

MPSC Performance Monitoring Series: Bayesian Methods for Assessing Transplant Program Performance

September 23, 2014

2 p.m. – 3 p.m. ET

Audio Information:

Dial-in Number: 1-866-901-6455

Access Code: 727-560-237

Audio PIN: Shown after joining the training

OPTN



[Chad Southward]

Hello and welcome to today's event: Bayesian Methods for Assessing Transplant Program Performance. This is the first event in an MPSC Performance Monitoring instructional series. I am Chad Southward, a Curriculum Development Instructor here at UNOS, and I will be helping run today's event. I also have Jayson Cooke, one of our E-learning instructors, in the room to help with the technical details.

Reminder

Technical Difficulties

Citrix Global Support (800) 263-6317

Recording in Progress

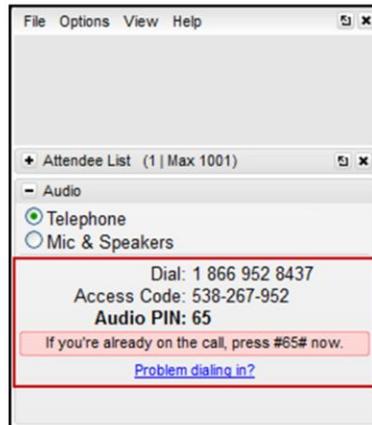
OPTN



[Chad Southward]

Let's get started with a few Citrix webinar reminders. If you experience any technical difficulties, try logging out of Citrix and logging back in. If the problem persists, contact Citrix Global Support at (800) 263-6317. Also please note that this event is being recorded and will be available within the next two weeks on the OPTN site and Transplant Pro.

Citrix GoToWebinar Tools



[Chad Southward]

If you have a microphone and speaker available on your computer, you can use the Mic and Speakers option. This will leave your phone available for calls. If you have dialed in using the telephone, make sure the telephone option is selected on your Audio Panel. Also, enter your Audio Pin now if you have not already done so. The Pin is highlighted in red on your Audio Panel.

Presenters



Jon Snyder, Ph.D., MS



**Sharon Shepherd, J.D.,
MSN, RN**

OPTN



[Chad Southward]

Our presenters for today's event are:

Dr. Jon Snyder – Director of Operations and Senior Epidemiologist at the Scientific Registry of Transplant Recipients.

Sharon Shepherd – Transplant Systems Performance Manager in the Membership Department at UNOS.

Objectives

After this session, you will be able to:

- Describe the reasoning for moving to Bayesian methodology
- Compare the previous algorithm for identifying programs that may be underperforming to the new Bayesian algorithm
- Summarize how the Bayesian methodology affects large, medium, and small volume programs

OPTN



[Chad Southward]

Jon will shortly fill you in on what he will be presenting but here are the objectives for today's session.

After this session, you will be able to:

Describe the reasoning for moving to Bayesian methodology

Compare the previous algorithm for identifying underperforming programs to the new Bayesian algorithm

Summarize how the Bayesian methodology affects large, medium, and small volume programs

SRTR Resources

Publications	<i>A Scientific Registry of Transplant Recipients Bayesian method for identifying underperforming transplant programs</i> <i>Bayesian methods for assessing transplant program performance.</i>
Presentations	<i>Improving the SRTR Methodology Used to Identify Potentially Underperforming Transplant Programs in the United States</i>
Other Documents	SRTR Bayesian Frequently Asked Questions (FAQ)

OPTN



[Chad Southward]

There are several resources available relevant to this event. The resources include Bayesian publications, presentations and a Frequently Asked Questions document. All of these resources are located on the SRTR website, srtr.org. You should have received an e-mail before the event pointing you to these useful resources.

Today's Session

Introduction

Bayesian Methodology

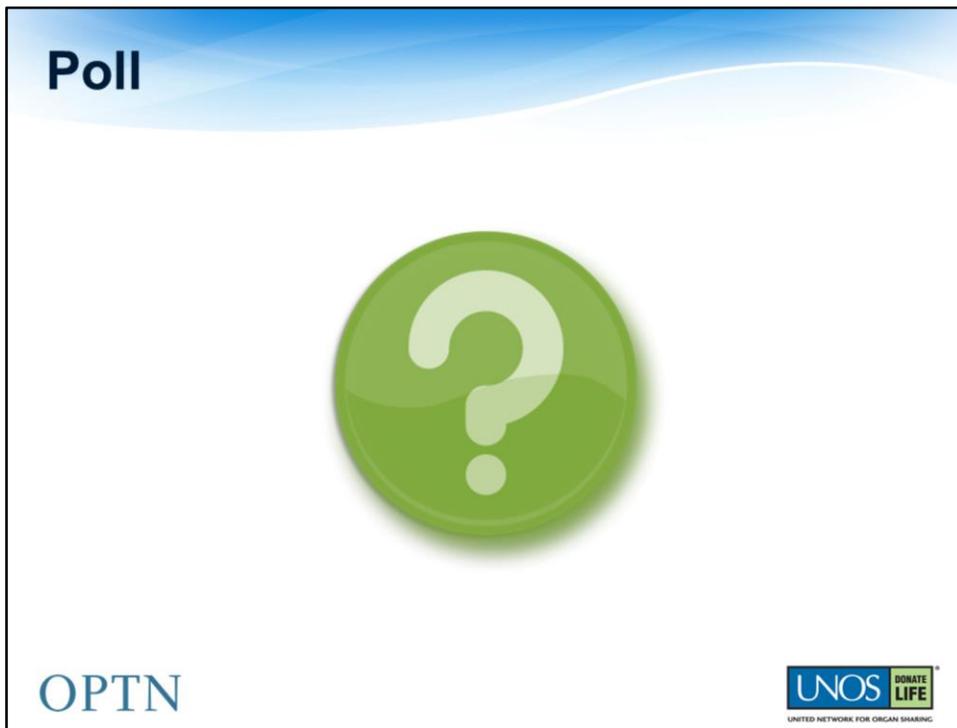
Questions & Answers

OPTN

UNOS UNITED NETWORK FOR ORGAN SHARING **DONATE LIFE**

[Chad Southward]

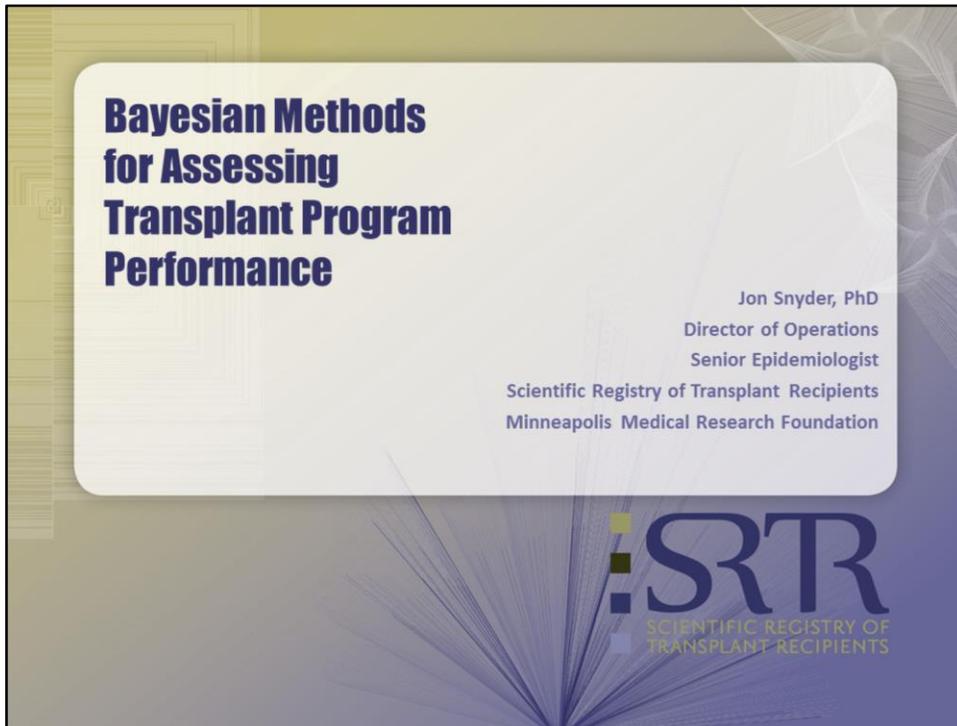
During today's event, we will discuss in-depth the Bayesian methodology for assessing transplant program performance. Throughout today's session, we will open up our expert presenters to questions from the audience. At any point during today's event, feel free to type your questions for our presenters into the Citrix Questions panel. There is no need to wait to submit your questions until the end.



[Chad Southward]

OK, folks, it is time for our attendance polling question. You will see a question appear on your screen. Please select the number range of people watching in the room, including yourself. This gives us an idea of the number of attendees that prefer to watch instructional events in groups. We also get a better idea of the actual number of attendees watching this offering.

We appreciate you taking the time to answer. The polling is now complete.



[Chad Southward]

So now it is my pleasure to hand the reigns over to Dr. Jon Snyder. We appreciate your patience as we switch the Presenter function over to Jon.

[Jon Snyder]

Good afternoon. My name is Jon Snyder and I serve as Director of Operations and Senior Epidemiologist for the Scientific Registry of Transplant Recipients.

What I'll cover today...

- High-level review of the difference between the current methodology for program assessment and the Bayesian methodology
- How will the Membership and Professional Standards Committee (MPSC) of the OPTN identify programs for review using the Bayesian method?
- Describe the timeline for transition

10



[Jon Snyder]

Here is a list of what I'll cover on today's webinar. I'll provide a high-level review of the difference between the current methodology for program assessment and the Bayesian methodology.

I'll cover how the Membership and Professional Standards Committee (MPSC) of the OPTN will identify programs for review using the Bayesian method. And I'll describe the timeline for the transition to the Bayesian methodology for program assessment.

The Charge and the Challenge...

The Charge to the SRTR

- The reporting requirements in §121.11(b) of the Final Rule state that the SRTR contractor shall disseminate for free over the internet timely and accurate program-specific information on the performance of transplant programs.
- These reports shall be updated no less frequently than every six months.
- The reports shall include **risk-adjusted** graft and patient survival following the transplant.

The Challenge of the MPSC

- Organ Procurement and Transplantation Network (OPTN) Bylaws, Appendix D, Section D.10: The MPSC will review a transplant program if it has a **low** survival rate compared to the **expected** survival rate for that transplant program.

