

Response to Beta Site Feedback

Beta Site Comment Period 5/14/2018 - 7/13/2018

Introduction

This document provides a summary of, and responses to, comments received by the SRTR during the formal comment period from 5/14/2018-7/13/2018 on changes to the SRTR's [beta website](#). SRTR extracted themes from the comments and this report contains responses to those themes, rather than responding to individual comments. The SRTR reviewed all comments with the SRTR's Visiting Committee (SVC) at its meeting on September 11, 2018. If the SVC recommended further changes to the site in response to the comments, those changes are detailed in this report. The full listing of comments is supplied in [Appendix A](#). Pending completion of the changes recommended by the SVC, the SVC voted to move the updated beta site to the main public SRTR website at which time the beta site will be retired until the SRTR wishes to pilot a new display of information. HRSA reviewed these changes and approved moving the updated beta site to the main SRTR website effective February 5, 2019.

SRTR wishes to thank all who took time to review and comment on the beta website. You may continue to provide feedback at any time by contacting us at SRTR@SRTR.org.

SRTR Responses to Feedback

- 1. Comment/Theme: Transplant programs should be compared only to programs within their region, not to the nation as a whole.**
 - a. Response: During development of the website, SRTR explored implementing a system that would allow the user to perform a "regional" search vs. a "national" search. Certain areas of the country may have a low transplant rate for reasons of supply/demand imbalances. In those regions, all programs may have a low transplant rate tier rating when comparing to a national standard, but if the tiers were recalculated based only on programs within a defined region, or with regional adjustment, differentiation between those programs may be more evident. This concept was presented to the SVC for consideration in January 2017, prior to the formal comment period. The SVC felt it important to continue to provide data based on national experience so that patients could understand that certain areas of the county may be more or less advantaged than others under the current allocation systems. The committee was supportive of continuing to explore the possibility of

allowing for region-specific search results. The SRTR brought the idea back to the committee at its July 2017 meeting at which time the committee recommended only presenting the tiers based on national experience rather than making results available with regional adjustment. The committee felt that transplant rates based on the national experience better informed patients about what their experience could be at programs within their region. The program-specific transplant rates are presented in the detailed program-specific reports that allow for comparisons of programs within specific regions. Finally, the committee felt that the website's current functionality allowing for only displaying search results within a certain state or radius from a user-supplied ZIP code also helped to facilitate regional searches.

- b. SRTR Action: SRTR is exploring a web application that would allow users to tailor search results to a region of their choosing. The current website does allow for regional searches (by state or a distance from a ZIP code), but the new application may allow for recalculation of tier assignments only based on programs within the search parameters.

2. Timeframes covered by the various metrics should be noted on the search results page.

- a. Response: SRTR has attempted to find a balance between presenting too much information on the page, risking cluttering the presentation, and presenting more detail for each of the metrics. We currently present phrases such as "in a year" on some column headers. We agree with the commenters that it is potentially helpful to know which metrics cover which timeframe. We currently have a "Click [here](#) to learn about the information provided on this page" allowing users to find more information about the data presented in the search results. We have added the specific cohort dates to that document.
- b. SRTR Action: We have added specific cohort dates to the explanatory document.

3. The combined deceased and living donor transplant rate should be eliminated from the website.

- a. Response: The transplant rate tier (labeled "getting a transplant faster") currently displays the deceased-donor only transplant rate for kidney and liver programs, censoring candidate follow-up at living donor transplant if applicable. In a previous version of the beta site, the transplant rate was presented as the raw all-donor transplant rate, including both deceased donor and living donor transplants in the numerator of the rate calculation. Following feedback received following the initial launch of the beta site in 2016, SRTR changed the basis for the tier metric from the all-donor to the deceased-donor-only transplant rate. This was recommended because the all-donor transplant rate can potentially mislead patients as to the wait time for a deceased donor organ. The counter argument is that program variability likely exists in how well programs assist their patients in finding a living donor. The

SVC debated this topic during the January and July 2017 meetings and recommended basing the tier rating on the deceased-donor only transplant rate, although both rate formulations contain potentially useful information. Ultimately, the committee felt that the all-donor transplant rate was potentially gameable based on when transplant programs place patients with a living donor on the transplant waitlist. Listing patients with an intended living donor early or late in the process can increase or decrease the all-donor transplant rate. While the search results now present a tier based on the deceased-donor only transplant rate, the program summary data page continues to display the raw all-donor transplant rate, and the full program-specific report for kidney and liver programs contain both formulations.

- b. SRTR Action: In response to this feedback, SRTR added the deceased-donor transplant rate to the program's summary data page, along with the all-donor transplant rate. The tier in the main search results continues to be the deceased-donor transplant rate as recommended by the SVC.

4. Living donation transplant rates should be highlighted on the site.

- a. Response: We received both written and verbal feedback that we should better highlight programs that perform a large number of living donor transplants. This version of the beta site attempted to make strides in this area by splitting the "transplant volume" column into both a deceased-donor and a living-donor transplant volume. This allows the user to see which programs do a large number of living donor transplants, and the sort order can be changed to sort large living donor volume programs to the top of the list. The feedback we received also pointed out that we currently state "getting a transplant faster is most important to patient survival" for both kidney and liver search results, whereas getting a living donor may be even more beneficial. We agree with this feedback and are continuing to explore alternatives to only including the living donor volume. One idea currently being explored is to present the living donor transplant rate based on the competing event formulation, i.e., the transplant rate censoring at deceased-donor transplant rather than at living-donor transplant.
- b. SRTR Action: Until a better alternative for living donor "rate" is developed, we will change the note to read "For kidney transplant candidates, this measure has the largest impact on survival after listing among these three measures."

