

DR JON KOBASHIGAWA (Orcid ID : 0000-0001-9308-3172)

DR SHELLEY A HALL (Orcid ID : 0000-0002-4894-396X)

DR MARYJANE FARR (Orcid ID : 0000-0003-1066-9368)

DR EMILY A BLUMBERG (Orcid ID : 0000-0002-5193-6170)

DR HANNAH COPELAND (Orcid ID : 0000-0002-3753-5194)

PROFESSOR HOWARD J. EISEN (Orcid ID : 0000-0003-1149-4344)

DR KIRAN KAUR KHUSH (Orcid ID : 0000-0001-9697-5926)

DR JESSE DYLAN SCHOLD (Orcid ID : 0000-0002-5341-7286)

Article type : P - Meeting Report

Proceedings from the metrics forum in heart transplantation for performance monitoring

Jon Kobashigawa¹, Shelley Hall², Maryjane Farr³, Emily Blumberg⁴, A. Michael Borkon⁵, Monica Colvin⁶, Hannah Copeland⁷, Howard Eisen⁸, Maryl Johnson⁹, Ulrich Jorde¹⁰, Kiran Khush¹¹, James K. Kirklin¹², Jignesh Patel¹, Sean Pinney¹³, Hannaneh Saadaeijahromi¹, Jesse D. Schold¹⁴, Josef Stehlik¹⁵ on behalf of the Consensus Conference Participants

¹Cedars-Sinai Smidt Heart Institute, Los Angeles, CA; ²Baylor University Medical Center, Dallas, TX; ³Columbia University, New York, NY; ⁴University of Pennsylvania, Philadelphia, PA; ⁵Saint Luke's Hospital, Kansas City, MO; ⁶University of Michigan, Ann Arbor, MI; ⁷Lutheran Hospital,

This is the author manuscript accepted for publication and has undergone full peer review but has not been through the copyediting, typesetting, pagination and proofreading process, which may lead to differences between this version and the [Version of Record](#). Please cite this article as [doi: 10.1111/AJT.16901](https://doi.org/10.1111/AJT.16901)

This article is protected by copyright. All rights reserved

Indiana University Fort Wayne, Fort Wayne, IN; ⁸Pennsylvania State University, State College, PA; ⁹University of Wisconsin, Madison, WI; ¹⁰Montefiore Medical Center, New York, NY; ¹¹Stanford University, Palo Alto, CA; ¹²University of Alabama at Birmingham, Birmingham, AL; ¹³University of Chicago, Chicago, IL; ¹⁴Cleveland Clinic, Cleveland, OH; ¹⁵University of Utah, Salt Lake City, UT

Correspondence:

Jon A. Kobashigawa, Cedars-Sinai Smidt Heart Institute, 127 S. San Vicente Blvd, Los Angeles, CA 90048

Email: Kobashigawaj@cshs.org

Abbreviations

American Society of Transplantation (AST)

Centers for Medicare/Medicaid Services (CMS)

Collaborative Innovation and Improvement Network (COIIN)

Donation after Circulatory Death (DCD)

end-stage heart disease (ESHD)

Interagency Registry for Mechanically Assisted Circulatory Support (INTERMACS)

Observed/Expected (O/E)

Organ Procurement and Transplantation Network (OPTN)

Scientific Registry for Transplant Recipients (SRTR)

Thoracic and Critical Care Community of Practice (TCC COP)

United Network for Organ Sharing (UNOS)

ABSTRACT

Regulatory oversight for heart transplant programs is currently under review by the United Network for Organ Sharing (UNOS). There is concern whether 1-year patient and graft survival truly represent heart transplant center performance. Thus, a forum was organized by the Thoracic and Critical Care Community of Practice (TCC COP) of the American Society of Transplantation (AST) for the heart transplant community to voice their opinions on matters involving program performance monitoring by UNOS. A TCC COP work group was formed to review outcome metrics for adult heart transplantation and culminated in a virtual community forum (72 participants representing 61 heart transplant programs) on November 12-13, 2020. One-year post-transplant survival is still considered an appropriate and important measure to assess program performance. Waitlist mortality and offer acceptance rate as pre-transplant metrics could also be useful measures of program performance, recognizing that outside factors may influence these metrics. In depth discussion of these metrics and other issues including auditing thresholds, innovations to reduce risk-averse behavior and personally designed program scorecards are included in this meeting proceedings.

