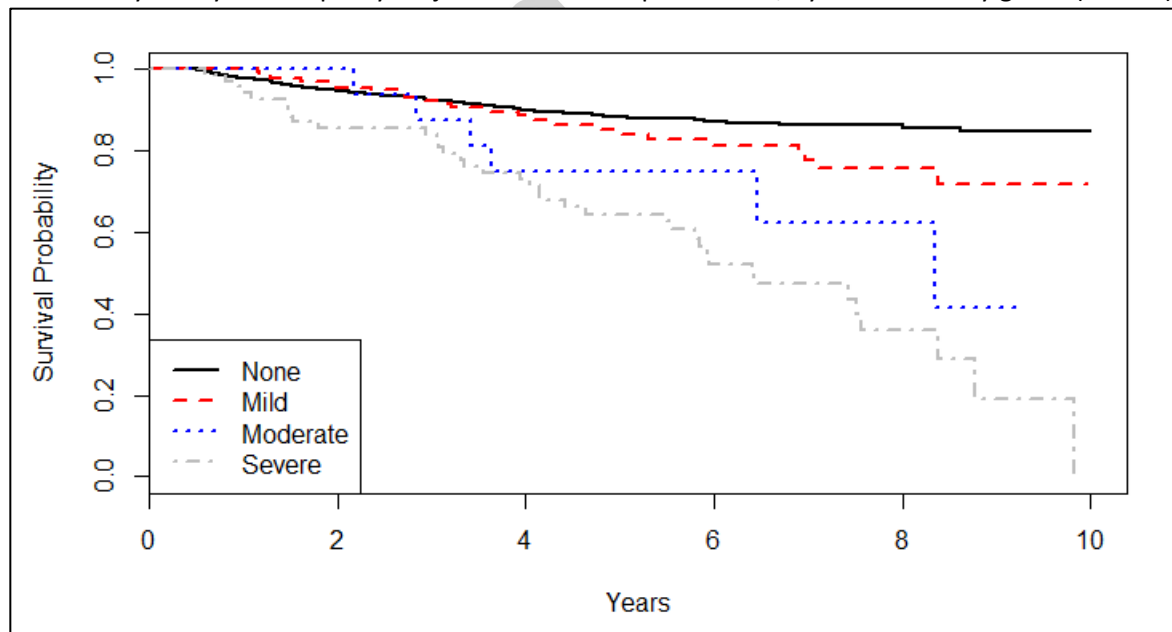
FIGURE 1. Kaplan-Meier freedom from graft loss by major adverse transplant event.

A. Chronic kidney disease (CKD) Major Adverse Transplant Event, by CKD Stage (P=0.04).



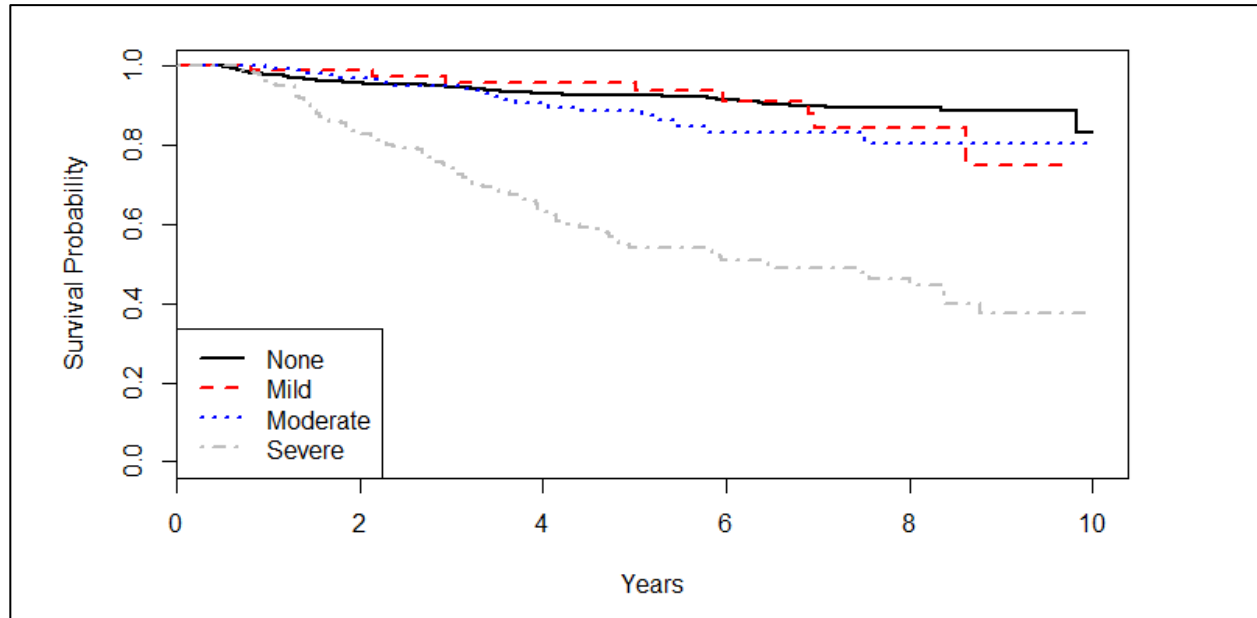
Mild, moderate and severe CKD are defined in **TABLE 2**.

