

Proposed Changes to the OPTN Transplant Program Outcomes Review System

OPTN/UNOS Membership and Professional Standards Committee

*Prepared by: Chad Waller
UNOS Member Quality Department*

Executive Summary	1
Is the sponsoring Committee requesting specific feedback or input about the proposal?	1
What problem will this proposal solve?	2
Why should you support this proposal?	2
Which populations are impacted by this proposal?	16
How does this proposal impact the OPTN Strategic Plan?	17
How will the OPTN implement this proposal?	17
How will members be evaluated for compliance with this proposal?	18
How will the sponsoring Committee evaluate whether this proposal was successful post implementation?	18
Policy or Bylaws Language.....	19

Proposed Changes to the OPTN Transplant Program Outcomes Review System

Affected Bylaws: OPTN Bylaws Appendix D.11.A (Transplant Program Performance)
Sponsoring Committee: Membership and Professional Standards
Public Comment Period: August 15, 2016 – October 15, 2016

Executive Summary

There is a perception in the transplant community that the OPTN's current method for monitoring transplant program outcomes has contributed to members' increasingly conservative behavior. Specifically, in order to avoid being reviewed for transplant outcomes performance, members are reportedly overly selective in patients they will list for transplant and the organs they will accept to transplant. This proposal aims to change this perception and associated behavior by modifying the system the OPTN uses to monitor and review transplant program outcomes for each organ type (excluding vascular composite allografts). The proposed system entails four tiers: the highest tier identifies programs with high hazard ratios that will initiate automatic Membership and Professional Standards Committee (MPSC) review; the middle two tiers represent routine, quality improvement program reviews and include a random selection component (50% and 10% probability, respectively, of MPSC review) for all other programs with worse than expected post-transplant outcomes; and the bottom tier represents programs performing as expected or better that the MPSC will not engage in outcome reviews. Based on transplant programs' 1-year graft survival and 1-year patient survival hazard ratios, programs will be placed in the tier that corresponds with the worse result of these two analyses.

The top tier in this proposed system, which prompts automatic MPSC review, is defined by a higher hazard ratio threshold than what is currently used. Due to an increased likelihood of transplant program underperformance at this higher hazard ratio threshold, programs identified in this tier will initially be requested to complete a more detailed and expansive survey than what is initially requested of programs currently identified for MPSC review. The middle two tiers of this proposed system, which include a random selection component for determining which programs the MPSC will review, reflect a validation of quality improvement efforts that are already being undertaken by transplant programs prior to MPSC involvement. A transplant program's recognition of the issues that may have led to more events than what is expected, and the quality improvement efforts enacted to address those issues, will be a primary focus of the MPSC's reviews that are prompted by the random selection of programs within these tiers. Programs that are able to demonstrate ongoing and appropriate quality improvement efforts will be released from further MPSC review.

Changing transplant programs' perception of the OPTN's outcome review system is intended to minimize the unintended consequence of increasingly conservative selection of patients and organs for transplant, thereby supporting the OPTN's strategic plan goal to increase the number of transplants.

Is the sponsoring Committee requesting specific feedback or input about the proposal?

Please note that the MPSC is also sponsoring a proposal this public comment cycle that specifically impacts how kidney transplant programs are reviewed, and that the MPSC developed independently from this proposal. Reviewing both of these proposals, do you recognize advantages or concerns with moving both forward in parallel?

What problem will this proposal solve?

There is a perception in the transplant community that the OPTN's current method for monitoring transplant program outcomes has contributed to members' increasingly conservative behavior. This is thought to have created a behavior in which transplant programs are becoming more selective in the types of patients listed for transplant and the quality of organ accepted to transplant.

Why should you support this proposal?

This proposal is expected to increase the utilization of high-risk organs by changing the screening algorithm used by the MPSC to identify underperforming programs, and the MPSC review process that is initiated by these identifications. This result is expected because the proposed changes are intended to address concerns in the transplant community regarding unintended consequences of the current OPTN transplant program outcomes review system.^{1, 2}

One of the primary changes included in this proposal is an increase in the hazard ratio thresholds that prompt automatic MPSC review. Based on one-year graft and survival outcomes, the MPSC currently reviews programs if the probability is greater than 75 percent that the hazard ratio is greater than 1.2, or the probability is greater than 10 percent that the hazard ratio is greater than 2.5. The “review space” defined by these analyses are illustrated in Figure 1. If a program’s 1-year graft survival or 1-year patient survival outcomes are within the shaded area, then it meets the MPSC’s current review criteria.³

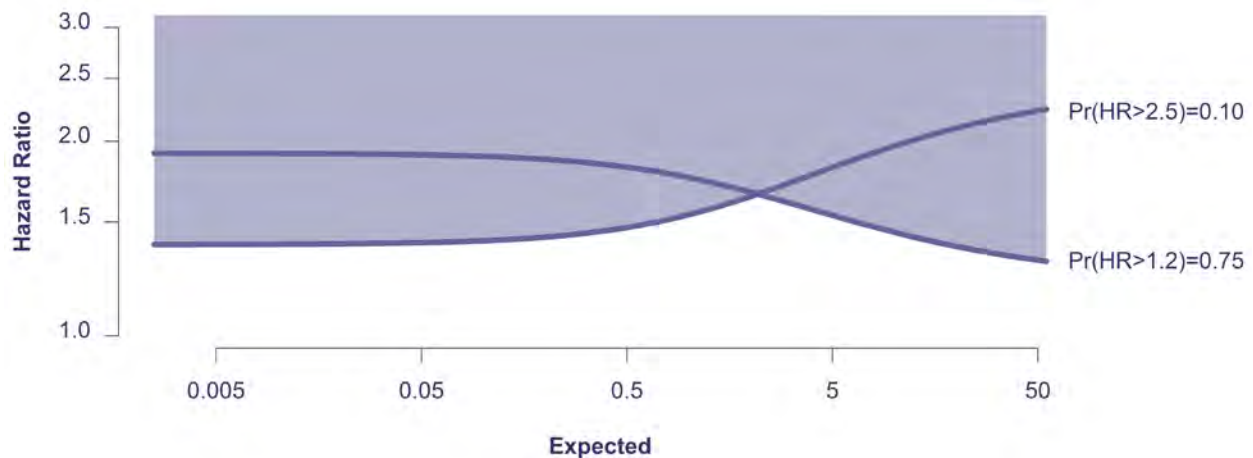


Figure 1: Current MPSC review criteria.

This proposal recommends changing the Bylaws so that automatic MPSC review is driven by a higher hazard ratio threshold than what is currently used. The MPSC believes that raising the hazard ratio threshold that prompts automatic MPSC review should increase confidence that programs above this threshold are more likely to be underperforming with respect to their transplant outcomes, i.e., a higher true-positive rate. Specifically, in a plot of observed hazard ratio versus the expected number of events, the proposed hazard ratio threshold to prompt automatic MPSC review is defined by a greater than 60%

¹ Schold, J. D., Buccini, L. D., Srinivas, T. R., Srinivas, R. T., Poggio, E. D., Flechner, S. M., Soria, C., Segev, D.L., Fung, J., & Goldfarb, D. A. (2012). The association of center performance evaluations and kidney transplant volume in the United States. *American Journal of Transplantation*, 13(1), 67–75.

² Roberts, J. P. (2012). Impact of outcomes monitoring on innovation and risk in liver transplantation. *Liver Transplantation*, 18(S2), S59–S63. In-line Citation: (Roberts, 2012)

³ Salkowski, N., Snyder, J. J., Zaun, D. A., Leighton, T., Edwards, E. B., Israni, A. K., & Kasiske, B. L. (2014). A Scientific Registry of Transplant Recipients Bayesian Method for Identifying Underperforming Transplant Programs. *American Journal of Transplantation*, 14(6), 1310-1317.

probability that the program’s graft or patient survival hazard ratio is greater than 1.75. This “tier” in the proposed OPTN outcomes review system is referred to as “Expanded Program Performance Review Tier 1” (Figures 2 and 3). Along with increasing the hazard ratio threshold that prompts automatic MPSC review, the MPSC’s response to the programs identified in this tier will also change. Programs with graft or patient survival hazard ratios that are above the proposed threshold will go through an expanded review that is commonly reserved currently for programs that remain under MPSC review for an extended period of time (usually longer than a year). With this approach, the MPSC hopes it will be able to better identify those programs having real difficulties with their transplant outcomes, and that it will be able to help those programs more quickly discover and address the root of those issues.