1. INTRODUCTION

The work of the Ad Hoc System Performance Committee of the Organ Procurement and Transplantation Network (OPTN) culminated in a public meeting in March 2019 to review performance metrics for all organ transplantation¹. There was a concern whether 1-year patient and graft survival truly represent transplant center performance. Thus, a forum was organized by the Thoracic and Critical Care Community of Practice (TCC COP) of the American Society of Transplantation (AST) for the heart transplant community to voice their opinions on matters involving program performance monitoring by UNOS. A TCC COP work group was formed to review outcome metrics for adult heart transplantation and was launched in November 2019. This group discussed various performance monitoring topics in a series of conference calls, culminating in a virtual community forum (72 participants representing 61 heart transplant

programs) on November 12-13, 2020. This forum was endorsed by the AST and was supported financially by a grant from the California Heart Center Foundation in Los Angeles.

1.1. Meeting Objectives:

- Explore additional independent metrics for program performance
- Determine practical “flagging” thresholds for each individual metric
- Discuss potential innovations to avoid risk averse behavior to increase the number of heart transplants performed and save more lives
- Discuss the purpose of a program “scorecard” which would be designed for programs to assess their own performance, and to enable comparison to peer institutions, but which would not be subject to regulatory oversight.

2. REVIEW OF THE COMMUNITY SURVEY: OUTCOME MEASURES IN ADULT HEART TRANSPLANTATION

Prior to the Forum Meeting, a survey was sent to transplant cardiologists (~90% returned responses) representing 61 programs. Programs with different transplant volumes were represented (~50% performed 21-40 transplants/year). All 11 UNOS regions were represented. Below are responses from select questions (the complete survey results are included in supplementary materials):

1. Most programs (~80%) were not satisfied with the current metrics for center performance evaluation
2. Most programs were not satisfied with the current flagging threshold for 1-year patient and graft survival, citing the following reasons:
 - It results in risk-averse decisions (75% of responses)
 - It does not reflect true center performance (67% of responses)
 - It is too strict (23% of responses)

- It does not affect transplant decisions (7% of responses)
3. The OPTN 1-year post transplant survival metric was more concerning for centers due to perceived comparisons with competing programs, as opposed to the OPTN flagging threshold for auditing (79% vs. 21%)
 4. Given that the current death rate Observed/Expected (O/E) hazard ratio is adjusted for both recipient and donor risk factors, programs were asked what death rate O/E hazard ratio would lessen risk-averse decision making at their center: 5% responded O/E >1.2 (current), 33% responded O/E >1.5, 13% responded O/E >1.8, 16% responded O/E >2, and 33% did not know
 5. 85% of programs would prefer that death rate O/E hazard ratio in the non-flagging range be expressed as “acceptable” rather than “worse” or “better”
 6. More than 90% of programs agreed with a “buffer zone” for death rate O/E ratio prior to reaching a flagging threshold, to provide an opportunity for improvement measures before formal flagging
 7. Considering very low survival rates in urgent status groups (status 1, 2, and 3) and an 80-90% 1-year survival benefit for these urgent patients, 75% of programs supported a lower threshold of unadjusted 80% 1-year survival for flagging
 8. Among currently available metrics in the SRTR database, patient/graft survival at multiple time points, and survival from listing/intent to treat, were the two top ranked metrics regarding center performance evaluation, followed by transplant rate/volume, waitlist mortality, and offer acceptance rate
 9. Among currently available data in the SRTR database with no current metric in use, time to transplant and length of hospitalization post-transplant were highly ranked as potential metrics for center performance evaluation, followed by active vs. inactive status on the waiting list and offer acceptance rate
 10. Among potential metrics with no current data collected, survival benefit (expected post-transplant survival for acuity less the expected survival without transplant) was ranked highest, followed by survival rates among end-stage heart disease (ESHD) patients,

pretransplant readmission/complication rate, quality of life, transplant rates among ESHD patients, and listing rate for referred patients

11. Almost 85% of the programs believe longer-term outcomes, such as 5-year survival data, should be included in adult heart transplant program evaluation

3. BREAKOUT DISCUSSIONS:

Forum participants were divided into groups to discuss four main meeting objectives:

