

**OPTN Membership and Professional Standards Committee
Performance Monitoring Enhancement Subcommittee
Meeting Summary
September 11, 2020
Conference Call**

Richard Formica, M.D., Chair

Introduction

The Performance Monitoring Enhancement Subcommittee (the Subcommittee) of the Membership and Professional Standards Committee (MPSC) met via Citrix GoToTraining teleconference on September 11, 2020 to discuss the following agenda items:

1. Overview of Project Plan and Timeline
2. Scope of Project
3. Scientific Registry of Transplant Recipients Presentation: Scorecard Development and Dashboard
4. Next Steps

The following is a summary of the Subcommittee's discussions.

1. Overview of Project Plan and Timeline

Staff gave an overview of the Performance Monitoring Enhancement Project and a timeline of projected events and deliverables. The Subcommittee's goal is to provide a recommendation of possible metrics to the MPSC by the October 2020 meeting in order to receive feedback on next steps for the Subcommittee. A proposal will need to be developed for MPSC review and approval at its June 2021 meeting. The goal is to have a proposal ready for public comment in August 2021 and submit the proposal to the OPTN Board of Directors for consideration at the December 2021 meeting.

2. Scope of Project

The MPSC has expressed an interest in developing a better performance monitoring methodology and process and to do so quickly. In order to realistically meet both of those goals, the MPSC's work on performance monitoring will need to be an iterative process. As part of any OPTN project work, the sponsoring committee needs to be able to clearly describe the OPTN authority that supports the actions being proposed. In light of the timeline for this proposal, the MPSC needs to initially focus on the transplant phases that the OPTN can confidently support are within the legal authority of the OPTN under the OPTN Final Rule. The OPTN has clear authority for transplant phases from listing through post-transplant. It would require a longer time frame to evaluate and determine what authority the OPTN has to look at phases of transplant prior to listing. As part of the iterative process, the MPSC can, in addition to evaluating the effectiveness of the proposal post-implementation, further investigate the ability of the OPTN to collect additional data and propose metrics for the pre-listing phases of transplant.

3. Scientific Registry of Transplant Recipients Presentation: Scorecard Development & Dashboard

The director of the SRTR gave a presentation to the Subcommittee in which he addressed factors to consider in choosing appropriate metrics, the differences between system metrics and program metrics, and absolute (unadjusted) versus relative (risk-adjusted) evaluations. The SRTR director presented a diagram of system metrics and program/OPO performance metrics to demonstrate that system metrics

are good metrics to illustrate how the overall system is working but span multiple phases of transplant and contain multiple factors including some that are outside the control of the transplant program. The SRTR director suggested that the MPSC may wish to focus on performance metrics directly under the program's control rather than metrics that describe system performance which include aspects that are beyond the control of the program. The SRTR director used transplant rate as an illustration. The transplant rate is a good metric for patients since it describes how quickly a patient may be transplanted at that program but when looking more closely at what contributes to the transplant rate, it may not be a good metric to evaluate program performance. The factors that affect the transplant rate include the waitlist mortality rate, active vs. inactive status of candidates, deceased donor availability, living donor availability, OPO death to donor conversion, allocation policy, and the offer acceptance rate. Overall survival from listing includes a total of 10 underlying metrics, some of which may not be under the control of the program, which would make it difficult for the MPSC to focus on the areas for improvement for a program. If all of the program performance and OPO performance metrics are improving, then the systems metrics overall should also improve.

One subcommittee member noted that system metrics can be useful but the underlying program metrics are more useful in meeting the goals of the MPSC because they allow the MPSC to focus on how the transplant programs and OPOs are achieving those overall system metrics. So in presenting the proposal, it will be important to note that the MPSC is interested in the overall system metrics; however, in order to be practical and for the MPSC to work with individual programs and OPOs, the MPSC needs to focus on the more discrete program and OPO performance metrics to support program and OPO improvement and ultimately, system improvement. The SRTR director noted that the system metrics are informative since they do reflect how well the system is working, but if the MPSC goal is to help individual programs or OPOs improve, looking at system performance metrics in isolation without understanding the underlying metrics could be a missed opportunity.

Another subcommittee member noted that there are many outside factors that affect the system metrics and some of the program and OPO metrics which will be a limiting factor to keep in mind when we look at the metrics in more detail. The Subcommittee member noted that we also must keep in mind how we support increasing transplants while also looking at oversight. This factor will be an important consideration in how we define the boundaries for triggering a review.

The chair asked whether all metrics must compare one program to another or could we use a raw measure such as raw waitlist mortality rates. The SRTR director responded that looking at unadjusted raw measures is a reasonable approach if programs are looking at their own performance. If the MPSC is looking at this data as a performance measure, then looking at raw unadjusted metrics is probably not helpful. With unadjusted metrics, programs do not get credit for the acuity of their patients or for using higher risk donor organs. In order to take into account the acuity of an individual program's patients, we must look at the observed waitlist mortality compared to how many deaths we would expect. This is not necessarily directly comparing one program to another program, it is comparing a program to what we would expect to have happened had their waitlist mortality rate been in line with what is happening around the country for the types of patients the program has listed.