This proposal also introduces two other review tiers to identify programs that have observed more events than expected via the methodology used by the OPTN to evaluate program performance. These new tiers include programs that have observed hazard ratios greater than 1.00, but that are not included in Expanded Program Performance Review Tier 1. “Routine Program Review Tier 2” is defined by a greater than 60% probability that the hazard ratio is greater than 1.25, and the program is not included in Expanded Program Program Performance Review Tier 1. “Routine Program Review Tier 3” is defined by a hazard ratio that is greater than 1.0, and the program is not included in Expanded Program Performance Review Tier 1 or Routine Program Review Tier 2. The final tier introduced in this proposal, “As Expected or Better Tier 4,” is defined by an observed hazard ratio of 1.0 or less.

Programs included in Routine Program Review Tier 2 and Routine Program Review Tier 3 will not be automatically reviewed; rather, these programs will only be reviewed by the MPSC if randomly selected. Of those programs in each tier, 50 percent in Routine Program Review Tier 2 and 10 percent in Routine Program Review Tier 3 will be randomly selected for MPSC review. Programs included in As Expected or Better Tier 4 will not prompt an MPSC review. Figure 2 illustrates the MPSC review space for each of these four tiers and Figure 3 illustrates the MPSC review space for Tiers 1-3, with programs in the As Expected or Better Tier 4 not shown (but represented as those below 1 on the y-axis) for the sake of simplicity and to focus on the details included in the top three tiers. In these figures, the y-axis represented the program’s observed hazard ratio and the x-axis represents the program’s number of expected graft loss events; “x” represents an expanded program performance review; “+” represents a routine program review; and “o” represents programs that would not be reviewed through this system. Please note that there are a few programs shown with a “+” in tier 4; this happens because the figure displays results for the graft survival outcome and these programs had higher levels of underperformance for patient survival. In these few instances, the patient survival outcomes caused their program component to receive a higher-level review.

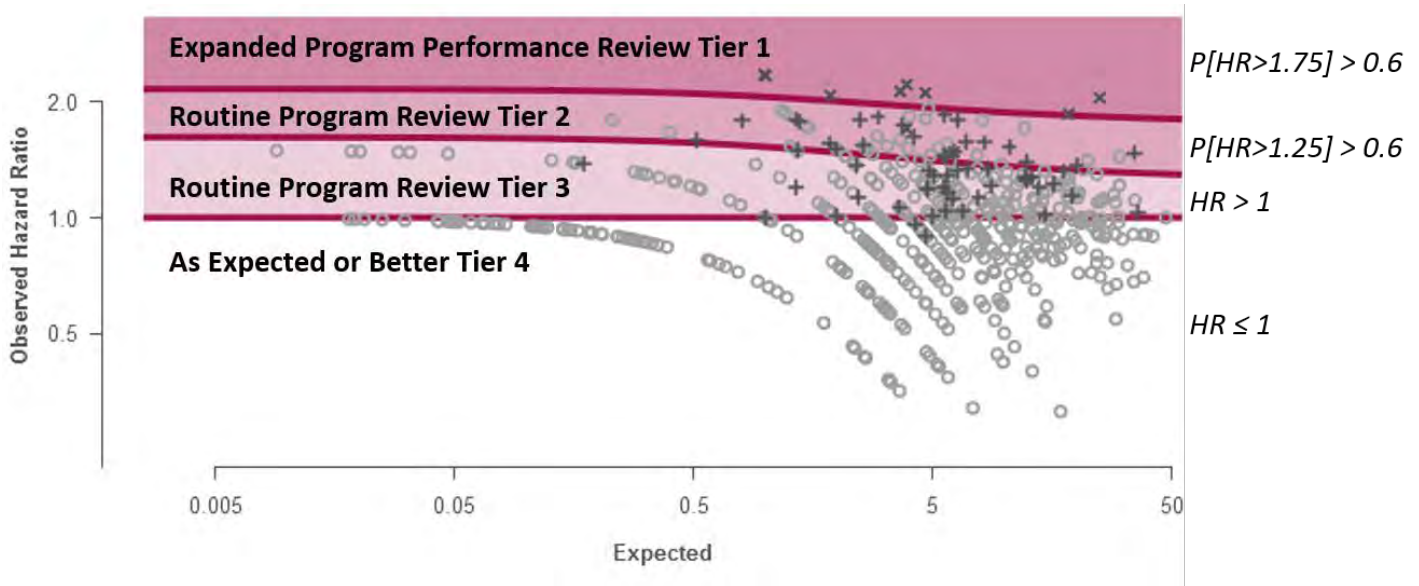


Figure 2: Proposed four-tier system for OPTN transplant outcomes review

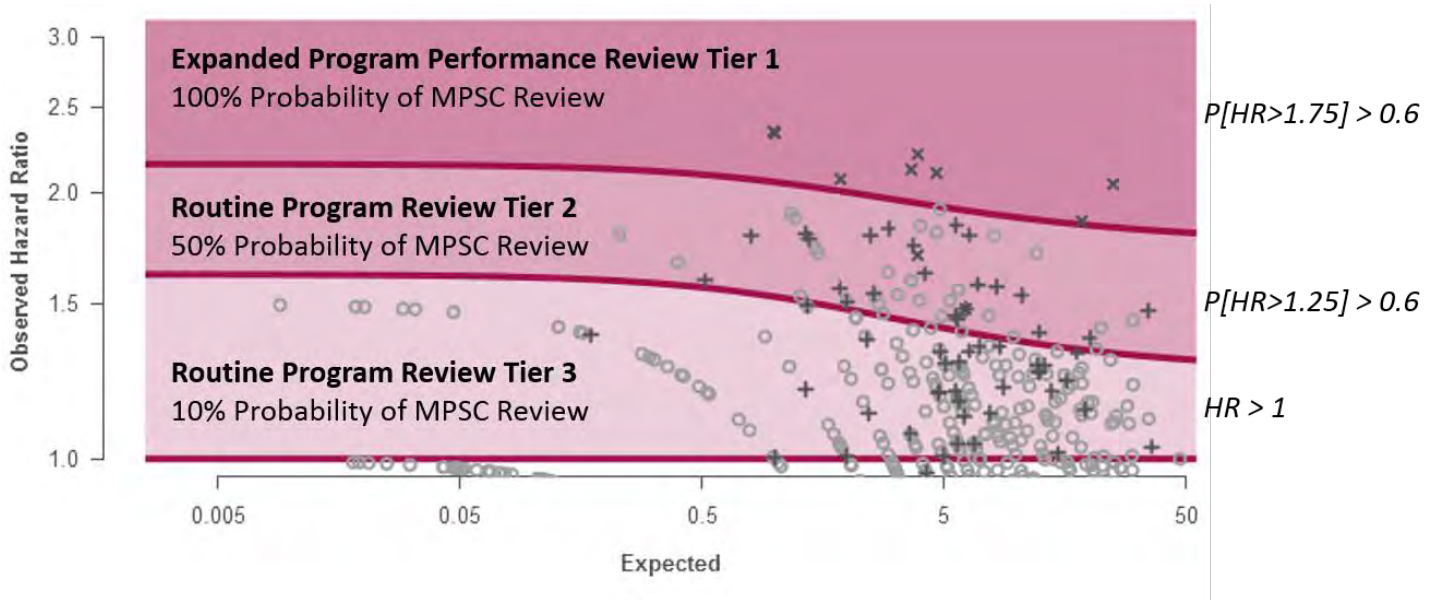


Figure 3: Proposed four-tier system for OPTN transplant outcomes review, focusing on the top three tiers.

