Model performance is strongly correlated with measures of rurality and socioeconomic disadvantage

Katherine E. Brown, PhD, Sharon E. Davis, PhD

Department of Biomedical Informatics, Vanderbilt University Medical Center, Nashville TN

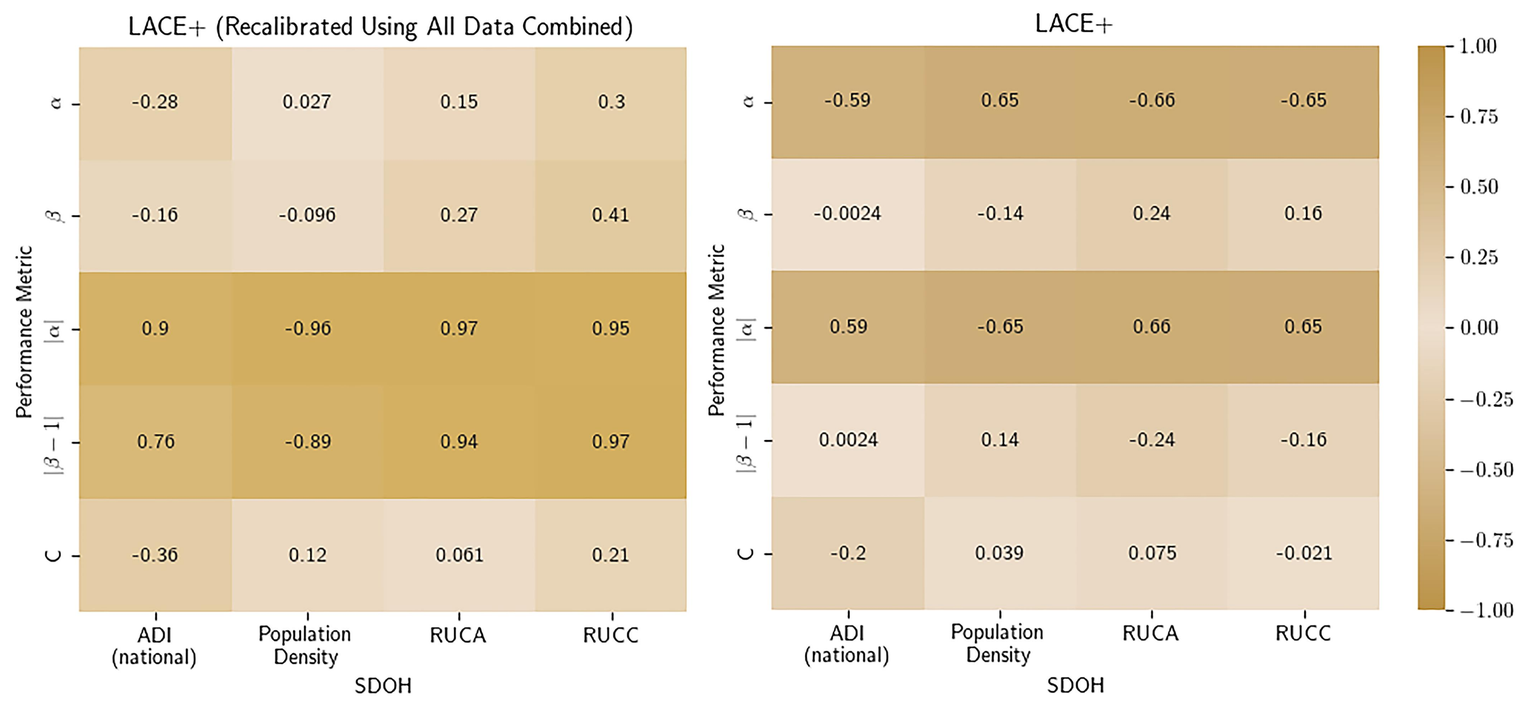
INTRO

- Clinical prediction models trained at one site can deteriorate when applied at other sites
- There is evidence that model performance is correlated with clinical factors
- **Hypothesis:** Socioeconomic factors associated with rurality of hospital location are correlated with model performance

METHODS

1. Collected 30-day unplanned hospital readmission status from 3 Middle TN hospitals: S_0, S_1, S_2
2. Predicted likelihood of readmission using the LACE+ prediction model
3. Analyzed a) model predictions without recalibration (LACE+) and b) recalibrated models using all data (LACE+ Recalibrated)
4. Evaluated Pearson's correlation coefficient between socioeconomic measures and model performance

Site	Rural	Area Deprivation Index	Population Density	Rural-Urban Continuum Codes	Rural-Urban Commuting Area Codes
S_0		< 100	< 1500	1	1
S_1	✓	≥ 100	< 100	5.5	5.4
S_2	✓	≥ 100	< 150	4	4.5



Discussion

- LACE before recalibration has moderate correlation to socioeconomic factors of rurality
- Upon recalibration, this correlation sharply increases
- The sparsity of catchment areas for S_1 and S_2 may reduce data available and overall magnify model miscalibration at these sites
- Understanding the impact of these socioeconomic factors on model performance is imperative to ensuring equitable model performance for all populations.

VANDERBILT UNIVERSITY MEDICAL CENTER

Department of Biomedical Informatics

Take a picture to download the full text

This research was funded by the National Institutes of Health grants T15LM007450.

katherine.brown@vumc.org