5. Rates should be labeled "per person-year".

- a. Response: Earlier versions of the website did label the rates as "per 100 person-years." Patient focus groups lead by Dr. Cory Schaffhausen and Dr. Ajay Israni determined that patients did not understand what was meant by "per 100 person-years" and favored the language "per 100 years of waiting." While we agree this is not the technically correct terminology, we responded to patient feedback that this made the data easier to understand. The program summary infographic also

contains the language “per 100 people per year” which we have modified to make consistent with language in other parts of the site.

- b. SRTR Action: We modified the infographic to mirror the language on the search results page, so rates are displayed as “X.X people [are transplanted] [die] per 100 years of waiting [at this hospital] [nationally].”

6. A 5-tier system over-emphasizes 1-2% differences in outcomes across tiers, making differences appear larger than they are.

- a. Response: The previous version of the website did not provide direct context around what differences in tiers represented. In response to feedback that this would be helpful, we added the “Show National Rates” table above the search results. This table provides context as to what likely outcomes are for an average patient at an average program within the tier. For the July 2018 PSR release, this table contained the following range of outcomes:

	KIDNEY (TIER 1 – TIER 5)	LIVER (TIER 1 – TIER 5)	HEART (TIER 1 – TIER 5)	LUNG (TIER 1 – TIER 5)
TRANSPLANT RATE PER 100 PY	5.2 – 30.4 Range: 25.2	16.2 – 130.7 Range: 114.5	34.7 – 156.6 Range: 121.9	50.5 – 460.3 Range: 409.8
WAITLIST MORTALITY RATE PER 100 PY	7.4 – 3.3 Range: 4.1	21.7 – 9.5 Range: 12.2	18.6 – 6.8 Range: 11.8	31.5 – 12.3 Range: 19.2
FIRST-YEAR GRAFT SURVIVAL (%)	93 – 97 Range: 4	85 – 94 Range: 9	85 – 96 Range: 11	80 – 94 Range: 14

Smallest differences across tiers are seen within kidney program outcomes, with a range of approximately 4% from an average tier-1 program to an average tier-5 program. Actual outcomes for programs within those tiers vary even more because these values are estimated for an “average” program within the tier. While the “National Rates” table was meant to address this concern, it has also caused confusion as seen in the following comment.

The comment assumes that differences in survival probabilities are the preferred way to compare program performance. This is not appropriate because the range of predicted survival probabilities is a function of both the overall survival percentage and the variability in the relative event rates among programs. If every kidney program were to become more risk tolerant, the average event rate would increase and the range of predicted survival probabilities would widen. If every kidney program were to become more risk averse, the average event rate would decrease and the range of predicted survival probabilities would narrow. For these reasons, the range absolute differences in program outcomes is confounded by the overall

survival percentage and may not provide information on the magnitude of differences between programs.

This pattern exists when applied to an individual program, as well. If a single kidney program with lower graft survival than expected were to become more risk averse, the difference between its observed and expected graft survival percentages would decrease, but that would not indicate that the program had improved. The SRTR evaluates programs based on their relative event rates precisely because this approach is not affected by the overall graft survival percentage, and program evaluations are risk adjusted to minimize the impact of differences in program risk tolerance on program evaluations. The variability in relative event rates for kidney programs is comparable to the variability observed for other organs:

	KIDNEY (TIER 1, TIER 5)	LIVER (TIER 1 – TIER 5)	HEART (TIER 1 – TIER 5)	LUNG (TIER 1 – TIER 5)
RELATIVE TRANSPLANT RATE	0.42, 2.47	0.34, 2.75	0.49, 2.22	0.32, 2.93
RELATIVE WAITLIST MORTALITY RATE	1.38, 0.61	1.48, 0.65	1.60, 0.58	1.77, 0.69
FIRST-YEAR RELATIVE RISK OF GRAFT FAILURE	1.60, 0.53	1.58, 0.58	1.77, 0.47	1.79, 0.51

Here we see that first-year graft failure rates are approximately doubled in tier 1 programs and approximately halved in tier 5 programs relative to expected outcomes based on national experience.

The comment also assumes that a difference of 1-2% is ignorable. The SRTR also disagrees with this assumption. In many medical settings, less severe events are often more frequent than more severe events, yet more severe events are rarely considered to be unimportant. A kidney graft failure is a severe event with a significant impact on a recipient’s quality of life and risk of death. Precisely because graft failures are severe events, a difference of 1% (or perhaps even less) in the risk of graft failure is a meaningful difference. In addition, when considering the transplant system as a whole, if improvements to program quality were able to achieve a 1% increase in the average graft survival rate, that improvement would be a substantial achievement. A program that is able to achieve a higher graft survival percentage of 1-2% without becoming more risk averse is a program that ought to be recognized with a higher tier.

b. SRTR Action: No modifications made at this time.

7. Comparison of observed to expected rather than comparing observed to the national rate is confusing. The National Rates table has added some confusion to the search results.