The random selection component of Tier 2 and Tier 3 is intended to address concerns in the community about the ability to assess the true level of program underperformance at these lower hazard ratio thresholds. Although the mathematics of the model used to review transplant program outcomes necessarily means that a hazard ratio greater than 1.0 indicates that the program experienced a greater number of events than what would be expected, the transplant community has indicated concerns with the model’s ability to discriminate significant differences between programs plotted close together as a result of these analyses and the lack of available data to describe all patient risk profiles. The reviews that are prompted by these random selections will be similar to the initial reviews currently undertaken by the MPSC, including a focus on the program’s quality improvement efforts. Additionally, as OPTN Bylaws Appendix D.3 (Quality Assessment and Performance Improvement (QAPI) Requirement) requires that all transplant hospitals, “develop, implement and maintain an ongoing, comprehensive and data-driven QAPI program,” reviews driven by the random selection of programs in these tiers will be considered opportunities for the MPSC to communicate with members about their QAPI plans as they apply to transplant program outcomes performance. From this perspective, minimal resources should be required by transplant programs to respond to these reviews as programs will primarily be reporting on ongoing efforts at their institutions.

Additionally, the Task Force to Reduce Disincentives to Transplantation (the Task Force) that participated in the development of these recommendations believed that, over multiple reporting cycles, most programs will commonly fluctuate between Routine Program Review Tier 2, Routine Program Review Tier 3, and As Expected or Better Tier 4. As such, the Task Force suggested that most programs will eventually be randomly selected and engaged by the MPSC, and that all transplant programs should expect to go through this process at some point. Because it is expected that almost all programs will eventually go through this MPSC review, this process should be viewed as routine of all transplant programs and not an action limited to programs with transplant outcome performance problems. Engaging the majority of programs in this manner will also make it more likely that the MPSC becomes aware of particularly successful quality improvement efforts, which the MPSC can use to better help members in need of process improvements and can share as appropriate to help create educational opportunities within the OPTN.

Another advantage of the system detailed in this proposal is that it eliminates Bylaws that establish a different system for reviewing programs performing nine or fewer transplants in a two and a half year period. Eliminating these Bylaws simplifies how transplant programs are reviewed for outcomes by establishing a system that is consistent and applicable to all transplant programs, regardless of size.

How was this proposal developed?

Task Force Creation and Preliminary Discussions

At its December 2015 meeting, the OPTN/UNOS Board of Directors (Board) discussed concerns raised throughout the community that the MPSC review process has yielded an unintended consequence of transplant programs becoming increasingly conservative in their patient and organ selections for transplantation. The Board suggested that this behavior shift has resulted from programs trying to avoid MPSC review because of the resources those reviews entail and perceptions about the quality of a transplant program that is under review by the MPSC. These perspectives and behaviors are to the detriment of transplant candidates who would likely be better served by a transplant with a greater number of risk factors, as compared to remaining on the waiting list. At the conclusion of this discussion, the Board adopted the following resolution to begin addressing this problem promptly:

RESOLVED, that the MPSC is tasked over a period of 6 months to provide the Board with a proposal for an improved program specific reporting system that identifies substantive clinical differences in patient and graft outcomes.

FURTHER RESOLVED, that the President will appoint a working group consisting of 10 members; 3 from the UNOS/OPTN Board, 3 from the societies of the AAAU, and 3 from the MPSC, and 1 ad hoc member from CMS - this working group will, over a three month period, identify objective measures that define clinically relevant outcome differences - this work group will then submit their findings to the MPSC for approval, and by the June 2016 board meeting, present that proposal to the Board for action.

In response to this resolution, OPTN leadership formed the Task Force to Reduce Disincentives to Transplantation (the Task Force).

The Task Force convened regular conference calls from January 2016 through May 2016 to develop a proposal that could be shared with the MPSC for review. Data analyses presented by the Scientific Registry of Transplant Recipients (SRTR) and UNOS supplemented and supported the Task Force's discussions. After reviewing the system currently used by the MPSC, the Task Force suggested a number of ideas for consideration:

- Establish a review threshold for all programs at a set measure of graft/patient survival
 - e.g., review any kidney program under 91% survival
- Adjust hazard ratio thresholds, lower false positive targets, set an absolute minimal survival threshold, etc., to capture a more appropriate number of programs that the MPSC should review
- Add a second requirement to the current system where the difference between observed and expected outcomes must also be above a set threshold before a program is reviewed by the MPSC
 - e.g., reviewed if: (expected outcomes % - observed outcomes %) \geq 5%
- Add a second requirement to the current system that also evaluates the outcomes of a more recent cohort. Only flag programs if their standard 2.5 year cohort AND a more recent cohort are below expected
 - e.g., substandard outcomes demonstrated in a 2.5 year and 1 year cohort
- A combined patient AND graft survival metric
 - Programs only reviewed if both metrics are below established thresholds
- Periodic review of every transplant program
 - Frequency of review influenced by outcomes
- A "warning" tier

- Alerts program as they are nearing a review, but does not entail further action
- Establish an MPSC review threshold that corresponds to the potential benefit of a transplant opportunity as compared the risk of remaining on the waiting list
- A combination of the ideas above

Task Force members understood concerns about establishing a review threshold for all programs at a set measure of graft/patient survival, noting that a threshold with no risk adjustment may perpetuate the risk-averse behaviors that the group is trying to curtail.

The Task Force's discussions proceeded to focus on the possibility of changing the system to one where a greater number of programs are periodically reviewed by the MPSC, with the frequency and focus of the reviews being driven by the programs' outcomes. The Task Force acknowledged that part of the benefit of MPSC reviews is driving programs to look at their own performance and evaluate their quality improvement metrics to make sure that they're responding to the things that need to be changed.

The Task Force also discussed the possibility of evaluating programs based on a combined measure of their graft and patient survival. The underlying intent of this idea is to reduce the number of reviews conducted by the MPSC. Similarly, adding a second requirement that also evaluates the outcomes of a more recent cohort has the potential to reduce the number of flags while also accounting for quality improvement opportunities that may have already been recognized and implemented by the transplant program.

Ultimately, the Task Force decided to focus on the possibility of using several different approaches to define a new transplant program outcomes review system. Specifically, the Task Force agreed to present the following ideas for the MPSC's consideration:

- Adjust hazard ratio thresholds, lower false positive targets, set an absolute minimal survival threshold, etc., to capture a more appropriate number of programs that the MPSC should review
- A combined patient AND graft survival metric
- Add a second requirement that also evaluates the outcomes of a more recent cohort that could potentially account for quality improvement activities that may have been implemented by the program
- A "warning" threshold tier that would identify a program that might be trending towards an adverse outcome

Preliminary MPSC Review of Task Force Concepts

The MPSC Vice Chair, who was also a member of the Task Force, presented these preliminary recommendations for the MPSC at its March 2016 meeting. Some MPSC members expressed disappointment that the proposed changes represented small changes to the current system, and not a more substantive overhaul that might be more likely to result in the behavior change that the Board was seeking. MPSC members noted the need for better risk adjustment, that more complicated transplants incur a greater expense that isn't factored into reimbursement, that CMS reviews and insurance companies' assessments are a larger driver in transplant program behavior, and the need for pre-transplant metrics.

The OPTN Vice President responded to the Committee's discussion stating that two things are accomplished by the current OPTN flagging system: programs that are in real trouble with their outcomes are identified and the MPSC uses its tools (peer review, interviews, etc.) to work with these identified programs to promote change. Secondly, it also prods programs' self-assessment to explore opportunities for improvement. Both of these considerations come with a stigma of being flagged by the MPSC, and this prompted the idea of creating a "warning tier" to nudge programs to improve without the stigma that is associated with it. In an effort to avoid the "flagging stigma," suggestions were made that the OPTN should explore establishing a process that looks at all programs periodically. Regarding payer structures, it seems unlikely at this time that the OPTN can significantly impact the decisions made by those groups.

