Validating Prediction Models of Kidney Transplant Outcome Using Local Data

Hongying Tang, MS¹, Alexander S. Goldfarb-Rumyantzev, MD, PhD¹, Cheri Hunter¹, Mollie R. Poynton PhD, APRN¹, Ming Tu, MS¹, Bradley Baird, MS, M.Stat ¹, Sergey Krikov, MS¹, James K. Koford, PhD¹, John Hurdle, MD, PhD¹
¹ University of Utah, Salt Lake City, UT

Background

Prognosis of kidney transplant outcome, while clinically important, represents a challenging problem. Existing models to predict acute rejection or other transplant outcomes use predictors available in the post-transplant period. While the real value of the model is to predict the outcome prior to the transplantation, there is little experience with predicting graft survival using pre-transplant variables.¹ Furthermore, those prediction models developed using national registry data have not been validated in the local clinical environment. This is the first study to compare a model derived from aggregate national data to real-world local dataset for validation.

Methods

Five classification tree models predicting 1, 3, 5, 7, and 10 year graft survival were generated based on the United States Renal Data System (USRDS) data between 1/1/1990 and 12/31/2002. The models only included the predictors available prior to kidney transplantation. Local clinical data for model validation were obtained from the University of Utah Health Science Center (UUHSC) Enterprise Data Warehouse and the United Network for Organ Sharing (UNOS) reports from the transplant program at the UUHSC between 1/1/1990 and 12/31/2004. The data characteristics of kidney transplant recipients were compared between the USRDS and the UUHSC datasets. The outcome prediction models were evaluated against both the USRDS testing dataset and the UUHSC dataset. The discrimination of the prediction models was measured using the area under the ROC curve (AUC).

Results

In the USRDS dataset, the number of patients who had known graft outcomes at 1, 3, 5, 7, and 10 years were 854, 635, 462, 325, and 213 respectively. We found that UUHSC data demonstrated significantly higher graft survival rate (94% vs. 86%, 87% vs. 72%, 77% vs. 54%, 61% vs. 36%, and 33% vs. 8% at 1, 3, 5, 7, and 10 years, all p<0.0001) than the USRDS data. In the Utah population, the majority of recipients (>93%) and donors (> 95%) were white. In addition, the UUHSC data showed a significantly higher proportion of living donors (42% vs. 24%, 42% vs. 24%, 41% vs. 22%, 40% vs. 20%, and 39% vs. 18% at 1, 3, 5, 7, and 10 years, all p<0.0001). The AUC of the models predicting 1, 3, 5, 7, and 10 year graft survival on USRDS testing was 0.59, 0.63, 0.76, 0.91, and 0.97, respectively. However, the AUC of models predicting 1, 3, 5, 7, and 10 year graft survival in the UUHSC dataset was only 0.54, 0.58, 0.58, 0.61, and 0.70 respectively. When the models were tested on the UUHSC data, sensitivity, specificity, and positive predictive value are 97%, 3%, and 94% for 1-yr; 76%, 35%, and 89% for 3-yr; 29%, 81%, and 84% for 5-yr; 12%, 94%, and 75% for 7-yr; and 0,100%, and N/A for 10-yr respectively.

Conclusion

The prediction models derived from the national dataset (USRDS) performed better on its own data than on a local dataset (UUHSC), and this is almost certainly due to the differences in data characteristics between the two datasets. Researchers routinely extrapolate the results from national studies to local circumstances, but this study is one of the first to show the potential dangers of doing so with real-world data. Adopting wholesale a prediction model developed on a large national dataset for local purposes should be done with caution.

References