PERSPECTIVE

Best practices for statistical analysis of pretransplant medical urgency

Grace R. Lyden, PhD,[a] Daniel Y. Johnson, BA,[b] Jon J. Snyder, PhD, MS,[a,c,d] Jessica R. Golbus, MD, MS,[e,f,1] and William F. Parker, MD, MS, PhD[g,h,i,1]

From the [a]Scientific Registry of Transplant Recipients, Hennepin Healthcare Research Institute, Minneapolis, Minnesota; [b]Pritzker School of Medicine, University of Chicago, Chicago, Illinois; [c]Department of Medicine, Hennepin Healthcare, University of Minnesota, Minneapolis, Minnesota; [d]Department of Epidemiology and Community Health, University of Minnesota, Minneapolis, Minnesota; [e]Division of Cardiovascular Diseases, Department of Internal Medicine, University of Michigan, Ann Arbor, Michigan; [f]Michigan Integrated Center for Health Analytics and Medical Prediction (MiCHAMP) and Division of Cardiovascular Diseases, Department of Internal Medicine, University of Michigan, Ann Arbor, Michigan; [g]Department of Medicine, University of Chicago, Chicago, Illinois; [h]Department of Public Health Sciences, University of Chicago, Chicago, Illinois; and the [i]MacLean Center for Clinical Medical Ethics, University of Chicago, Chicago, Illinois.

Deceased donor organs suitable for transplant are exceptionally scarce in the United States. Accurately estimating each candidate's medical urgency is essential for organ allocation systems to save the greatest number of lives and meet federal performance goals, as stipulated in the Organ Procurement and Transplantation Network Final Rule.[1] Despite these considerable social and policy implications, the optimal statistical approach for quantifying medical urgency is unresolved, with wide variation in the analytic methods published in leading transplantation journals and federal regulatory or monitoring reports.

In October 2018, a new heart allocation policy replaced the previous 3-tiered system of medical urgency with a 6-tiered system, from status 1 (most urgent) to status 6 (least urgent). If a candidate does not meet standard criteria for a status, their transplant center can request an exception, which allows the candidate to be listed at a status due to perceived illness severity. Although the 2018 policy change was intended to decrease exception requests, the opposite has occurred, with centers now requesting exceptions for 1 in 5 adult heart candidates at initial listing.[2]

Working with 2 independent research teams, the authors of this perspective sought to determine if exception candidates are as medically urgent as standard-criteria candidates. Our Scientific Registry of Transplant Recipients (SRTR) cohort analyses, both published by *The Journal of Heart and Lung Transplantation*,[2,3] had overlapping methods, results, and conclusions. However, differences in key statistical choices led to differences in our results. Motivated by this example, we highlight the challenges in statistical analysis of pretransplant medical urgency and collectively propose a set of best analytic practices that are generalizable to all deceased donor organ allocation.

Statistical practices to follow when analyzing the medical urgency of transplant candidates

Risk adjustment should reflect the study question

In both of our papers, pretransplant mortality was estimated by exception status, adjusting only for candidates' medical urgency status. Neither analysis adjusted for other factors that predict mortality, such as hemodynamics, comorbidities, or mechanical circulatory support device history. This was intentional. Medical urgency status should fully capture the risk of pretransplant mortality of an adult on the heart waiting list. Adjusting for additional covariates is often inappropriate as the current heart allocation policy is blind to these additional prognostic indicators, using medical urgency status alone.[4] We recommend that researchers consider their research question when deciding whether and how to adjust statistical models, especially in the setting of policy analysis.