Refinement of Task Force Concepts

The Task Force responded to the MPSC's feedback regarding the magnitude of these changes stating that these concepts could be established such that they yielded changes that are too extreme (e.g., top tier defined by a 95% probability that the program's hazard ratio is above 2.0). The challenge is finding an appropriate balance. The Task Force proceeded to consider what a system might look like that includes multiple tiers of review, with reviews in the lower tiers being random and the highest tier of review reflecting those programs that are currently engaged in the more severe MPSC outcome reviews actions (beyond the MPSC's expanded survey, and including peer visits, MPSC interviews, etc.). The Task Force discussed that the probability of a random review would be impacted by a programs' outcomes (closer to hazard ratio of 1.0, less likely to be reviewed). Hazard ratios to define the highest, most urgent tier of review would be increased so that there is increased confidence that the programs with the most deficient outcomes receive sufficient MPSC attention and assistance. The Task Force agreed that this flagging threshold should be as high as comfortably possible, and yield a number of programs comparable to the current number of programs that are flagged and involved in more in-depth engagement with the MPSC. Raising the hazard ratio to this extent for the highest tier would result in low- and medium-volume programs rarely being flagged, if ever, due to a lack of statistical power that creates an inability to say definitively that those programs reached that level of underperformance. To accommodate this, the Task Force suggested program volume also be a factor in determining the probability for random review. To investigate these ideas further, the Task Force requested that the SRTR model some of the ideas that had been discussed.

In response to this request, the SRTR developed a "strawman" concept considering the following goals:

- A top-tier focused review threshold: The threshold for this zone would be high such that only programs with the most extreme outcomes would be identified. The Task Force suggested that the number of programs that would fall into this zone should be relatively small and commensurate with the number of programs that the MPSC has historically had to take more in-depth action against, perhaps 10-20 programs a year.
- A quality improvement focused review zone or zones: All programs would be subject to random MPSC performance reviews. We could consider different review zones where the probability of being selected for review is influenced by the program's observed outcomes. Programs with worse outcomes would have a higher likelihood of being reviewed by the MPSC.

With the above goals in mind, the SRTR created the following strawman system for the Task Force's consideration:

1. A two-tiered quality control focused review zone:
 - a. Rapid Review: Greater than 60% probability that the program's hazard ratio is greater than 1.75
 - b. Standard Review: Greater than 60% probability that the program's hazard ratio is greater than 1.50
2. A three-tiered quality improvement review zone:
 - a. High-frequency Review: Greater than 60% probability that the program's hazard ratio is greater than 1.25. Review probability = 50%.
 - b. Medium-frequency Review: Greater than 60% probability that the program's hazard ratio is greater than 1.00. Review Probability = 20%.
 - c. Low-frequency Review: All other programs. Review probability = 10%.

Based on the December 2015 program specific report data, this strawman system is displayed in Figure 4 for the graft survival assessment of adult program components, using the same notation as explained above for Figures 2 and 3. The two quality control zones are delineated by the "Rapid" and "Standard" review lines. The three quality improvement zones are delineated by the "Hi Freq" (50% review probability) and "Med Freq" (20% review probability) lines. Programs in the white zone below the "Med Freq" line would be subject to the low frequency review (10% review probability).

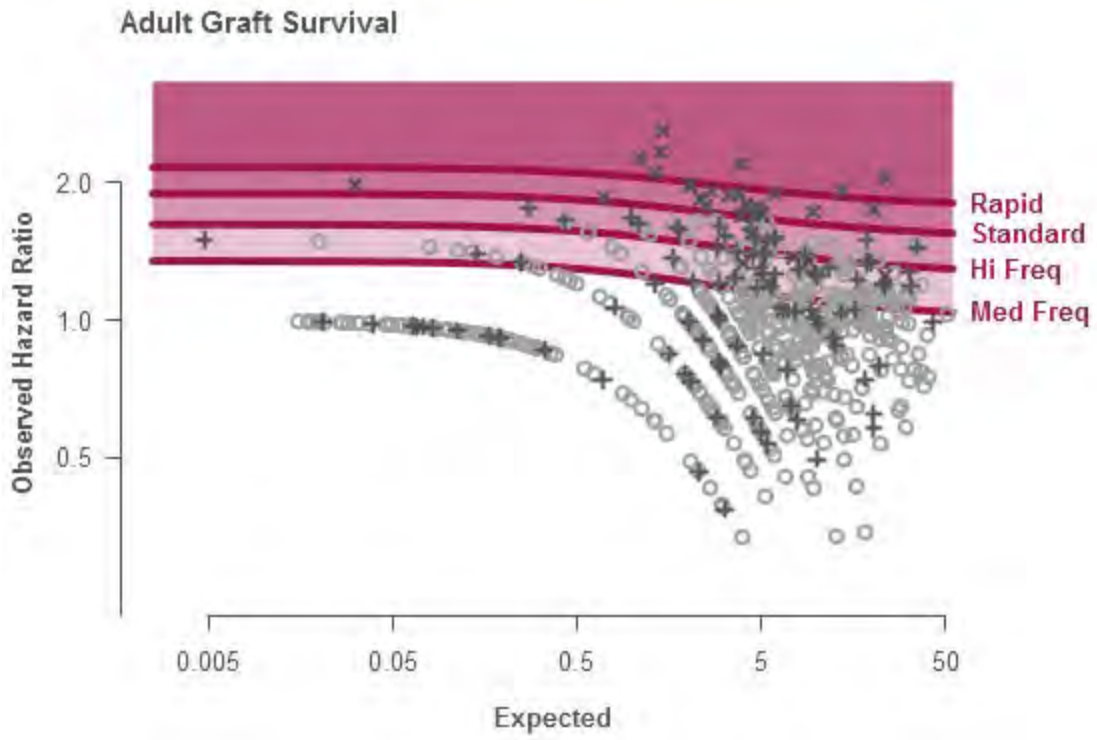


Figure 4- Example of the strawman five-tiered review system for adult graft survival outcomes

Task Force discussion of these SRTR analyses highlighted the following considerations:

- Task Force members questioned what value would be realized from engaging transplant programs that perform as expected or better. Others suggested that there may be some benefit to obtain perspectives of what successful programs look like and how they behave. The Task Force considered that uncovering, understanding, and disseminating best practices of successful programs will ideally be an outcome of the Collaborative Innovation and Improvement Network (COIIN) project that the OPTN is also engaged. Furthermore, program performance fluctuates such that every program may occasionally not perform as expected making it unlikely that any program will go for an extended period without MPSC engagement. Ultimately, the Task Force agreed that any program performing as expected or better should be excluded from MPSC review.
- The Task Force expressed a desire to reduce the number of flags.
 - The first suggestion was to raise the hazard ratios for each tier.
 - Regarding the lower quality improvement tiers, the number of randomly reviewed programs should be reduced so as to not overburden the MPSC, or members, with these reviews. The Task Force recommended reducing the review probability of each quality improvement tier by half. Doing so would reduce the random review probability for each tier as follows: High-frequency review- 25%, Medium-frequency review- 10%, Low-frequency review- 5%.
 - Call participants asked if there is a particular reason why reviews occurred every six months, as compared to once a year. Changing the review period from every six months to every year could also reduce the number of flags as a longer review period may allow programs to recognize and fix problems, before the need for MPSC engagement.
 - Alluding to previous Task Force discussions, the Task Force also suggested exploring that a program would only be flagged if their graft *and* patient survival are below expected, instead of being flagged if either metric was below expected.

Subsequent Task Force discussions focused on further refining and clarifying its recommendations. First, the Task Force considered the number of tiers and the boundaries to define those tiers. The Task Force reiterated that programs performing as expected or better should not be reviewed by the MPSC as a function of this system. To simplify this system, the Task Force recommended condensing the top two quality control tiers into one. This tier should represent the identification of programs that are significant outliers and identification of these programs would prompt a more intensive MPSC review due to a high level of confidence that these programs are performing below what is expected.

