Seeking new answers to old questions about public reporting of transplant program performance in the United States

Bertram L. Kasiske¹,² | Andrew Wey¹ | Nicholas Salkowski¹ | David Zaun¹ | Cory R. Schaffhausen² | Ajay K. Israni¹,²,³ | Jon J. Snyder¹,³

¹Scientific Registry of Transplant Recipients, Hennepin Healthcare Research Institute, Minneapolis, MN, USA
²Department of Medicine, Hennepin Healthcare Systems, Minneapolis, MN, USA
³Department of Epidemiology and Community Health, University of Minnesota, Minneapolis, MN, USA

Correspondence: Bertram L. Kasiske, Scientific Registry of Transplant Recipients, Minneapolis Medical Research Foundation, Minneapolis, MN. Email: bkasiske@cdrg.org

Funding information
Health Resources and Services Administration, Grant/Award Number: HHS250201000018C

The Scientific Registry of Transplant Recipients (SRTR) is mandated by the National Organ Transplant Act, the Final Rule, and the SRTR contract with the Health Resources and Services Administration to report program-specific information on the performance of transplant programs. Following a consensus conference in 2012, SRTR developed a new version of the public website to improve public reporting of often complex metrics, including changing from a 3-tier to a 5-tier summary metric for first-year post-transplant survival. After its release in December 2016, the new presentation was moved to a “beta” website to allow collection of additional feedback. SRTR made further improvements and released a new beta website in May 2018. In response to feedback, SRTR added 5-tier summaries for standardized waitlist mortality and deceased donor transplant rate ratios, along with an indicator of which metric most affects survival after listing. Presentation of results was made more understandable with input from patients and families from surveys and focus groups. Room for improvement remains, including continuing to make the data more useful to patients, deciding what additional data elements should be collected to improve risk adjustment, and developing new metrics that better reflect outcomes most relevant to patients.

KEYWORDS
editorial/personal viewpoint, Organ Procurement and Transplantation Network (OPTN), organ transplantation in general, Scientific Registry for Transplant Recipients (SRTR)

INTRODUCTION

The Scientific Registry of Transplant Recipients (SRTR) is mandated by the National Organ Transplant Act, the Final Rule, and the SRTR contract with the Health Resources and Services Administration (HRSA) to report program-specific information on transplant program performance. Following many recommendations from a consensus conference in 2012,¹ SRTR made major changes to the program-specific reports (PSRs) published on the SRTR website every 6 months. In December 2012, SRTR converted the tabular data reports to a more graphical presentation with the goal of improving readability. In addition to the full PSRs, the public website had historically displayed a 3-tier summary of first-year graft survival, categorized as higher, lower, or as expected. Following a recommendation from the 2012 consensus conference that “PSRs should be better suited to the needs of all users, particularly patients,” SRTR launched a new website in December 2016 that displayed a 5-tier rating summary...
for first-year graft survival and included other changes designed
to make the information more understandable to patients and
the general public. The new website was generally well received
among patients, but has been controversial among transplant pro-
grams and professionals. For 18 months, SRTR collected feedback
from many sources and worked with the SRTR Visiting Committee
and with HRSA on a number of changes, including the develop-
ment of new 5-tier rating systems for pretransplant metrics. The
new version of the website was released on the beta SRTR web-
site in May 2018 (Figure 1).

2 | WHY DOES THE SCIENTIFIC REGISTRY
OF TRANSPLANT RECIPIENTS PRODUCE
PROGRAM-SPECIFIC REPORTS?

The National Organ Transplantation Act (NOTA; Public Law 98-507)
of 1984 provides the legal authority for the US Department of
Health and Human Services (HHS) to regulate solid organ transplanta-
tion in the United States. Regarding data collection, NOTA states:

“The Secretary shall, by grant or contract, develop
and maintain a scientific registry of the recipients
of organ transplants. The registry shall include such
information respecting patients and transplant pro-
ductures as the Secretary deems necessary to an ongo-
ing evaluation of the scientific and clinical status of
organ transplantation.”

The Organ Procurement and Transplantation Network (OPTN)
Final Rule operationalizes how HRSA of HHS enforces the provisions
of NOTA. The OPTN Final Rule states:

“121.11(b) Reporting requirements. (1) The OPTN and
the Scientific Registry, as appropriate, shall ... iv. Make
available to the public timely and accurate program-
specific information on the performance of transplant
programs. This shall include free dissemination over
the Internet, and shall be presented, explained, and
organized as necessary to understand, interpret, and
use the information accurately and efficiently. These
data shall be updated no less frequently than every
six months (or such longer period as the Secretary
determines would provide more useful information
to patients, their families, and their physicians), and
shall include risk-adjusted probabilities of receiving
a transplant or dying while awaiting a transplant,
risk-adjusted graft and patient survival following the
transplant, and risk-adjusted overall survival following listing..."