- a. Response: This comment represents a fundamental misunderstanding of how the metrics are computed. Metrics are formed as a comparison between observed events and expected events based on case-mix at the program rather than a comparison to national averages. We believe this confusion is exacerbated by the inclusion of the "National Rates" table above the search results. This table seems to have caused some confusion because users are looking at the tier assignment, looking up the "estimated survival for an average patient transplanted at an average program within the tier" as shown in the National Rates table above the search results, and then comparing that value to the program's actual survival outcomes (or waitlist mortality or transplant rate). Some users have then been confused because their actual survival percentage exceeds that of the average tier 5 program, but the program only achieved a tier 4 rating. This can happen for a number of reasons, including small volume or a lower risk patient population. We added the National Rates table above the search results in response to feedback that we needed to provide context around expected outcomes within the tiers. This was recommended specifically in response to feedback similar to that in comment #6 above that some tier differences represent small differences in absolute failure rates. Because the National Rates table seems to have created some confusion while at the same time satisfying prior feedback, we attempted to enhance the educational material explaining the National Rates table.
- b. SRTR Action: We created an educational page explaining in more detail how the national rates were estimated and how users should interpret the table. A separate educational page was created for each organ type. Furthermore, in the accompanying educational material, we included performance ranges within each cell of the table in addition to the predicted value for an average patient at an average program within the tier.

8. Distance for searches should be more granular.

- a. Response: Following feedback received from the first iteration of the beta site launched in 2016, we added more options for the circle radii, which now include 50, 100, 250, 500, and 750 miles. The commenter specifically requested an additional circle between 100 and 250. We believe the current choices represent a good number of choices and note that the site allows the user to sort the results by the distance from the ZIP code, which is the default sort if the user chooses to search by a ZIP code.
- b. SRTR Action: No modifications made at this time.

9. Colors of the tier icons should be the same rather than different for different tiers.

- a. Response: The AHRQ guidelines for best practices in public reporting of healthcare quality data specifically recommend using color to aid the user in discerning differences across programs.¹ We use brighter colors for higher-rated programs. Dr. Schaffhausen and Dr. Israni piloted versions of the site to patient focus groups that used varying colors or constant colors. Feedback was mixed, with some patients preferring the changing colors, and some preferring one color. Given we have heard opinions on both side of this issue and have not heard clear consensus, we have chosen to stick with the AHRQ recommendation to use color to aid the user in visualizing top-performing programs.
- b. SRTR Action: No modifications made at this time.

10. Programs with 100% survival should be rated a 5-tier program.

- a. Response: The tier system is based on a comparison of observed outcomes to expected outcomes after adjustment for case-mix at the program. The methodology used to assign a tier is described in detail [here](#). In brief, the tier methodology accounts for both the observed. vs. expected outcomes at the program and the amount of information we have based on the number of transplants performed. For example, it would not be unusual for an average kidney program performing 10 transplants to have 100% survival given national first-year success rates of approximately 95%. Given 95% success rates nationally, we may not expect to see a failure until the program transplants at least 14 patients. Therefore, the tier system may show a 4-tier rating, which is somewhat better than expected, but we do not have enough information to conclude the program is truly a tier 5 program. More successful transplants would need to be done to boost the program into the tier 5 rating. This also works in the other direction such that we need stronger information to conclude a program is a tier 1 program.
- b. SRTR action: Added more educational material about this within the guide to the rating system.

11. Transplant rates are not meaningful for intestinal transplant candidates and outcomes assessments should be performed for intestinal transplant recipients

- a. Response: SRTR will be implementing outcomes assessments for all transplant recipients, including intestinal transplant recipients. Historically, the SRTR has not performed outcomes assessments if the data were not adequate to build a risk adjustment model. For the reasons the commenter has highlighted, the SRTR debated this with the SVC and concluded that it would be preferred to include outcomes assessments for all transplant recipients, even if the outcomes were unadjusted and simply based on national success rates. SRTR is in the process of implementing this change. The result will be to include intestinal transplant outcome assessments on the search results next to the waitlist mortality tier and the

transplant rate tier. While the commenter may be correct that the transplant rate is perhaps a different construct in intestinal transplantation because there is no shortage of organs, we believe it still represents a valid construct from the patient's perspective because it gives a sense for how quickly patients at the program receive a transplant compared with other programs. We believe the outcomes assessment will provide the context the commenter is advocating for.

- b. SRTR Action: SRTR is working to implement the outcomes assessments for all organ types and will include these on the website as they become available.

12. Inclusion of waitlist mortality is misleading because centers have no control over waitlist mortality.

- a. Response: This comment is generally made in reference to kidney programs, which tend to be unique in that patients are often cared for by community nephrologists while on the waiting list rather than by physicians within the transplant program. The SVC debated this issue at the July 2017 meeting and recommended Dr. Schaffhausen and Dr. Israni study the issue with the patient focus groups they are conducting and report back to the committee. Dr. Schaffhausen and Dr. Israni conducted focus groups with kidney recipients and specifically asked the focus group participants whether or not they would like to see a tier rating for waitlist mortality. Results were mixed. Some patients indicated they would find it valuable, and others indicated they would not consider it. The SVC considered the findings of the focus group and heard Dr. Schaffhausen's recommendation that there did not appear to be a strong argument to exclude it for kidney programs. The committee did not disagree that kidney programs often are not directly caring for patients on their waiting list and recognized that kidney programs may object to the metric, but felt that the waiting list mortality metric was a valuable construct of what happens to patients that are listed at the program and worthy of showing to patients. Therefore, the beta site as it went out for public comment continued to display the waitlist mortality tier for kidney programs. In response to continue feedback on the issue, the SVC again debated the issue at its meeting on September 11, 2018. At this time the committee voted to remove the waitlist mortality tier for kidney transplant programs from the search results. The waitlist mortality data can still be found on the program summary page and in the full PSR, but has been removed from the search results.
- b. SRTR Action: SRTR has removed the waitlist mortality tier from the website for kidney transplant programs.