- Explore additional independent metrics for program performance. (Current metrics seen in Table 1)
- Determine practical flagging thresholds for each individual metric
- Discuss innovations to avoid risk averse behavior
- Discuss the purpose of a program scorecard

As seen in Table 1, the Centers for Medicare/Medicaid Services (CMS) has removed the requirement that states that all transplant programs must meet all data submission, clinical experience, and outcome requirements in order to obtain Medicare re-approval. However, the above requirements do apply to those transplant center applying for initial Medicare approval². Each breakout session was led by 2 moderators, a heart transplant clinician (member of the Forum organizing committee) and a SRTR representative (to provide statistical and registry expertise). Each objective above was discussed in the breakout groups with consensus reached by verbal agreement. Following the breakout session, all attendees gathered in a reconvened session with a moderator from each group summarizing their discussions. From this reconvened session, consensus for the Forum Statements was reached by verbal agreement.

3.1. Explore additional independent metrics for program performance

Prior to the breakout sessions, discussion centered on three tables describing the various metrics for program performance which included: Currently available metrics (Table 2), data exists but no current metric (Table 3) and OPTN data not available and no metric exists (Table

4). Each table described the strengths and weaknesses for each metric. In general, there was broad agreement during the breakout sessions to consider waitlist mortality and offer acceptance rate as pre-transplant metrics, and 1-year graft and patient survival as the main post transplant metrics. There was moderate acceptance to add both short term (30 days and 90 days) and longer-term (3 and 5-year survival) outcomes. There was an overall willingness to include quality of life metrics; however, the participants recognize that this data is challenging to obtain and may not be practical.

Other metrics discussed in the breakout sessions included socioeconomic disparities and race distribution as important measurements to better understand the type of patients being waitlisted and transplanted at individual programs. Socio-economic disparities have been described as a neighborhood socioeconomic deprivation index-linked via patient home ZIP code-with a range of 0-1 (values nearing 1 indicate neighborhoods with greater socioeconomic deprivation)³. This “deprivation index” may be a useful tool for risk adjustment and for disparity research. Another metric discussed included “days alive out of the hospital” which may serve as a surrogate for quality of life. In addition, reassessment of risk adjustment for frailty, cardio-oncology (e.g. prior chest radiation), and congenital heart disease should be pursued.

Cardiac allograft vasculopathy as a metric was not embraced due to lack of standardized definition despite the International Society for Heart and Lung Transplantation Cardiac Allograft Vasculopathy grading scale. Hospital length of stay was also not supported due to multiple factors that may not reflect patient disease severity or success of the transplant procedure.

Finally, it was also suggested that donor offer sequence number could be a potential new metric to compare programs. Acceptance of donors with higher sequence numbers may identify programs that are willing to use more extended criteria donor hearts. The article by Baran et al⁴, however, showed no significant difference in survival between recipients of low and high donor sequence numbers, suggesting that only a small risk adjustment would be needed for a program accepting higher donor sequence number hearts.

There was a noted distinction made to separate a *system* metric from a *program* metric. The system metric is affected by outside factors (such as geographical variation in donor supply and demand which are not controlled by the program), as opposed to a program metric (such as donor organ acceptance rate) which has potentially controllable factors within a program. Another example is transplant rate which is more of a system process (due to external factors affecting this metric) rather than a program process and should not be included as a program metric.

There was consideration for overall survival from listing through 1-year post transplant as this would provide a "total patient experience" perspective. However, this may be more of a system metric (such as local Organ Procurement Organization performance, donor organ availability and characteristics). Thus, from a regulatory standpoint this would not be useful as a program metric. Transplant survival benefit is also important but is listing status dependent.

3.2. Determine practical flagging thresholds for individual metrics

Risk adjustment for both donor and recipient characteristics was cited as the most important factor in determining flagging thresholds. Several groups emphasized that all risk factors are not adequately captured, such as healthcare disparities in income, education, race, and geographic distance. Overall, it was agreed that the focus of flagging should be on program improvement rather than punitive measures.

Many believe that risk-adjusted 1-year survival is appropriate for flagging purposes. There was overall agreement that a higher (more lenient) O/E flagging threshold should be considered. There was also consideration for unadjusted 1-year survival > 80% to mitigate risk averse behavior and, therefore, increase heart transplant rates. This may be driven, in part, by the very high waitlist mortality seen in status 1-3 patients, such that even 80% 1-year survival would be a huge improvement in outcome.