The Task Force proceeded to discuss the level of engagement the MPSC should have with programs identified in this tier. Currently, programs identified for review are sent an “initial survey.” The MPSC reviews the results of this survey, and then responds to the member. If a program is not released from review, and the desired changes have not been seen by the MPSC (usually after about a year of MPSC review), the MPSC may request the completion of an “expanded survey.” Expanded surveys are more involved than the initial survey, seeking greater details about the program, each graft failure or patient death, the steps taken to review each of these incidents, etc. Upon reviewing the results of the expanded survey, the MPSC will engage the program as it believes is necessary. If a program is not released from review at this point, additional MPSC interactions may include an informal discussion with the MPSC’s Performance Analysis and Improvement Subcommittee (PAIS), an MPSC-directed peer visit, and in extreme cases, the MPSC may ask the program to inactivate. (Requesting inactivation is often reserved until after a peer visit or informal discussion, and if the MPSC is still not convinced that the program has achieved satisfactory changes). The Task Force recommended that programs identified in the highest, quality control tier, should initially be requested to complete something similar to the MPSC’s current “expanded survey.” Based on the program’s responses to this survey, the MPSC would proceed in its review as it does currently, including releasing programs from review as appropriate. The Task Force believed this was reasonable because there is a higher level of confidence that these programs are

performing below what would be expected, and it would allow the MPSC to more promptly uncover and address the issues that may be impacting a program's performance.

Continuing with this approach, and focusing on those programs in the quality improvement tiers that may be randomly selected for MPSC review, the Task Force suggested that the MPSC request the completion of something similar to the current "initial survey" when programs are identified in these tiers. As with the quality control tier, the MPSC would then proceed in its review as it does currently, including additional follow-up as believed to be necessary and releasing programs from review as appropriate. With the philosophically different purpose of the quality control tiers versus the quality improvement tiers, the Task Force suggested that, ideally, the MPSC's threshold for release in the quality improvement tiers should be less stringent than it currently is.

The Task Force then focused on refining the boundaries to define the four tiers it had tentatively agreed to. The Task Force evaluated possibilities using an interactive tool developed by SRTR. The tool developed by SRTR for these Task Force discussions allowed the Task Force to evaluate the number of programs that could be anticipated to fall within each tier as the probabilities and hazard ratios that would define each of these tiers were altered. This tool also allowed the Task Force to evaluate the impact of reviewing programs based on their graft *and* patient survival, or their graft *or* patient survival. Working with this tool, the Task Force agreed on the following framework:

- **Quality Control Tier:** >60% probability that hazard ratio > 1.75
 - 100% probability of MPSC engagement
- **Quality Improvement Tier 2:** >60% probability that hazard ratio > 1.25, and the program is not included in the Quality Control Tier
 - 50% probability of MPSC engagement
- **Quality Improvement Tier 1:** hazard ratio > 1.0, and the program is not included in the Quality Control Tier or Quality Improvement Tier 2
 - 10% probability of MPSC engagement
- **Performing As Expected or Better Tier:** hazard ratio ≤ 1.0
 - 0% probability of MPSC engagement

Further Refinement of Task Force Concepts by MPSC Task Force Work Group

With general agreement on this framework, the Task Force requested the formation of an MPSC work group to further consider these recommendations and additional details that should be incorporated in this system. This MPSC Task Force work group was formed, and agreed to review and comment on:

- Total number of tiers
- The hazard ratios to define the boundaries of each of these tiers
- Whether a program's tier placement should be determined by its graft **and** patient survival, or by its graft **or** patient survival.
- How to address small volume programs with this system
- The resulting MPSC and member actions that can be expected for each tier

Total Number of Tiers

The MPSC work group unanimously agreed that the four tiers discussed by the Task Force is reasonable and appropriate. The work group noted that programs will likely move between tiers as time progresses, agreeing with similar sentiments raised by the Task Force. For those programs in the quality improvement tiers, the proposed system provides an opportunity to address any issues it may be having prior to extensive MPSC engagement. If the programs are unaware of their problems, or unable to address them effectively, then it will likely move to a higher tier with an increased likelihood of MPSC engagement. Additionally, the MPSC Task Force work group felt that the process should be as transparent as possible. This will hopefully allow programs to anticipate what tier they may be in so that they may pre-emptively

act to correct any problems that are identified. As such, the relative simplicity of four tiers was thought to be appropriate.

Boundaries to Define Each Tier

The MPSC Task Force work group next discussed the boundaries that define the four tiers. Members commented that the boundaries set by the Task Force yield approximately the same number of identified programs as the current system (72 versus 70, respectively and across all organs); however, the proposed tier system only identifies 12 programs that would then initiate a more involved MPSC interaction. These boundaries also capture the majority of programs prior to CMS involvement; however, it is possible that CMS could issue a condition-level citation before a program is reviewed by the MPSC.

Graft And/Or Patient Survival?

The work group then considered if a program's tier placement should be determined by its graft *and* patient survival, or by its graft *or* patient survival. If programs are reviewed based on its graft or patient survival, its tier placements is dictated by the worst of these two measures; if reviews are based on graft and patient survival, its tier placement is dictated by the better of these two measures. The work group initiated its discussion on this topic with focus on kidney transplants. Due to potential disparity between graft and patient survival at kidney programs (i.e., a program could have less than expected graft survival that may not necessarily be reflected in its patient survival metrics), and the lost transplant opportunities that may persist if substandard kidney graft survival is not addressed, the work group believed that kidney programs should be evaluated based on its graft survival or patient survival.

The work group acknowledged that the disparity between these two metrics is much less for the other organs, and it debated which approach would be most reasonable for monitoring the outcomes of liver, heart, and lung transplants. Work Group members suggested that a similar approach should be taken for liver programs considering that liver patients who experience primary graft failure automatically qualify as Status 1A. Additionally, failure to focus on a liver program's graft survival may result in missed warning signs that a program is stepping outside its capabilities in terms of patient and organ selection, and at the expense of multiple transplantable livers being allocated to a single patient. Graft loss and patient loss are both important events, and from the MPSC's quality perspectives, asking questions about graft survival and patient survival assess separate issues. If this proposed system for transplant program outcomes reviews yields the desired behavior changes such that transplant programs are performing more transplants with organs and patients that have higher risk factors, it will be difficult to recognize programs that may be struggling due to its efforts to perform more transplants if programs' tier placement is decided by graft and patient survival. Ultimately, the work group agreed that liver programs should be evaluated by graft or patient survival.

Regarding heart and lung programs, the work group agreed that either approach would probably render the same results. As such, and for the sake of simplicity through consistency, the work group also agreed that heart and lung programs should be evaluated based on its graft or patient outcomes.

Small Volume Program Considerations

The work group proceeded to consider how small volume programs would be impacted by this proposed system, and if additional considerations for these programs were necessary. Discussion highlighted that the elevated hazard ratios of the proposed system, combined with the statistical power that results from evaluating the lower number of transplants at small volume programs, means that the top, quality control tier would rarely include small volume programs. The only way to increase the statistical power to assess small volume programs is to accept a greater number of false positives. Alternatively, to make it more likely that small volume programs might fall in the quality control tier, the hazard ratios to define each boundary or the probability associated with those hazard ratios could be lowered; however, this will increase the total number of programs in the quality control tier.

This understanding prompted additional questions regarding if additional considerations should be applied in evaluating outcomes at small volume programs. Fundamental to this discussion is how much time and energy the OPTN wants to spend reviewing programs that account for a small percentage of the total transplants performed.

OPTN Bylaws Appendix D.11.A (Transplant Program Performance) currently states that - in addition to the Bayesian analysis - small volume programs are reviewed as follows:

For programs performing 9 or fewer transplants in a 2.5 year period, the MPSC will review a transplant program if the program has one or more events in a 2.5 year cohort.