The chair also referenced articles by Jesse Schold, noting that post-transplant survival for all kidney programs are clumped together so there is not a big difference between the best and the worst. Even the worst kidney program's outcomes are so much better than remaining on dialysis. Can we use a threshold for post-transplant survival that is some amount better than the survival expected for patients similar to program's patients with end-stage kidney disease? So can risk adjust for the acuity of a program's patients and determine the rate of survival without transplant, which for sake of the question we will use 67% survival, and then determine a threshold such as 90% survival. As long as the program's

patient survival is above that threshold, they are a healthy functioning program. In response, the SRTR director opined that we are then switching the question from what is the program's survival experience as compared to the survival experience at other programs around the country to how much did the program benefit the patient as compared to dialysis or if the patient did not receive a transplant. The question is then did the program benefit the patient by doing a transplant. In answer to this question, the programs that have the highest adjusted post-transplant survival rates would also have the highest change between patient survival on dialysis and patient survival with a transplant. With regard to the question about the slight differences between survival at kidney transplant programs, the MPSC can explore this issue when considering the triggers for review. The chair then further asked whether we should compare programs against each other or compare programs' performance against the disease, which is measuring the benefit to patients. The SRTR director opined that if a program is performing well on the discrete program metrics, then the program would be benefiting their patients.

A subcommittee member noted that once a scorecard is created, it will be used by everyone, payers included, to compare programs to programs. So we should be very specific and stay focused on metrics that programs have some degree of control over. The subcommittee member then noted that we hear repeatedly that post-transplant graft and patient survival is not good enough and we need more metrics. The numbers of transplants in the US were stagnant for a number of years, and then, there was a decrease in regulatory evaluation and the transplant numbers significantly increased. So do we really need more metrics? The chair noted that the community has expressed dissatisfaction with the emphasis placed on post-transplant patient and graft survival rates. The idea is to be more holistic and identify the components of what is a well-performing program.

Another subcommittee member noted that some of the metrics included on the same slide as program performance metrics affect each other as well, and those factors need to be considered as we design a scorecard. The SRTR director noted that the MPSC should consider whether a metric is measuring a distinct part of the process in determining which metrics to include in a scorecard.

The SRTR director then addressed the concept of risk adjustment and provided examples of risk adjusted versus unadjusted metrics. Unadjusted metrics may actually create a stronger disincentive to transplant higher risk patients or use higher risk donor organs. The SRTR director chose one of the risk adjusters, use of DCD donors, to illustrate the point. In the risk-adjusted model, DCD donor transplant outcomes are being compared to other DCD donor transplant outcomes and not to brain dead donor transplant outcomes. If the MPSC began looking at unadjusted metrics with a 90% survival floor, the perceived disincentive to use DCD donors would be a real disincentive because then the easiest way to reflect good outcomes would be to not take risks. Jesse Schold frequently presents on this topic and makes the point that often programs that have an overall risk adjusted lower rate of survival will cut back on use of DCD donors. This practice actually often hurts the programs rate of survival because the raw survival of DCD donor transplants may be lower than brain dead donor transplants but the program's DCD donor transplant survival compared to national DCD donor transplant survival is actually higher. The director further noted that if the MPSC accepts the need for risk adjustment, the metrics would need to assess observed outcomes versus expected outcomes. The chair followed up with the statement that programs get credit for taking risk under risk-adjusted metrics, and if we were to use an unadjusted metric, programs would move to the lowest risk, easiest transplants.

At request of the chair, the SRTR director touched briefly on the understandability of hazard ratios and whether there are other ways to present the same data with the same amount of precision that is more understandable. Hazard ratios tend to be in the 0.5 to 2.0 range with 2.0 being double the number of expected failures, 0.5 being half the number of expected failures, and 1.0 being right at the expected number of failures. SRTR has gotten questions from the transplant community about how to interpret

the hazard ratios and how to put it into a clinical context. One of the things SRTR is working on is how to translate the hazard ratios into the expected survival at a program to give programs information on what was the expected survival and what was the actual survival for a program. SRTR presented some preliminary information to the SRTR Visiting Committee on this effort. The SRTR can work with the MPSC on this issue as well. If it is easier to work with the metrics using a scale of survival rather than a scale of hazard ratios, the SRTR can provide that information. The chair noted that this issue should be addressed by the MPSC in the context of how much uncertainty in a measurement are we willing to tolerate to gain more understandability and utility for the members. Hazard ratios are developed in the context of a bell shaped curve of probability. To produce a scale of survival, the SRTR would be converting the bell shaped curve into a different scale. In that process, there could be some error in the conversion but that error could be very slight. The MPSC could consider whether the level of error that would occur would be counterbalanced by the ability to make the metrics more easily understood by programs and increase the ability to determine if there is a clinically meaningful difference in survival. The SRTR could work with the MPSC to design that type of system.

