



# HHS Public Access

Author manuscript

*Am J Transplant.* Author manuscript; available in PMC 2020 February 01.

Published in final edited form as:

*Am J Transplant.* 2019 February ; 19(2): 391–398. doi:10.1111/ajt.15038.

## Comparing Scientific Registry of Transplant Recipient posttransplant program-specific outcome ratings at listing with subsequent recipient outcomes after transplant

Andrew Wey, PhD<sup>1</sup>, Nicholas Salkowski, PhD<sup>1</sup>, Bertram L. Kasiske, MD<sup>1,2</sup>, Melissa Skeans, MS<sup>1</sup>, Cory R. Schaffhausen, PhD<sup>2</sup>, Sally K. Gustafson, MS<sup>1</sup>, Ajay K. Israni, MD<sup>1,2,3</sup>, Jon J. Snyder, PhD<sup>1,3</sup>

<sup>1</sup>Scientific Registry of Transplant Recipients, Minneapolis Medical Research Foundation, Minneapolis, Minnesota

<sup>2</sup>Department of Medicine, Hennepin County Medical Center, University of Minnesota, Minneapolis, Minnesota

<sup>3</sup>Department of Epidemiology and Community Health, University of Minnesota, Minneapolis, Minnesota

### Abstract

To improve accessibility of program-specific reports to patients, the Scientific Registry of Transplant Recipients released a 5-tier system for categorizing 1-year posttransplant program evaluations. Whether this system predicts subsequent posttransplant outcomes at the time patients are wait-listed has been questioned. We investigated the association of tier at listing and the corresponding continuous score used for tier assignment, which ranges from 0 (poor outcomes) to 1 (good outcomes), with eventual 1-year posttransplant graft survival for candidates listed July 12, 2011-June 16, 2014, who underwent transplant before December 31, 2016. One additional tier at listing was associated with better 1-year posttransplant outcomes in liver (hazard ratio [HR], 0.89<sub>0.93</sub><sup>0.97</sup>) and lung transplantation (HR, 0.84<sub>0.90</sub><sup>0.97</sup>), but not kidney (HR, 0.92<sub>0.96</sub><sup>1.01</sup>) or heart transplantation (HR, 0.93<sub>1.02</sub><sup>1.10</sup>). In liver and lung transplantation, longer time between listing and transplant was associated with stronger protective effects for high-tier programs. In kidney, liver, and lung transplantation, posttransplant evaluations at listing had non-linear associations with eventual posttransplant outcomes: relatively flat for 5-tier scores below 0.5 and decreasing for scores above 0.5. After adjusting for measured recipient and donor risk factors, posttransplant evaluations at listing predicted differences in eventual outcomes in liver and lung transplantation, providing useful information to patients.

---

Correspondence: Andrew Wey, PhD, [awey@cdrg.org](mailto:awey@cdrg.org).

Disclosure

The authors of this manuscript have no conflicts of interest to disclose as described by the *American Journal of Transplantation*.

Supporting Information

Additional Supporting Information may be found online in the supporting information tab for this article.

## Introduction

The Organ Procurement and Transplantation Network (OPTN) Final Rule requires the Scientific Registry of Transplant Recipients (SRTR) to publish program-specific reports (PSRs) that can be “accurately and efficiently” used and understood. In December 2016, SRTR released a 5-tier system to improve the accessibility of 1-year posttransplant program evaluations already included in the PSRs. The statistical summary measures included in the PSRs (hazard ratios [HRs] and confidence intervals) were considered too challenging for non-technical stakeholders, e.g., patients and families, to understand.<sup>1;2</sup> Instead, a more accessible approach would transform the statistical summary measures into a limited number of categories. SRTR previously used a 3-tier system based on statistical hypothesis testing. The 3-tier system poorly differentiated program evaluations and may have obscured potentially relevant differences, especially for small-to-moderately sized programs.<sup>3</sup> The 5-tier system was designed to better differentiate posttransplant evaluations and, in kidney transplantation, it reduced the variability of program-specific HRs within a tier by nearly 80%.<sup>3</sup>

Since the 5-tier system narrowed the outcome differences across tiers, programs were more likely to change tiers over time than they were under the previous 3-tier system.<sup>4</sup> The higher variability in tier assignment over time and long duration from listing to transplant, especially for kidney candidates,<sup>5</sup> could attenuate the association between tier assignment when a candidate joins the waiting list and the posttransplant outcomes when the candidate eventually undergoes transplant. If tier assignment at listing is not associated with eventual posttransplant outcomes, then the 5-tier system would not necessarily convey the relative survival experience after transplant across programs. However, there are other considerations for public reporting of posttransplant outcomes including, for example, creating incentives for quality improvement at all programs. Further, current regulatory review could attenuate an association between tier assignment at listing and eventual posttransplant outcomes, because programs that approach or cross regulatory thresholds have incentives to improve.<sup>6</sup> Regardless, the association of tier assignment at listing and, more generally, posttransplant program evaluations at listing with eventual posttransplant outcomes has not been evaluated.