B. Coronary artery vasculopathy Major Adverse Transplant Event, by ISHLT severity grade (P<0.01).



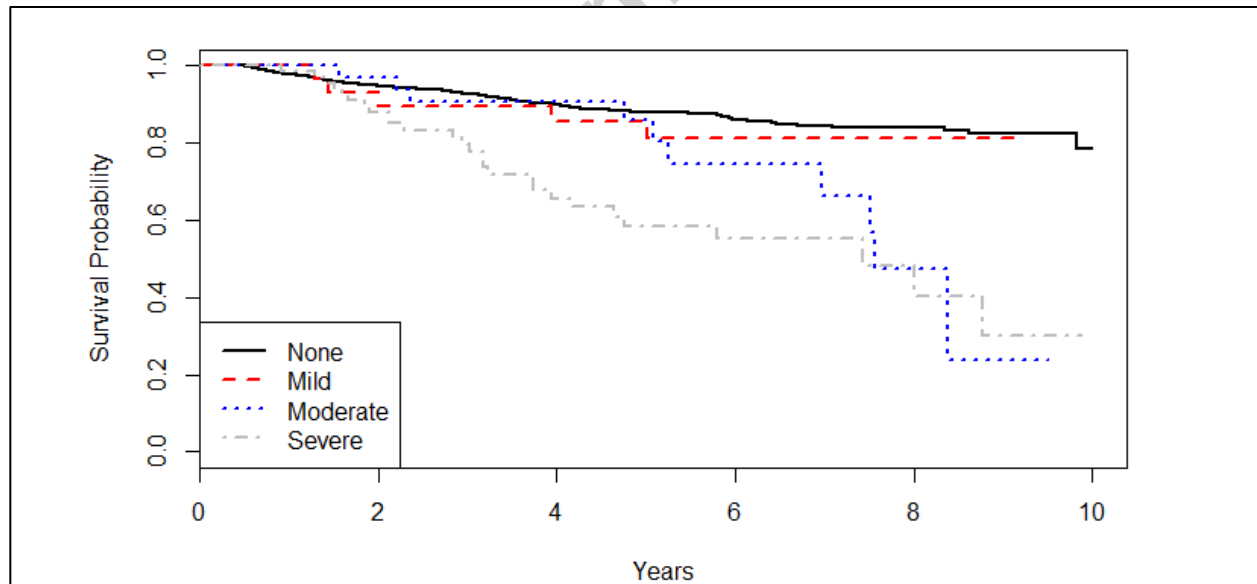
Mild, moderate and severe CAV are defined in **TABLE 2**.

