# A Bayesian method for identifying transplant programs for further peer review

Nicholas Salkowski, PhD, Jon Snyder, PhD, Tabitha Leighton, MS, David Zaun, MS, Ajay Israni, MD, and Bertram Kasiske, MD

SCIENTIFIC REGISTRY OF TRANSPLANT RECIPIENTS

## Introduction

- The SRTR is charged with assessing transplant program performance relative to expectation, and OPTN's Membership and Professional Standards Committee (MPSC) with deciding which programs warrant further peer review due to substandard outcomes. Following a recommendation of the Consensus Conference on Transplant Program Quality and Surveillance, the SRTR will begin using a Bayesian statistical method to assess program performance.
- The Bayesian method gives a posterior probability distribution of a program's hazard ratio. Screening criteria can then be developed to identify programs for peer review if evidence of underperformance is sufficient.

## Methods

- As MPSC's current screening methodology has high false-flagging rates for small-volume programs and low power for mid-volume programs, the SRTR performed a simulation study to identify optimal thresholds to best achieve these goals: 1) increase statistical power to identify underperforming programs, particularly mid-sized, and 2) hold the probability of false flagging to approximately 5% regardless of program volume.
- We simulated outcomes 2500 times for all kidney, heart, liver, and lung programs included in the July 2012 program-specific reports assuming, for all programs, 1) performance was consistent with the national average, and 2) mortality rates were twice the expected rate of failure. In each simulation, we examined 57,915 screening criteria to find the optimal criteria that maximize true positives while holding false positive rates to approximately 5% (Figure 1). The optimization criteria penalized an algorithm 0.05 points for every percent the false-positive rate differed from 5% and 0.01 point for every percent the true-positive rate differed from 100%. The false-positive and true-positive rates were defined as the probabilities of identification when a simulated program was performing precisely as expected and when a simulated program had twice the expected event rate, respectively.

## Results

- The optimal criteria were:
  - 1) Greater than 75% probability that the program's hazard ratio is greater than 1.2, or
  - 2) Greater than 10% probability that the program's hazard ratio is greater than 2.5 (Figure 2).
- Figures 3 and 4 compare the current and the optimal Bayesian flagging systems for false-positive and true-positive rates, respectively.

## Conclusions

- The optimal Bayesian-based criteria achieve the goal of holding the false positive rate relatively low regardless of program volume while maximizing the ability to identify true underperformance. The OPTN Board of Directors has approved a variation on these criteria for use by the MPSC to identify programs for further review.

Figure 1: Distribution of scores for tested algorithms.

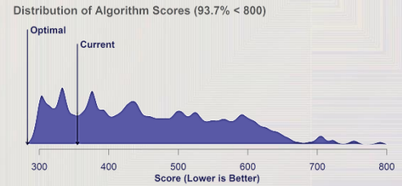



Figure 2: Graphical representation of the optimal criteria, based upon the simulation study.

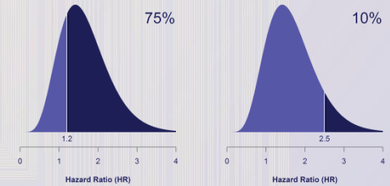



Figure 3: Simulated false-positive rates (identification of average programs) for the optimal Bayesian criteria (top) and the current algorithm (bottom).

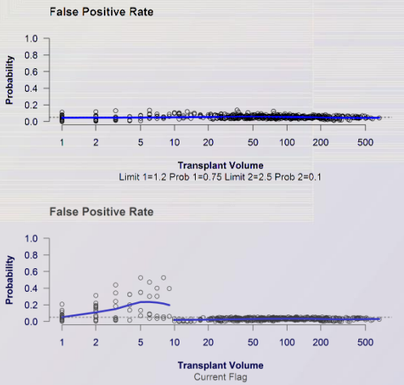
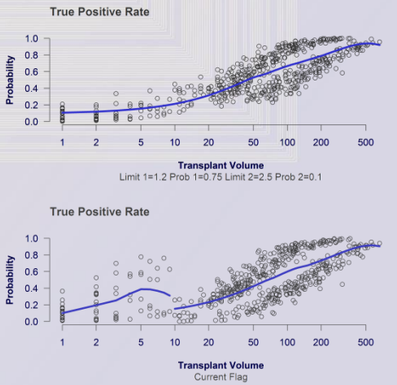



Figure 4: Simulated true-positive rates (identification of programs with twice the expected event rate) for the optimal Bayesian criteria (top) and the current algorithm (bottom).




Presented July 29, 2014 at the World Transplant Congress conference held in San Francisco, CA.