In terms of reporting 1-year survival, there was support for a change to a descriptive designation of “acceptable” rather than “better/worse” results for public reporting, which will reduce focus on specific numbers, and minimize competition or comparison of programs. Acceptable outcomes would include a range both above and below national averages. This would be reasonable as most programs have a confidence range that straddle the national average.

There was concern about the presentation of the 5-tier bar graphs on the SRTR public website (see figure 1 example) and whether they truly represent a program’s functional status. Program bar graphs could potentially be misleading, as the outcomes for a given patient are not necessarily correlated with the bar graphs. For example, even a program that is not “top tier” may be best for a given patient who has a need for expertise such as for the highly sensitized patient. Additionally, there were concerns that the patient-facing 5-tier metric may have caused undue competition between geographically close programs which may have resulted in some high-risk patients not being offered transplant at any center.

3.3. Discuss potential innovations to avoid risk averse behavior

Many participants suggested that the Collaborative Innovation and Improvement Network (COIIN) project concept might be applied for high-risk donor profiles⁵. The COIIN project in renal transplantation is a three-year study exploring an innovative approach to transplant performance monitoring that aims to reduce the risk-avoidance behaviors associated with the current monitoring system, and develop and test an alternative, data-rich quality monitoring framework. It supports a collaborative approach toward performance improvement and effective practices.

The overall aim of the project is to increase transplantation, with a particular focus on utilization of deceased donor kidneys with a KDPI score greater than 50 percent. The COIIN methodology can be applied to the heart transplant community to offer a different approach to quality assessment and utilization metrics. For example, Donation after Circulatory Death

(DCD) donors and the combination of older age/longer distance donors could be analyzed separately or excluded from overall program statistics. The COIIN project concept could encourage centers to use extended criteria donor hearts and organs traditionally believed to be suboptimal without incurring the risk of flagging. As noted previously, acceptance of high donor sequence number hearts could also be used as a metric to identify programs willing to take on additional risk. Finally, sensitized candidates could also receive separate designation. It was noted, however, that standardization of laboratory testing across the country would be needed to provide comparable immunogenetic assessment of sensitized patients. Several participants suggested an organized mentorship program, whereby a low performing center would partner with a high-performing center for program improvement.

Others advocated for public reporting that also reflects a program's efforts to increase access to transplant. An example is the inclusion of survival benefit as a modifier to one-year survival reporting. It was noted that, while there is no accepted model at this time to calculate survival benefit, potential models could be developed based on waitlist survival or using Interagency Registry for Mechanically Assisted Circulatory Support (INTERMACS) data for patients on mechanical circulatory support devices. Periodic prevalent assessment which includes older transplants but only focus on the follow-up during a set (2-year) evaluation window might also be a potential solution for longer-term survival reporting (see figure 2).

Participants were also supportive of a shorter update cycle (2.0 years rather than 2.5 years), which may facilitate changes in practice to improve outcomes more expediently. However, it was noted that the reporting period is set by the criteria of collecting enough data to make the measurement meaningful. An example is in post-transplant outcomes as smaller programs may need more time to accumulate a meaning number of data points.

There was also support for the establishment of a national consortium to test innovation strategies and assess their utility and feasibility. At this time, it may be important to prioritize

innovation strategies that address healthcare disparities and expand access to transplantation to underserved or disadvantaged patients.

3.4. Discuss the purpose of a program scorecard

Many participants suggested that the primary purpose of a program scorecard should be program self-improvement, and that the metrics should not be public or subject to regulatory review. In lieu of detailed metrics, there was a desire for graphical representation of performance, employing a red, yellow, and green scale to reflect the bottom 10%, middle 80%, and top 10% of the country, respectively, to enable comparison to other programs of similar size and/or geography. The program scorecard could also be used for quality improvement purposes. Standardization of these metrics across the country would enable programs to better assess their performance in the context of their peers. The metric “time alive out of hospital” would be an important measure to include in the scorecard, which would also ideally incorporate future quality of life metrics.