C. Acute cellular rejection Major Adverse Transplant Event, by ISHLT severity grade ($P<0.01$).



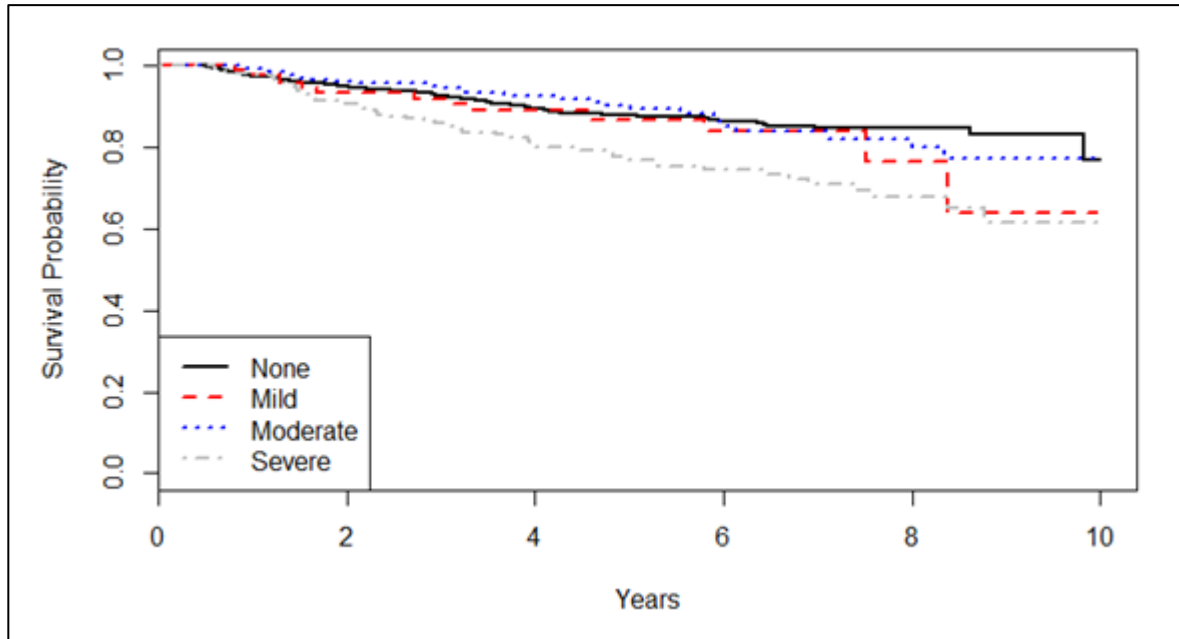
Mild, moderate and severe ACR are defined in **TABLE 2**.

D. Antibody-mediated rejection Major Adverse Transplant Event, by ISHLT severity grade ($P<0.01$).



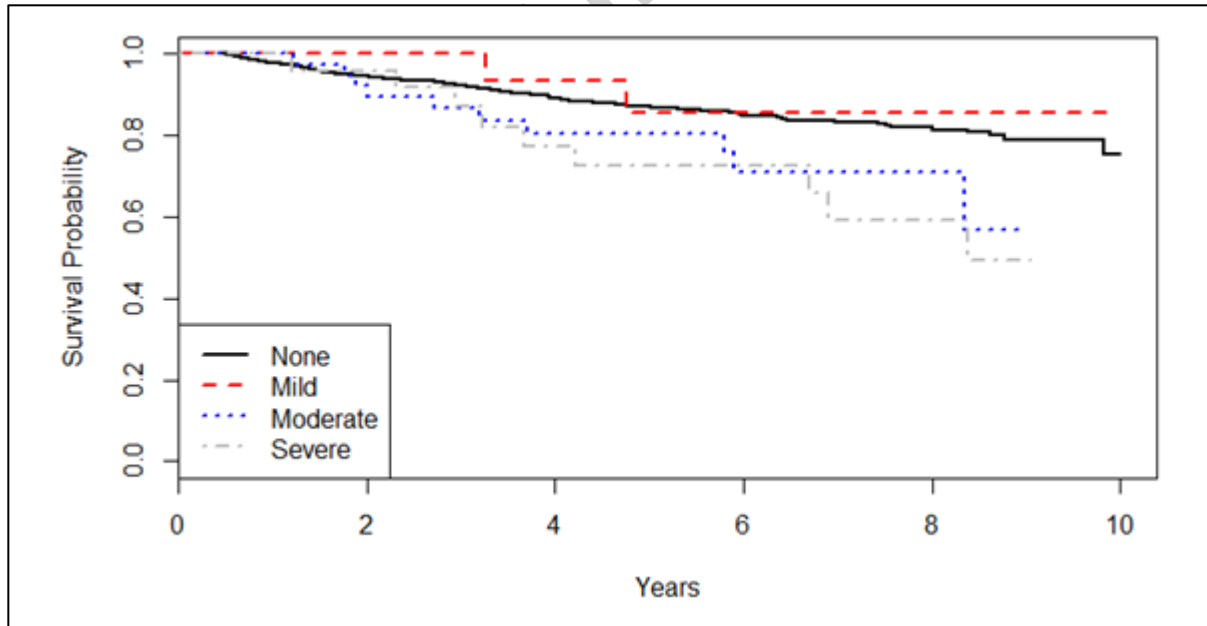
Mild, moderate and severe AMR are defined in **TABLE 2**.