11



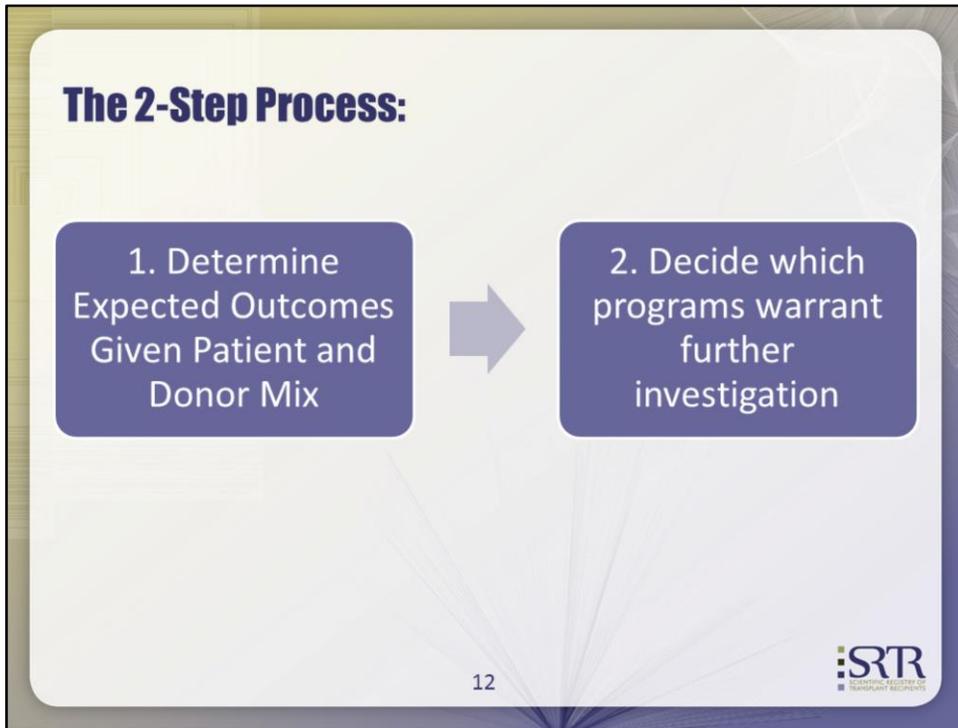
[Jon Snyder]

Let's begin by considering both the charge to the SRTR and the challenge the MPSC faces.

The reporting requirements of the Final Rule state that the SRTR contractor shall disseminate for free over the internet timely and accurate program-specific information on the performance of transplant programs. These reports shall be updated no less frequently than every six months, and the reports shall include **risk-adjusted** graft and patient survival following the transplant.

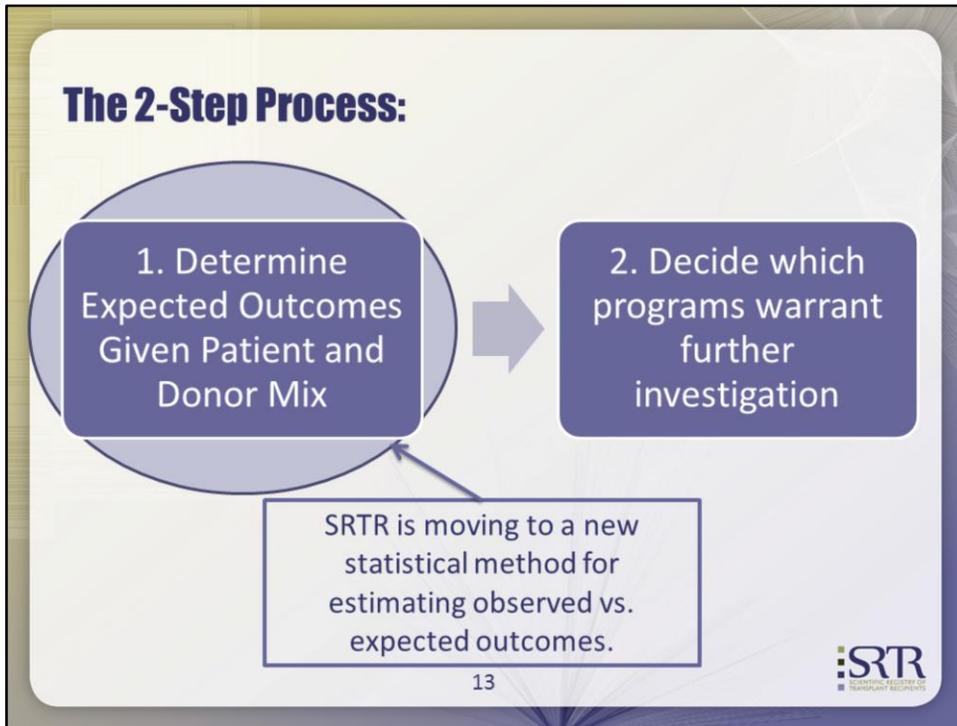
The OPTN bylaws state that the MPSC will review a transplant program if it has a **low** survival rate compared to the **expected** survival rate for that transplant program.

You'll notice I've highlighted a few key words on this slide. The SRTR must identify expected risk-adjusted outcomes, and the MPSC must decide on the definition of "low". In other words, we first must decide on a measure that attempts to account for different mixes of patients at different programs, and then we must decide which programs appear to be underperforming according to that measure.



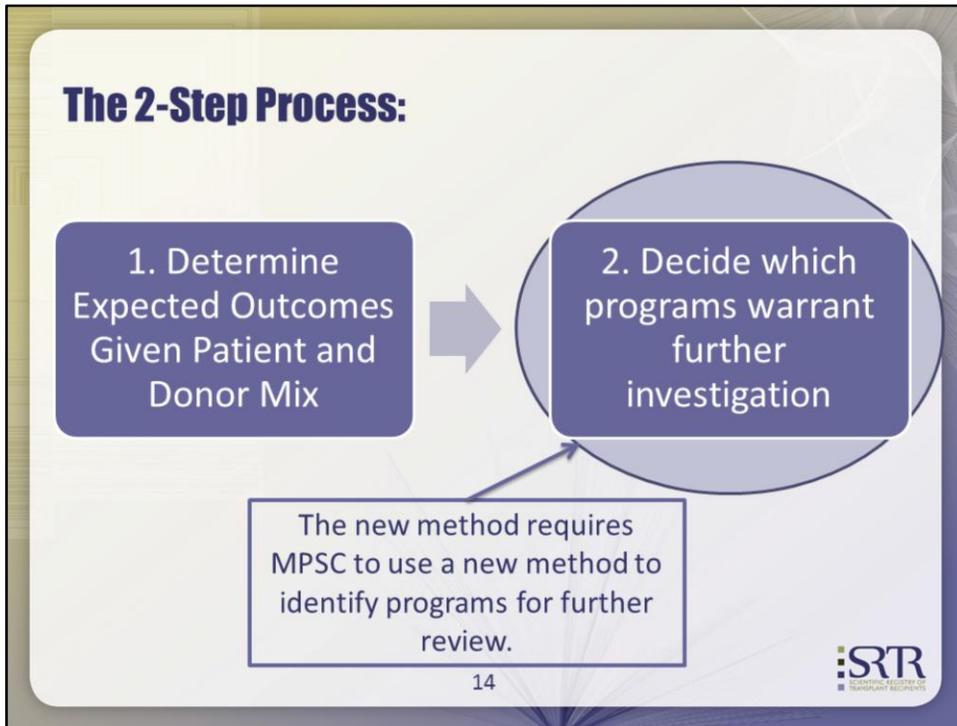
[Jon Snyder]

Here we more clearly see this two-step process. We must first determine expected outcomes given patient and donor mix at each transplant program. Then the MPSC must decide which programs warrant further investigation.



[Jon Snyder]

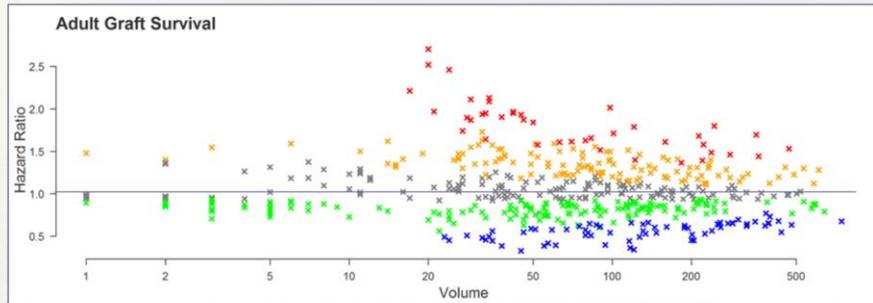
SRTR is changing the statistical methodology used to estimate observed vs. expected outcomes at each transplant program. This is the change to a Bayesian statistical methodology.



[Jon Snyder]

This change will require the MPSC to use a new method to identify programs for further review.

The Challenge: Which Programs to Review?



15



[Jon Snyder]

This figure illustrates the challenge before the MPSC. Each point in this cloud of points is a transplant program's performance evaluation for first-year adult graft survival outcomes. The x-axis shows program volume with larger programs to the right and smaller programs to the left. The y-axis shows each program's estimated hazard ratio. A hazard ratio of 2 would indicate graft failure rates at twice the expected rate, whereas a hazard ratio of 0.5 would indicate graft failures at half the expected rate. The horizontal line at 1 indicates where programs are performing exactly as expected given their patient and donor mix. The points are color-coded simply for illustrative purposes. The red points are programs that have relatively poor outcomes and the blue points are programs with relatively good outcomes. The MPSC's challenge is to decide which points are worthy of further evaluation.

MPSC Criteria Goals

Correctly identify as many underperforming programs as possible without . . .