Specific Responses to the AST/ASTS Metrics Task Force White Paper

Here we paraphrase key themes from the ASTS/AST Metrics Task Force White Paper which was submitted to the SRTR as formal feedback by the AST and the ASTS.

1. Unintended consequences of public reporting are detrimental to the field, resulting in reduced organ supply, limiting patient access, and creating barriers to innovation.
 - Response: This is a critique of performing any program evaluations and not a specific critique of the current presentation of the metrics on the SRTR website. SRTR is contractually obligated to report various performance metrics publicly.
 - SRTR Action: No modifications made at this time.
2. Concerns about risk adjustment, data validity, and poor predictive performance of SRTR risk adjustment models, i.e., low c statistics.
 - Response: SRTR uses a sophisticated statistical approach to building risk adjustment models which has been published.² SRTR continues to advocate for more accurate and more relevant data collection. SRTR has recently conducted a systematic literature review in partnership with the Evidence Review Practice Center at the Mayo Clinic in Rochester, MN. Those findings have been provided to the OPTN's Data Advisory Committee. In addition, the SRTR has presented findings of data quality and data governance to the Data Advisory Committee and has advocated for changes to be made to improve the data collection. In addition, SRTR has recently performed an analysis showing that c statistics are not a valid way to judge the predictive ability of a risk adjustment model, building on previously published research.³ Those findings have been accepted for publication in the American Journal of Transplantation.
3. Medicare conditions of participation (CoPs) and private payer Centers of Excellence hinge on these data
 - Response: SRTR is charged with publicly reporting transplant program evaluations. We do this to the best of our ability given the data provided and are attempting to present the data in a manner that is accessible to multiple audiences. Medicare has recently announced a plan to discontinue ongoing monitoring of transplant program performance following initial approval of the program.
4. Patients have little/no choice of programs
 - Response: It may be true that certain patients have limited choice due to various factors, including payer restrictions or inability to travel. It is the opinion of the SRTR that this does not constitute a valid reason to not publish the data as we are contracted. Dr. Schaffhausen conducted a survey of patients in cooperation with the National Kidney Foundation (NKF), Transplant Families (TF), and Transplant Recipients International Organization (TRIO), and 73% of combined respondents indicated they would consider 2 or more transplant programs, with 83% of kidney patients indicating they would consider 2 or more programs.

	NKF	TF	TRIO
NUMBER OF PROGRAMS REASONABLE TO CONSIDER; % (N) *	342	26	34
1 CENTER	17% (59)	46% (12)	29% (10)
2-3 CENTERS	62% (211)	42% (11)	56% (19)
4-5 CENTERS	9% (30)	4% (1)	6% (2)
5 OR MORE CENTERS	6% (20)	4% (1)	0% (0)

5. High rates of false positive and false negative flags (referring specifically to CMS CoPs).
 - Response: This is a critique of the CMS flagging algorithm which is not used by the SRTR nor the OPTN, and CMS has recently announced a plan to stop ongoing monitoring of transplant outcomes. The SRTR recommended that the OPTN move away from the CMS flagging algorithm in favor of a screening algorithm based on the new Bayesian performance assessments being produced by the SRTR. The Bayesian system improved the false positive rate, holding it at approximately 5% across the range of program volumes, while also improving the true positive rate of the screening algorithm. The screening algorithm is owned by the Membership and Professional Standards Committee (MPSC) of the OPTN. SRTR can help to modify the algorithm if the MPSC feels the false positive rate is unacceptably high. Whether or not a program is flagged by the MPSC or the CMS algorithm is not presented on the public website. In addition, the tier system is not based on the flagging algorithm used by either CMS or MPSC.
6. Little difference in failure rates across programs (citing kidney outcomes on the absolute survival scale)
 - Response: See response above (in the general response section) to comment #6.
7. Need to include patient-reported outcomes
 - Response: Inclusion of patient-reported outcomes has not been considered by the SRTR to-date and we are not aware of available data sources. We welcome the opportunity to discuss the idea with the AST, ASTS, and the transplant community to explore ways of including patient-reported outcomes such as quality of life evaluations or patient satisfaction surveys.
8. Programs may not be responsible for graft failures or deaths
 - Response: It is true that cause of death and reported cause of graft failure are not considered when the SRTR performs the evaluations. Therefore, patients dying in a

traffic accident are included alongside patients dying more directly as a result of a failed allograft. SRTR cannot adjudicate which deaths or graft failure were or were not directly related to the care provided by the program.

In addition to the themes above, the white paper made a number of specific recommendations:

1. Rank order performance
 - Response: The website rank orders by performance rather than the alphabetical ranking used in the previous version of the SRTR website (prior to 2016). In addition, the website bases the performance ranking on the metric with the largest impact on a patient's overall survival following listing. Users can choose to rank order search results based on any available metric of their choosing.
2. Use symbols rather than numbers
 - Response: The website includes icons that convey meaning through the symbol and the color as recommended in the AHRQ best practices document. In addition, the new version of the website removed the numerical transplant rate in favor of an icon.
3. Provide an overall summary measure
 - Response: SRTR is working on a summary measure of overall survival following listing and hope to be able to roll this metric out in the future.
4. Include fewer reporting categories
 - Response: This recommendation was made in the AHRQ best practices report specifically with the example to favor 5 reporting categories over a more detailed system that would include, e.g., 9 categories. We attempted to strike a balance between too few categories that do not do a good job of stratifying performance within each category, to having too many categories. The tier system substantially reduces intra-tier variability in comparison to the previous 3-tier system.⁴ Therefore, patients can be more assured that programs within a given tier assignment are similar to each other.
5. Combine data over multiple years to increase precision
 - Response: The waitlist mortality tier and the transplant rate tier use data over a recent two-year period, and the outcomes evaluations use data over a recent 2.5-year period. We must strike a balance between using evaluation windows that are too narrow and therefore lack statistical precision, with those that are too wide and possibly contain irrelevant data from years past. Transplant rate estimates are very precise given the relatively large number of transplant events compared with deaths or graft failures, which are less frequent. The tier assignment methodology takes into account how certain we are of each estimate.
6. Aggregate data from different measures