This additional consideration was included in the Bylaws when the change to the Bayesian system was made because MPSC members expressed strong concerns that small volume programs would be overlooked. SRTR referenced past analyses that indicated small volume programs with one event were unlikely to be true positives. The work group replied that this aligned with their assumptions, and suggested that the OPTN should be careful about special considerations for small volume programs. It doesn't seem particularly valuable to spend a disproportionate amount of time reviewing programs that account for a small percentage of the total number of transplants performed, combined with the likelihood of false positives. Along these same lines, there is more power in the statistics that highlight the underperformance of larger volume programs.

The MPSC Task Force work group considered multiple suggestions for the review of small volume programs, including;

- Placing small volume programs that meet current Bylaws outcomes review requirements in the top quality improvement tier
- Excluding small volume programs from the tier review system and define a number of events that would initiate a quality improvement review
- Increasing the probability for review of small programs to make it more likely that they would come under review (e.g., in the lowest quality improvement tier, increase the probability of MPSC review from 10% to 20% for low volume programs)
- Add a narrow, 5th tier for the purpose of focusing on a few more small volume programs

The work group ultimately felt that the first two alternatives could not be supported, but they were interested in studying the impact of the concepts expressed in the last two bullets above. In response to these requests, the SRTR analyzed the impact of doubling the quality improvement tier sampling frequency for those programs that performed nine or fewer transplants in the 2.5 year cohort; and the impact of adding a narrow, 5th tier. The work group reviewed these results, but was comfortable with its decision to keep the system relatively simple with four total tiers for review, and that increasing the number of MPSC engagements in the quality improvement tiers would not necessarily result in an overall greater value to the system. Ultimately, the work group preferred the simplicity of a system that did not include additional considerations for small volume programs and that treated all programs similarly.

MPSC and Member Actions Corresponding to Each Tier

With agreement around the number of tiers, the boundaries that define each of those tiers, and whether special considerations were necessary for small volume programs, the MPSC Task Force work group considered MPSC and member actions that would correspond to each tier. During this discussion, the Work Group referenced a flow chart that illustrated the MPSC's current approach for outcomes review (Figure 5).

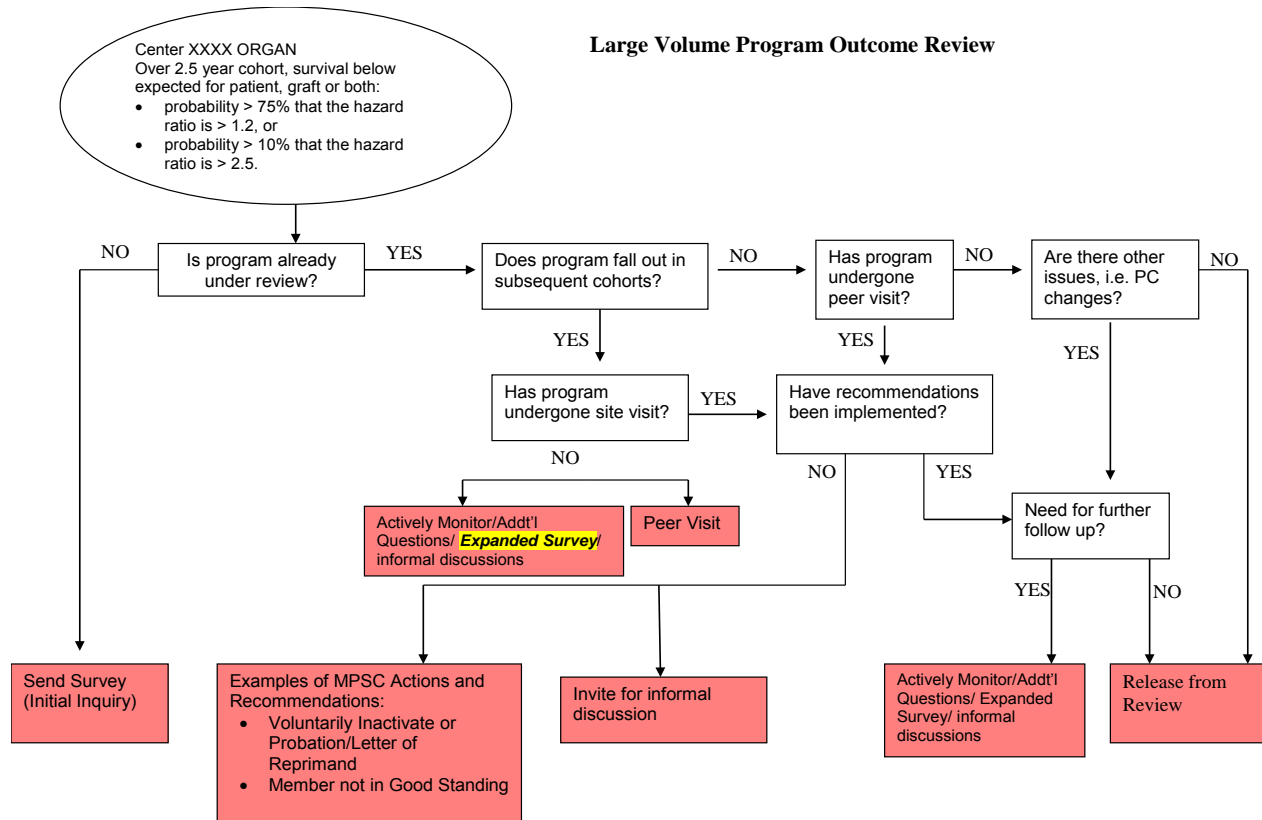


Figure 5- Current MPSC program outcomes review flowchart.

The work group agreed with the Task Force’s preliminary recommendation that programs captured by the highest, quality control tier should enter this process at the point where expanded surveys currently exist. The results of the expanded survey are reviewed by PAIS, and PAIS review would then determine the next steps to be taken, including releasing the program from review. The work group also agreed with the Task Force’s recommendation that programs randomly selected from the quality improvement tiers should be sent an initial survey. All programs engaged by the MPSC would then proceed through this process as done currently.

Work group discussion also indicated that the current operational rule for programs that have been released in the last review cycle should also apply in this new system. That is, if the MPSC releases a program from review, then the program will be excluded from further review through the proposed system for the two subsequent cohorts immediately following its release from MPSC review.

Tier Names

Finally, work group discussion about this process prompted considerations about how each tier should be labeled. These are critical considerations because the overarching purpose of these efforts is to shift the transplant community’s perspective of the OPTN’s outcome review process in hopes that this will lead to a behavioral shift whereby transplant programs will feel comfortable pursuing transplants that involve patients and organs with greater risk factors. The work group agreed that it is critical that the labels selected for these tiers align with and further support this shift in perspective.

For the top, quality control tier, the work group opined that the label should reflect that all of these programs will be reviewed by the MPSC. Additionally, the top tier label should not reference a speed at which these reviews might take place (e.g., “rapid review”), nor should it reference “quality” for the sake of avoiding the need to differentiate between a top tier “quality review” and the quality improvement focus of the lower tiers. The work group believed that the label for the top tier should reflect an active intervention and the labels for the lower tiers should reflect the routine quality improvement processes that this

proposed system is aiming to establish. The work group ultimately agreed to use “Expanded Program Performance Review” for the highest tier, and “Routine Program Review” for the lower, quality improvement tiers. “Performance” was excluded from the quality improvement tiers to encourage more focus on the routineness and quality improvement aspects of these tiers.

Review of and Action on MPSC Task Force Work Group Recommendations

After the MPSC Task Force work group worked through these decisions, a joint Task Force/MPSC Task Force Work Group teleconference was held. During this call, the MPSC Task Force work group presented its recommendations for the Task Force’s consideration, and to assure its recommendations aligned with the Task Force’s intended direction. Summarizing the main recommendations presented by the MPSC Task Force work group:

- The work group agreed with the four-tier system developed by the Task Force
- The work group agreed with the Task Force’s placement of the hazard ratio lines to define the boundaries of each tier
- A program’s tier placement should be determined by its graft *or* patient survival, as compared to its graft *and* patient survival
- No separate, additional considerations for evaluating small volume programs
- Label quality control tier as “Expanded Program Performance Review”
 - Programs in this tier will be expected to complete an expanded survey, the results of which will determine any further MPSC action
- Label the quality improvement tiers as “Routine Program Review”
 - Programs randomly selected in these tiers will be expected to complete an initial survey, the results of which will determine any further MPSC action
 - A greater focus on program’s quality improvement systems & efforts will drive the MPSC’s decisions to release from review

Figures 2 and 3 further illustrates these recommendations, and what each tier space looks like.