Task 3.9.1 of the contract from HRSA to SRTR states:

“The Contractor shall develop PSRs on the performance of transplant programs and organ procure- ment organizations (OPOs).

- The Contractor shall disseminate for free over the internet the timely and accurate program-specific information on the performance of transplant programs according to 121.11(b) of the OPTN Final Rule.
- The transplant program information shall include waitlist data, pretransplant outcomes, acceptance and utilization of organs, and posttransplant outcomes.
- Transplant programs and OPOs with better or worse outcomes shall be identified.”

3 | ARE THE DATA GOOD ENOUGH TO ASSESS THE PERFORMANCE OF PROGRAMS?

It is possible that collecting additional or better data would improve SRTR risk-prediction models. Studies have shown that data from other sources can improve risk-prediction models, but whether collecting these data would yield substantive changes in the PSR models or substantially affect expected program outcomes is unclear. In addition, because a data element can be added to a model does not mean it should be (for example, a debate in healthcare centers on whether analyses should be adjusted for socioeconomic variables, eg, race/ethnicity). Some argue that such adjustments allow providers to avoid taking extra steps to provide care that our most challenging patients require. OPTN has established a Data Advisory Committee (DAC), and one of its charges is to develop procedures for adding data elements to the OPTN data collection system.

4 | ARE THE STATISTICAL MODELS GOOD ENOUGH TO ASSESS PROGRAM PERFORMANCE?

A common criticism is that the C-statistics for models predicting posttransplant outcomes are low. However, the C-statistics for PSR models generally compare favorably to those for other healthcare outcomes prediction models. For survival models, the C-statistic measures how well the ranks of the predicted risks match the order of the observed events. A C-statistic of 1.00 implies that a model perfectly ordered the observed failure times with predicted survival. A C-statistic of 0.50 implies that the model-based survival predictions had no association with observed graft failure times. For example, in 1-year posttransplant survival, a C-statistic of 1.00 would imply that the model assigned less risk to every recipient whose graft survived to 1 year than to every recipient whose graft failed before 1 year. The C-statistics for the PSR models ranged from 0.66 to 0.76 for kidney models and from 0.67 to 0.83 for heart models. The C-statistic, however, is not necessarily a good way to assess how well models predict transplant program performance. To perform well, the models must predict the aggregate number of events for each program, not distinguish between pairs of patients.

SRTR builds the risk-adjustment models using a process designed to achieve the highest possible predictive performance given the data as measured by minimizing the cross-validation error. These models examine multiple variables collected in the OPTN data and attempt to identify the appropriate association of all predictors with posttransplant outcomes by considering flexible linear splines for continuous predictors.

5 | ARE 1-YEAR PATIENT AND GRAFT SURVIVAL ADEQUATE MEASURES OF PROGRAM PERFORMANCE?

There are several arguments against using outcomes with longer than 1 year of follow-up to assess program performance: (a) The risk of graft failure is highest in the first year. (b) The longer the follow-up, the more likely that practices at a transplant program have changed since the initial transplant; grafts that fail after several years may have been transplanted and treated by caregivers and protocols no longer active at the program. (c) The longer the time after transplant, the less control transplant programs may have over patient care, as patients may return home and be followed by local caregivers. Therefore, it may not be fair to hold programs accountable for the quality of care they do not provide. Nevertheless, SRTR also reports 3-year posttransplant outcomes, and continues to explore the possibilities and methods for reporting outcomes after even longer follow-up.

For most patients, undergoing transplant is more important for survival and quality of life than what happens after transplant. Receiving a kidney at a program with lower than expected posttransplant graft survival is usually better than remaining on the waiting list. Therefore, SRTR is adding new measures that reflect the likelihood that patients who are listed at a program will undergo transplant, so-called pretransplant metrics. In particular, the new beta website includes 5-tier assessments of waitlist mortality rates and transplant rates based on the standardized mortality rate ratio and standardized transplant rate ratio as presented in the full PSRs. Programs may rightly note that transplant rates depend on local supply of and demand for organs, and a program may have little direct control over these factors. However, a high transplant rate is a benefit to patients irrespective of the underlying reasons. Thus, prominently reporting this information is in patients’ interest.
WHY DOES SRTR USE TIERS AS A SUMMARY METRIC RATHER THAN OTHER STATISTICAL MEASURES?