We investigated three specific dimensions of the relationship between posttransplant program evaluations at listing and eventual posttransplant outcomes. First, we investigated the overall association between tier at listing and eventual posttransplant outcomes, because tier assignment is a discrete number and could influence patient decision-making. However, the association may be small due to the potentially long delay between transplants included in the evaluation and the actual time of listing. Second, we investigated the dependence of the association on the time from listing to transplant. If the association attenuates with longer time between listing and transplant, then long waiting times could reduce the utility of the 5-tier system, especially in kidney transplantation. Lastly, we investigated the association of posttransplant program evaluations at listing without categorization into tiers, i.e., the continuous score used to categorize programs into tiers, with eventual posttransplant outcomes. The role of posttransplant evaluations in regulatory review could create a non-linear relationship with eventual posttransplant outcomes due to the incentives to improve for programs close to or violating regulatory thresholds. Understanding these associations

will help inform further discussion of the role of posttransplant program evaluations in public reporting, and we evaluated each dimension in kidney, liver, lung, and heart transplantation to ensure relevance to the broader transplant community.

## Materials and Methods

This study used SRTR data. The SRTR data system includes data on all donors, waitlisted candidates, and transplant recipients in the US, submitted by the members of OPTN, and has been described elsewhere.<sup>7</sup> The Health Resources and Services Administration, US Department of Health and Human Services, provides oversight of the activities of the OPTN and SRTR contractors.

Adult recipients (age ≥ 18 years at transplant) were included in the study if they were listed between July 12, 2011 (release date for the 2011 summer PSR cycle), and June 16, 2014 (day before the release of the 2014 summer PSR cycle), and underwent transplant before December 31, 2016. The appropriate PSR release for, e.g., tier assignment at listing, was determined by the release date of archived PSRs. In 2013 and 2014, the PSR release schedule was temporarily postponed due to problems related to determining dates of death. In these situations, the PSR release date was approximated by the prior release of the corresponding biannual PSR cycle.

The primary outcome was 1-year posttransplant graft survival; the Supplementary Materials provide the specific definitions of graft survival for each organ. Follow-up was censored at 1 year posttransplant or December 31, 2016, whichever occurred first. The posttransplant evaluations at listing were the 1-year posttransplant graft survival outcomes included in the PSR at the time the recipient was listed. The tier assignment was determined by the algorithm for the 5-tier system. The algorithm depends only on the number of observed and expected graft failures and can therefore be calculated with archived PSRs.<sup>3,8</sup>

A linear trend estimated the association of tier at listing with eventual posttransplant graft survival. The corresponding interpretation is the average change in the hazard for one additional tier at listing. To determine if time from listing to transplant modified the association, recipients were separated into three groups: those who underwent transplant in less than 1 year, 1–2 years, and 2 years or longer after listing. The linear trend for tier at listing was then estimated for recipients in each group.

To identify potential non-linear associations, penalized splines estimated the association of eventual posttransplant graft survival with the underlying continuous score used to categorize programs into the 5-tier system (referred to throughout as the 5-tier score). The 5-tier score ranges from 0 to 1; a 5-tier score close to 1 corresponds to above average outcomes, while a 5-tier score close to 0 corresponds to below average outcomes. Penalized splines have wider confidence intervals than linear trends due to the additional flexibility.

The linear and non-linear associations with 1-year posttransplant graft survival were estimated with separate Cox proportional hazards models that adjusted for recipient and donor factors. A recipient or donor factor was included if the corresponding SRTR 1-year adult graft survival model for the January 2018 PSR release had a non-zero effect for the

given factor; documentation for SRTR posttransplant risk-adjustment models is accessible at <https://www.srtr.org/reports-tools/risk-adjustment-models-posttransplant-outcomes/>. The Supplementary Materials list the specific donor and recipient factors included in each model. For kidney and liver transplantation, living and deceased donor organs were integrated into the same model with an indicator for living donor transplant. Factors with non-zero effects in either the living or deceased donor PSR models were included. The value of deceased-donor-specific factors for living donor transplants was set to the median for continuous factors and the reference level for categorical factors. Multiple imputation with 10 iterations accounted for missing data.<sup>9</sup>

The effects of the continuous factors were estimated with penalized splines. Robust standard errors accounted for correlation among transplants at the same program. The Supplementary Materials include a sensitivity analysis for the effect of an additional tier in the traditional 3-tier system based on statistical hypothesis testing.

All analyses were completed in R v3.3.3.<sup>10</sup> Cox proportional hazard models were estimated with the “survival” package,<sup>11</sup> and multiple imputation was completed by the “mice” package.<sup>12</sup>

## Results

A significant proportion of kidney transplant recipients (46%) underwent transplant within 1 year of listing, although the time between listing and transplant was strongly associated with donor type and metrics of allocation priority, e.g., dialysis duration at transplant (Table S1). Most liver (80%; see Table S2), lung (88%; see Table S3), and heart (79%; see Table S4) recipients also underwent transplant within 1 year of listing.

### Kidney Transplantation (Figure 1)

An additional tier at listing had a 4% lower hazard of kidney graft failure 1 year after transplant, although the association was not significant (HR: 0.920.96<sub>1.01</sub>). The association was not modified by time from listing to transplant, although, likely due to the smaller sample sizes, the confidence intervals were wider for each category of time from listing to transplant; e.g., one additional tier at listing for recipients who underwent transplant within 1 year was associated with a 4% reduction in the hazard of graft failure (HR: 0.900.96<sub>1.04</sub>).

