Objective

- The impact of the number of events per variable (EPV) on measures of survival model performance is examined through simulation.
- The sensitivity of model performance measures to weighting approaches and the omission or inclusion of important model variables is also explored.

Introduction

- The impact of EPV on the estimation of regression coefficients has been well described in the literature.
- When developing clinical prediction models, investigators are interested in measures of model performance besides estimation of association parameters.
- Survival prediction models can be evaluated based on their ability to:
  1. discriminate between high-risk and low-risk patients (Concordance-index).
  2. accurately predict the risk of event by time \( t \) (Brier Score).

Methods

Measurement of Model Performance

- Simulation study was employed to investigate the performance of the discrimination or concordance index (C-index) and overall model performance (Brier score).
- Various mechanisms that influence EPV in a model were examined:
  1. Underlying survival time was generated based on Weibull Cox proportional hazard model with shape parameter \( \nu = 0.5, 1, \) and \( 2 \).
  2. Censoring times were generated from random exponential with rate parameter \( \lambda = 1, 2, 5, 8, \) and \( 11 \).
  3. 6 continuous covariates and 2 binary variables were simulated from standard normal and Bernoulli distributions with \( p = .1, .6 \).
  4. 480 unique combination of parameters was utilized and approximately 1000 simulations were run for each parameter combination.
  5. Within each simulation, optimism was assessed using 10-fold cross-validation and Bootstrap CV for the Brier score and Concordance measures, respectively.

- Impact of model misspecification on the relationship of EPV to the model performance metrics is assessed by omitting 1 continuous and 1 binary variables.
- Impact of inclusion of unassociated or noise variable is examined by including noise variable simulated from an independent standard normal distribution.
- Inverse probability censored weighting (IPCW) is employed to examine the optimism of the C-index.

Results

- Impact of model misspecification on the relationship of EPV to the model performance metrics is assessed by omitting 1 continuous and 1 binary variables.
- Impact of inclusion of unassociated or noise variable is examined by including noise variable simulated from an independent standard normal distribution.

<table>
<thead>
<tr>
<th>Event Rate</th>
<th>N</th>
<th># Censored</th>
<th># Covariates (Associated + Noise)</th>
<th>EPV</th>
<th>Brier Score/C-index</th>
</tr>
</thead>
<tbody>
<tr>
<td>Fixed</td>
<td>Fixed</td>
<td>Fixed</td>
<td>Fixed</td>
<td>Fixed</td>
<td>Fixed</td>
</tr>
</tbody>
</table>

Performance of Brier Score and C-Index as a function of EPV

Conclusion

- Optimism of C-index and Brier score has inverse relationship with EPV.
- Optimism of the C-index and Brier score depends on the underlying shape of the hazard function as well as the censoring rate.
- Inclusion of noise variable (and consequent lowering of EPV) had a marginal increase in the apparent C-index.
- Omission of associated predictors can lead to a dramatic drop in the C-index depending on the strength of association.
- The apparent weighted C-index appears to be less optimistic than the apparent unweighted C-index.
- Time specific reporting of scaled Brier-score should take into account the EPV and the time-horizon of the clinical study as this score is lower at the beginning and end of the study.
- Principled development and validation of prediction models should utilize corrected model performance metrics.