Supplementing these recommendations, SRTR and UNOS analyzed the number of program components that could be expected in each of the proposed tiers, and how many programs would likely be eligible for review in any given cohort after the MPSC applies the following operational rules, or filters, for conducting reviews of transplant program outcomes (Table 1):

- Is the program active?
- Is the program already under review?
- Was the program released in the last 2 meeting cycles?
- Has the program had a death or graft failure since the most recent release from review?

Organ	New CMS (O/E>1.85)	New CMS Condition-Level	Current MPSC	Post MPSC filters	Expanded Reviews Tier 1	Post MPSC filters	Routine Reviews Tier 2	Post MPSC filters	Routine Reviews Tier 3	Post MPSC filters	Total Routine Reviews	Post MPSC filters	Total Reviews - All Tiers	Post MPSC filters
HR	6	2	12	7	2	1	7	2	4.5	3.7	11.5	5.7	13.5	6.7
KI	17	8	32	12	8	2	17	9.5	8.1	7	25.1	16.5	33.1	18.5
LI	3	2	14	4	1	0	10	5	4.9	4.5	14.9	9.5	15.9	9.5
LU	3	1	12	7	1	0	6.5	4	2	1.4	8.5	5.4	9.5	5.4
ALL	29	13	70	30	12	3	40.5	20.5	19.5	16.6	60	37.1	72	40.1

Table 1- Comparison of Counts Using Proposed System: Adult Program Components w. MPSC Operational Rules

The Task Force, along with OPTN leadership on the call, did not raise any concerns with these recommendations, and indicated their support for presenting these recommendations for the MPSC and Board’s consideration. Following this discussion, the full-MPSC received the same update with the Task Force’s feedback. MPSC members raised similar concerns as before, but the MPSC indicated its support for presenting the work group’s recommendations for the Board to obtain additional feedback.

The MPSC Vice Chair presented these recommendations for the Board at its June 2016 meeting. Board members raised similar questions and concerns that came up during MPSC and Task Force discussions. Upon the conclusion of the Board’s discussion, it suggested that the MPSC consider the best approach for obtaining the community’s feedback on these recommendations.

Following the June 2016 Board meeting, the MPSC met to discuss the best path forward on these efforts. Members raised similar concerns as before, but ultimately, the committee supported a motion to distribute these recommendations for public comment as compared to presenting them during regional meetings for informal feedback (17 support, 6 oppose, 0 abstain). Subsequent to that vote, the committee supported a motion to distribute for public comment the modified Bylaws included in this proposal (22 support, 1 oppose, 0 abstain).

Additional Considerations

The MPSC is sponsoring another proposal this public comment cycle that also impacts how the OPTN reviews transplant program outcomes- *Transplant Program Performance Measures (Outcome Measures)*. The MPSC developed these proposals independently of one another, and each of these proposals separately focus on different aspects of the OPTN transplant program outcomes review system. This proposal, *Proposed Changes to the OPTN Transplant Program Outcomes Review System*, is recommending systematic changes to how programs (for every organ type) are identified for MPSC review. The *Outcome Measures* proposal recommends excluding certain types of kidney transplants from hazard ratio analyses to determine whether the program should be reviewed by the MPSC. These proposals can be implemented together to supplement the impact each proposal may have, but as these proposals are not interconnected, there is nothing in either proposal that would prohibit them from being implemented separately or in isolation. Additionally, both of these proposals are also separate from the COIIN pilot project currently being undertaken by UNOS.

Finally, there are some additional considerations worth noting when reviewing the content of this proposal. Some MPSC and Board members have questioned if this approach makes substantial enough changes to achieve the desired goal of more transplants. Summarizing some of the points that have been raised:

- Many discussion participants noted that improved risk-adjustment analyses are critical to improving the OPTN’s outcomes review system. Imperative to improved risk-adjustment is the collection of more data. All agreed that discussing and determining the additional data elements necessary for improved risk-adjustment was outside the scope and timeframe of the Task Force’s charge; however, the

OPTN will need to undertake these discussions to make greater improvements to the transplant outcomes review system and to further increase the community's confidence in this system.

- Relying on random selection may not necessarily yield reviews of the programs that could benefit from those reviews the most. Considering multiple review cycles over several years, it is possible that a program with hazard ratio metrics near the very top of Routine Program Review Tier 2 could maintain this level of performance and never be reviewed, while during the same period another program that remains near the very bottom of the Routine Program Review Tier 3 could be reviewed multiple times.
 - Patient safety must remain in focus as the impact of these changes are evaluated. The MPSC has had legitimate concerns with a number of the programs it reviews, and some of these programs would not come to the MPSC's attention in the same immediate manner with this new system.
- Factors other than OPTN performance reviews have a large impact on transplant program behavior such that they may limit the impact these changes ultimately have.
- Behavior modeling studies would be helpful to better understand the impact these changes may have.
- The OPTN should do more to evaluate multi-organ outcomes, and this proposal does not address or consider these deficiencies in the current outcomes review system.
- Pre-transplant metrics considerations are absent from this proposal, and are thought to be critical.

In considering this proposal, and in response to some of these comments, it is important to recognize the context in which it was developed. This proposal is intended to make changes that could be realized in a short time frame, and while the OPTN begins to undertake larger systematic changes, namely the COIIN project. Working within that context, the Task Force and the MPSC believe that the proposed outcomes review system has the potential to streamline the MPSC's reviews so that they are more focused on programs having real difficulties with their outcome results, while also establishing regular quality improvement discussions, from which some lessons may be able to be leveraged across the entire transplant community.

How well does this proposal address the problem statement?

The MPSC believes that increasing the hazard ratio thresholds that dictate its reviews of transplant program performance, along with routine, quality improvement reviews of transplant programs that are driven by the random selection of programs with a hazard ratio greater than 1.0, will change transplant hospitals' perspective of the current OPTN transplant outcomes review system. As a result, programs will have greater confidence that transplanting patients and organs with increased risk factors will not yield a punitive MPSC review. The MPSC acknowledged that perceptions will not change immediately. The MPSC also acknowledged that reviews by CMS and insurance companies also impact the transplant programs' decisions, and that successes realized from modifying the OPTN's transplant program review system may be limited by the reviews (and the corresponding implications) conducted by these organizations.

Which populations are impacted by this proposal?

This proposal is intended to increase the overall number of transplants, and to encourage safe, but less-conservative, selection of patients and organs for transplant. If this intent is realized, all transplant candidates are expected to be impacted by this proposal, particularly those with greater risk factors who remain on the waiting list for an extended period or who may not be listed for transplant today because of the complexity of their particular diagnosis and the somewhat lower level of transplant success that might be expected.

How does this proposal impact the OPTN Strategic Plan?

Increase the number of transplants: Improved confidence in how transplant outcomes are reviewed is intended to promote the transplantation of organs that are often discarded today.

Improve equity in access to transplants: Improved confidence in how transplant outcomes are reviewed is intended to promote the addition of candidates to the waitlist that would be suitable for transplant that may not be offered that opportunity today because of the complexity and somewhat lower level of transplant success with their particular diagnosis.

Improve waitlisted patient, living donor, and transplant recipient outcomes: Improved confidence in how transplant outcomes are reviewed is intended to expand the type of donors that would be accepted for certain patients, thereby giving waitlisted patients a greater opportunity for transplant, and therefore an improved outcome as compared to remaining on the waitlist.