E. Infection Major Adverse Transplant Event ($P=0.01$).



Mild, moderate and severe Infection are defined in **TABLE 2**.

F. Post-transplant lymphoproliferative disorder Major Adverse Transplant Event ($P<0.01$).



Mild, moderate and severe PTLD are defined in **TABLE 2**.

FIGURE 2. CPE statistics

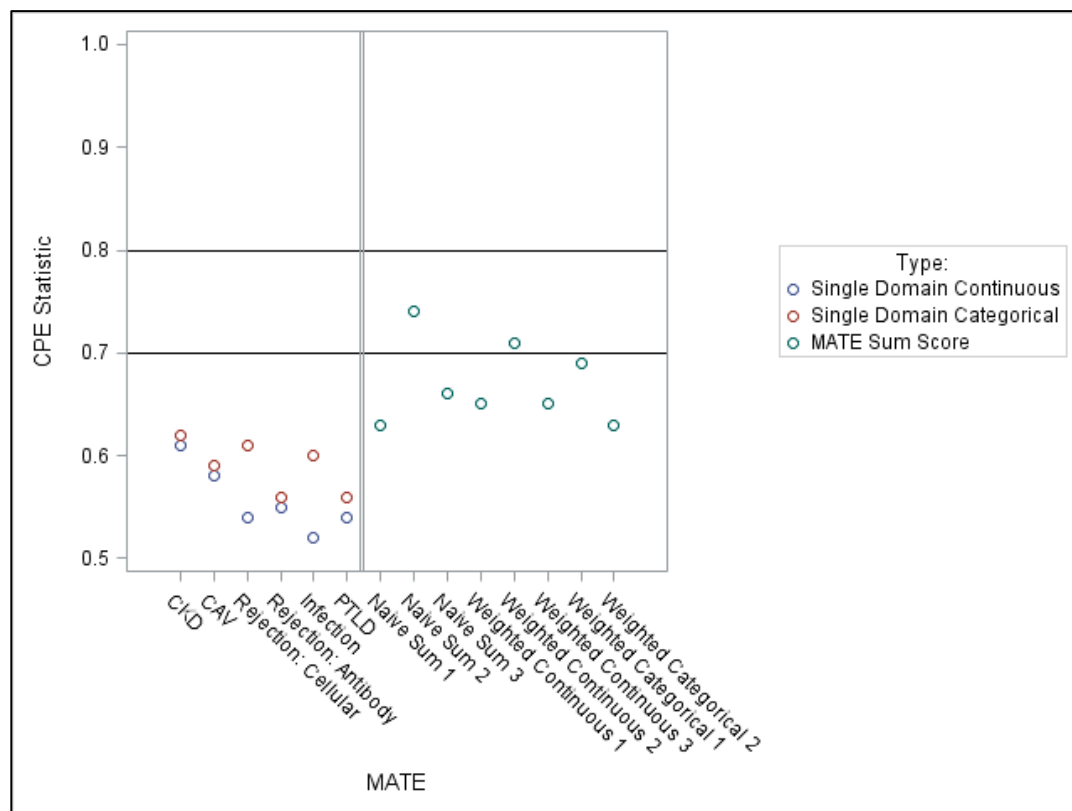
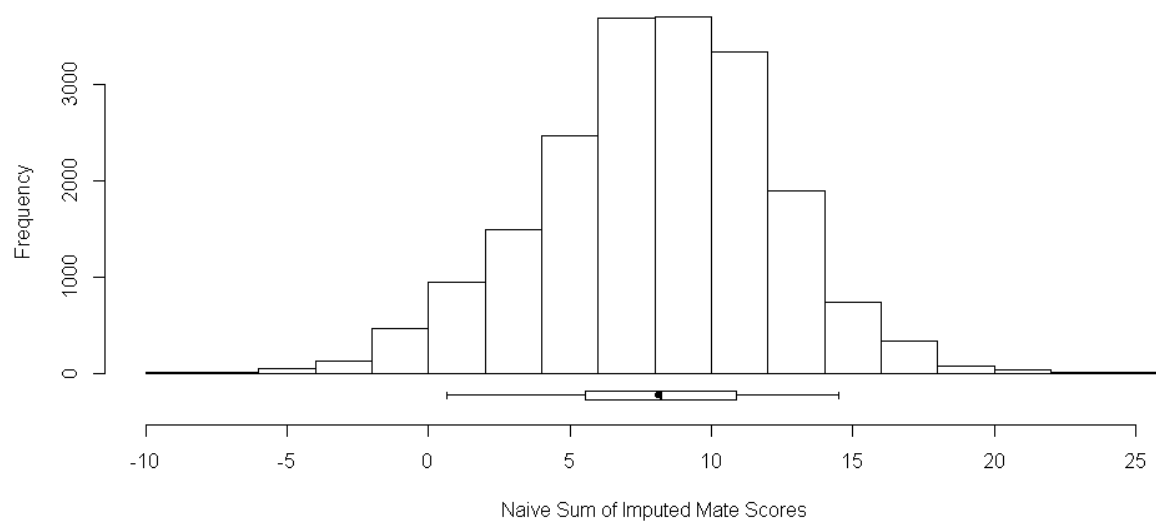