1. Falsely identifying programs that do not need review
2. Identifying too many programs overall

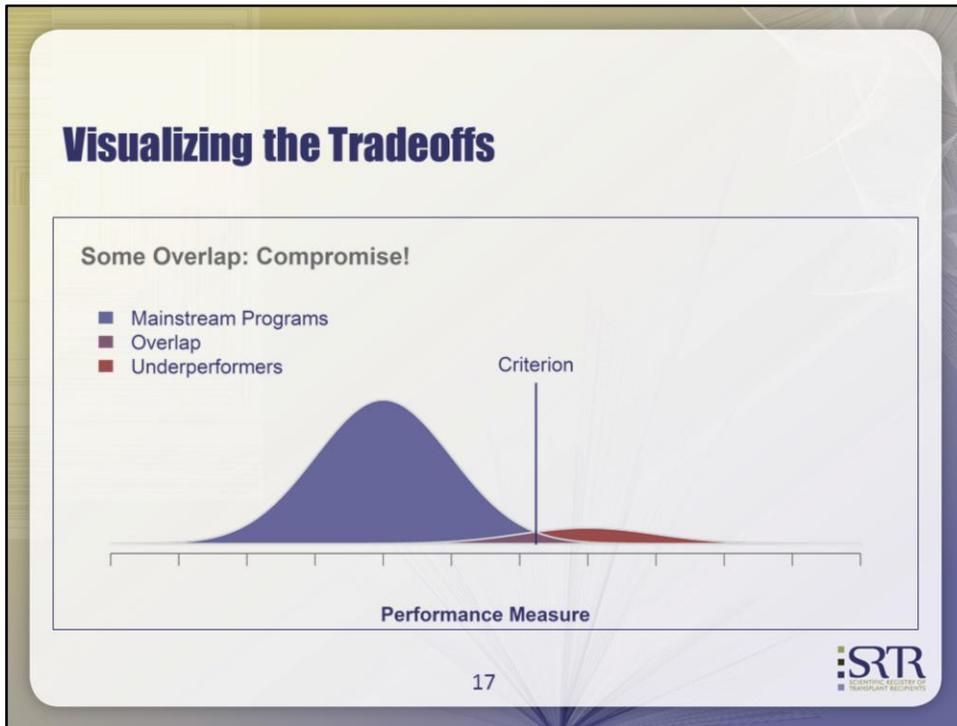
Only the MPSC can determine the correct balance between correctly identifying underperforming programs and avoiding falsely identifying programs.

16



[Jon Snyder]

The goals of a good set of criteria for identifying underperforming programs include not falsely identifying programs, and not identifying too many programs overall. Only the MPSC can determine the correct balance to achieve these goals as I'll illustrate on the following slides.



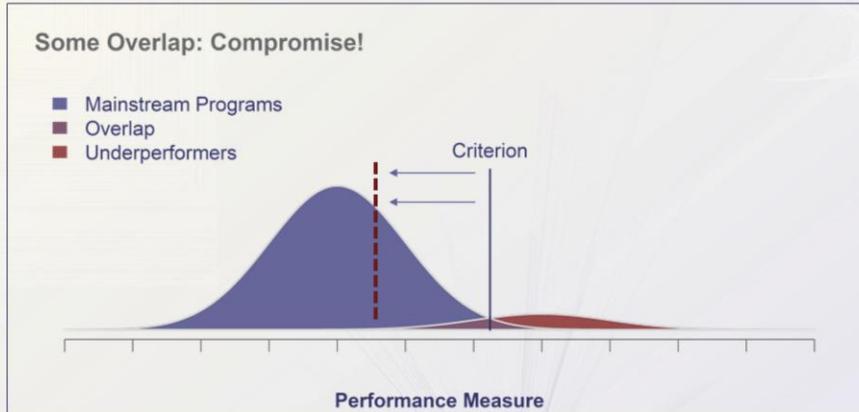
[Jon Snyder]

When we evaluate program performance, we can never know with absolute certainty whether a program truly has a performance issue or whether worse than expected outcomes may just happen due to chance or some other factors not included in the risk adjustment models. We can only observe each program for a period of time, analyze their data, and make a decision as to whether the program warrants further investigation.

In this illustration, let's assume we know for certain there are 2 groups of programs: programs without performance issues that we've labeled "Mainstream Programs" shown by the blue bell-shaped curve and programs that are underperforming shown by the smaller red curve. We believe that most programs do not have performance problems, which is why the blue curve is larger than the red curve. The x-axis is a hypothetical performance measure, with worse performance indicated by a shift to the right. Remember that we cannot know absolute truth when we evaluate performance, we can only observe the program for a period of time. Therefore, some of the mainstream programs may look better or worse simply by chance as indicated by the spread in the blue bell-shaped curve. Similarly, programs that are truly underperforming may have outcomes that look OK over the time-period observed simply by chance, so we see some spread in their assessment as well.

Notice that these two bell-shaped curves overlap. The challenge becomes where to draw the threshold line, or the criterion to identify underperforming programs. If we draw the line as you see here, it appears we will only identify about 2/3rd of the underperforming programs as shown by the amount of the red curve to the right of the criterion line. You'll notice however, that this would also incorrectly identify a small group of mainstream programs (the blue curve to the right of the criterion line) and would miss about 1/3rd of the underperforming programs as shown by the red curve to the left of the criterion line.

Visualizing the Tradeoffs



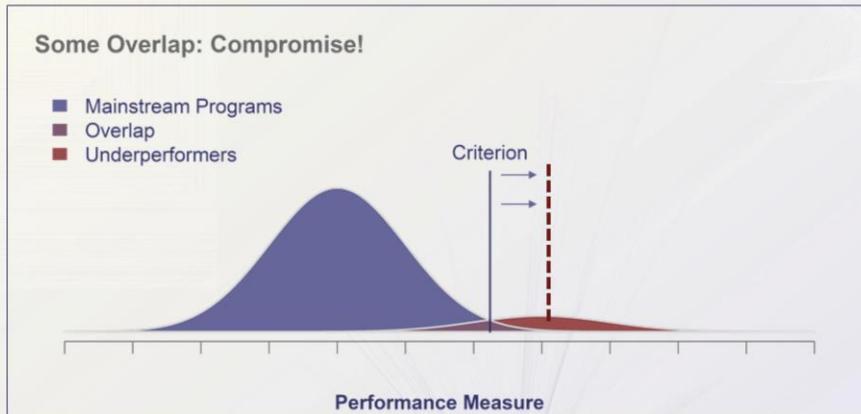
18

SRR
SOUTHERN REGIONAL ORGANIZATION OF
TRANSPLANT RECIPIENTS

[Jon Snyder]

We could move the criterion line to the left, or “lower the bar”. If we moved the line to this location we will be almost certain to identify all true underperforming programs, but we’ll also identify about 1/3rd of mainstream programs incorrectly. This creates a situation that is likely unacceptable to the transplant community and to the MPSC.

Visualizing the Tradeoffs



19



[Jon Snyder]

If we raise the bar as shown in this example, we can be fairly certain we won't falsely identify any programs, but as you can see, we will only identify approximately half of the truly underperforming programs. MPSC's challenge was to determine the optimal location to place this criterion so that they can identify as many programs as possible that are truly in need of review while avoiding identifying programs that truly do not have any performance issues.

How does MPSC currently identify programs for further review?

- For **small-volume programs** (<10 transplants in 30 months): Identify if >0 events.
- For **medium- and large-volume programs** (10 or more transplants in 30 months): Identify if **all 3 criteria** (below) are met.



SRR
SURVIVAL RISK RATIO
TRANSPLANT RECORDS

[Jon Snyder]

This slide details the current algorithm used by the MPSC. Currently, for small volume programs, those programs performing <10 transplants over a 2.5-year period, MPSC will take an initial review if 1 or more events (graft failures or deaths) are observed. For larger programs, MPSC will review if:

1. There are at least 3 more events observed at the program than what the models would expect ($O - E > 3$)
2. There are at least 50% more events observed at the program than what the models would expect ($O/E > 1.5$)
3. And the O/E is statistically significantly greater than 1.0 as assessed by the one-sided p-value < 0.05

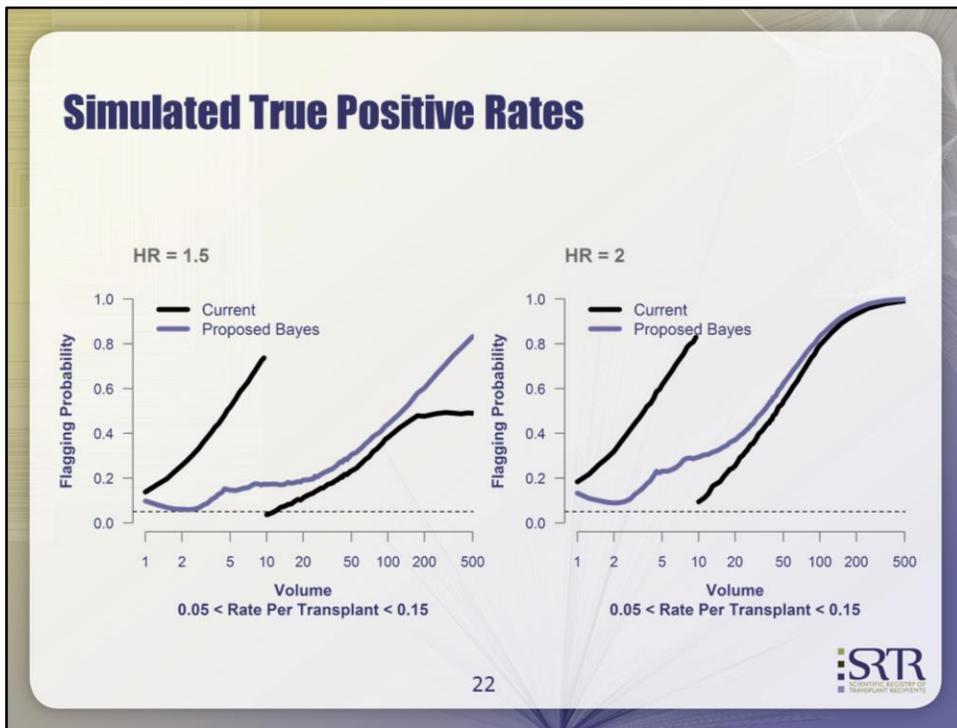
Programs meeting all three of the above criteria will be reviewed by MPSC.



[Jon Snyder]

This figure illustrates an important feature of the current system. Presented here are results of a computer simulation assessing the false positive rate across the range of transplant program volumes. The x-axis displays the program volume and y-axis displays the probability that a program is falsely identified for review under the current system (black line) and the Bayesian system (light blue line). As you can see, for programs in the small volume range (<10 transplants), there is a high probability that they will be falsely identified for having at least one event occur during the first year even if the program is not underperforming. Above a volume of 10, the probability of being falsely identified drops to 5% or below (the horizontal dashed line is drawn at 5%).

The Bayesian system was designed to attempt to hold the probability of false positives at about 5% across the range of program volumes (more on this later).



[Jon Snyder]

This set of slides presents the current system's ability to correctly identify truly underperforming programs. Again we are seeing results of computer simulations. In the left panel, all transplant programs were simulated to have event rates that were 50% higher than expected (HR = 1.5), and in the right panel all programs were simulated to have event rates that were 100% higher than expected (HR = 2).