Promote living donor and transplant recipient safety: Programs that have been under MPSC review for a considerable time under the current system may not necessarily be reviewed with the proposed system. Some may argue that this would be to the detriment of transplant recipient safety at those program. The MPSC and the Task Force believes that these programs would likely progress upwards through tiers, and eventually be reviewed by the MPSC if the underlying causes of the program's underperformance are not identified and addressed. Ultimately, this potential downside is thought to be of a smaller magnitude than the gains this proposal aims to achieve, and thus, necessary to realize those gains.

Promote the efficient management of the OPTN: There is no impact to this goal.

How will the OPTN implement this proposal?

Implementation of this proposal will require the SRTR to revise the process for running the MPSC Outcomes reports used to identify programs for outcomes review. If the proposal is approved by the OPTN Board of Directors in December 2016, the MPSC would implement the new process for identification of programs for outcomes review in the June 2017 MPSC Outcomes Reports. The SRTR produces these reports at the same time as the public program specific reports.

The MPSC has not explored all of the implications of the four tiered process on the current MPSC process of review of outcomes; however, much of the current MPSC review process will remain the same. The expectation is that those programs that are randomly selected from Tier 2 and Tier 3 will be required to complete the current Initial Outcomes Questionnaire ("initial survey") or something very similar. The MPSC will request completion of the more detailed Expanded Outcomes Questionnaire ("expanded survey") for those programs that are identified in Tier 1. Any additional review following the initial inquiry will be determined by the unique circumstances of the program under review.

The MPSC currently has established operational rules to further evaluate which of the programs that have been flagged under the MPSC criteria should receive an inquiry. Currently, the MPSC does not send an initial inquiry for graft or patient outcomes if the program is not active at the time of receipt of the SRTR reports, is already under review, or has been released from review for outcomes within the last two MPSC meeting cycles. The application of these operational rules to the proposed system such as whether the operational rules will be applied prior to or after the random selection of programs in Tier 2 and Tier 3 has not yet been determined. In addition, the MPSC has not determined how a program review that was initiated under Tier 2 or Tier 3 will be affected by movement of the program to Tier 1 in a subsequent SRTR report. The MPSC will continue to evaluate these questions.

UNOS will provide additional educational opportunities for members to become more familiar with the OPTN's new performance monitoring system. This proposal will not require programming in UNetSM.

How will members implement this proposal?

Transplant Hospitals

Members will not be required to submit additional data or make any changes to current processes under this proposal. Transplant hospitals are expected to be responsive to the MPSC if engaged in an outcome performance review.

Will this proposal require members to submit additional data?

No, this proposal does not require additional data collection.

How will members be evaluated for compliance with this proposal?

This proposal establishes a new process by which the MPSC will conduct transplant program performance reviews. As with current transplant program performance reviews, members are expected to be responsive to the MPSC if engaged in an outcome performance review. The proposal does not directly create any new member compliance implications.

How will the sponsoring Committee evaluate whether this proposal was successful post implementation?

The committee will use this system for a period of three years, at which time they will review all data available to them, and decide whether it is appropriate to continue with this system or if additional changes are necessary. The MPSC will also review a number of factors, in as near to real time as possible, to evaluate the criteria's effect on behavior and outcomes. If unacceptable changes are observed, the MPSC will readdress these proposed before the three year trial ends. The following are among the factors to be reviewed:

- The recovery and discard rates pre and post implementation, stratified by organ-specific donor quality measures (e.g., KDPI groupings).
- Post-transplant graft and patient survival to monitor for significant drops in survival rates to unacceptable levels. The MPSC recognizes that if higher risk transplants are performed more often, the average national survival rate may decrease, and the MPSC will review these data closely.
- Changes in transplant programs' listing behaviors.
- Number and length of MPSC reviews.

Policy or Bylaws Language

Proposed new language is underlined (example) and language that is proposed for removal is struck through (~~example~~).

D.11 Additional Transplant Program Requirements

A. Transplant Program Performance

Appendix D.11.A does not apply to VCA transplants.

~~The MPSC will conduct reviews of transplant program performance to identify underperforming transplant programs and require the implementation of quality assessment and performance improvement measures. One measure of transplant program performance is triggered through a review of the one-year graft and patient survival rates. The MPSC utilizes performance metrics produced by the Scientific Registry of Transplant Recipients (SRTR) as the principal tool to identify transplant programs that have lower than expected outcomes.~~

~~For programs performing 10 or more transplants in a 2.5-year period, the MPSC will review a transplant program if it has a higher hazard ratio of mortality or graft failure than would be expected for that transplant program. The criteria used to identify programs with a hazard ratio that is higher than expected will include *either* of the following:~~

- ~~1. The probability is greater than 75% that the hazard ratio is greater than 1.2.~~
- ~~2. The probability is greater than 10% that the hazard ratio is greater than 2.5.~~

~~For programs performing 9 or fewer transplants in a 2.5-year period, the MPSC will review a transplant program if the program has one or more events in a 2.5-year cohort.~~

~~The MPSC review will be to determine if the higher hazard ratio or events can be explained by patient mix or some other unique clinical aspect of the transplant program. If a program's performance cannot be explained by patient mix or some other unique clinical aspect of the transplant program, the program, in cooperation with the MPSC, will adopt and promptly implement a plan for quality improvement. The member's failure to adopt and promptly implement a plan for quality improvement will constitute a violation of OPTN obligations.~~

The MPSC primarily uses performance metrics produced by the Scientific Registry of Transplant Recipients (SRTR) to identify transplant programs for performance and routine quality improvement reviews.

Transplant programs will be placed in one of four tiers based on their one-year post-transplant mortality and graft failure hazard ratios produced by the SRTR. Comparing a transplant program's tier placements as determined by the analyses of its mortality hazard ratio and graft failure hazard ratio, the MPSC may review a transplant program depending on the highest tier it is placed. The four review tiers are defined as follows:

1. Expanded Program Performance Review Tier 1, the highest tier, includes transplant programs with a greater than 60 percent probability that the program's mortality or graft failure hazard ratio is greater than 1.75. Every program in this tier will be reviewed by the MPSC.
2. Routine Program Review Tier 2 includes transplant programs with a greater than 60 percent probability that the program's mortality or graft failure hazard ratio is greater than 1.25, excluding those programs that were included in the Expanded Program Performance Review Tier 1. Fifty percent of the programs in this tier will be randomly selected for MPSC review.

50 3. Routine Program Review Tier 3 includes transplant programs with a mortality or graft failure
51 hazard ratio greater than 1.00, and that were not included in the Expanded Program
52 Performance Review Tier 1 or Routine Program Review Tier 2. Ten percent of the programs
53 in this tier will be randomly selected for MPSC review.

54
55 4. As Expected or Better Tier 4, the lowest tier, includes transplant programs with a mortality or
56 graft failure hazard ratio less than or equal to 1.00. Placement in this tier will not prompt
57 MPSC review.

58
59 As part of these reviews, the MPSC may, but is not limited to, any of the following:
60

- 61 • Request that a member adopt and implement a plan for quality improvement. The member's
62 failure to adopt and implement a plan for quality improvement will constitute a violation of
63 OPTN Obligations. As part of this process the MPSC may c
- 64 • Conduct a peer visit to the program, at member expense. The MPSC may also r
- 65 • Require, at its discretion, that the member participate in an informal discussion. The informal
66 discussion may be with the MPSC, a subcommittee, or a work group, as determined by the
67 MPSC. The informal discussion will be conducted according to the principles of confidential
68 medical peer review, as described in Appendix L of these Bylaws. The informal discussion is
69 not an adverse action or an element of due process. A member who that participates in an
70 informal discussion with the MPSC is entitled to receive a summary of the discussion. The
71 MPSC may r
- 72 • Recommend that a member inactivate a program or a component of a program or withdraw
73 its designated transplant program status based on patient safety concerns arising from
74 review of the program's graft and patient survival. If the program fails to inactivate or
75 withdraw its designated transplant program status when the MPSC recommends it do so, the
76 MPSC may recommend that the Board of Directors take appropriate action as defined in
77 Appendix L: Reviews, Actions, and Due Process of these Bylaws.

78 #