SDOH in Health Care Predictive Models

Methods

Population Studied:
Person-month panel data set spanning 35 months
- 465,749 Maryland Medicare fee-for-service beneficiaries who received primary care at an MDPCP-affiliated practice
- 59.5% female
- 69.8% White; 22.7% Black
- 14.9% dually eligible for Medicaid

Study Design:
Six discrete time survival models estimating risk for an AH – SAS (v 9.4)
- Different combinations of utilization and SDOH risk factors
- 144 claims-based features
- 37 SDOH features from 11 publicly available sources (e.g., American Community Survey) – Census Tract and ZCTA versions

Outcomes:
Calculated agreement between Census Tract and ZCTA indicators

Compared fit, performance, and interpretation across models:
- Akaike Information Criterion (AIC)
- C-statistic
- Gini index
- Percentage of AH events incurred by top 10% riskiest beneficiaries
- Salient SDOH predictors

SDOH in Health Care Predictive Models

Research Question: Does enhancing the granularity of SDOH predictors from ZIP code tabulation area (ZCTA) to Census Tract strengthen an existing predictive model for avoidable hospitalizations (AH)?
- Model used by primary care practices participating in the Maryland Primary Care Program (MDPCP)
- 144 claims-based features
- 37 SDOH features
- 14.9% dually eligible for Medicaid
- 59.5% female
- 69.8% White; 22.7% Black
- 14.9% dually eligible for Medicaid

Conclusions & Policy Implications

Making risks more granular did not dramatically improve fit or predictive performance when estimating risk for AH.

Different SDOH variables were retained depending on the granularity level, which does influence model interpretation.
- Because Census Tract measures are likely a better representation of a person’s proximal environment, model interpretation may be more meaningful.

Differences in interpretation will be critical if models are used to inform the distribution of resources.
- Especially salient as funds become available to address drivers of health that exist beyond the bounds of traditional health care.

References & Funding


Stepwise variable selection process retained different SDOH depending on geographic level.

Census Tract-level models fit better (lower AIC) and had stronger predictive ability than models with ZCTA-level or no SDOH risk factors.
- Differences in fit and predictive performance minimal with utilization-based risk factors in model

Granularity Had Small Impact on Predictive Performance but Led to Differences in Interpretation

Different SDOH variables were retained depending on the granularity level, which does influence model interpretation.
- Because Census Tract measures are likely a better representation of a person’s proximal environment, model interpretation may be more meaningful.

Differences in interpretation will be critical if models are used to inform the distribution of resources.
- Especially salient as funds become available to address drivers of health that exist beyond the bounds of traditional health care.