- Response: The SRTR attempts to present metrics that target different aspects of patient experience and are therefore not necessarily correlated with each other. For example, programs that have low waitlist mortality rates do not necessarily also have low post transplant mortality rates. If metrics were highly correlated, there would be limited value in presenting more than one metric since they would all point in the same direction. Because the three metrics we currently present (waitlist mortality, transplant rate, and posttransplant outcomes) are measuring different aspects of the patient experience at the program and they are not highly correlated, we believe there is value in presenting the metrics separately rather than combining them into a composite or aggregate measure. Composite measure necessarily require decisions on how much weight to give each component metric. Furthermore, composite metrics inherently excuse low performance on one metric if performance is high on another metric. Given the metrics are measuring different aspects of the experience at the program, it is the opinion of the SRTR that a composite metric is not useful. However, we are pursuing a metric that attempts to encompass the patients' full experience at the program: overall survival from listing. This metric is under development and is being considered by the SVC. While this is not a composite metric by definition, we believe it does capture the nature of what many commenters have argued for: a metric that captures how likely a patient is to be alive following listing at the program.
 - SRTR Action: Continued development of a metric measuring overall survival from listing.
7. Incorporate patient reported outcomes:
 - Response: See response above
 8. Incorporate process measures:
 - Response: SRTR currently does not have access to process measures. Some argue that process measures are only good for public reporting if they improve patient outcomes, which we can and do measure directly.
 9. Better reimbursement
 - Response: The paper argues for increased resources for programs to accommodate improved, and perhaps more, data collection to support better assessments. SRTR cannot influence these decisions directly, but encourage all efforts that result in better data not only to improve metrics, but to improve the field of transplantation as a whole.
 10. Better collection of important data
 - Response: See response above with regards to risk adjustment and c statistics.
 11. Work more closely with the transplant stakeholders when developing these reports
 - Response: The recent initiative to improve the public reports began in early 2012 when SRTR hosted a consensus conference to engage the community in improving

the reports. This was an effort to engage the community, and indeed many of the changes were driven by consensus recommendations stemming from the conference.⁵ SRTR has been developing the new version of the website since the conclusion of the consensus conference under the guidance of the SVC. SRTR has also greatly bolstered its communications strategy through targeted newsletters to programs and all societies and a strong social media presence, all with the goal of better informing the community of developments within the SRTR. We will continue to explore ways to better engage the transplant community and look forward to partnering with the professional societies in continued development whenever possible.

¹ Hibbard J, Sofaer S. Best Practices in Public Reporting No. 1: How To Effectively Present Health Care Performance Data To Consumers.

² Snyder JJ, Salkowski N, Kim SJ, Zaun D, Xiong H, Israni AK, Kasiske BL. Developing statistical models to assess transplant outcomes using national registries: The process in the United States Transplantation. 2016;100:288-294.

³ Austin PC, Reeves MJ. The relationship between the c-statistic of a risk-adjustment model and the accuracy of hospital report cards: A Monte Carlo study. *Med Care*. 2013 Mar; 51(3): 275–284. doi: 10.1097/MLR.0b013e31827ff0dc.

⁴ 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;53:1979-1991.

⁵ Kasiske BL, McBride MA, Cornell DL, et al. Report of a consensus conference on transplant program quality and surveillance *Am J Transplant*. 2012 Aug;12(8):1988-96.

Appendix A: Feedback Received 5/14/2018 – 7/13/2018

Feedback was received primarily through comments provided through the SRTR website "[Contact Us](#)" form. SRTR staff kept a running log of all feedback received and collated feedback into positive, negative, or neutral. The SRTR Visiting Committee had an initial discussion about the feedback received at its regularly-scheduled meeting on 7/16/2018. Here we present the themes received and SRTR responses.

Positive Feedback

Positive feedback generally fell into the following themes:

- User was grateful for the content
- Easier to read/understand than before
- Presenting this tiered system and the supporting data is valuable

Specific quotations from positive feedback included:

- "When I logged on I saw that there were some new metrics and some beautiful design elements that I hadn't seen before. Plus I was able to pull the data that I needed to provide the context I needed."
- "You guys are doing wonderful job. Because of this database I was able to help my brother locate a transplant center that has shortest wait-time. He got his transplant in a month after being listed. God bless you guys!"
- "I just perused the existing and beta sites and am pleased with the changes"
- "I like the changes you have made to the website...it looks good and is easy to read."
- "What an incredible compilation of information! I haven't read through all the different areas yet, and I know this stuff changes quite a bit, but how great it would have been to have this around 5 or 10 years ago! Your team is amazing! Thank you for everything you do to help guide people through this crazy system."
- "I was asked a question by a clinician today and I didn't have quite the right [Hospital X]-specific data. I realized that the beta site might be useful. When I logged on I saw that there were some new metrics and some beautiful design elements that I hadn't seen before. Plus I was able to pull the data that I needed to provide the context I needed. Thank you for all the work you've put onto the beta site. I look forward to seeing it used more widely."
- "You guys are doing wonderful job."
- "The SRTR's 5-tier system Beta website highlights the performance data and benchmarks our leadership not only requires but also values."