The 5-tier score at listing had a non-linear association with eventual posttransplant graft survival (Figure 1, right panel). The 5-tier score had an association that was relatively flat until a score of 0.5, after which the risk of graft failure slowly decreased. In summary, outcomes did not differ for recipients who listed at programs with average or below average evaluations, but recipients who listed at programs with above average evaluations had improved posttransplant graft survival.

### Liver Transplantation (Figure 2)

An additional one tier at listing was associated with a 7% lower hazard (Figure 2, left panel) of 1-year posttransplant liver graft failure (HR: 0.890.93<sub>0.97</sub>), and the effect was larger for recipients who waited longer than 2 years. For example, one additional tier at listing for

recipients who underwent transplant within 1 year was associated with a 6% lower hazard of graft failure (HR: 0.90<sub>0.94,0.98</sub>), while one additional tier for recipients who underwent transplant more than 2 years after listing was associated with a 32% lower hazard of graft failure (HR: 0.56<sub>0.68,0.81</sub>). Posttransplant evaluations in liver transplantation were associated with eventual posttransplant graft survival, and a longer time between listing and transplant strengthened the association.

Posttransplant liver graft survival had a non-linear relationship with the 5-tier score at listing (Figure 2; right panel). The hazard of graft failure for the 5-tier score at listing gradually decreased for programs with scores above 0.2. Thus, recipients who listed at programs with better posttransplant evaluations tended to have better eventual 1-year posttransplant graft survival, although the differences attenuated for programs with 5-tier scores below 0.2.

### Lung Transplantation (Figure 3)

An additional tier at listing was associated with a 10% lower hazard (Figure 3, left panel) of 1-year posttransplant graft failure (HR: 0.84<sub>0.90,0.97</sub>), and the effect was stronger for recipients who waited longer after listing. For example, one additional tier at listing for recipients who underwent transplant within 1 year was associated with an 8% lower hazard of graft failure (HR: 0.85<sub>0.92,1.00</sub>), while one additional tier for recipients who underwent transplant between 1 and 2 years after listing was associated with a 24% lower hazard (HR: 0.59<sub>0.76,0.97</sub>). Posttransplant evaluations in lung transplantation were associated with eventual posttransplant graft survival, and a longer time between listing and transplant strengthened the association.

The effect for the 5-tier score at listing was relatively constant until a score of approximately 0.6, then sharply decreased for 5-tier scores at listing above 0.6, corresponding to programs with better posttransplant evaluations at listing (Figure 3, right panel). Thus, recipients who listed at programs with above average posttransplant evaluations experienced gradually better outcomes, while recipients who listed at programs with average or below average posttransplant evaluations experienced similar outcomes.

### Heart Transplantation (Figure 4)

Tier assignment at listing (Figure 4, left panel) and the 5-tier score at listing (Figure 4; right-panel) were not associated with eventual posttransplant outcomes in heart transplantation. Thus, heart recipients who listed at programs with better posttransplant evaluations did not experience better or worse outcomes than heart recipients who listed at programs with worse posttransplant evaluations.

## Discussion

For liver and lung recipients, posttransplant evaluations at listing were moderately associated with eventual posttransplant outcomes, and the associations did not attenuate with longer durations between listing and transplant but became stronger. Interestingly, posttransplant evaluations had a non-linear association with eventual posttransplant outcomes in kidney, liver, and lung transplantation: notable differences in eventual outcomes between programs with above average to average evaluations but attenuated or no differences between

programs with average to below average evaluations. In heart transplantation, a better tier assignment at listing was not associated with better or worse outcomes, and it is not clear that any categorization system would be associated with posttransplant outcomes in heart transplantation because the 5-tier score at listing had no association with eventual outcomes.

We performed a sensitivity analysis that repeated the primary analysis for the traditional 3-tier system based on statistical hypothesis testing. The analysis for the 3-tier system showed larger effects but also higher variability for one additional tier at listing than the associations for the 5-tier system (Table S5). The larger effect size is not surprising because one additional tier in the 3-tier system corresponds to approximately two additional tiers in the 5-tier system, while the higher variability is likely explained by the smaller number of recipients in tiers 1 and 3.

There are at least two frameworks for understanding the utility of posttransplant evaluations: first, the ability of posttransplant evaluations to predict eventual outcomes, and second, creating incentives for quality improvement. The former framework has a direct role in patient decision-making because an association with eventual posttransplant outcomes would create an incentive for patients to list at programs with good posttransplant evaluations. In contrast, the latter framework is only indirectly related to patient outcomes through the potentially better outcomes at all programs due to the incentive for quality improvement.

Unfortunately, these frameworks may work in opposition to each other; that is, an incentive for quality improvement could attenuate or remove the ability of posttransplant evaluations to predict eventual outcomes. For example, regulatory agencies (i.e., the Centers for Medicare & Medicaid Services [CMS] and OPTN) use 1-year posttransplant evaluations to identify programs for regulatory intervention. Regulatory agencies may then force transplant programs that consistently violate the thresholds to institute quality improvement measures.<sup>6</sup> Additionally, transplant programs in violation of, or close to, the regulatory thresholds have strong financial incentives to improve posttransplant outcomes due to the financial burden associated with regulatory interventions. This incentive structure could reduce the differences in eventual survival outcomes among programs with below average posttransplant evaluations, although the incentive structure may not be strong enough for programs with average or better evaluations to pursue similar, but potentially costly, quality improvement measures. The trends that could occur within this incentive structure are consistent with the non-linear associations observed in kidney, liver, and lung transplantation; that is, the association would be relatively flat for hazard ratios above 1 and decrease for hazard ratios below 1. The 5-tier system for categorizing posttransplant outcomes could create an incentive for all programs, not only programs with below average evaluations, to pursue quality improvement measures,<sup>13</sup> and the incentive could hypothetically eliminate *any* association with eventual posttransplant outcomes.<sup>14</sup> Yet, in this hypothetical situation, public reporting would still have an important role despite less direct utility in patient decision-making.