The information on the SRTR website is designed to be accessible and understandable to patients and their families. SRTR has followed advice from the Agency for Healthcare Research and Quality (AHRQ) Best Practices in Public Reporting. AHRQ recommends using visual icons rather than numbers, and especially avoiding technical reporting with confidence intervals and other statistical measures. In addition to following the AHRQ recommendations, some SRTR investigators obtained a grant from AHRQ and have been conducting surveys and focus groups to determine optimal presentation of transplant program information. Prior to the use of 5 tiers, SRTR received many inquiries from patients asking for explanations as they struggled to interpret SRTR data reports.

WHY DID SRTR CHANGE FROM A 3-TIER RATING SYSTEM TO A 5-TIER RATING SYSTEM?

The old 3-tier system did not adequately differentiate transplant program performance. For example, in the January 2017 PSRs, the 3-tier system rated “as expected” 98.4% of heart, 93.6% of kidney, 97.6% of liver, and 97.0% of lung programs (Table 1). Under the 3-tier system, program graft failure rates varied almost 4-fold within the “as expected” tier. The 5-tier system better achieves the goals of the Final Rule by better differentiating program performance, ie, reducing the intra-tier variation in outcomes by 80% (reduced sum-of-squares).

ARE DIFFERENCES BETWEEN THE TIERS CLINICALLY RELEVANT?

What is or is not clinically relevant is subjective. However, for heart, kidney, liver, and lung programs, respectively, there are 3.4- and 3.3-fold differences in the rates of first-year graft failure between tier 1 and tier 5 programs (Table 2). Translating these relative risks to the absolute risk scale, predicted first-year graft survival for kidney transplant recipients ranges from 93% to 97% at tier 1 and tier 5 programs, respectively, a difference of 4 percentage points. For other organs, the differences on the absolute scale are more pronounced given higher overall failure rates. For example, the difference between tier 1 and tier 5 programs for predicted

---

**TABLE 1** Comparison of the former 3-tier with the new 5-tier outcomes rating system

<table>
<thead>
<tr>
<th>Organ</th>
<th>Tier 1</th>
<th>Tier 2</th>
<th>Tier 3</th>
<th>Tier 4</th>
<th>Tier 5</th>
</tr>
</thead>
<tbody>
<tr>
<td>Heart</td>
<td>8</td>
<td>16</td>
<td>44</td>
<td>47</td>
<td>8</td>
</tr>
<tr>
<td>Kidney</td>
<td>12</td>
<td>52</td>
<td>78</td>
<td>61</td>
<td>30</td>
</tr>
<tr>
<td>Liver</td>
<td>5</td>
<td>32</td>
<td>40</td>
<td>37</td>
<td>10</td>
</tr>
<tr>
<td>Lung</td>
<td>3</td>
<td>17</td>
<td>22</td>
<td>20</td>
<td>5</td>
</tr>
</tbody>
</table>

**TABLE 2** Different mean outcomes in each tier by type of transplant

<table>
<thead>
<tr>
<th>Type of transplant</th>
<th>Tier 1</th>
<th>Tier 2</th>
<th>Tier 3</th>
<th>Tier 4</th>
<th>Tier 5</th>
</tr>
</thead>
<tbody>
<tr>
<td>Waitlist survival (deaths per 100 y of waiting)</td>
<td>Kidney 7.4</td>
<td>6.3</td>
<td>5.3</td>
<td>4.6</td>
<td>3.3</td>
</tr>
<tr>
<td>Kidney-pancreas 13.3</td>
<td>9.5</td>
<td>7.0</td>
<td>5.0</td>
<td>3.0</td>
<td></td>
</tr>
<tr>
<td>Liver 21.7</td>
<td>17.4</td>
<td>14.5</td>
<td>13.0</td>
<td>9.5</td>
<td></td>
</tr>
<tr>
<td>Heart 18.6</td>
<td>15.9</td>
<td>11.9</td>
<td>9.8</td>
<td>6.8</td>
<td></td>
</tr>
<tr>
<td>Lung 31.5</td>
<td>24.5</td>
<td>18.5</td>
<td>14.0</td>
<td>12.3</td>
<td></td>
</tr>
<tr>
<td>Faster transplant (transplants per 100 y of waiting)</td>
<td>Kidney 5.2</td>
<td>8.7</td>
<td>12.7</td>
<td>18.0</td>
<td>30.4</td>
</tr>
<tr>
<td>Kidney-pancreas 18.1</td>
<td>29.8</td>
<td>45.0</td>
<td>75.8</td>
<td>110.6</td>
<td></td>
</tr>
<tr>
<td>Liver 16.2</td>
<td>32.4</td>
<td>48.1</td>
<td>72.3</td>
<td>130.7</td>
<td></td>
</tr>
<tr>
<td>Heart 34.7</td>
<td>48.7</td>
<td>74.9</td>
<td>105.3</td>
<td>156.6</td>
<td></td>
</tr>
<tr>
<td>Lung 50.5</td>
<td>100.5</td>
<td>167.6</td>
<td>244.8</td>
<td>460.3</td>
<td></td>
</tr>
<tr>
<td>1-y graft survival (percentage with functioning graft at 1 y)</td>
<td>Kidney 93.0</td>
<td>94.0</td>
<td>95.0</td>
<td>96.0</td>
<td>97.0</td>
</tr>
<tr>
<td>Kidney-pancreas 92.0</td>
<td>94.0</td>
<td>96.0</td>
<td>97.0</td>
<td>98.0</td>
<td></td>
</tr>
<tr>
<td>Liver 85.0</td>
<td>88.0</td>
<td>90.0</td>
<td>92.0</td>
<td>94.0</td>
<td></td>
</tr>
<tr>
<td>Heart 85.0</td>
<td>87.0</td>
<td>91.0</td>
<td>92.0</td>
<td>96.0</td>
<td></td>
</tr>
<tr>
<td>Lung 80.0</td>
<td>85.0</td>
<td>88.0</td>
<td>90.0</td>
<td>94.0</td>
<td></td>
</tr>
</tbody>
</table>
first-year liver survival is 9% (85% and 94% for tier 1 and tier 5 programs, respectively); differences are 11% (85% and 96%) for heart programs and 14% (80% and 94%) for lung programs. Although some differences may seem minor, when all else is equal, minimizing any risk of posttransplant complications is likely relevant to patients.