Do not ignore deaths during inactive periods

Another analytic consideration is how to handle inactive time, when a candidate remains on the waiting list but temporarily stops receiving organ offers. Many pre-transplant deaths occur during inactive periods, as candidates who become too sick for transplant are often made inactive instead of delisted in the hope that they will recover. There is no existing best practice recommendation on how to encode medical urgency status while a patient is inactive. One approach is to treat "inactive" as a unique status in heart allocation (i.e., separate from medical urgency statuses 1-6).[5] However, we believe this approach can obscure important information about the mortality risk of a given status, which might be mediated by inactive time. That is, moving to inactive status does not change the nature of the urgency status previously met and may even be on the causal pathway to a poor waitlist outcome. For future research, we recommend the approach of carrying forward both medical urgency and exception status into inactive periods, so that mortality is attributed to the most recently requested status and most recent criteria met.

Consider time-varying covariates and avoid "immortal time bias"

In both of our papers, a time-dependent covariate approach to handle frequent changes in candidates' medical urgency and exception status over time was used (see Supplemental Table 1 for an example). Two alternatives to time-dependent covariates are frequently used. First, candidates can be analyzed in groups defined by their initial covariates at baseline (e.g., listing or a landmark start time). This approach is also statistically correct and has been used to create medical urgency scores, such as the model for end-stage liver disease.[6] However, this approach estimates mortality *after having* a particular covariate value at baseline, while our models estimated mortality *while having* a particular covariate value; this distinction may be important, depending on the policy goals. Second, a researcher could "look into the future" and create a static covariate that summarizes a candidate's entire waitlist history. For example, a researcher could code a binary "ever exception" flag that is positive if the candidate ever received an exception during their wait for transplant. This approach, however, is statistically inappropriate and should not be used. In particular, this approach leads to "immortal time bias" due to creating periods where a candidate cannot die by construction and would underestimate mortality at exception status.[7]

Use cause-specific hazards, not subdistribution hazards, to assess medical urgency

The cause-specific hazards model and the subdistribution hazards (i.e., Fine-Gray) model are 2 options for analysis of competing events (e.g., heart transplant and death before heart transplant). Both of our papers used a cause-specific hazards model, and we recommend this approach for analyses that aim to compare the relative medical urgency of different candidates on the waiting list. The Fine-Gray model targets the cumulative incidence of pretransplant mortality, which is directly affected not only by medical urgency but also by covariates that solely affect the rate of transplant. For example, transplant candidates with blood type O might have a higher cumulative incidence of pretransplant mortality due to lower access to transplant,[8] but blood type O candidates are not more likely to die without a transplant compared with similar candidates with non-O blood types. In contrast, a cause-specific hazard captures the instantaneous risk of death among candidates who have not yet undergone transplant and, under certain conditions, can approximate the risk of death in a world without transplant.

It is important to recognize that cause-specific hazard models are potentially susceptible to informative censoring bias when candidates who are more likely to die are also more likely to get a transplant (and be censored). To mitigate this bias, causal inference methods, such as marginal structural models,[9] could be considered.

## Primary outcome for medical urgency remains controversial

A key difference between our analyses was in our choice of outcome. Golbus et al[2] used a composite outcome of waitlist removal due to death or deteriorated condition, whereas Johnson et al[3] used an outcome of pretransplant death within 60 days of waitlist removal. The latter choice, to use deaths after delisting, mirrors the pretransplant mortality metric in SRTR program reports, which also follows patients after removal to capture pretransplant deaths.[10] Considerations for this type of analysis include the completeness of data on deaths after delisting, impact of removal without transplant on patient quality of life, and how to define a candidate's medical urgency or exception status when no longer on the waiting list. Johnson et al carried forward candidates' last medical urgency and exception status for up to 60 days after delisting. Beyond 60 days, it might be more difficult to justify assigning a postremoval death to a candidate's last medical urgency and exception status. Another consideration is when to censor patients without a recorded death: at removal or at some point after removal (e.g., 60 days). The latter approach has been recommended to avoid the "ascertainment bias" that may result when follow-up is extended only for patients who experience the event (i.e., death).[11,12] Of course, this analysis would ideally have access to complete death data after removal, as candidates are "assumed alive" in the absence of a death record.