4. FORUM SUMMARY

One-year post-transplant survival is still considered an appropriate and important measure to assess program performance; however, the flagging threshold for program review should be more lenient. Waitlist mortality and offer acceptance rate as pre-transplant metrics could also be useful measures of program performance, recognizing that many outside factors may influence these metrics. There are several other metrics involving socioeconomic and race disparities that should be considered.

Flagging thresholds were felt to result in risk averse behavior and should be more lenient. A “buffer zone” should be considered, in which programs that fall within the buffer zone could request help (mentorship) by higher performing programs to enact quality improvement measures. It was also suggested that the unadjusted 1-year survival threshold be lowered to

80% to avoid risk averse behavior and subsequently increase transplantation and save more lives.

Most innovations included a COIIN type program by which extended criteria donors, DCD donors, and research subjects would be excluded from regulatory oversight. This could stimulate innovation and increase transplantation in general.

Participants were in favor of a scorecard designed for programs to assess their own performance, and to enable comparison to peer institutions, but which would not be subject to regulatory oversight.

Although CMS has removed data submission, clinical experience, and outcome requirements in order to obtain Medicare re-approval, initial applicants must meet these requirements. Despite this change by CMS, it is believed that oversight by UNOS is appropriate for patient safety and efficacy.

5. FORUM STATEMENTS:

1. Most programs support implementation of different metrics that would more appropriately reflect their program's performance. These could include pre-transplant metrics such as waitlist mortality and offer acceptance, in addition to post-transplant survival at various time points.
2. The threshold for flagging should be more lenient to avoid risk averse behavior and thereby increase the number of heart transplants performed.
3. Potential innovations revolve around COIIN type processes or policies that would avoid regulatory oversight for extended criteria donors, highly sensitized patients, and patients from underrepresented socioeconomic and demographic groups.
4. Overall survival benefit should be considered but would require more data and model simulations.

5. A program scorecard with quality measures could be created for program quality improvement and not for regulatory oversight.

ACKNOWLEDGMENTS

The authors express their appreciation and gratitude to Christine Sumbi and Venise Strand for their assistance in organizing this forum.

DISCLOSURE

The authors of this manuscript have conflicts of interest to disclose as described by the *American Journal of Transplantation*. S Hall is the Chair of the UNOS Heart Committee. K Khush serves on the SRTR Review Committee. The other authors have no conflicts of interest to disclose.

Figure Legends

Figure 1: Key describing the meaning of the icons and providing actual numbers expected within each tier for each outcome assessed.

Figure 2: Periodic prevalent assessment for long-term outcomes.

REFERENCES:

1. OPTN Ad Hoc Systems Performance Committee Meeting Chicago, Illinois March 11-12, 2019 https://optn.transplant.hrsa.gov/media/3015/201906_spc_boardreport.pdf
2. 51732 Federal Register / Vol. 84 / No. 189 / Monday, September 30, 2019 / Rules and Regulations. <https://www.govinfo.gov/content/pkg/FR-2019-09-30/pdf/2019-20736.pdf>

3. Wadhvani SI, Beck AF, Bucuvalas J, Gottlieb L, Kotagal U, Lai JC. Neighborhood socioeconomic deprivation is associated with worse patient and graft survival following pediatric liver transplantation. *Am J Transplant*. 2020;20(6):1597-1605. doi:10.1111/ajt.15786
3. Baran DA, Copeland H, Copeland J. What Number Are We?. *Circ Heart Fail*. 2019;12(5):e005823. doi:10.1161/CIRCHEARTFAILURE.118.005823
4. Tosoc-Haskell H, Sisaithong K, Carrico R. The Collaborative Improvement and Innovation Network project to drive quality improvement. *Curr Opin Organ Transplant*. 2019;24(1):73-81. doi:10.1097/MOT.0000000000000596
5. Jay C, Schold JD. Measuring transplant center performance: The goals are not controversial but the methods and consequences can be. *Curr Transplant Rep*. 2017;4(1):52-58. doi:10.1007/s40472-017-0138-9

Supporting Information

Additional supporting information may be found online in the Supporting Information section at the end of the article.