9 | DO PSR RESULTS CAUSE PROGRAMS TO AVOID HIGH-RISK DONORS AND PATIENTS?

Programs are not more likely to be ranked lower if they perform transplants with higher measured risk. Of course, programs that have poor outcomes with high-risk transplants should seek ways to improve these outcomes. But programs that are capable of performing higher-risk transplants can do so without fear of negatively affecting their performance evaluations. Nevertheless, an informal poll of transplant management personnel at a national meeting purported to show risk aversion caused by low PSR 3-tier performance ratings. Of 63 responders, 55% had had low or near-low performance ratings at their programs in the past 3 years, and personnel at low-performing programs were more likely to indicate that they restricted selection of candidates (81% vs 38%, P < .001) and donors (77% vs 31%, P < .001). Schold et al also reported that among 23% of transplant programs with at least 1 low PSR performance rating over a 3-year period, there was a mean decline of 22.4 transplants compared with an increase of 7.8 transplants among the other programs.

Overall, the number of transplants has been increasing in the United States despite concerns that reporting program outcomes causes programs to reduce their transplant numbers. Over the past several years, the number of kidneys retrieved for transplant that have not been transplanted (ie, have been “discarded”) has increased. In particular, the number of discards increased in the 2 years since implementation of the new kidney allocation system (KAS). Discards of kidneys with Kidney Donor Profile Index (K DPI) above 85% also increased more rapidly in the 2 years since KAS implementation. However, over 80% of the increase in the discard rate can be explained by changes in the donor population rather than changes in program acceptance practices. The proportion of transplants using high-KDPI kidneys declined from 10.7% in 2005% to 7.9% in 2016. Clearly, more should be done to discourage discards, many of which are likely suitable for transplant, and transplant programs should be instructed that PSR adjustment models protect their outcomes from adverse effects due to transplanting high-KDPI kidneys.

10 | DO CHANGES IN TIERS OVER TIME INDICATE THAT TIERS DO NOT PREDICT OUTCOMES?

In a recent brief communication, Schold et al examined changes over time in 5-tier program ratings based on 1-year graft survival. Due to a substantial lag between listing and transplant, the tier assigned to a program at listing may not accurately reflect survival after transplant. Over 10 consecutive PSRs from June 2012 to December 2016, these authors found that ratings were “highly volatile,” and that this implied that the ratings were of limited utility. However, they did not report how well the ratings actually predicted subsequent outcomes after transplant.

The 5-tier ratings were explicitly designed to better differentiate program outcomes by narrowing the variance in outcomes between programs within the same tier. In addition, with 4-tier boundaries rather than 2, it is not surprising that 5-tier ratings would be more likely to change over time than 3-tier ratings. In addition, if low ratings encourage programs to improve their outcomes, a change (improvement) in ratings would be an intended, desirable outcome. SRTR is currently examining how well the 5 tiers predicted posttransplant graft survival, arguably a better approach than examining how often programs move from 1 tier to another tier over time. The new version of the SRTR beta site also provides data indicating which outcome (waitlist mortality, transplant rate, or posttransplant graft survival) is most relevant to prospective patient survival after listing, and demonstrates that the ratings correlate with survival after listing.