A number of subcommittee members expressed support for risk-adjustment. One member noted that it is not the metric itself but the flagging using the metric that makes programs uncertain about the metric. The subcommittee member suggested that the committee needs to be careful about the language used and how we design the system of identification and review to remove the current perception around flagging. Staff responded that there has been significant time included in the plan for this project to discuss and determine the MPSC performance monitoring process. The process the MPSC designs could look very different from the current process, for example, it could incorporate routine review of all programs' performance.

A subcommittee member asked whether the MPSC should consider working with programs that are performing only the easiest transplants from the perspective of encouraging more transplants and serving the disease population. The SRTR director noted that if the risk adjustment is working, really good survival on "easy" transplants should only give a program a hazard ratio of about 1. Within the realm of adjusted outcomes, the way to get really good outcomes is to do well on the more high risk transplants. It is a complex question. Are more complex patients not being referred to that program? Is the program not listing complex patients or if they are listing complex patients, are they not accepting offers for that patient? One has to ask what it is that is causing that program not to perform more complex transplants. The current data point for expected survival is a measurement of the risk tolerance of a program. Another subcommittee member responded that there is almost certainly programs that are risk averse and only willing to do "easy" transplants but looking at that behavior may have unintended consequences since some programs may not have the resources or the expertise to perform higher risk transplants safely. We would not want to penalize a program for not doing something that the program feels they do not have the resources or expertise to perform.

One subcommittee member noted that the unadjusted rate is available on the public website which creates incentives for risk averse behavior because programs will use those unadjusted rates to promote their program and can be used by insurers so we need to consider being consistent in the data available.

A subcommittee member commented that risk adjustment might not be appropriate in every field, noting as an example in pediatric heart transplant, a condition, Fontan circulation, is associated with highest risk but is not currently captured separately from congenital heart disease in OPTN data. Pediatric programs tend to be smaller volume and so doing these high risk transplants with higher mortality rates can significantly drop the survival rate without appropriate risk adjustment. The chair also mentioned the lack of cardiovascular data collected for kidney transplant recipients. The SRTR director responded that this point supports the collection of better data for risk adjustment and the

SRTR is working with the OPTN Data Advisory Committee to consider some of the gaps in data collection. The SRTR published an article that addressed some of the data that is not captured that might help the risk adjustment calculation.

Finally, the SRTR director touched briefly on the use of composite scores or independent metrics in the context of a scorecard. Ideally, the metrics chosen to be included in the scorecard would be uncorrelated measuring distinct aspects of transplant. Need to consider that if the metrics are uncorrelated, may want to have a threshold for each independent metric rather than a composite. This is a question the MPSC would need to consider. Composite scores may allow poor performance on one metric to be counterbalanced by high performance on another metric. To have high performance on the system metrics, need to have high performance on all the component metrics.

The SRTR director then demonstrated a tool that allows the Subcommittee to explore the relationship or correlation between different metrics. The tool includes waitlist mortality rate ratio, offer acceptance rate ratio, one-month graft survival hazard ratio, one-year graft survival hazard ratio, and three-year graft survival hazard ratio. The tool provides a scatterplot that shows the correlation between two chosen metrics and provides the correlation in a rate ($r = x$) as well. When exploring using the tool, the SRTR director noted that metrics that measure distinct aspects of the process would have low correlation. If metrics are highly correlated, then only need to look at one of the metrics since the two metrics are not measuring distinct aspects of the process. The Subcommittee was provided a link to access the tool so they can explore the correlations prior to the next Subcommittee meeting.

Upcoming Meetings

- September 29, 2020: MPSC Conference Call
- October 9, 2020: Performance Monitoring Enhancement Subcommittee Conference Call
- October 27-29, 2020: MPSC meeting (virtual)
- November 20, 2020: Performance Monitoring Enhancement Subcommittee Conference Call

Attendance

- **Subcommittee Members**
 - Richard Formica
 - Adam Frank
 - Alice Gray
 - Catherine Frenette
 - Christy Keahey
 - Didier Mandelbrot
 - Errol Bush
 - Jennifer Prinz
 - John Gutowski
 - Jon Kobashigawa
 - Jules Lin
 - Lisa Stocks
 - Mary Killackey
 - Matthew Cooper
 - Matthew O'Connor
 - Michael Gautreaux
 - Nicole Berry
 - Sanjeev Akkina
 - Stephon Thelwell
 - Steven Potter
 - Willscott Naugler
- **HRSA Representatives**
 - Arjun Naik
 - Marilyn Levi
- **SRTR Staff**
 - Andrew Wey
 - Jon Snyder
- **UNOS Staff**
 - Amanda Gurin
 - Amy Minkler
 - Betsy Warnick
 - Danielle Hawkins
 - Gabe Vece
 - Ginny McBride
 - Jaqueline O'Keefe
 - Leaf Slife
 - Liz Robbins Callahan
 - Nadine Drumn
 - Robyn DiSalvo
 - Sally Aungier
 - Sharon Shepherd
 - Tameka Bland
- **Other Attendees**
 - None