The Link Between Mental Health Provider Density and Poor Mental Health at the County Level

Joann Jun
University of Connecticut

Lily Fransen
University of Illinois Urbana-Champaign

Karim Zaggoti
Washington University in St. Louis

Carnegie Mellon University
Statistics & Data Science

Background

Motivation:

- According to the Centers for Disease Control and Prevention (CDC), more than 1 in 5 adults in the U.S. is affected by mental health conditions.
- Only about half of those affected receive treatment, highlighting a significant gap in mental healthcare access and utilization.
- Understanding regional patterns in mental health can help guide policies aimed at improving mental health outcomes, inform better resource distribution, and support mental health efforts at the community level.

Main Question:

Do the number of mental health professionals per county affect the number of poor mental health days?

Data Source:

2025 County Health Rankings Dataset

Variables:

- County FIPS code
- Frequent mental distress
- Mental health providers
- Poor mental health days
- Lack of social and emotional support
- Suicide rate

EDA

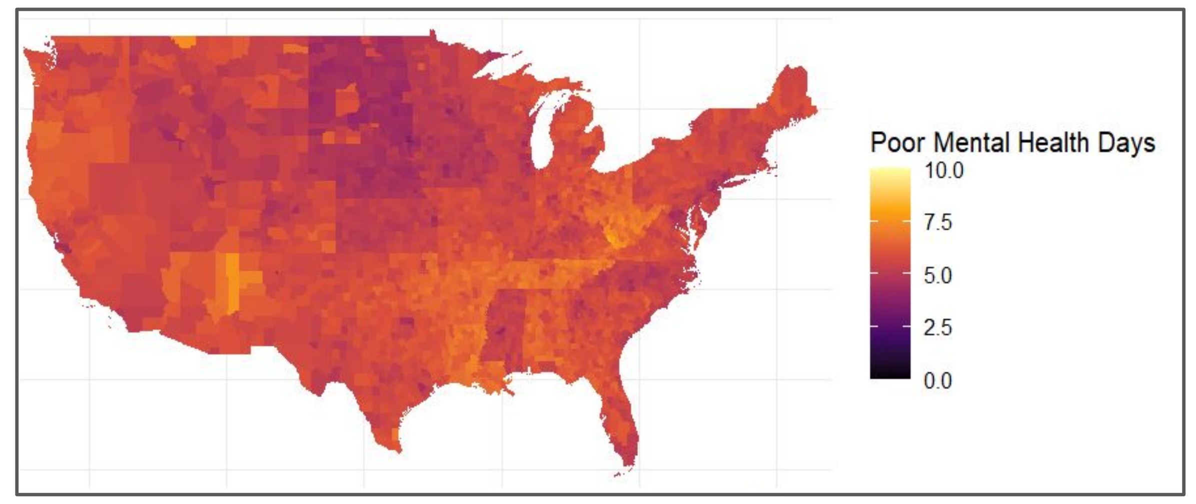

Figure 1: Choropleth map of the U.S plotting the average number of poor mental days.

The northern United States, particularly the north-central region (e.g. North Dakota, South Dakota, Minnesota), tends to report fewer poor mental health days compared to the southern U.S. On the map, these counties appear darker, signifying better mental health.