Again we see a large disconnect in the current system (black lines) across a volume of 10. Above a volume of 10, the current system has a low ability to detect true underperformance. The ability to identify these programs improves as volume increases (we have more data, so we are more certain of our assessment). Again, the Bayesian system improves the ability to identify the truly underperforming programs as shown by the light blue line being higher than the black line .

Moving to a Bayesian Statistical Method

American Journal of Transplantation 2012; 12: 1988-1996
Wiley Periodicals Inc.

© Copyright 2012 The American Society of Transplantation
and the American Society of Transplant Surgeons

doi: 10.1111/j.1600-6143.2012.04130.x

Meeting Report

Report of a Consensus Conference on Transplant Program Quality and Surveillance

B. L. Kasiske^{a,b,*}, M. A. McBride^c, D. L. Comell^d,
R. S. Gaston^e, M. L. Henry^f, F. D. Irwin^g,
A. K. Israni^{a,b,h}, N. W. Metzlerⁱ, K. W. Murphyⁱ,
A. I. Reed^k, J. P. Roberts^l, N. Salkowski^b,
J. J. Snyder^{b,h} and S. C. Sweet^m

^aDepartment of Medicine, Hennepin County Medical Center, University of Minnesota, Minneapolis, Minnesota

^bScientific Registry of Transplant Recipients, Minneapolis

^cM

^dL

^eU

^fU

sess outcomes at small-volume transplant programs should be developed. More data on waiting list risk and outcomes should be provided. Monitoring and reporting of short-term living donor outcomes should be enhanced. Overall, there was broad consensus that substantial improvement in reporting outcomes of transplant programs in the United States could and should be made in a cost-effective manner.

Key words: Program-specific reports, OPTN, SRTR

1.3. The potential benefits of hierarchical and mixed effects methods should be studied.

[Jon Snyder]

The move to a Bayesian statistical methodology was one of the recommendations made at the Consensus Conference on Transplant Program Quality and Surveillance held by the OPTN and the SRTR in 2012. You can read the full meeting report in this article in the American Journal of Transplantation.

The COPSS Report: January 27, 2012

STATISTICAL ISSUES IN ASSESSING HOSPITAL PERFORMANCE

Commissioned by the Committee of Presidents of Statistical Societies

The COPSS-CMS White Paper Committee:

Arlene S. Ash, PhD; Stephen E. Fienberg, PhD; Thomas A. Louis, PhD
Sharon-Lise T. Normand, PhD; Thérèse A. Stukel, PhD; Jessica Utts, PhD

Original report submitted to CMS on November 28, 2011

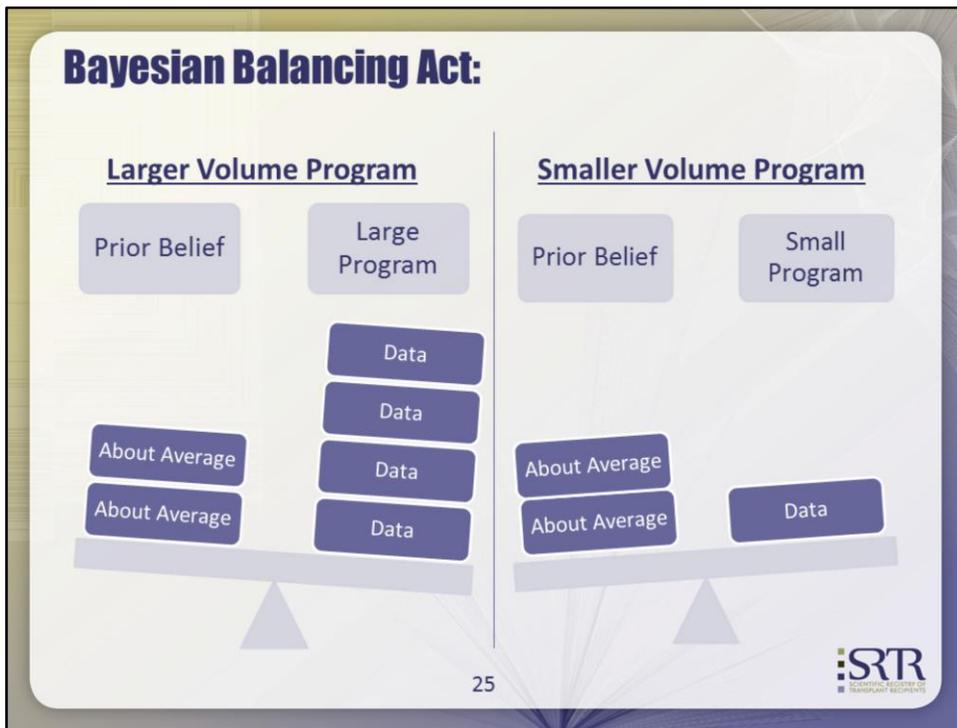
Revised on January 27, 2012

24



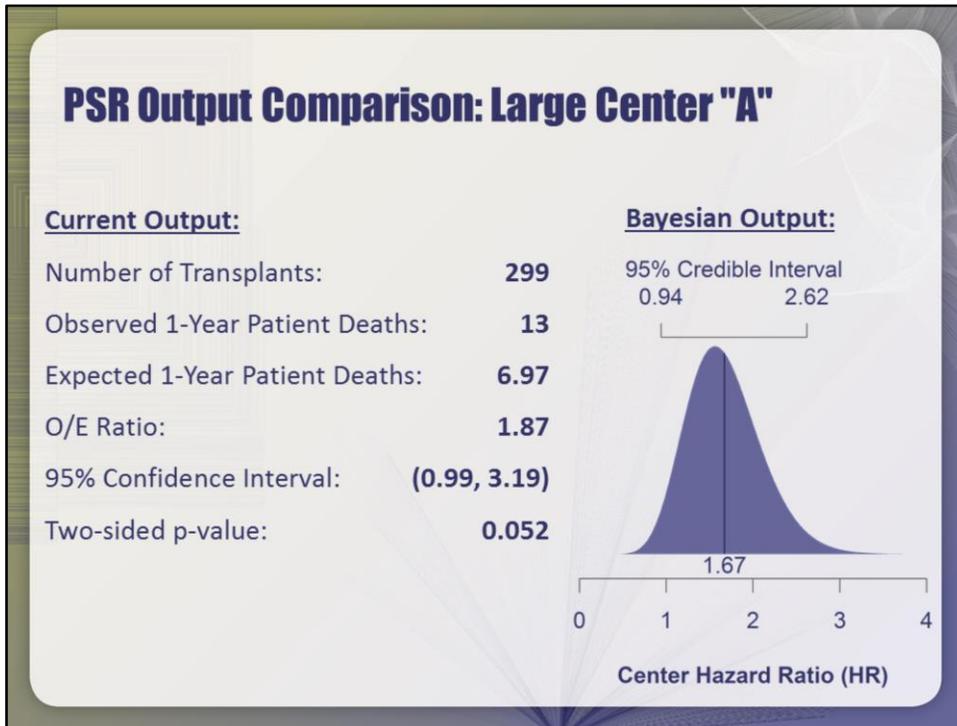
[Jon Snyder]

Moving to a Bayesian statistical methodology was also recommended in a report to CMS by the Presidents of Statistical Societies (COPSS) entitled “Statistical Issues in Assessing Hospital Performance.” This report was published at approximately the same time as the Consensus Conference in 2012. The COPSS report and the consensus conference recommendations were reviewed by the SRTR’s Technical Advisory Committee and the SRTR was charged with exploring implementing a Bayesian framework for program assessments.



[Jon Snyder]

Bayesian statistics make use of what is called a prior probability. Before looking at a program's data, SRTR has a prior belief about the likely location and spread of hospital performance in the nation. The prior belief was developed by the SRTR Technical Advisory Committee and is detailed in the Salkowski publication in the resource list. In short, the prior belief is that most programs in the country have hazard ratios between about 0.25 and 2.5 with few programs performing worse than 2.5 (150% higher event rates than expected). We then collect and observe the data and weigh the data against the prior belief. This weighting yields a bell-shaped probability distribution for each program. As shown here, large programs with a lot of data will tend to shift our conclusion towards the observed data, whereas smaller programs without a lot of data will tend to shift the conclusion towards the prior belief about that program. Let's consider a few examples on the following slides.

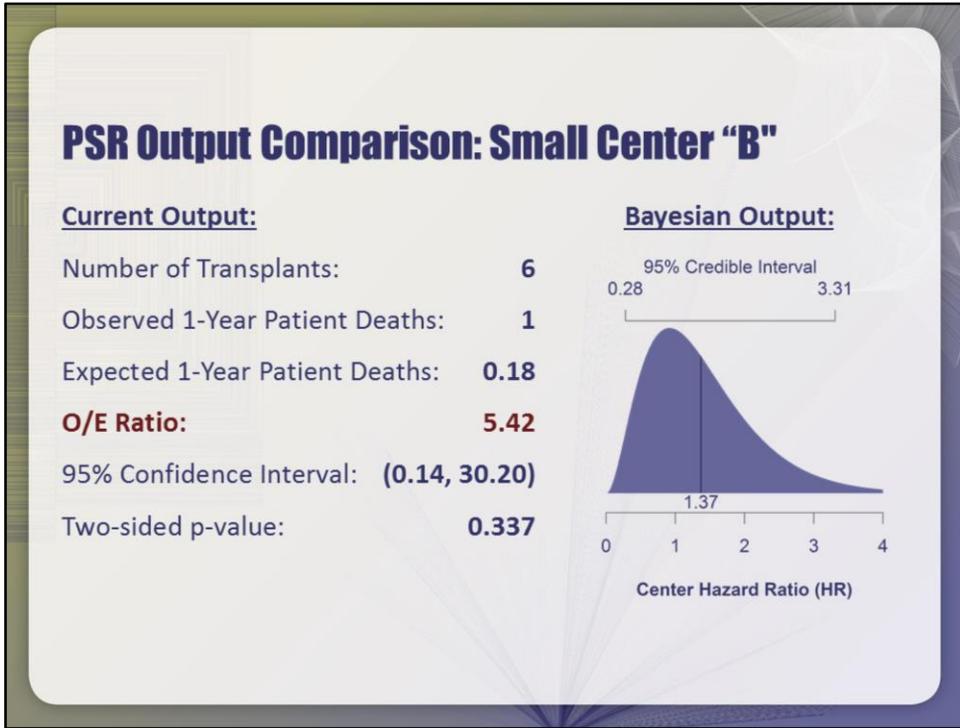


[Jon Snyder]

In this example, we have a large program that performed 299 transplants and experienced 13 first-year patient deaths. The models predicted we would see 6.97 patient deaths. Under the current system, the O/E for this program is 1.87.