Both frameworks can justify further public reporting of posttransplant evaluations. For example, an association with eventual posttransplant outcomes would motivate prominent

and accessible reporting of posttransplant evaluations. In contrast, in the absence of an association with eventual posttransplant outcomes (e.g., heart transplantation), the approach to public reporting would be more nuanced. Specifically, the reporting must sufficiently motivate quality improvement efforts,<sup>13</sup> while minimizing the risk that patients over-emphasize posttransplant evaluations at the cost of other important components of transplant program care, e.g., access to transplant as measured by transplant rate. In this situation, the presentation of posttransplant evaluations must balance accessible and adequately differentiated observed outcomes, while ensuring that stakeholders, e.g., patients, understand that the transplant rate evaluation may predict their long-term survival better than the posttransplant evaluation. A potential approach to balancing these issues is integrating into the public reporting a plain-language description of the association of posttransplant evaluations with eventual posttransplant outcomes.

The focus on transplant recipients is an inherent limitation of our analysis because many patients, especially in liver and kidney transplantation, die or are removed from the waiting list before transplant.<sup>5;15</sup> While the association of posttransplant evaluations with eventual recipient outcomes is information sought by patients,<sup>2</sup> posttransplant evaluations may not be particularly informative of the experience of most candidates after listing and, instead, the transplant rate ratio evaluation may have a stronger association with survival. In fact, posttransplant evaluations were associated with overall mortality after listing for liver and lung transplantation but not for kidney or heart transplantation. Further, in kidney transplantation, the transplant rate evaluation at listing was associated with overall mortality after listing (SRTR, unpublished data, March 2018). These associations suggest that the process of undergoing a kidney transplant is more important than posttransplant outcomes for survival after listing. Thus, SRTR plans to better integrate transplant rate and waitlist mortality evaluations into the public reporting and will only emphasize posttransplant evaluations for organs with a strong association with candidate mortality after listing.

Our analysis helps inform the role of posttransplant evaluations in public reporting, not in regulatory review. SRTR is mandated to provide program-specific information that can be used “accurately and efficiently,” and understanding the importance of program-specific posttransplant evaluations in eventual patient outcomes is important for accurate use by patients and their families. In contrast, CMS developed conditions of participation to “improve quality and protect the health and safety of beneficiaries,”<sup>16</sup> which suggests greater priority on incentivizing quality improvement in transplantation, not informing patients of potential outcomes. The different role of posttransplant evaluations in regulatory review was also evident in the justification of recent revisions to CMS’ criteria for review: a more difficult standard to violate because patient survival improved over the past decade.<sup>17</sup> Thus, the association of posttransplant evaluations with prospective graft survival is less important for informing the role of posttransplant evaluations in regulatory review.

For each organ, the differences in eventual posttransplant outcomes were relatively small, especially compared with the differences suggested by the posttransplant evaluations. However, smaller differences were expected because the predictive performance of statistical models is almost always worse for future outcomes than for outcomes used to estimate the model.<sup>18</sup> For example, a tier-5 program had an 81% probability of a truly better HR than a

tier-3 program within a PSR cycle.<sup>3</sup> A prospective, rather than within-PSR cycle, probability would naturally be lower because program care can change due to quality improvement efforts and/or staff turnover. However, the association in liver and lung transplantation suggests that the 5-tier system remained predictive despite these challenges.

Public reporting of posttransplant evaluations may create unintended consequences. CMS' introduction of regulatory review of posttransplant evaluations was associated with lower transplant volume and higher rates of waitlist removals for programs under regulatory review.<sup>6;19;20</sup> Better public reporting with, e.g., the 5-tier system, may result in similar risk-averse behavior because programs may fear, for example, loss of referrals and/or pressure from private insurers. The unintended consequences of public reporting may have an association different from that of regulatory review due to indirect rather than direct interaction with programs. That is, CMS directly interacted with programs under review and sometimes required quality improvement efforts. In contrast, public reporting indirectly interacts with programs through, for example, the potential loss of referrals and/or pressure from private insurers. The unintended consequences of public reporting deserve further investigation, especially with respect to long-term survival of candidates awaiting transplant, and risk-averse behavior may be reduced through further education about the role of risk adjustment in posttransplant evaluations.<sup>21</sup>

While these results may partially reflect changes in programs over time, the variability in tier assignment over time is not an inherently informative metric for evaluating the categorization of posttransplant outcomes.<sup>4</sup> The 5-tier system was explicitly designed to improve the differentiation of program performance at the cost of a higher misclassification rate.<sup>3</sup> As a direct consequence, the 5-tier system would be expected to be more variable in tier assignment over time because better differentiation was achieved by narrowing the differences in outcomes between tiers. More importantly, variability in tier assignment over time does not answer the clinical question of interest: that variability in posttransplant evaluations over time may mislead patients about their eventual posttransplant survival experience due to long waiting times.<sup>4</sup> Yet, longer waiting times did not modify the association of posttransplant evaluations at listing with eventual outcomes for kidney recipients, and the association was stronger for liver and lung recipients with longer waiting times, despite the relatively shorter organ-specific waiting times. Finally, in heart transplantation, posttransplant evaluations had no association with eventual recipient outcomes, and waiting times did not modify the association. Thus, tier assignment at listing was associated with eventual recipient outcomes in kidney, liver, and lung transplantation, and the time from listing to transplant either did not modify the associations or the associations were modified in the direction *opposite* what was expected.