APPENDIX: FORUM PARTICIPANTS:

Jon Kobashigawa (Chair), Shelley Hall (Co-Chair), Eric Adler, Rami Alharethi, Amrut Ambardekar, Francisco Arabia, David Baran, Maya Barghash, Emma Birks, Emily Blumberg, Michael Borkon, João Roberto Breda, Patrick Campbell, John Chin, Monica Colvin, Hannah Copeland, Jennifer Cowger, Mani Daneshmand, Teresa De Marco, Eugene Depasquale, Mark Drazner, Iona Dumitru, Howard Eisen, Maryjane Farr, David Feldman, Richard Formica, Alan Gass, Kambiz Ghafourian, Michael Givertz, Eman Hamad, Charles Hoopes, Ajay Israni, Daniel Jacoby, Maryl Johnson, Mark Jones, Ulrich Jorde, Kiran Khush, James Kirklin, David Klassen, Andrew Kolodziej, Joann Lindenfeld, Brian Lowes, Donna Mancini, H. Todd Massey, Mandeep Mehra, Nader

Moazami, Asum Mohammed, Ali Nsaïr, Chetan Patel, Jignesh Patel, Michael Pham, Sean Pinney, Barbara Pisani, Nirav Raval, Ashwin Ravichandran, Alex Reyentovich, Debbie Rinde-Hoffman, Maria Rodrigo, Joseph Rogers, Hannaneh Saadaeijahromi, Nicholas Salkowski, Kelly Schlendorf, Jesse Schold, Jeffrey Scott, Sharon Shepherd, Melissa Skeans, Jon Snyder, George Sokos, Randall Starling, Josef Stehlik, Johannes Steiner, Koji Takeda, Jose Tallaj, Inna Tchoukina, Ajay Vaidya, J. David Vega, Mary Norine Walsh, Daniel Westerdahl, Andrew Wey

Table 1. First year patient and graft survival metrics and criteria used for regulatory monitoring of transplant centers according to SRTR, MPSC, and CMS.⁵

	Criteria for current regulatory monitoring
SRTR Program Specific Reports	<p>If beyond lower threshold of 95% CI for probability that HR>1 based on Bayesian method, “lower than expected” survival</p> <p>If beyond upper threshold of 95% CI for probability that HR<1 based on Bayesian method, “higher than expected” survival</p>
MPSC (OPTN/UNOS)	If probability >75% that HR>1.2 OR probability > 10% that HR>2.5
CMS²	<p>For Existing Programs: Removed ongoing performance for Medicare re-approval as of 2019</p> <p>For New Programs: Standard-level deficiency: O-E>3, O/E>1.5, AND 1-sided p<0.05</p> <p>Conditional-level deficiency: O-E>3, O/E>1.85, AND 1-sided p<0.05</p>

SRTR Scientific Registry for Transplant Recipients, CI Confidence Intervals, HR Hazard Ratio, MPSC Membership and Professional Standards Committee, OPTN Organ Procurement and Transplant Network, UNOS United Network for Organ Sharing, CMS Center for Medicare/Medicaid Services

Table 2: Currently OPTN Available Metrics

Transplant Program Metric	Description	Existing Metric	Strengths	Weaknesses
Pre-transplant: Offer Acceptance Rate (risk stratified, DRI, KDPI, etc.)	Offer acceptance ratio indicates whether the program is more or less likely to accept offers than the average program. If the offer acceptance ratio is >1.0, then the program tends to accept more offers than average; if the offer acceptance ratio is <1.0, then the program tends to accept fewer offers than average.	SRTR reports on offer acceptance in the PSR and CUSUMs	Measures a distinct aspect of waitlist management practice that is within the program's control.	May incentivize behavior targeted at improving the metric rather than patient care. Some centers have overly broad criteria to receive all offers to screen out a majority. The data can be challenging to work with.
Pre-transplant: waitlist mortality rate	Ratio of observed vs. expected number of deaths once candidate is listed. Observed rate calculated by dividing the number of deaths (as reported to OPTN as removal reason & identified through other data sources) by number of person-years. Expected rate is based on national experience and takes into account various patient characteristics in an attempt to adjust for differences among programs	SRTR waitlist mortality rate	Measures a specific, important outcome	Programs, particularly kidney, may not provide care for patients on the waiting list. This metric may incentivize avoidance of high-risk candidates despite risk adjustment.
Pre-transplant: Transplant rate/volume	Ratio of observed vs. expected rate of transplants. Observed calculated by dividing the number of removals due to transplant at a program by the total number of person-years on the waiting list at this program, multiplied by 100. Expected rate is based on national experience and takes into account various patient characteristics in an attempt to adjust for differences among programs.	SRTR transplant rate	Directly measures the ability to get candidates transplanted; important to patients.	Reliant on issues outside of a program's control, for example geographic variability in donor supply and demand. May not be useful for evaluation of a program's performance. May be useful as a secondary metric when program under review.
Pre-transplant: Survival from listing/intent to treat metric	Overall survival from time of listing	SRTR will start reporting on survival from listing in July 2020*	Relevant to patients; Closely aligned with mission of increasing the # of transplants, improving waiting list and post-transplant outcomes.	Combines pre- and post-transplant outcomes into a single metric that is less actionable for improvement. Dependent on some factors that are outside a program's control.
Post-transplant: Patient/graft survival at multiple time points	Ratio of program's graft failure or death rate compared to the expected graft survival or death rate based on the national experience for patients similar to those at the program. A hazard ratio is used to estimate a program's graft failure rate or death rate based on donor and recipient characteristics.	SRTR 1 year & 3 year patient & graft survival	Able to implement. Adjusted for donor and candidate characteristics. Offers increased precision, combined with risk adjustment, to encourage more transplants, and to take on adjusted risks.	Reporting lag for longer term outcomes. Current metrics may be unable to detect troubling performance for smaller programs. Discourages programs from taking unadjusted risks. Prone to cause programs to be flagged after the few occurrences of events (e.g., X events = no flag, while X+1 events = flag)