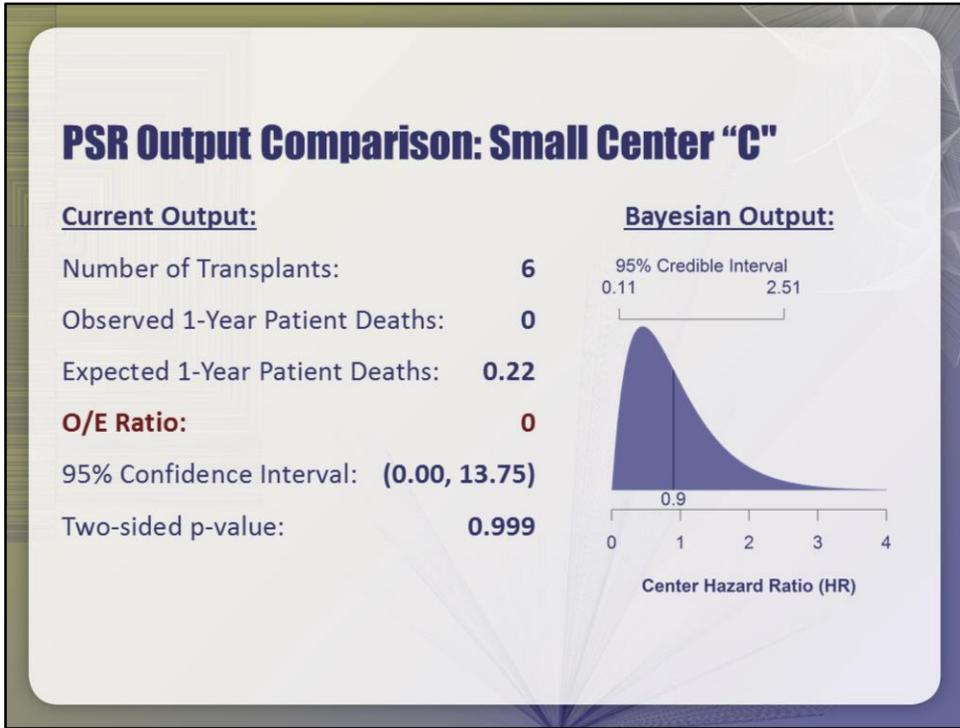
The Bayesian assessment for this program is depicted by the bell-shaped curve on the right. Here we see the likely location of this programs hazard ratio (analogous to the O/E). Our new estimate for this program is a hazard ratio of 1.67. Notice that this is closer to 1 than the original estimate of 1.87. This is a result of weighing our prior belief that the program was likely about average (HR=1) with the observed data. This is a large program, so the data carry more weight. The estimate of 1.67 moves closer to 1, but not by much.

The Bayesian analysis also provides a 95% credible interval depicted above the bell curve. This gives the location of the program's true hazard ratio with 95% certainty given the assumptions that went into the modeling.



[Jon Snyder]

Let's now consider a small program. This program performed only 6 transplants and observed 1 patient death. The O/E for this program is 5.42. You'll notice that the Bayesian analysis yields an estimate of 1.37, again weighting the prior belief with the available data. Since this program is small, the estimate is moved much closer to 1. While this program is very small, it illustrates a benefit of the Bayesian analysis. After observing only 6 transplants, are we willing to say that we believe this program's mortality rate is over 5 times higher than what the models would suggest? The Bayesian estimate does show a higher than expected mortality rate, but only 37% higher with a wide credible interval because we just do not have a lot of data to base our inference on.



[Jon Snyder]

As a final example, again consider a program that performed 6 transplants, but now had zero deaths. The O/E in this case is precisely 0, but again it strains credulity to believe that there is 0 risk of death post-transplant at any program. The Bayesian analysis yields an estimate of 0.9. The data support that this program’s event rate may be 10% lower than expected, but again we have a wide credible interval.

What I'll cover today...

- High-level review of the difference between the current methodology for program assessment and the Bayesian methodology.
- How will the Membership and Professional Standards Committee (MPSC) of the OPTN identify programs for review using the Bayesian method?
- Describe the timeline for transition.

29



[Jon Snyder]

Now, let's look at how the MPSC developed the new algorithm for identifying programs for review using the results of the new Bayesian analyses.

MPSC Criteria Goals

Correctly identify as many underperforming programs as possible without . . .

1. Falsely identifying programs that do not need review
2. Identifying too many programs overall

Only the MPSC can determine the correct balance between correctly identifying underperforming programs and avoiding falsely identifying programs.

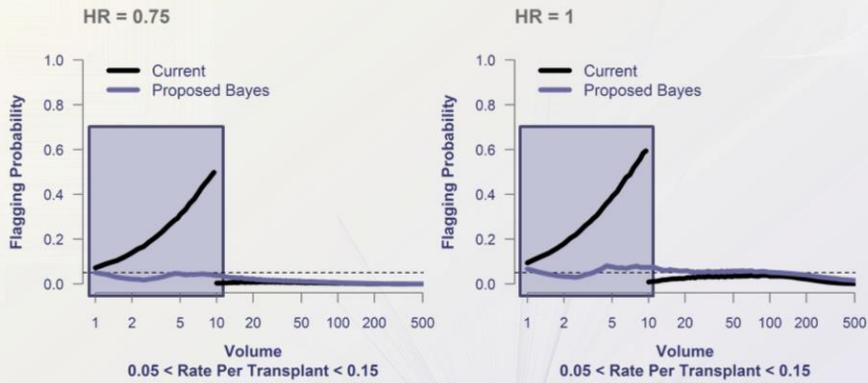
30



[Jon Snyder]

Recall that a good algorithm will perform well at identifying truly underperforming programs while avoiding falsely identifying programs or identifying too many programs.

High Probability of False Positives under the Current System among Small Volume Programs



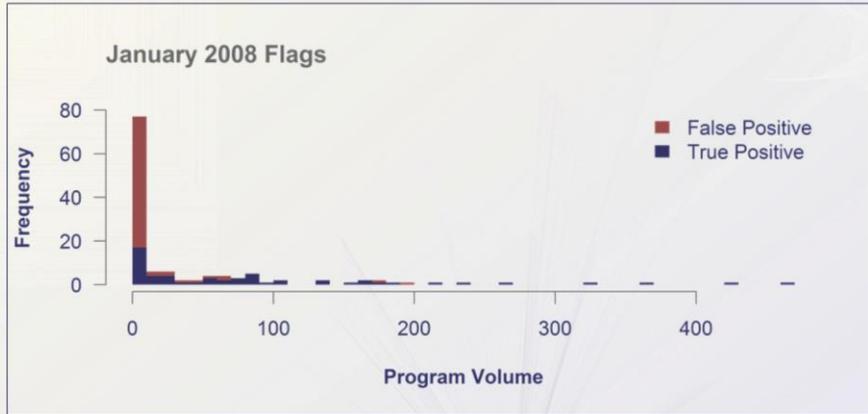
31



[Jon Snyder]

And recall that the current algorithm has a high probability of falsely identifying small volume programs as shown in the shaded regions in these simulations.

Review by MPSC of Programs Identified in January 2008



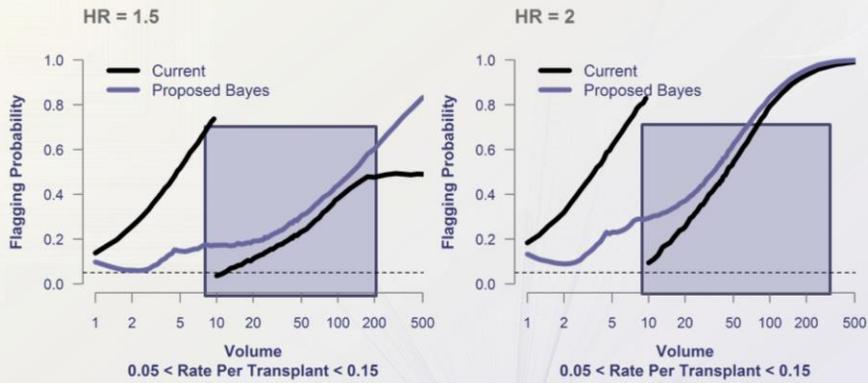
32



[Jon Snyder]

The simulations were supported by results from a historic review of MPSC actions for programs that were identified during the January 2008 review cycle. In this historic review, MPSC determined for all programs that were identified if it was a true positive or a false positive by assessing whether or not the MPSC took any significant action against the program following its initial review. Here we see that for the small volume programs, the majority of programs identified were deemed to be false positives as indicated by the red shaded areas of the bars.

The current algorithm is likely missing underperformance in the mid-volume range.



33



[Jon Snyder]

We also know that the current algorithm is likely missing truly underperforming programs in the mid volume range as shown by the low probabilities in the shaded regions.

To improve upon these shortcomings, the SRTR worked jointly with the MPSC to develop a new algorithm based on the Bayesian assessments.

Bayesian Models : Posterior Probability Distribution

Bayesian models allow us to show the distribution for the performance of a particular center, which can be compared to identified thresholds or national norms

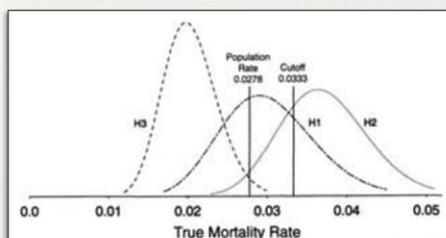


Figure 2. True mortality rate probability graphs for three hospitals (H1, H2, H3) in New York State (1). Vertical lines indicate the population rate and the chosen standard; curves represent the probability densities that determine the chance that the mortality rate at each hospital exceeded the 3.33% standard.

Christiansen CL, Morris CN. *Ann Intern Med.* 1997;127:764.



[Jon Snyder]

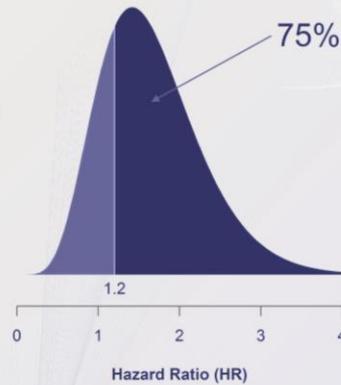
Using a Bayesian methodology in hospital performance assessments is not a new concept. In this paper published in 1997, Christiansen and Morris demonstrated how it could be applied. In this example, we see Bayesian performance evaluations for 3 theoretical hospitals (H1, H2, and H3). The Bayesian method yields probability distributions, or bell-shaped curves, showing the likely location of the hospital's performance metric, in this case the hospital's mortality rate. The population average mortality rate is denoted by the vertical line in the middle, and an cutoff is indicated to the right of the population rate. A decision rule can then be developed to decide which hospitals warrant further investigation. In this case, we need to determine how much of the bell curve is to the right of the proposed cutoff to cause a sufficient level of concern. SRTR worked with MPSC to determine 2 values: the location of the cutoff and the proportion of the bell curve that needs to be to the right of the cutoff to warrant review.