Negative Feedback

Negative feedback generally fell into the following themes:

- Concern over how transplant rate is explained for patients' level of understanding
- Confusion over the National Rates table meaning
- Misunderstanding the metric for the tiered assessments

Specific negative comments included (note that some are paraphrased if feedback was received via telephone):

- "The 'access to transplant' information is not helpful for patients. You compare programs to national data. The vast majority of dialysis patients can't access national programs due to multiple logistical and financial constraints. Most patients are limited to their immediate region. To be helpful to patients, programs should be compared to other transplant programs in their region. This is the only way that patients can reasonably compare their options for listing. National data has no practical utility for patients. If the SRTR's mission is to help patients make useful choices, the comparisons should be made that actually reflect the patient's realistic possible decisions."
- "On Beta Site, as someone deciding which facility to go to, the bars do not tell me if these are percentages, number of people or what. Website is cumbersome. There needs to be a distance somewhere between 100 and 250 (too wide). I really don't think someone who just found out they had a fatal disease would take the time to figure this out."
- "Our transplant program leadership team want to let the SRTR team know that we feel strongly that the transplant rate most visible to the public needs fixing. We are wanting your advice on getting this changed and willing to discuss in person, file a complaint, or do whatever we need to do, please let us know. We greatly appreciate the addition of transplant rates that only include deceased donor transplant, which is by far a more accurate picture of the time to transplant rate at programs, however, it is buried in the data. The SRTR continues to display the combined rate (deceased and living donor rate) which still misleads the public and payers. It is unfair and bias to programs who have cultural or financial barriers, thus not as many living donors. We are aware that the SRTR may have experienced pressure from centers who are benefiting from this, however, when looking at deceased transplant rates only, the data leads to a very different picture. We are still experiencing patients being told they should go somewhere else by their payers based on this data. It is a burden for patients to re-locate for transplant, pay for housing, etc., when in reality it is no benefit and based on poor information. We feel strongly that this should be amended further by removing the combined time to transplant and ask that you bring this problem forward as a very strong complaint in the transplant community. We know the SRTR team wants to do the right thing and that this data is meant to represent truth so that informed decisions can be made."

Continuing to display the time to transplant rate that includes living donor cases continues to place programs with underserved patients at a disadvantage. Thanks for your help and advice on this!"

- "The Beta Website, while improved, still has a major issue in my mind, and that is the transplant rate. I recognize that it is based on person-years, so someone who is listed and transplanted in a few weeks doesn't contribute to a person-year. However, the current reporting doesn't say person-years per se, and to the average reader, it seems non-sensical to have more than 100 out of 100 people per year receive a transplant. There needs to be a way to provide this in some understandable format, because "people per year" is not person-years, and this will very much confuse people."
- "For center outcomes, differences in graft survival between the different tiers are generally 1-2% yet a scale of only 5 tiers implies much larger differences between programs. The scale that is presented should be one that reflects the differences more precisely"
- "The question from our providers upon seeing the 5 tier assessment is clarity on the time frame. The summarized PSR page does a good job of showing the time frame associated with the data. Currently the 5-tier do not show the same level of clarity."
- "I am a practicing nephrologist and I reviewed this site as a patient, so I could assess it before recommending to patients. I typed in my home zipcode, and the organ of interest, "Kidney", and then looked at centers within 50 miles. I think the site will be very confusing to patients. They are going to look at the national data as normative, thus 3 bars for a given center will be viewed as being the stated national average. However, if you look at the data for the programs, 3 bars for a given outcome does not track to the 3 bar national average. If you look at the program at my hospital ([Hospital X]), there is 1 bar for survival per 100 pt years-- the national average is 5.3 deaths/100 patient years. [Hospital X] is given as having a 3.5 deaths/100 patient years, in the detailed description of outcomes. It has been explained that [Hospital X] would have an expected outcome much better than the national average (or its present outcome), due to the demographics of those on the list, and that is why it is given 1 bar. This is going to be totally confusing to patients--I'm confused myself. Are you including Status 7 patients? If you look at the rate of transplant, the national average is 12.7 transplants/100 pt years (3 bars), but on the detailed information set, you indicate the national average is 18/100 pts years. [Hospital X] has 21.7/100 patient years, and gets 2 bars--I have no idea why. It should get at least 3. I think you need to leave off the national averages, if you are not using them to norm the data. And then, you need to explain, in 8th grade language, how you developed the ratings, and what the bars mean."
- "[Hospital X] has 100% graft survival for LD recipients but on the summary data section they have a 4 bar rating instead of a 5... If volume is the reason for the 4 bar ranking instead of the 5 bar ranking, maybe there should be a note that states this program has 100% graft survival and would be rated at 5 bars but isn't, because of lower volume."