Table 3: OPTN data exists but no current metric

Transplant Program Metric	Description	Data available but no metric exists	Strengths	Weaknesses
Pre-transplant: active vs. inactive status on waiting list	Ratio of active to inactive candidates on the waiting list	Data available, no metric	Simple metric. May be too simple.	Not aligned with specific goal or outcome. Whether a candidate should be active or inactive and whether a program should have fewer or greater number of inactive candidates is not clear. Use may incentivize undesirable behavior to effect the metric such as maintaining candidate as active when not currently appropriate for transplant. Does not account for offer screening information which may result in no offers.
Pre-transplant: time to transplant	Average time from time of listing to transplant	Data available, no metric; Possible no metric could be consistently applied to each program.	Very simple metric, important to patients.	May not be estimable for every program. Combines transplant and waitlist mortality rates.
Offer response time	Time from offer to acceptance of offer	Collect offer time stamps, some data quality questions that could raise questions on how the data should be used.	Measure of important phase of transplant process.	Some factors may be out of control of the program.
Post-transplant: Length of stay (risk adjusted)	Comparing length of stay to expected and national experience	Data available, no metric	Length of stay may reflect quality of care; common hospital metric exists today	More of an intermittent metric rather than an evaluation metric. Length of stay varies by circumstances and no clear goal for optimal length of stay. Must also incorporate deaths prior to discharge.

Table 4: OPTN data not available and no metric exists

Transplant Program Metric	Description	No data available, No metrics available	Strengths	Weaknesses
Post-transplant: Quality of life post-transplant	How do we measure quality of life? Do we have good proxy data?	No data available	Quality of life is an important indicator of benefit to patients.	Limitations and potential inaccuracies of patient self-reported data. May require direct data collection by the OPTN. How "quality of life" would be incorporated for deceased recipients.
Post-transplant: Readmission/ complication rate	Rate of recipients that require readmission due to complications	Follow-up form includes a field for whether a patient has been hospitalized since last follow-up but no additional detail is collected		
Listing rate for referred patients	Percentage of referred patients that get listed	No data available since do not collect data on patient referrals; would require additional data collection	One measure of how the program is serving the community of ESOF patients.	Do not currently collect data on referrals. Unclear what the desired outcome. Variability in number and type of referrals between programs.
Survival rates among end-stage organ failure (ESOF) patients	Rate of survival using the larger patient cohort of end-stage organ failure patients	No OPTN data available; may be other data sources for certain organs such as ESRD report for kidney	Broader measure of benefit transplant is providing.	Cannot assign ESOF patients to a provider (transplant program)
Transplant rates among ESOF patients	Rate of transplant using the larger patient cohort of end-stage organ failure patients	No OPTN data available	Broader measure of benefit transplant is providing.	Cannot assign ESOF patients to a provider (transplant program)

Figure 1



Figure 2