Example: Quality Oversight Using A Bayesian Approach

“These two values, the **standard for acceptable care** and the **minimum probability for compliance**, should ideally be determined with the collaboration of all knowledgeable and interested parties.”

~ Christiansen & Morris, Ann Int Med 1997

Example: Review a program if the **probability** that a center's **hazard ratio** is greater than **1.2** is greater than **75%**.



[Jon Snyder]

This concept was nicely illustrated in the Christiansen and Morris paper. The authors stated “These two values, the **standard for acceptable care** and the **minimum probability for compliance**, should ideally be determined with the collaboration of all knowledgeable and interested parties.”

In this example, MPSC would review a program if at least 75% of the bell curve was to the right of the cutoff of 1.2. In other words, if we are at least 75% certain that the program’s hazard ratio is greater than 1.2, MPSC will review the program.

Goals of a Screening Algorithm as Stated by the MPSC:

1. Attempt to avoid high false positive rates in the small volume range and better capture true underperformance in the mid-volume range.
2. Keep the false positive rate relatively low (approximately 5%) regardless of program volume.
3. Maximize our ability to identify truly underperforming programs while holding to #1-#2 above.

36



[Jon Snyder]

In working with MPSC to determine the cutoff and minimum probability of compliance, we attempted to achieve these three goals:

1. Attempt to avoid high falsely identified rates in the small program volume range (<10 transplants in 2.5 years) and better capture true underperformance in the mid-volume range.
2. Keep the falsely identified rate relatively low (approximately 5%) regardless of program volume.
3. Maximize our ability to identify truly underperforming programs while holding to #1-#2 above.

Computer Simulations Determined the Optimal Boundaries Used to Screen Programs for Review

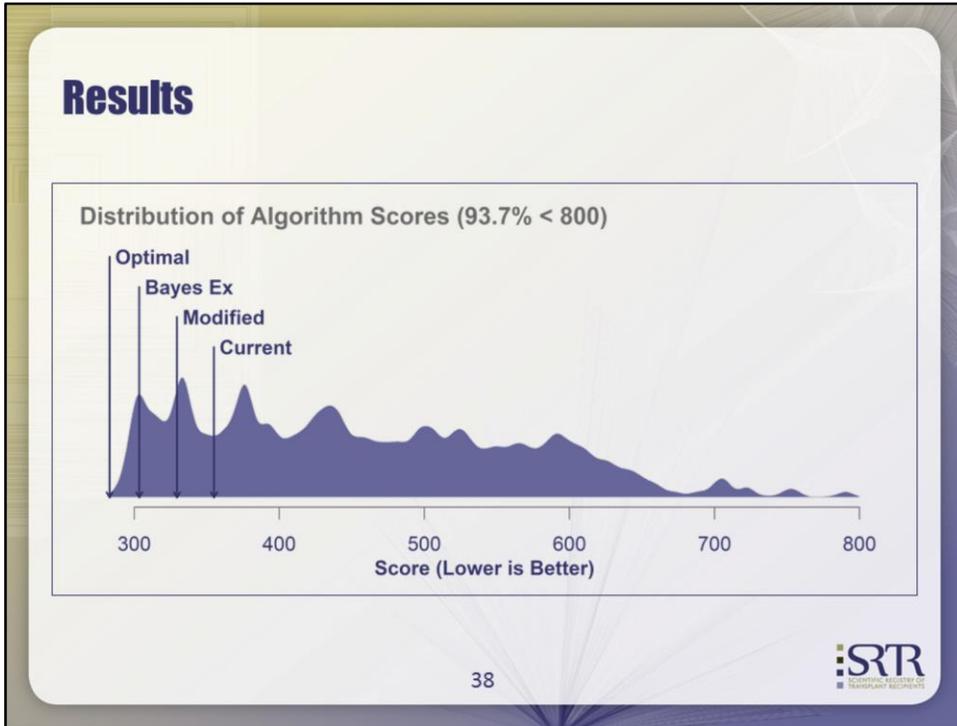
- The SRTR calculated the performance of 57,915 possible Bayesian algorithms, then calculated the score for each algorithm.
- The score reflected how well the algorithm performed against the MPSC goals.

37



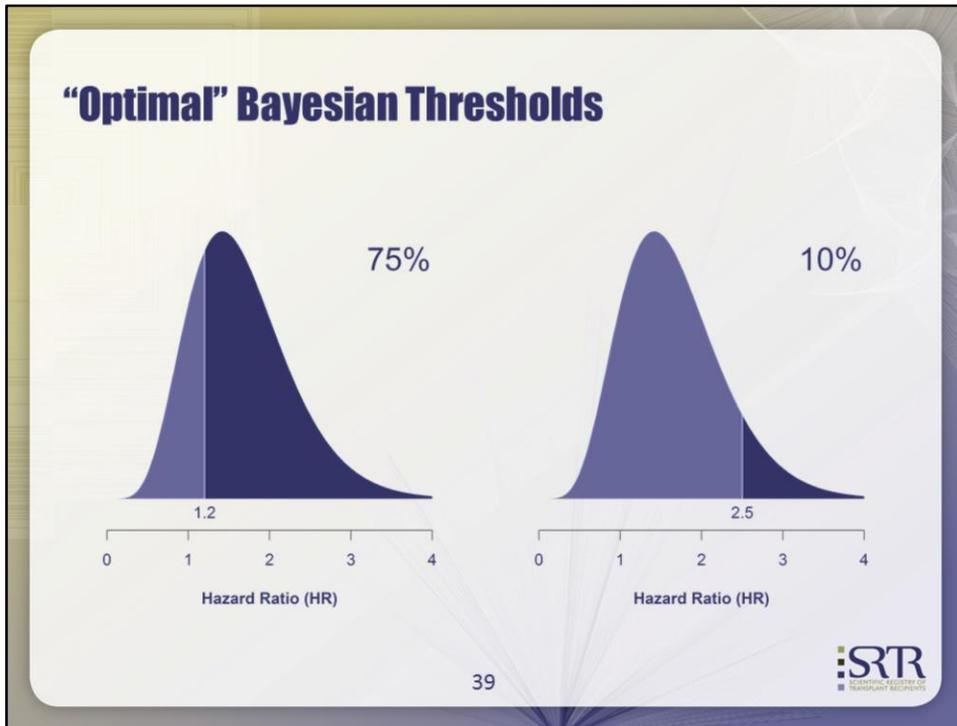
[Jon Snyder]

We did this through a complex computer simulation where we simulated programs and outcomes and evaluated nearly 58,000 different algorithms... each moving the cutoff and the minimum probability for compliance. Given that these were computer simulations, we knew with 100% certainty which were the programs that should be identified and which were the programs that should not be identified. For each algorithm, we were able to calculate a score that captured how well each algorithm did at identifying the right programs.



[Jon Snyder]

This shows the distribution of the scores for each of the nearly 58,000 algorithms evaluated. The optimal algorithm was the one with the lowest score, shown here on the far left. The Bayes Ex algorithm was a preliminary algorithm reviewed by the MPSC during some initial exploratory work. The Modified algorithm was also an algorithm being considered by the MPSC to work within the current framework. Finally, the Current MPSC algorithm is shown. While the current algorithm is certainly not the worst algorithm of all of the 58,000 explored (it ranked 11,987th), the optimal Bayesian algorithm was found to be superior at achieving the MPSC's goals.

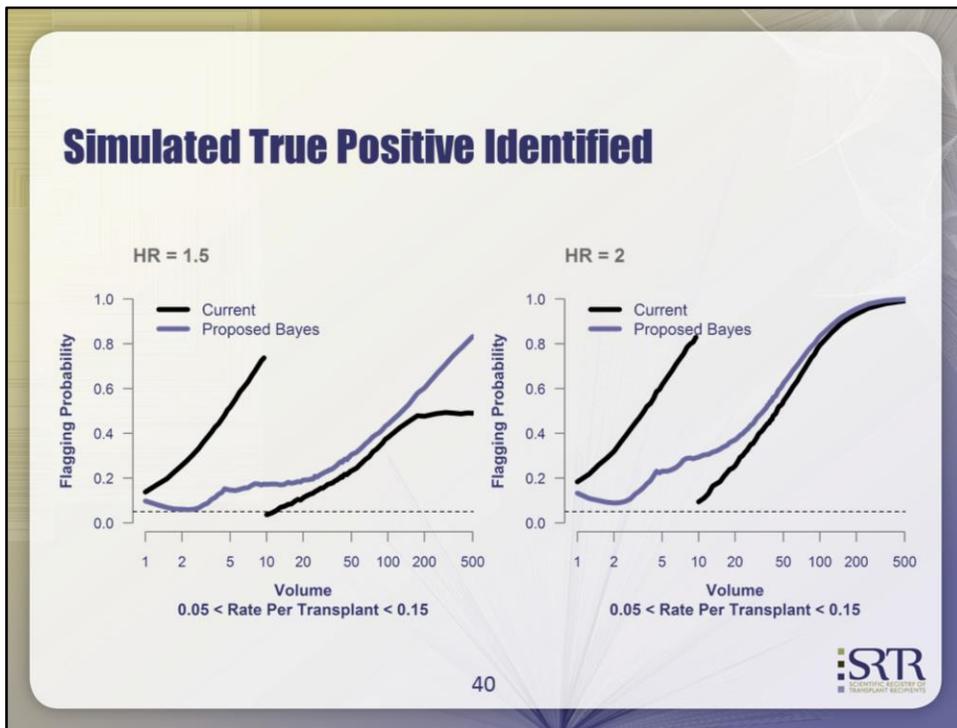


[Jon Snyder]

Here we see the optimal Bayesian algorithm. The algorithm employs two sets of criteria with two different cut-points and two different minimum probabilities for compliance. Using this algorithm, MPSC will review a program if either statement is true:

1. There is a greater than a 75% probability that the program’s hazard ratio is greater than 1.2; or
2. There is greater than a 10% probability that the program’s hazard ratio is greater than 2.5.

The first criterion (the left bell curve) will be better at identifying mid-volume or large-volume programs, while the second criterion (the right bell curve) will increase the chances of identifying underperforming small volume programs.



