SCIENTIFIC REGISTRY OF TRANSPLANT RECIPIENTS

The C-statistic Provides no Information on the Accuracy of Program Evaluations or the Presence of Unmeasured Confounders

Andrew Wey[1], PhD, Nicholas Salkowski[1], PhD, Bertram Kasiske[1,2], MD, Melissa Skeans[1], MS, Sally Gustafson[1], MS, Ajay Israni[1,2,3], MD, Jon Snyder[1,3], PhD
[1] Scientific Registry of Transplant Recipients, [2]Dept of Medicine, University of MN, [3]Hennepin County Medical Center; Minneapolis, MN

Introduction

Accurate risk-adjustment models are critical for ensuring that recipients with more observed comorbid conditions and/or lower-quality donors do not generate worse adjusted posttransplant evaluations.

The C-statistic is regularly used to assess the validity of individual risk-adjustment models. However, C-statistics can only compare the performance of two alternative risk-adjustment models and provides no information on whether a single model is good or bad.

A simulation study illustrates that the C-statistic of a single risk-adjustment model has no relationship with the accuracy of estimated program-specific evaluations.

Methods

The simulation study was similar to the SRTR process for estimating program-specific posttransplant hazard ratios (HRs). Further, it was designed with similar characteristics to the posttransplant evaluations of 1-year graft survival for deceased donor kidney-alone recipients from the January 2018 program-specific reports (PSRs).

The simulation study created synthetic data to illustrate the accuracy of transplant program evaluations across a range of C-statistics. This was accomplished by increasing the variability in recipient-level risk, which was equivalent to the linear predictor in a Cox proportional hazards model.

Methods (Cont'd)

A range of C-statistics was generated by scaling the level of variability in recipient-level risk. The observed standard deviation of recipient-level risk from the January 2018 PSRs was multiplied by a scaling factor, denoted throughout by s. When $s = 1$, the standard deviation of recipient-level risk was equal to the observed standard deviation from the January 2018 PSRs.

The effect of unadjusted risk factors was investigated by systematically increasing or deceasing the risk of every transplant within a program.

Three metrics measured the accuracy of the estimated program-specific posttransplant HRs:

- Mean-squared error: The averaged squared difference between the estimated and true HRs.
- Spearman's rho measured the strength of association between the true HR and the 5-tier assignment.
- Probability that a program identified for MPSC regulatory review had a true HR above 1.25.

Because the accuracy depends on sample size, the metrics were split into 3 categories of expected effects:

- <3 expected events
- 3-<10 expected events
- ≥10 expected events

The metrics were averaged over 1,000 iterations of the simulation to reduce the effect of sampling error.

This work was supported wholly or in part by HRSA contract 250201000018C. The content is the responsibility of the authors alone and does not necessarily reflect the views or policies of the Department of HHS, nor does mention of trade names, commercial products, or organizations imply endorsement by the U.S. Government.

Results

Table 1. A comparison of the C-statistic, expected events, and observed 1-year survival for scenarios with and without unadjusted risks. *s* was the scaling factor on the standard deviation of recipient-level risk.

Value of *s*	C-Statistic		Observed 1-Year Survival	
	Without unadjusted risks	With unadjusted risks	Without unadjusted risks	With unadjusted risks
0.5	0.57	0.57	95.0%	95.0%
1	0.64	0.63	95.0%	95.0%
2	0.75	0.75	95.0%	95.0%
4	0.89	0.89	95.0%	95.0%
8	0.97	0.97	95.0%	95.0%

Table 2. The 1-year graft survival for the different percentiles of risk. High C-statistics likely require unrealistic differences in graft survival across levels of patient-level risk. *s* was the scaling factor on the standard deviation of recipient-level risk.

Value of *s*	1-Year Graft Survival at a Percentile of Risk			
	75th	90th	95th	99th
0.5	94.5%	93.7%	93.1%	91.9%
1	94.1%	92.1%	90.7%	87.2%
2	94.1%	89.5%	85.3%	73.4%
4	96.1%	87.8%	76.7%	36.5%
8	99.2%	91.2%	68.3%	0.5%

Figure 1. The relationship of the C-statistics with MSE. A flat line indicates the C-statistic does not provide information on the MSE.

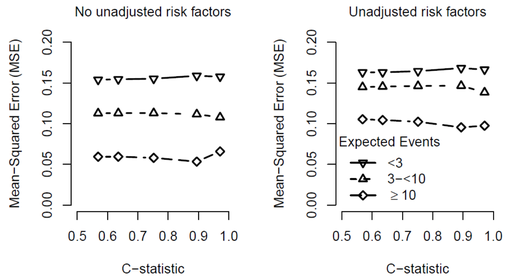


Figure 2. The relationship of the C-statistic with Spearman's rho between the true HRs and the 5-tier assignment. A flat line indicates the C-statistic does not provide information on the accuracy of 5-tier assignment.

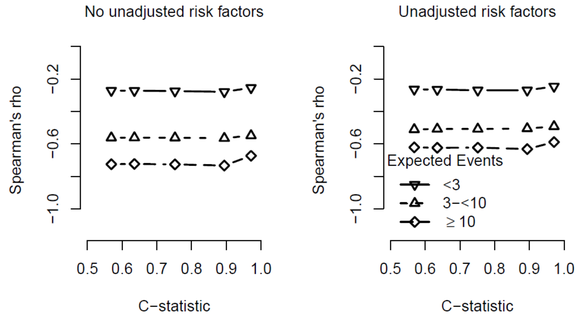
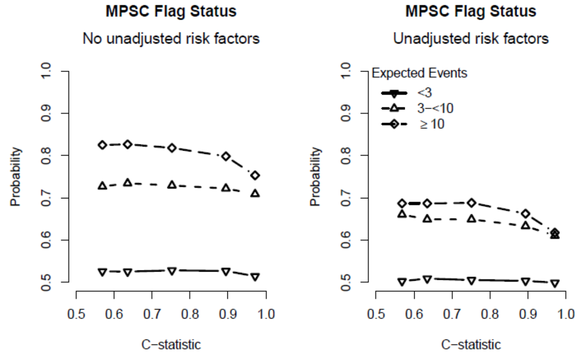


Figure 3. The relationship of the C-statistics with the probability that a program identified for regulatory review had a true HR above 1.25. A flat line indicates the C-statistic does not provide information on the accuracy of being identified for MPSC regulatory review.



Conclusions

The C-statistic provided no information on the accuracy of program-specific HRs, 5-tier assignment, or identification for regulatory review.

The C-statistic and any other metric of model performance depends on underlying and unknown characteristics of the data. Therefore,

- The C-statistic does not provide any information on the performance of a single risk-adjustment model.
- The C-statistic should not be compared across different data sets (e.g., bypass surgery) because differences may only identify the relative difficulty of the prediction problems.

The C-statistic and other metrics of model performance can provide information on the relative performance of different, alternative models within the same context. Although, such C-statistics *must* account for model complexity through, for example, cross-validation.

References

Wey et al., "The relationship between the C-statistic and the accuracy of program-specific evaluations," *American Journal of Transplantation*, 2019.

Austin and Reeves, "The relationship of the C-statistic of a risk-adjustment model and the accuracy of hospital report cards: Monte Carlo study," *Med Care*, 2013.

Austin and Steyerberg, "Interpretating the concordance statistic of a logistic regression model: relation to the variance and odds ratio of a continuous explanatory variable," *BMC Med Red Methodol.*, 2012.

Authors have no financial relationships to disclose.