Our analysis is subject to potential limitations. First, it is difficult to understand the causal mechanisms that modify the association of tier at listing with posttransplant outcomes across the different durations of time on the waiting list. For example, if liver recipients who waited for over 2 years were more likely to have unmeasured risk factors at programs in tiers 1–3 than at programs in tiers 4 and 5, then listing at a tier-4 or –5 program could have an apparent protective effect. This could occur, for example, if programs with good posttransplant evaluations more carefully monitored candidates on their waiting lists and



were therefore more likely than programs with poor posttransplant evaluations to identify candidates who develop risk factors after listing that are not collected by OPTN. If these risk factors are not contraindications to transplant, then the best approach for alleviating their impact is collection of additional data. Second, the posttransplant evaluations in this cohort used previous risk-adjustment models that incorporated fewer risk factors and did not consider flexible linear splines for continuous covariates.<sup>22</sup> In both the current PSR models and the risk-adjustment models described here, we adjusted for a much wider range of risk factors and included flexible splines that can identify non-linear effects of continuous covariates. The differences in risk adjustment could bias the association between posttransplant evaluations and eventual outcomes if, for example, the previous models inappropriately identified programs as performing poorly due to a high risk tolerance, while the new risk-adjustment models more accurately identify the program's risk tolerance. This hypothetical situation could cause programs with poor posttransplant evaluations to have average posttransplant outcomes, which would bias the relationship between the baseline evaluation and subsequent outcome. Lastly, patients likely choose transplant programs at the time of referral and evaluation rather than at listing, but listing date may be a reasonable approximation given lack of data on referral and evaluation.

Posttransplant evaluations at listing and, specifically, the 5-tier system were associated with eventual posttransplant outcomes in liver and lung transplantation, but not in kidney or heart transplantation. Thus, the 5-tier system differentiated eventual recipient outcomes in liver and lung transplantation that were unexplained by measured recipient and donor risk factors.

## Supplementary Material

Refer to Web version on PubMed Central for supplementary material.

## Acknowledgments

This work was conducted under the auspices of the Minneapolis Medical Research Foundation, contractor for the Scientific Registry of Transplant Recipients, as a deliverable under contract number HHSH250201500009C (US Department of Health and Human Services, Health Resources and Services Administration, Healthcare Systems Bureau, Division of Transplantation). As a US Government-sponsored work, there are no restrictions on its use. The views expressed herein are those of the authors and not necessarily those of the US Government. AKI was partially supported by R01 HS 24527. The authors thank SRTR colleague Nan Booth, MSW, MPH, ELS, for manuscript editing.

## Abbreviations:

<b>CMS</b>	Centers for Medicare & Medicaid Services
<b>HR</b>	hazard ratio
<b>OPTN</b>	Organ Procurement and Transplantation Network
<b>PSR</b>	program-specific report
<b>SRTR</b>	Scientific Registry of Transplant Recipients