For future work, we recommend that outcome and follow-up decisions be context specific and, ideally, explored in sensitivity analyses. Removal due to deteriorated condition is a patient-relevant outcome, and composite outcomes often have more events, which increases the precision of confidence intervals. Pretransplant death, however, might be more relevant for a mortality research question and does not depend on center-reported removal codes.

Table 1 Example Combinations of Pretransplant Mortality Analysis Choices

	Outcome definition		Censoring time			Covariate values while inactive	
Model	Death	Death or removed because too sick	Waitlist removal	7 days after removal	60 days after removal	Carry forward previous urgency status	Separate coefficient for inactive status
A		X		X			X
B		X	X			X	
C	X				X	X	

All 3 models used a cause-specific Cox model with 1 time-dependent covariate for medical urgency status. For models A and C, candidates removed for reasons other than transplant were assumed alive through the censoring time.

Analysis results are shown in Figure 1.
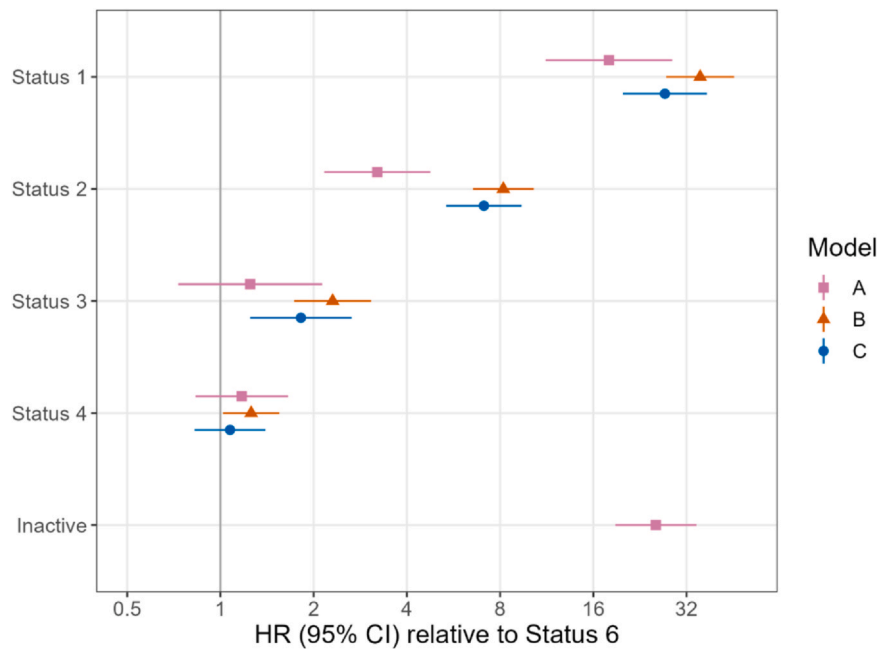
Figure 1   Pretransplant mortality cause-specific hazard ratios for medical urgency statuses 1 to 4 compared with status 6, among adult heart candidates listed under the 2018 heart allocation policy. Only model A estimated a hazard ratio for inactive status; for models B and C, the previous urgency status was carried forward into inactive time. Models A and B both had a composite outcome of death or removal due to deteriorated condition. For model C, the outcome was death through   60 days after delisting. Detailed analytic choices are shown in Table 1. CI, confidence interval; HR, hazard ratio.

Case study: Medical urgency statuses in heart transplant candidates

To illustrate the impact of these decisions on pretransplant mortality, we estimated hazard ratios by urgency status in adults listed for single-organ heart transplant from October 18, 2018, through   December 31, 2022, with follow-up through   March 2023. We considered 3 possible cause-specific Cox models with medical urgency status as a time-dependent covariate (Table 1, Figure 1). When status was not carried forward into inactive time, mortality while inactive was higher   than at any urgency status and confidence intervals for statuses 1 to 4 were wide. When   status was carried forward, hazard ratios were similar although   slightly closer to the null when the outcome was death only, compared with a composite outcome of death or deteriorated condition.

