Model-based ROC (mROC) Curve: A Method For Assessing The Effect Of Case-mix And Model Calibration on The ROC Curve

Mohsen Sadatsafavi[1], Paramita Saha- Chaudhuri[2], John Petkau[3]

[1]Faculty of Pharmaceutical Sciences, The University of British Columbia. 2405 Wesbrook Mall, Vancouver, BC, V6T1Z3, Canada
[2]Department of Epidemiology, Biostatistics & Occupational Health, McGill University. Purvis Hall, 1020 Pine Avenue West, Montreal, QC, Canada H3A1A2
[3]Department of Statistics, 2207 Main Mall, Vancouver, BC, Canada V6T1Z4

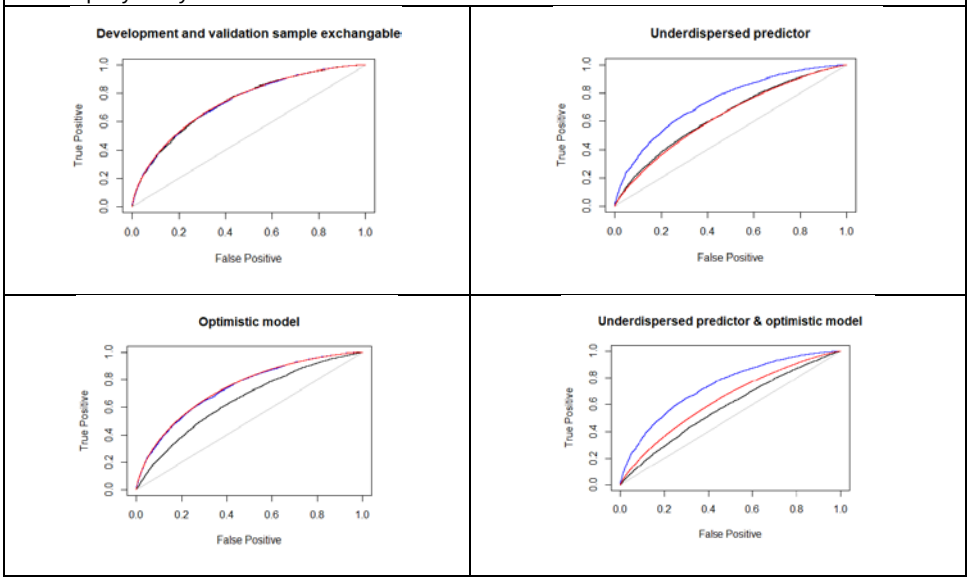
Background: The performance of risk prediction models is often characterized in terms of discrimination and calibration. The Receiver Operating Characteristic (ROC) curve is widely used for evaluating model discrimination. When comparing ROC curves between development and validation samples, the effect of case-mix makes the interpretation of discrepancies difficult. Further, compared to model discrimination, evaluating model calibration has not received the same level of attention in medical literature. The most commonly used graphical method for model calibration, the calibration plot, requires specification of smoothing parameters or number of groups.

Aims: This abstract introduces the 'model-based' ROC (mROC) curve, the ROC curve that should be observed if the prediction model is calibrated in the external population. Unlike the ROC curve, the mROC curve is affected by even monotonical transformations of predicted risks, thus is sensitive to model calibration. We show that moderate calibration (actual risk being p% among those with predicted risk of p%) is a sufficient condition for the convergence of mROC and ROC curves. Accordingly, we propose a novel test statistic for calibration that does not require any arbitrary parameterization.

Results: Through an example, we demonstrate how mROC separates the effect of case-mix and model mis-calibration when comparing ROC curves from different samples (***Figure***). We present the results of simulation studies that confirm the properties of the new calibration test. A case study puts the developments in a practical context.

Conclusion: mROC can easily be constructed and used to interpret the effect of case-mix and calibration on the ROC plot. This can facilitate interpretation of the ROC curve in external validation studies. Given the popularity of ROC curves among applied investigators, this framework can further promote assessment of model calibration.

Empirical ROC in the development (blue) and validation (black) samples, and mROC (red) curves for an exemplary study.



Keywords

Clinical Prediction; Risk Prediction; Receiver Operating Characteristic Curve