## References

1. Hibbard J, Sofaer S. Best practices in public reporting no. 1: How to effectively present health care performance data to consumers. Agency for Healthcare Research and Quality; 2010 Available at: <https://archive.ahrq.gov/professionals/quality-patient-safety/quality-resources/tools/pubrptguide1/pubrptguide1.pdf>. Accessed June 19, 2018.
2. Schaffhausen CR, Bruin MJ, Chesley D, et al. What patients and members of their support networks ask about transplant program data. *Clin Transplant*. 2017;31(12):1–8.
3. Wey A, Salkowski N, Kasiske BL, Israni AK, Snyder JJ. A five-tier system for improving the categorization of transplant program performance. *Health Serv Res*. 2018;53(3):1979–1991. [PubMed: 28608369]
4. Schold JD, Andreoni KA, Chandraker AK, et al. Expanding clarity or confusion? Volatility of the five-tier ratings assessing quality of transplant centers in the United States. *Am J Transplant*. 2018;18(6):1494–1501 [PubMed: 29316241]
5. Hart A, Smith JM, Skeans MA, et al. OPTN/SRTR 2015 Annual Data Report: Kidney. *Am J Transplant*. 2017;16(Suppl 2):S21–S116.
6. Hamilton TE. Regulatory oversight in transplantation: are the patients really better off? *Curr Opin Organ Transplant*. 2013;18(2):203–209. [PubMed: 23429659]
7. Leppke S, Leighton T, Zaun D, et al. Scientific Registry of Transplant Recipients: Collecting, analyzing, and reporting data on transplantation in the United States. *Transplant Rev*. 2013;27(2): 50–56.
8. Salkowski N, Snyder JJ, Zaun DA, Leighton T, Israni AK, Kasiske BL. Bayesian methods for assessing transplant program performance. *Am J Transplant*. 2014;14(6):1271–1276. [PubMed: 24787026]
9. Little R, Rubin D. *Statistical Analysis with Missing Data*. 2002, 2nd Ed.: Hoboken, New Jersey:Wiley;2002.
10. R Core Team. *R: A Language and Environment for Statistical Computing*. Vienna, Austria: R Foundation for Statistical Computing; 2015 Available at <http://www.R-project.org/>. Accessed June 19, 2018.
11. Therneau T. A package for survival analysis in S. version 2.38. 2015 Available at: <https://cran.r-project.org/web/packages/survival/index.html>. Accessed June 19, 2018.
12. Van Buuren S, Groothuis-Oudshoorn K. Mice: multivariate imputation by chained equations in R. *J Stat Software*. 2011;45(3):1–67. Available at: <http://www.jstatsoft.org/v45/i03/>. Accessed June 19, 2018.
13. Hibbard JH, Stockard J, Tusler M. Does publicizing hospital performance simulate quality improvement efforts? *Health Affairs*. 2003;22(2):84–94.
14. Hibbard JH, Stockard J, Tusler M. Hospital performance reports: Impact on quality, market share, and reputation. *Health Affairs*. 2005;24(4):1150–1160.
15. Kim WR, Lake JR, Smith JM, et al. OPTN/SRTR 2015 Annual Date Report: Liver. *Am J Transplant*. 2017;17(Suppl 1):174–251. [PubMed: 28052604]
16. Centers for Medicare & Medicaid Services. Conditions for Coverage (CfCs) & Conditions of Participations (CoPs). 2013 Available at: <https://www.cms.gov/Regulations-and-Guidance/Legislation/CFCsAndCoPs/index.html>. Accessed June 19, 2018.
17. Kasiske BL, Salkowski N, Wey A, Israni AK, Snyder JJ. Potential implications of recent and proposed changes in the regulatory oversight of solid organ transplantation in the United States. *Am J Transplant*. 2016;16(12):3371–3377. [PubMed: 27401597]
18. Hastie T, Tibshirani R, Friedman J. *The Elements of Statistical Learning: Data Mining, Inference, and Prediction*. 2nd ed.: New York, New York:Springer;2009.
19. Schold JD, Buccini LD, Srinivas TR, et al. The association of center performance evaluations with kidney transplant volume in the United States. *Am J Transplant*. 2013;13(1):67–75.
20. Schold JD, Buccini LD, Poggio ED, Flechner SM, Goldfarb DA. Association of candidate removals from the kidney transplant waiting list and center performance oversight. *Am J Transplant*. 2016;16(4):1276–1284. [PubMed: 26762606]

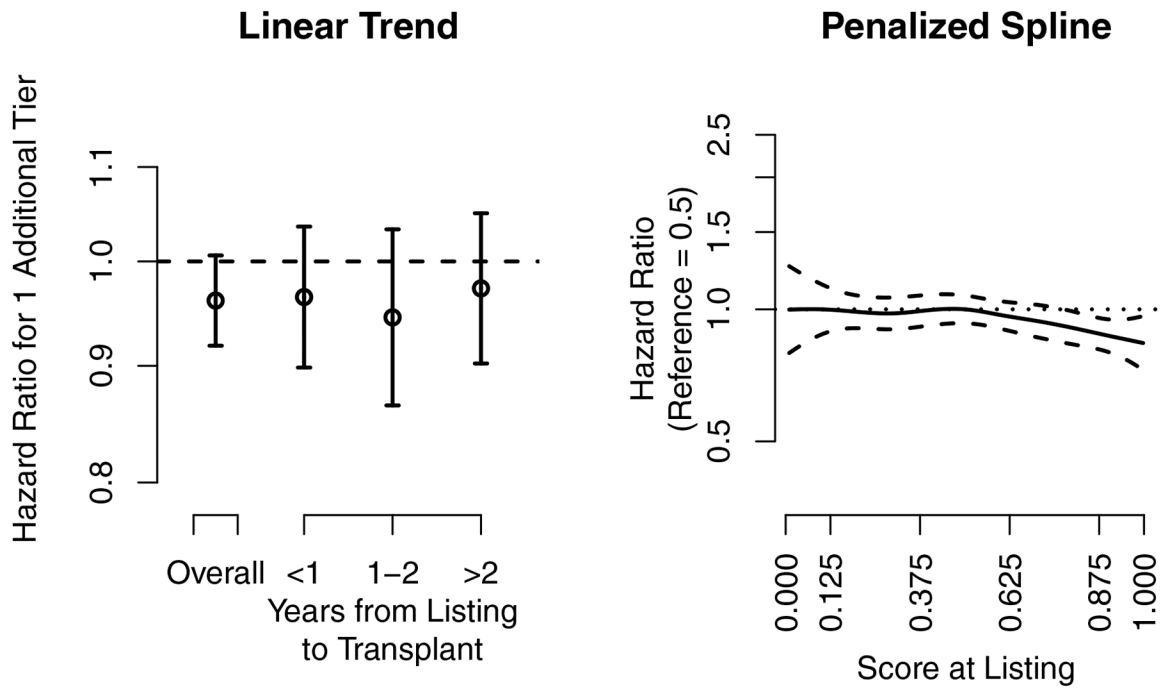
21. Snyder JJ, Salkowski N, Wey A, et al. Effects of high-risk kidneys on Scientific Registry of Transplant Recipients program quality reports. *Am J Transplant.* 2016;16(9):2646–2653. [PubMed: 26954720]
22. Snyder JJ, Salkowski N, Kim SJ, et al. Developing statistical models to assess transplant outcomes using national registries: The process in the United States. *Transplantation.* 2016;100(2):288–294. [PubMed: 26814440]