FIGURE 3. Observed distribution of imputed MATE scores using linear regression in the test cohort.



The mean naïve MATE score was 10 with a standard deviation of 4 units.

FIGURE 4A. Association between naïve MATE score categories and graft survival between 6 and 36 months post-heart transplant for all major adverse transplant events.

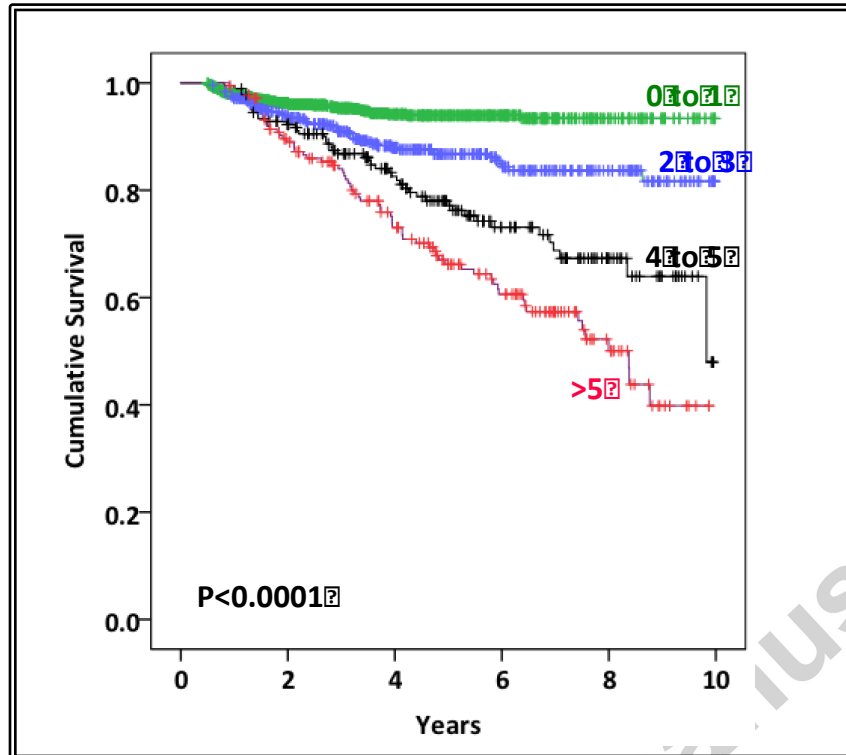


FIGURE 4B. Association between naïve MATE score categories and graft survival between 6 and 36 months post-heart transplant for three adverse events: renal dysfunction, coronary artery vasculopathy and cellular rejection.