11 | DOES SRTR TELL OVERSIGHT BODIES WHICH PROGRAMS NEED FURTHER SCRUTINY?

SRTR makes data available for everyone, restricting only identification of individual patients. The OPTN Membership and Professional Standards Committee (MPSC) and the Centers for Medicare & Medicaid Services use SRTR data, but decide independently which programs should undergo additional scrutiny. Insurance providers may use SRTR data to identify transplant programs of excellence for contracting, but they may use their own analysts to compare outcomes and “value.”

12 | WHAT CAN BE DONE TO FURTHER IMPROVE THE PROGRAM-SPECIFIC REPORTS?

SRTR investigators and others will continue to explore how reporting of transplant program outcomes can be made most understandable to patients and the general public. A long-term goal is to facilitate informed decision making about transplant care. This includes PSRs to communicate program performance metrics, and opportunities to place PSRs within a larger context of data relevant to patients and providers. SRTR is developing tools to understand potential waitlist outcomes based on individual patient characteristics, pilot programs to measure living donor outcomes, reports to assess offer acceptance behavior, and tools to understand the risks and benefits of individual donor offers. Over time, these tools can be refined and integrated to support the transplant community.

Ongoing assessment of what data elements should be included in OPTN data collection is necessary to optimize risk adjustment
in the PSR models.\textsuperscript{29} The DAC should work with OPTN's organ-specific committees and the rest of the transplant community to continually improve the OPTN database. New data elements should be added to adjust transplant risk in the PSR models in a way that discourages risk aversion. Old data elements that are no longer used should be eliminated. In addition, maintenance of the OPTN database should include methods to ensure the accuracy of data elements used in the PSRs.

There is discussion and debate about policies to remove selected transplants from the PSRs to discourage risk aversion. The Collaborative Innovation and Improvement Network (COIIN) is a 3-year study examining best practices to encourage transplanting moderate-to-high-risk deceased donor kidneys at select transplant programs (https://optn.transplant.hrsa.gov/resources/coiin/). COIIN removes participating programs from current performance review or “flagging” by OPTN’s MPSC.

New metrics to encourage transplant program and OPO quality improvement should be developed in conjunction with the transplant community. Measuring program acceptance of deceased donor organ offers, for example, may encourage programs to accept organs that are suitable for transplant and avoid unnecessary discards.\textsuperscript{16-18} Measuring survival after listing may encourage best practices in waitlist management.

Although the focus of this review has been on the PSRs, SRTR also produces reports on the 58 OPOs in the United States. Much work remains to better define eligible donors and efficiencies in organ procurement and utilization. Indeed, the growing rate of deceased donor organs procured for transplant but not transplanted (“discards”) is the subject of growing concern and scrutiny.

13 | SUMMARY

As mandated by NOTA, the Final Rule, and the SRTR contract with HRSA, SRTR has altered the way it publicly reports transplant program outcomes in the United States. It has sought feedback from the public for 18 months and has made changes as a result of this feedback. These include the addition of new pretransplant metrics, along with more detailed and hopefully more understandable explanations of the results that are reported. Nevertheless, work remains to improve the data and the reporting of program outcomes that matter to patients and their families. Patients, transplant centers, oversight agencies, third-party payers, OPTN/United Network for Organ Sharing, and SRTR should work together to develop systems that facilitate performance improvement and best outcomes while maximizing the number of transplants.

ACKNOWLEDGMENTS

This work was conducted under the auspices of the Hennepin Healthcare Research Institute, contractor for the Scientific Registry of Transplant Recipients, as a deliverable under contract no. HHSH250201000018C (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. The authors thank SRTR colleague Nan Booth, MSW, MPH, ELS, for manuscript editing.

DISCLOSURE

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

AUTHOR CONTRIBUTIONS

All authors participated in the conceptualization and writing of this manuscript.

REFERENCES


At ASTS, transplant is more than your CAREER. It’s your PASSION.

ASTS IS YOUR VOICE

01 Mingle with your peers at scientific meetings focused on transplant surgery

02 Have a voice in Washington and keep up with changes that affect your program

03 Access the transplant field’s best career resources, including the ASTS Compensation Survey

04 Keep transplant scientific research on the cutting edge with ASTS Research Grants

Enhance your career with leadership training and transplant-specific MOC credits

If you’re not a member, here’s what you’re missing!

Learn more at ASTS.org.

ASTS
American Society of Transplant Surgeons®