Conclusion

Estimating medical urgency is a cornerstone statistical task for organ allocation systems. Statistical methods that use time-varying covariates, cause-specific hazards, motivated risk adjustment, and deaths   during inactive periods are often best practice and should be adopted in the appropriate context.

Author contributions

G.L.: Conception and design, statistical analysis, drafting and revising of the manuscript, final approval of the manuscript submitted. D.J., J.S.: Revising of the manuscript, review of statistical analysis, final approval of the manuscript submitted. J.G., W.P.: Conception and design,

drafting and revising of the manuscript, final approval of the manuscript submitted.

Disclosure statement

Appendix A. Supporting information

Supplementary data associated with this article can be found in the online version at doi:10.1016/j.healun.2023.11.012.

References

1 US Department of Health and Human Services. Organ procurement and transplantation network: final rule. ⟨https://www.ecfr.gov/current/title-42/chapter-I/subchapter-K/part-121⟩; 1998. Accessed July 26, 2023.

2. Golbus JR, Ahn YS, Lyden GR, et al. Use of exception status listing and related outcomes during two heart allocation policy periods. J Heart Lung Transplant 2023;42:1298-306. https://doi.org/10.1016/j.healun.2023.05.004.

3. Johnson DY, Ahn D, Lazenby K, et al. Association of high-priority exceptions with waitlist mortality among heart transplant candidates. J Heart Lung Transplant 2023;42:1175-82. https://doi.org/10.1016/j.healun.2023.05.009.

4. Organ Procurement and Transplantation Network. OPTN policies. ⟨https://optn.transplant.hrsa.gov/media/eavh5bf3/optn_policies.pdf⟩. Published June 29, 2023. Accessed July 26, 2023.

5. Organ Procurement and Transplantation Network. Four-year monitoring of heart allocation proposal to modify the heart allocation system. ⟨https://optn.transplant.hrsa.gov/media/asdpqli5/data_report_heart_committee_4yr_rpt1.pdf⟩. 2023. Accessed July 26, 2023.

6. Kim WR, Biggins SW, Kremers WK, et al. Hyponatremia and mortality among patients on the liver-transplant waiting list. N Engl J Med 2008;359:1018-26. https://doi.org/10.1056/NEJMoa0801209.

7. Gaynor JJ, Tabbara MM, Ciancio G, et al. The importance of avoiding time-dependent bias when testing the prognostic value of an intervening event—two acute cellular rejection examples in intestinal transplantation. Am J Transplant 2023;23:815-30. https://doi.org/10.1016/j.ajt.2023.02.023.

8. Eapen S, Nordan T, Critsinelis AC, et al. Blood type O heart transplant candidates have longer waitlist time and higher delisting under the new allocation system. J Thorac Cardiovasc Surg 2022;S0022-5223:00826-1. https://doi.org/10.1016/j.jtcvs.2022.07.029. Epub ahead of print.

9. Robins JM, Hernán MÁ, Brumback B. Marginal structural models and causal inference in epidemiology. Epidemiology 2000;11:550-60. https://doi.org/10.1097/00001648-200009000-00011.

10. Scientific Registry of Transplant Recipients. Technical Methods for the program-specific reports.⟨https://srtr.org/about-the-data/technical-methods-for-the-program-specific-reports/⟩. Published 2023. Accessed July 26, 2023.

11. Lazenby KA, Narang N, Pelzer KM, Ran G, Parker WF. An updated estimate of posttransplant survival after implementation of the new donor heart allocation policy. Am J Transplant 2022;22:1683-90. https://doi.org/10.1111/ajt.16931.

12. Motter JD, Segev DL, Massie AB. Don't stop thinking about tomorrow: ascertainment bias in pre-post design transplant registry studies. Am J Transplant 2022;22:1513-4. https://doi.org/10.1111/ajt.16984.