- “I think the key describing the meaning of the icons is great, but providing the national numbers in this manner is confusing. It can easily be interpreted to mean that the center with 4 tier rating has a 1 year kidney survival of 96%, rather than understanding it as an expected outcome for an average patient at a typical program within the tier. Words like "expected", "average", "typical" are not easy to interpret/understand. I don't think the national numbers should be provided. Instead a button under each program's metrics that gives the exact number for the program will be helpful. National context is innate to the tier system.”
- Transplant rates are not understood [paraphrased]
- Explanation for “bar” (assessment) metrics were not easily found [paraphrased]
- "I appreciate the SRTR for allowing a comment period which provides member centers a voice on important issues pertaining to their field. 1) The usage of the phrase "getting a transplant faster" for the intestine graft is a misleading metric to the general public and should not be utilized for intestinal transplant recipients. There are two major and unique reasons why this should not be utilized for intestinal transplantations. a. This metric traditionally is used for patients to evaluate where they can get organs faster to potentially double-list for difficult to get organs, such as liver graft which currently have a predominantly regional sharing. In the case of evaluating a center's ability to "get organs faster," this metric makes sense for patients to evaluate liver transplant centers as there are wide disparities of organ donation among regions in the country for liver grafts. As we all know, there is a significant shortage of liver grafts in the country and knowing a center's ability to get liver grafts is an important piece of data to know in making a decision on where to list for transplant. In the case of intestinal transplants, there is a national listing and distribution whereby intestinal transplant candidates will come up on the match run on virtually every organ offer no matter of the location of the donor. In 2017 there were 2832 (liver) donors age less than 50 (the generally upper limit of acceptable intestine donor age) in the entire country. Assuming every one of these donors had an accompanying intestine graft, and also acknowledging that there were only 112 intestine grafts utilized, that means that 2,270 intestine grafts were discarded. There is no shortage of intestine grafts in the country. b. Because there is no shortage of intestine grafts, the main consideration for a center's decision on whether to utilize an intestine graft is based on multiple factors such as: HLA typing of the donor, CMV status, and donor size. HLA matching is a critically important component in intestine donor selection. Data have demonstrated the poor outcome of positive crossmatch intestinal transplant cases. With the use of virtual crossmatching, a negative crossmatch can be reliably predicted. If a center is highly selective on HLA pairing so as to avoid any potential of a positive crossmatch to obtain a superior outcome, naturally the center's ability to "get organs faster" will be prolonged. This, however, is not to the detriment to the patient, but rather a substantial benefit. Patients with high PRA will most likely be listed for a longer time, however, that longer wait will be offset by better outcomes

in the long term. Again, there is no shortage of intestine grafts, so one has the luxury of being highly selective in order to obtain better outcomes. CMV matching is also crucial for improved outcome in intestine transplantation. Data supports not performing high risk CMV positive donors to CMV negative recipients in intestinal transplantation. Since 60%--70% of the population is CMV positive, CMV negative intestinal transplant candidates will most likely have a longer wait time to transplant, reflecting negatively on the center's ability to "get organs faster." The same principals apply to donor size as a vast majority of intestinal recipients have short gut, and hence loss of abdominal domain. Smaller grafts are needed for a better outcome, potentially increasing the time to wait for a suitable intestine graft. Again, in order to obtain the perfect fit for an intestine graft, the center's ability to "get organs faster" may be prolonged. In summary, the metric "get a transplant faster" is well-intended and well-suited for patients evaluating transplant centers for other organs, but perhaps could be misleading for intestinal transplantation. There is no shortage of intestine grafts. A center could have an extremely favorable rating on its ability to get a transplant faster and perform a positive crossmatch, high risk CMV pair, poorly size matched (open abdomen) intestinal transplant in order to obtain a more favorable "rating," but be actually doing the patient a colossal disservice. There is a potential for significant unintended consequences by relying on this metric in intestinal transplantation. 2) Outcome Data for intestinal transplant is not readily presented in the SRTR center analysis but rather stated as "Not Assessed." It is a travesty that SRTR is not placing emphasis on the basic tenet of a center's analysis - that is, the quality of the program. Each center has graft and patient survival data available for both pediatric and adult cases. Because there are so few cases performed, I understand there is no acceptable "standard" for intestinal transplant outcome. Regardless, patients need to be informed in an easy obtainable manner, (i.e. the SRTR website) the center's outcome when comparing transplant centers. SRTR should make this data available for patients to truly evaluate a center's performance in an easily accessible and viewable manner. Perhaps patients would discover that centers with a favorable rating on "get a transplant faster" would have inferior outcomes for all the reasons I stated above. It is imperative that this obviously basic, simple and available data is published side by side with 'getting a transplant faster.'

- "The new 5-tier assessments for survival on the waitlist, getting a deceased-donor transplant faster, and 1-year organ survival are misleading. Regarding survival on the waitlist, transplant programs have little control over waitlist mortality, unless a program prioritizes removing candidates over preparing candidates for transplant. In the case of kidney transplant, programs don't directly care for patients on dialysis, therefore, kidney waitlist mortality is even less a measure of a transplant program's success. Due to geographic disparities, transplant programs also have little to no control over the speed at which a candidate can receive a deceased-donor transplant. Instead, speed of transplantation off the waitlist is almost completely influenced by Donor Service Area. Rating individual programs

on something they cannot control is unfair at best, and again, significantly misleading to the public. Donor acceptance rate is a more relevant metric of a program's performance that should be made clear to potential transplant candidates. In addition, while we applaud the SRTR for publishing live donor volume for kidney and liver transplantation, the number has no context. UNOS has a public goal to increase live donor kidney transplants. If the goal is to increase the number of transplanted patients, there should be a metric to drive patients to pursuing live donor transplant wherever possible, such as a comparison to the national average for ratio of live donor transplants to waitlist additions and kidney paired donations to waitlist additions. Clearly ranking programs based on live donor volume would also provide further context to the public. Finally, the five-tier metric for 1-year organ survival has an insignificant margin between bars, yet allows programs to be ranked against each other without disclosure of the small margins other than a comparison to national rates, which is hidden until users expand the comparison on their own. We appreciate the effort SRTR has made to educate the national community about program-specific information, but the changes noted above have limited context, will mislead the public, and will not fairly represent many of the transplant programs in the country."