[Jon Snyder]

You're familiar with these curves now, but they again demonstrate that the Bayesian system improves the MSPC's ability to identify truly underperforming programs for review above a volume of 10 when compared to the current algorithm. The improvement in the ability to identify these programs is most notable in the range of volume from 10 to about 150, or "mid-volume" range.

Note that in the <10 volume range, the Bayesian algorithm has less power to identify truly underperforming small programs, but recall that the majority of small volume programs identified under the current algorithm were false positives. We are sacrificing some ability to identify underperforming small volume programs in order to avoid falsely identifying many small volume programs.

This shift is illustrated on the following slide.

July 2012 PSR Cohort By Volume

Volume	Programs	Transplants	Current	Optimal Bayesian
[1,10)	223	799	54	15
[10,50)	270	7,519	22	44
[50,100)	126	9,139	11	19
[100,250)	147	23,694	11	15
[250,744]	61	23,977	4	4
Total	827	65,128	102	97

Bayesian identified algorithms should . . .

- Have fewer false positive identified for small volume programs
- Have more true positive identified for larger volume programs

41



[Jon Snyder]

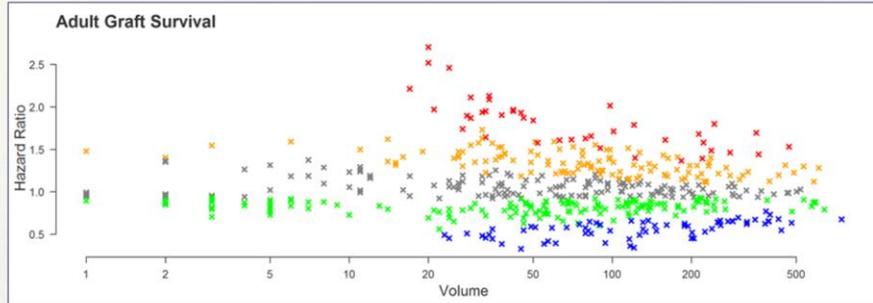
Here we compared which programs would be identified by the current algorithm and the Bayesian algorithm using data from the July 2012 PSR evaluation cycle. We see in the first row, for small volume programs the current algorithm identified 54 programs while the Bayesian algorithm reduced this to 15.

In the next row, in the volume range of 10-50, the current algorithm identified 22 programs whereas the Bayesian algorithm identified 44 programs. As volumes increase, the two algorithms perform more similarly since we have more data to base our decision on.

Some have interpreted this to be “unfair” to mid-volume programs. Recall that simulations suggest that the current algorithm is likely missing many underperforming programs in the mid-volume range and the Bayesian algorithm is identifying more programs in this range that are truly in need of review. Additionally, the Bayesian algorithm holds the false positive rate to approximately 5% or lower across the range of program volume, a feature that was distinctly unfair under the current algorithm given that the false positive rate was very high in small volume programs.

Additionally, programs in the small volume range performed a total of 799 transplants during this evaluation cycle, whereas programs in the 10-50 range performed nearly 10X the number of transplants. Shifting MPSC attention towards this volume range has the potential to impact many more transplant recipients.

The Challenge: Which Programs to Review?

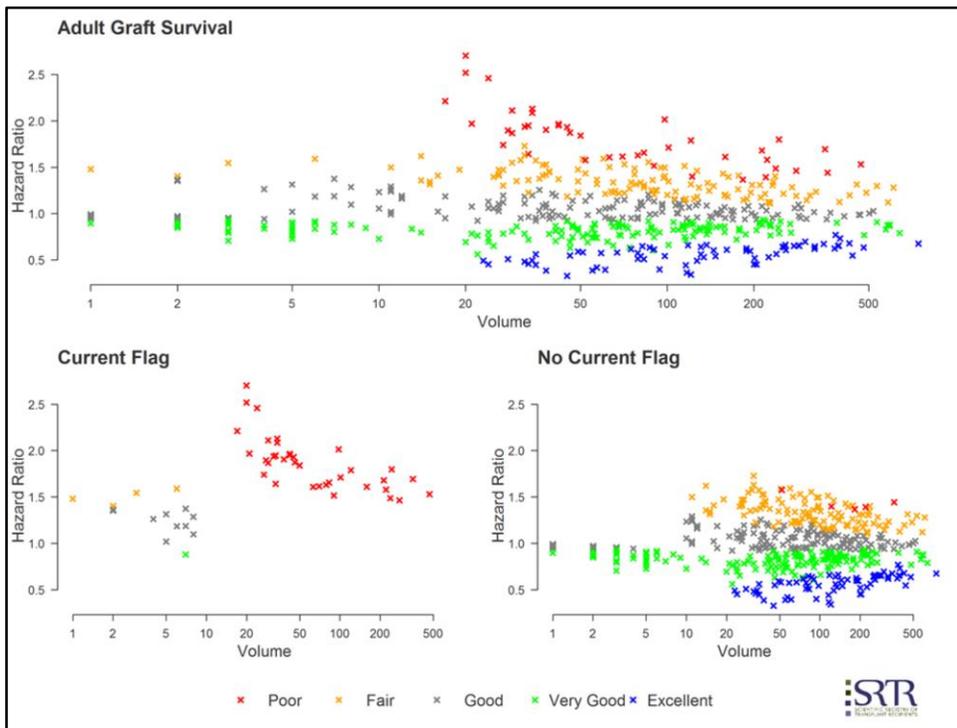


42



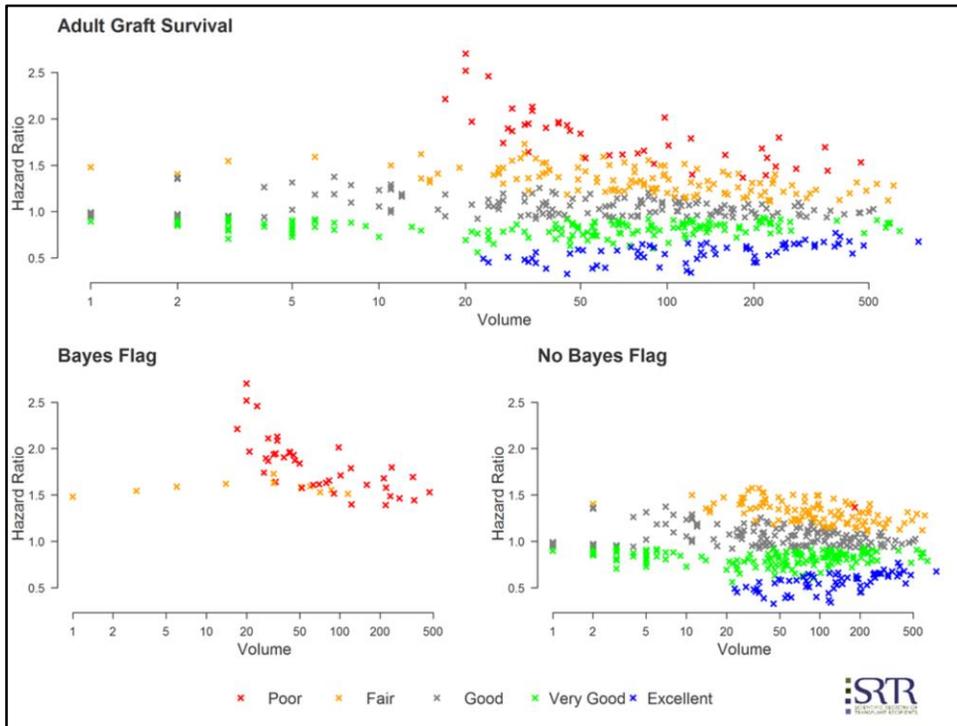
[Jon Snyder]

To better illustrate the performance of the Bayesian system, let's return to the cloud of performance assessments that MPSC is charged with reviewing.



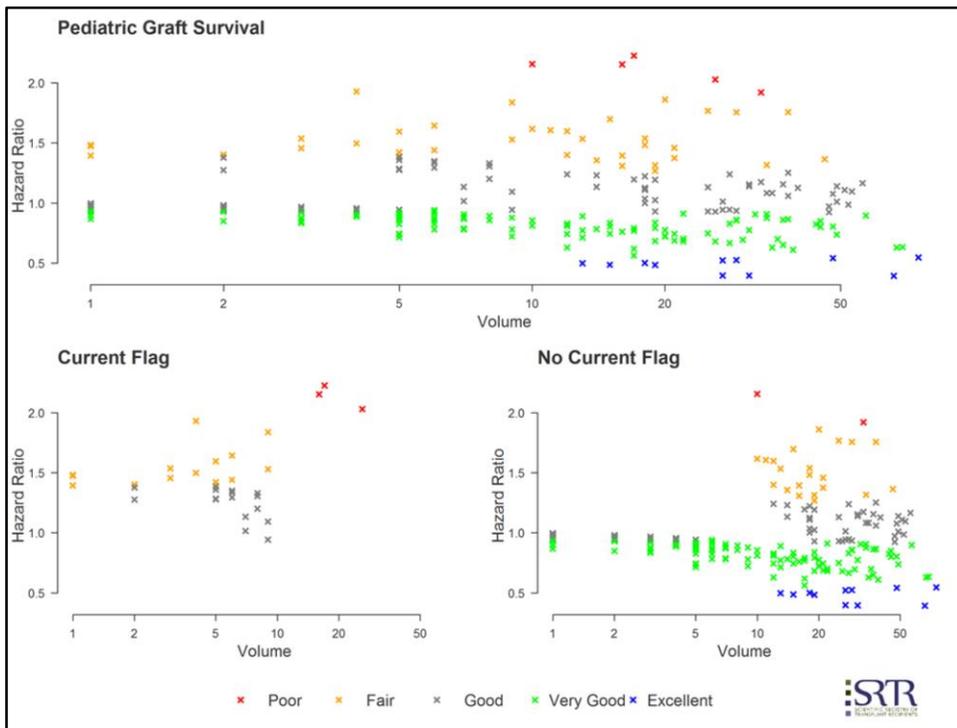
[Jon Snyder]

Here we see which programs were identified for review using the current algorithm. Note the programs in the small volume range that are identified that do not appear to have very high hazard ratios. Compare this with which programs are identified for review under the Bayesian algorithm on the next slide.