Author Manuscript

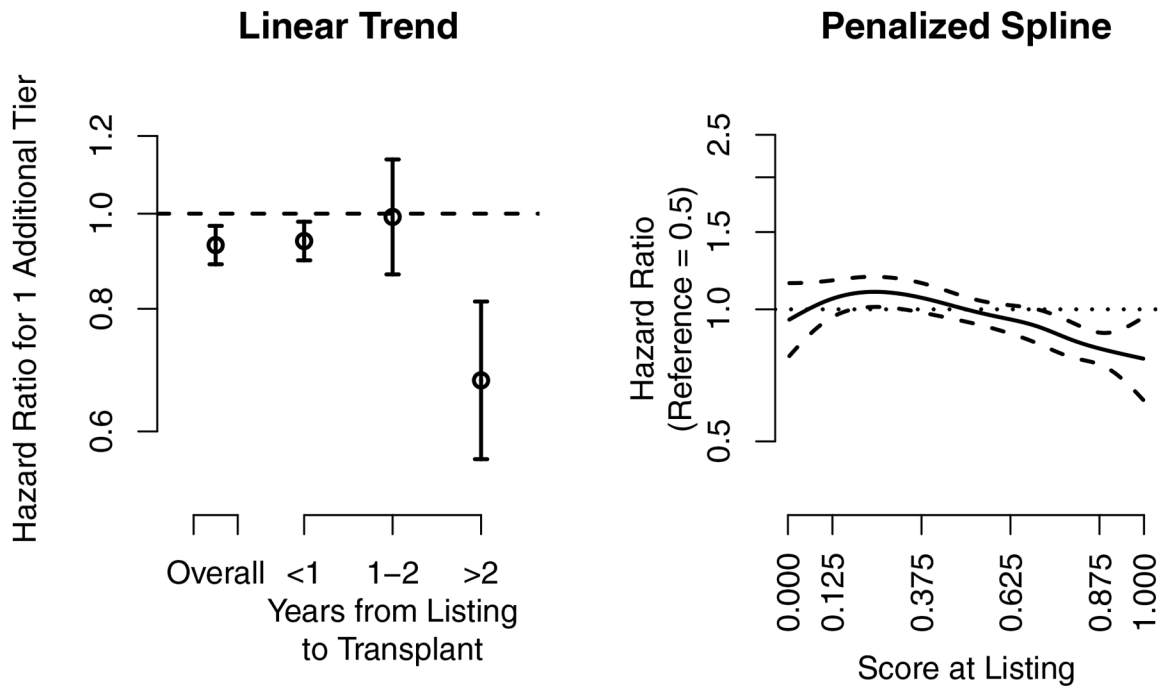
Author Manuscript

Author Manuscript

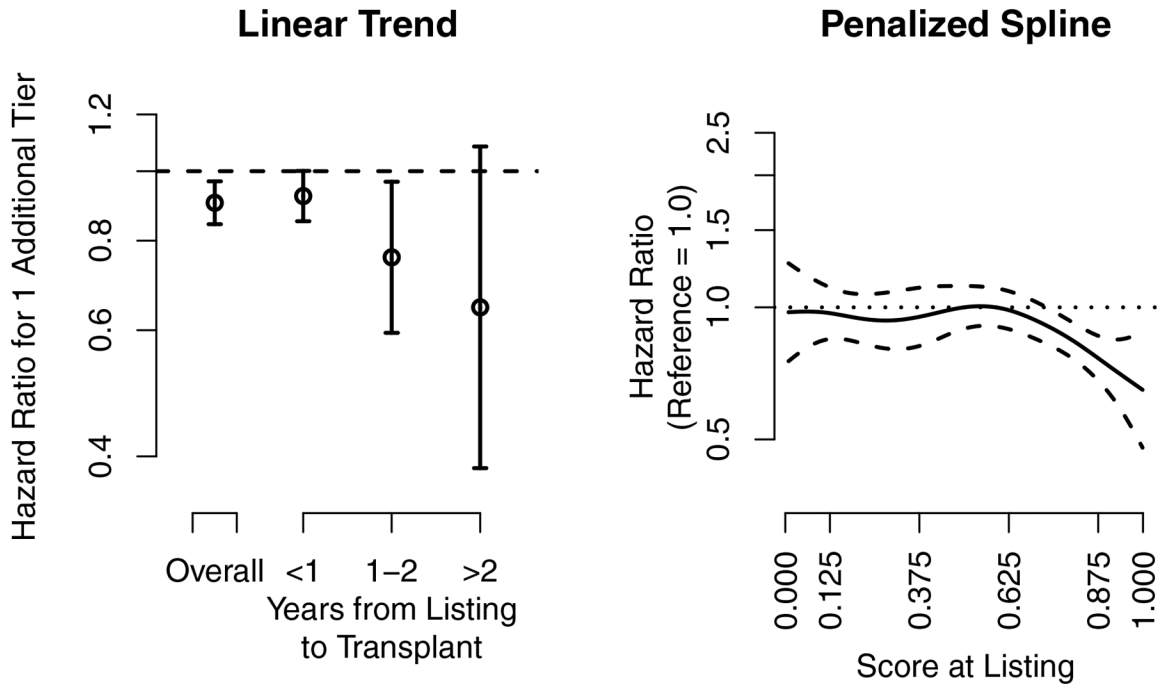
Author Manuscript



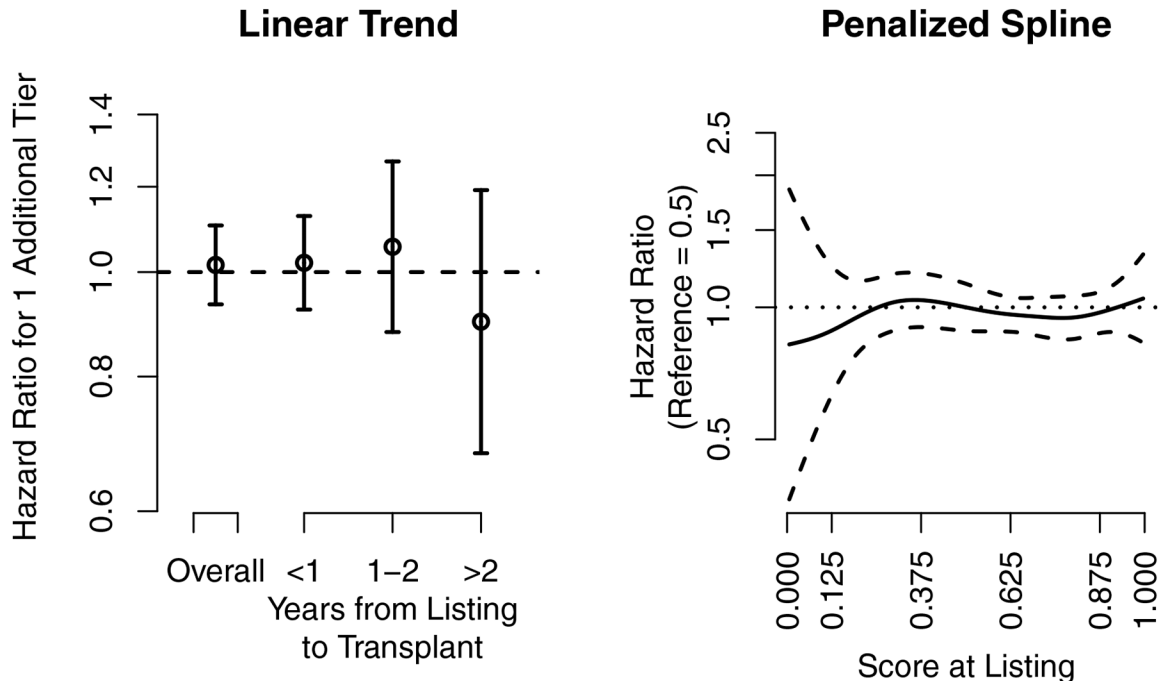
**Figure 1.** Kidney transplantation. The association of the hazard ratio at listing and the 5-tier score at listing with eventual posttransplant 1-year kidney graft survival. The dashed lines are the pointwise 95% confidence intervals, and the dotted line is a hazard ratio of 1. The 5-tier score is the underlying continuous score used to categorize programs into the 5-tier system. The x-axis tick marks for the 5-tier score correspond to the cut-points used to categorize programs into tiers; e.g., programs with a 5-tier score above 0.875 were categorized as tier 5.



**Figure 2.** Liver transplantation. The association of the hazard ratio at listing and the 5-tier score at listing with eventual posttransplant 1-year liver graft survival. The dashed lines are the pointwise 95% confidence intervals, and the dotted line is a hazard ratio of 1. The 5-tier score is the underlying continuous score used to categorize programs into the 5-tier system. The x-axis tick marks for the 5-tier score correspond to the cut-points used to categorize programs into tiers; e.g., programs with a 5-tier score above 0.875 were categorized as tier 5.



**Figure 3.** Lung transplantation. The association of the hazard ratio at listing and the 5-tier score at listing with eventual posttransplant 1-year lung graft survival. The dashed lines are the pointwise 95% confidence intervals, and the dotted line is a hazard ratio of 1. The 5-tier score is the underlying continuous score used to categorize programs into the 5-tier system. The x-axis tick marks for the 5-tier score correspond to the cut-points used to categorize programs into tiers; e.g., programs with a 5-tier score above 0.875 were categorized as tier 5.



**Figure 4.** Heart transplantation. The association of the hazard ratio at listing and the 5-tier score at listing with eventual posttransplant 1-year heart graft survival. The dashed lines are the pointwise 95% confidence intervals, and the dotted line is a hazard ratio of 1. The 5-tier score is the underlying continuous score used to categorize programs into the 5-tier system. The x-axis tick marks for the 5-tier score correspond to the cut-points used to categorize programs into tiers; e.g., programs with a 5-tier score above 0.875 were categorized as tier 5.