Recommendations for potential changes/improvements:

Here we list some specific comments and questions we received. We list them here because, while not necessarily positive or negative feedback, they do point to areas of potential improvement.

- "For the program summary at the top of a center's page and for the summary that appears in transplant center search results, we think it would be beneficial to clarify the time frames. Deceased and living donor transplants represent one time frame, and the survival data and transplant rate represent a different range of dates."
- "The question from our providers upon seeing the 5 tier assessment is clarity on the timeframe. The summarized PSR page does a good job of showing the timeframe associated with the data. Currently the 5-tier do not show the same level of clarity."
- "Education in the option of living kidney donation, especially paired donation, by also publishing living donation rates on the SRTR website. Likewise in the same spirit the annotated blue informational box on the current Beta website page should be edited from: 'For kidney transplant candidates, this measure has the largest impact on survival after listing' to: 'For deceased donor kidney transplant recipients, this measure has the largest impact on survival after listing.'"
- "Our kidney transplant program leadership appreciates the opportunity to review and provide feedback for the updated live SRTR Beta website and the published 5- tier system. Our program has one of the highest proportions of living kidney donation in the nation. We believe it is important looking forward to increase public awareness and patient education in

the option of living kidney donation, especially paired donation, by also publishing living donation rates on the SRTR website. Currently ~75% of our candidates eventually receive a living donor kidney transplant and we believe many programs adopting living donation best practices will someday achieve similar rates of living kidney donation. In the current era where the deceased kidney transplant rates remains largely unchanged our kidney transplant community will make progress in decreasing the constantly growing waitlist though increasing living kidney donation. The SRTR's 5- tier system Beta website highlights the performance data and benchmarks our SRTR leadership not only requires but also values. "Getting a living donor transplant faster" is currently an absent metric we believe needs to be transparently included in the 5-tier system page rather than have it truly buried within the many pages of programs' "summary data" and "complete report pdfs". Likewise in the same spirit the annotated blue informational box on the current Beta website page should be edited from: 'For kidney transplant candidates, this measure has the largest impact on survival after listing' to: 'For deceased donor kidney transplant recipients, this measure has the largest impact on survival after listing.'"

- "There needs to be a distance somewhere between 100 and 250 (too wide)."
- "The bars should not be different colored. I thought the difference in color represented something and was trying to look for a key on what the color meant."
- Some users wanted an easy way to predict where their center would fall in the tier and an easy way to know what metric they need to focus on to get into, or stay in the "good tiers". [paraphrased]
- "Mortality and transplant rates: Can you give me any other info as far as how they are calculated? I don't have a hazard ratio or expected value for these two measures, so I'm thinking is (the assessment) done differently. What are the # of deaths on the wait list and time to transplant compared to?"
- "Say a patient is listed for a combined organ (SLK) , dies after the liver transplant without receiving the kidney would that death be counted against the liver program?"
- "What does it mean when a transplant center performs, say, 500 transplants per 100 people? I didn't see an explanation in the links, but I may have missed it."
- "Could you clarify why ... our estimated probability of surviving with a functioning graft at 1-year was 97.4%, and is displayed as 4-bars. The legend at the top of the sort list indicates that 1-Year kidney Survival (% with functioning transplant at 1 year) 97% is 5-bars."
- "Please advise on how the bar metrics are calculated on the SRTR beta site. Please also confirm which 12-month timeline is reflected."
- "I read the guide to the five tier system, and my understanding is that none of the new 5 tier metrics are risk adjusted. It is just each centers raw metric compared to the national metric?"
- "The one year organ survival includes: Graft only? So not 1 year patient survival; For kidney/liver- is it overall living and deceased population or only deceased donor grafts?"

- “The “Percentage alive with a functioning transplant at 1-year” is unclear that it is inclusive of both graft failure and patient death. The third row of the same table cannot be interpreted any better than the previous interpretive text: “Worse than Expected,” “Somewhat Worse than Expected,” “Good (as Expected),” “Somewhat Better than Expected,” and “Better than Expected.” (i.e. three bars = Good (as expected)).”

AST/ASTS Metrics Task Force White Paper

In addition to the feedback noted above, the AST & ASTS submitted a white paper developed by the AST/ASTS Metrics Task Force as their formal feedback. This white paper has been provided to the SRTR Visiting Committee for consideration. Because SRTR understands the paper is being considered for peer-reviewed publication, we are not at liberty to include the full manuscript in this summary of feedback; however, specific themes included:

9. Unintended consequences of public reporting: reducing organ supply, limiting patient access, barriers to innovation.
10. Skepticism about risk adjustment, data validity, and c statistics
11. Medicare conditions of participation (CoPs) and private payer Centers of Excellence hinge on these data
12. Patients have little/no choice of programs
13. High rates of false positive and false negative flags (referring to CMS CoPs).
14. Little difference in failure rates across programs (citing kidney outcomes on the absolute survival scale)
15. Need to include patient-reported outcomes
16. Programs may not be responsible for graft failures or deaths

Recommendations contained with the white paper included:

12. Rank order performance
13. Use symbols rather than numbers
14. Provide an overall summary measure
15. Include fewer reporting categories
16. Combine data over multiple years to increase precision
17. Aggregate data from different measures
18. Incorporate patient reported outcomes
19. Incorporate process measures
20. Better reimbursement
21. Collection of important data
22. Work more closely with the transplant stakeholders when developing these reports