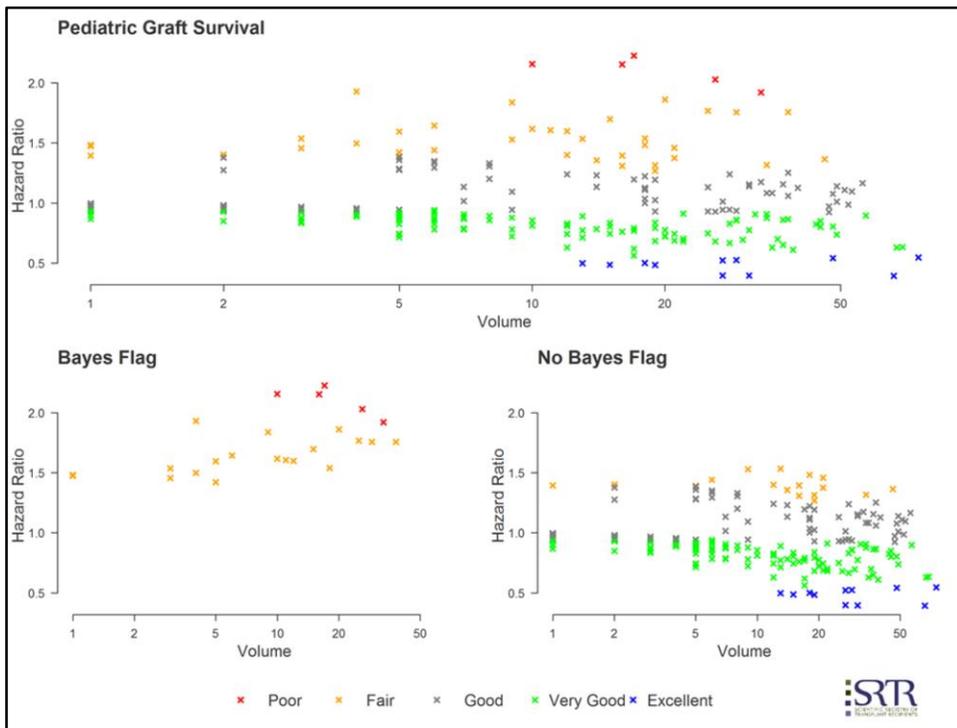
[Jon Snyder]

Here we see that the Bayesian algorithm avoids identifying the small volume programs that appeared not to be underperforming is better at identifying the “red” programs.



[Jon Snyder]

This figure shows the performance of the current algorithm within the pediatric programs. Again, we see many small volume programs identified while some of the “red” programs are missed.



[Jon Snyder]

Here are those same pediatric programs under the Bayesian system. We have now successfully identified the red programs while avoiding identifying the small volume programs with less indication of a problem.

What I'll cover today...

- High-level review of the difference between the current methodology for program assessment and the Bayesian methodology.
- How will the Membership and Professional Standards Committee (MPSC) of the OPTN identify programs for review using the Bayesian method?
- Describe the timeline for transition.

47



[Jon Snyder]

Finally, let's discuss the timeline for transition.

UNOS/OPTN Board of Directors

- On June 23, 2014, the UNOS/OPTN Board of Directors voted to approve use of the Bayesian identification algorithm, with continued scrutiny of small volume programs initially.

48

48



[Jon Snyder]

On June 23, 2014, the UNOS/OPTN Board of Directors voted to approve the use of the Bayesian identification algorithm, with continued scrutiny of small volume programs initially.

How the Proposal will Achieve its Goal

- Use of Bayesian methodology by SRTR
- New thresholds for identified programs that perform >9 transplants in 2.5 years
 - $P[HR > 1.2] > 75\%$, or
 - $P[HR > 2.5] > 10\%$
- Small volume <10 transplants – 1 event within 2.5 year cohort

[Jon Snyder]

Here we see the system as approved by the OPTN Board of Director's in June 2014.

Inclusion of the small volume identified for 1 year:

As stated during the Board presentation:

- The MPSC is convinced that the Bayesian methodology will achieve the stated goal
 - Volume greater than 9 transplants over a 2 and ½ year period
- Continue to evaluate how well the Bayesian methodology captures the true positives in small volume programs
- Committed to reevaluating the small volume methodology within one year of implementation

50

50



[Jon Snyder]

MPSC voted to continue the small volume identification method for a period of 1-year to assess how comfortable they were eliminating the additional identification method for small volume programs. As stated during the MPSC presentation to the Board of Directors in June 2014:

“Although the MPSC is convinced that the Bayesian methodology will achieve the stated goal with regards to programs with volume greater than 9 transplants over a 2 and ½ year period, we still want to examine additional data about how well the Bayesian methodology will identify underperforming small programs. Therefore, the proposal contains the current method for flagging those small volume programs, which is one event within the 2 and a half year cohort. The MPSC will continue to evaluate how well the Bayesian methodology captures the true positives in the small volume programs to determine if this additional method of flagging small volume programs can be eliminated in the future. The MPSC has committed to reevaluating the small volume methodology within one year of implementation.”

When will this change go into effect?

- The UNOS/OPTN Board of Directors voted to implement MPSC's change to the Bayesian algorithm effective January 1, 2015.
- Bayesian assessments will be included in the public PSRs for the December 16, 2014 release.
- MPSC will review the December 16 release of the PSRs at their March 2015 meeting and will use the Bayesian results for that review.

[Jon Snyder]

While the Board resolution stated an effective date of January 1, 2015, the MPSC has clarified that the intent is for the MPSC to use the Bayesian performance assessments at their spring 2015 meeting. This meeting will be reviewing the Fall 2014 PSR release. This release cycle will begin in October 2014 when programs begin to review their data. The Bayesian performance assessments will be included in the public PSRs set to be released in December 2014. MPSC will use these PSR evaluations at their spring 2015 meeting.

Preview Bayesian Screening Data Available to Programs Now!

- SRTR has released 4 versions of the PSRs for transplant programs to review showing both the current and Bayesian identified results.
- These are available for the following PSR cycles:
 - June 2014
 - January 2014
 - July 2013
 - January 2013
- Available on the SRTR's secure website and found in the "Expected Survival Worksheet" Excel file.

[Jon Snyder]

Note that SRTR has now provided Bayesian performance assessments to programs for the 4 previous PSR release cycles. These are provided so programs can begin to familiarize themselves with the methodology and the algorithm. These assessments are available to programs on the SRTR's secure website and are contained within the Expected Survival Worksheets that are provided as Excel files for programs to download.

Location of Expected Survival Worksheet on the SRTR Secure Website

Current Release Files

[Kidney](#) [Liver](#) [Heart](#) [Kidney/Pancreas](#) [Pancreas](#) [Lung](#) [Intestine](#) [Heart/Lung](#)

Kidney

[Spring 2014 Secure Preview Report: Traditional Format \(PDF\)](#) (posted June 3, 2014)
[Spring 2014 Secure Preview Report: Graphical Format \(PDF\)](#) (posted June 3, 2014)
[Spring 2014 1-Year Expected Survival Worksheets \(Excel\)](#) (posted June 3, 2014)

A guide to the primary source of risk-adjustment 1-year model covariates is available [here](#).

[Jon Snyder]

Here is an image of the SRTR secure website and the expected survival worksheet links.

That ends my portion of the presentation today. Thanks for listening and now let me hand over the presentation to Sharon.

MPSC Performance Monitoring Series

- Second event - the MPSC performance monitoring process
- Clarify current myths and misunderstandings about the current process
- January 2015

OPTN



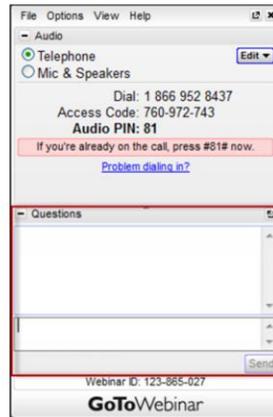
[Sharon Shepherd]

Thanks, Jon. As mentioned at the beginning of the webinar, this Bayesian event is the first in a planned instructional series on the MPSC Performance Monitoring process. Now that you have heard the background and foundation of how transplant programs are identified, our next event dives into the actual performance review process.

Our goal for the next event will be to help clarify and shed light on current myths and misunderstandings associated with the process. It is important to remember that although the MPSC is switching to the Bayesian methodology, the performance review process remains the same.

We are in the planning process to bring you the second event of this series in January 2015.

Ask Your Questions



OPTN



[Chad Southward]

As a quick reminder, you can type questions into the Question Panel and click Send. We will read the question and direct the question to one of our experts. Your questions will be answered in the order in which we receive them. As you are submitting your questions, please remember that the presenters may not be able to address all questions during this Q&A session, however we will address as many as possible. If you ask a question and we do not get to it today, we will make sure and send you an answer after the webinar. Due to the sheer volume of participants attending the webinar, we will need to stick to the submitted questions via the Question panel instead of unmuting hundreds of participants.

Panelists



Jon Snyder, Ph.D., MS



Sharon Shepherd, J.D., MSN, RN



Bob Carrico, Ph.D.



Nicholas Salkowski, Ph.D

OPTN



[Chad Southward]

Along with today's presenters, UNOS staff will help address questions during this question and answer session, they include:

Dr. Bob Carrico – Senior Biostatistician in the Research Department at UNOS.

Dr. Nicholas Salkowski – Senior Biostatistician at the Scientific Registry of Transplant Recipients.

With 5-10 minutes left in webinar, we will move on to CEPTC and Evaluation slides

Assessment

- Take the credit assessment to qualify for one CEPTC point:
- <https://www.surveymonkey.com/s/BayesianMethodCEPTC>
- The assessment will close: **March 23, 2015.**

OPTN



[Chad Southward]

If you wish to receive continuing education credit for your attendance on this webinar today, please use the web address on your screen to take the assessment.

Please note that you cannot take the assessment as a group; you are required to take the assessment individually to qualify for the credit.

Once you take the assessment, it will take about 4-6 weeks for you to receive your certificate.

This webinar was recorded and will be available soon for those that could not attend. They may also apply for credit after viewing the recording. The assessment will be open until March 23, 2015.

Please remember to write down the assessment address.

Evaluation

- Please take a few minutes to fill out an evaluation. Copy and enter this address into your browser address bar.
- <https://www.surveymonkey.com/s/BayesianEvaluation>
- We appreciate your feedback.

OPTN



[Chad Southward]

Please take a few minutes to fill out an evaluation. Copy and enter the address into your browser. We appreciate your feedback and constantly look for ways to improve instructional offerings. We would like participants to complete the evaluation individually. This allows us to receive more diverse and specific feedback. The evaluation link will be sent to all registrants following the webinar.

Again, thank you all for attending. We enjoyed answering your questions and we hope you found today's session full of useful information and worth your time.

Have a great rest of the day.