Alaska

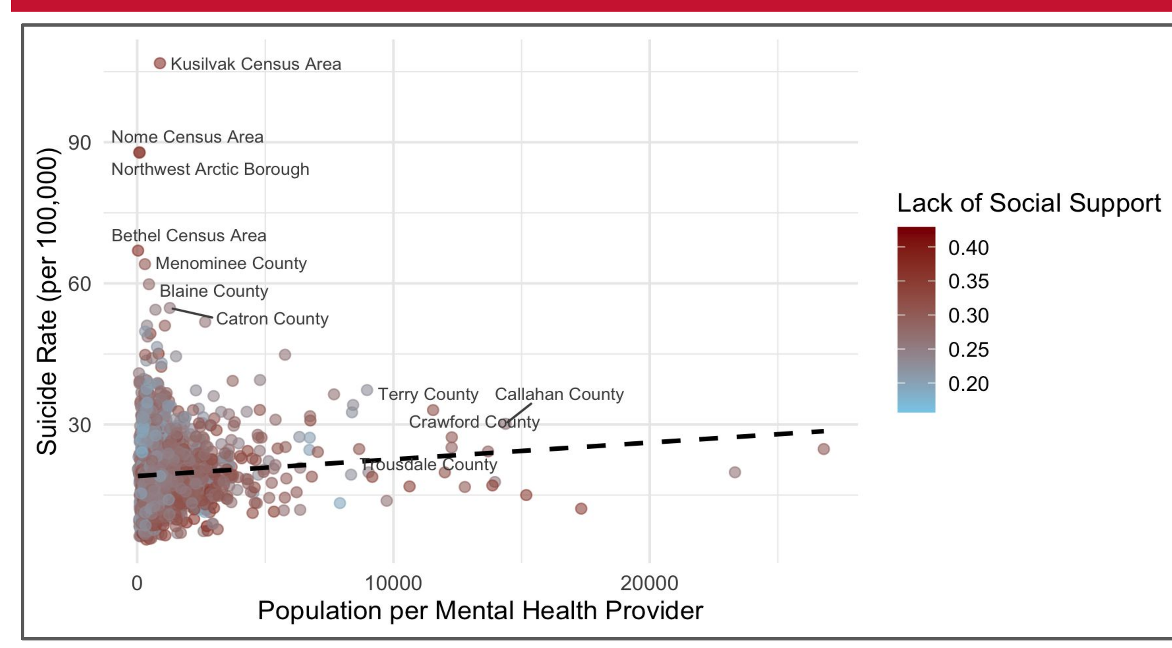

Figure 2: Plot of counties based on suicide rate and population per mental health provider ratio

Lower provider access is slightly linked to higher suicide rates, though Alaskan counties stand out with high suicide rates despite good access, possibly a reactive investment in care.

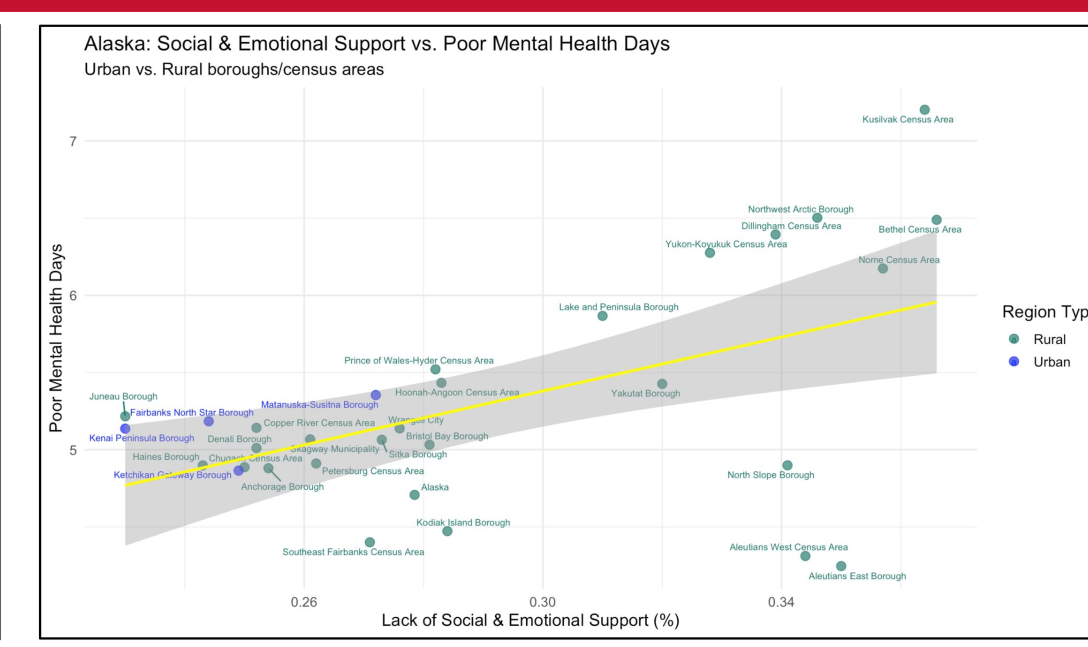

Figure 3: Scatter plot of rural and urban counties in Alaska

In Alaska's rural regions, poor mental health outcomes are strongly linked to low social support, highlighting the critical role of social determinants beyond clinical care.

Methods

Quasi-Poisson

- Used to identify significant predictors of poor mental health days, a count-based, non-negative outcome variable.
- Assumptions checked:
  - Normality of residuals using the Shapiro-Wilk test.
  - Independence of observations.
  - Equality of mean and variance
- Overdispersion was observed, indicating variance > mean, so a quasi-Poisson model was used to account for extra variance.
- Standardized all predictors before modeling to allow direct comparison of effect sizes.

Machine Learning Models

- With a large sample size, machine learning models were also implemented for prediction.
- Used to capture nonlinear patterns that Poisson did not capture.
- Models used:
  - Random Forest
  - XGBoost

Results

Quasi-Poisson Regression:

We used a Quasi-Poisson model to examine the relationship between key predictors and the number of poor mental health days. This approach accounts for overdispersion and interpretation in percentage terms.

Predictor	% Change in Poor Mental Health Days	95% Confidence Interval
Provider Ratio	+1.17%	[0.74%, 1.61%]
Loneliness	+3.54%	[2.98%, 4.09%]
Lack of Social Support	+0.92%	[0.34%, 1.50%]
Suicide Rate	+1.78%	[1.38%, 2.18%]

Figure 5: Significance of predictors in Quasi-Poisson regression

Loneliness had the strongest association, increasing poor mental health days by 3.54% per standard deviation. Suicide rate, provider ratio, and lack of social support also showed positive relationships, highlighting the impact of both clinical and social factors.

Random Forest & XG Boost:

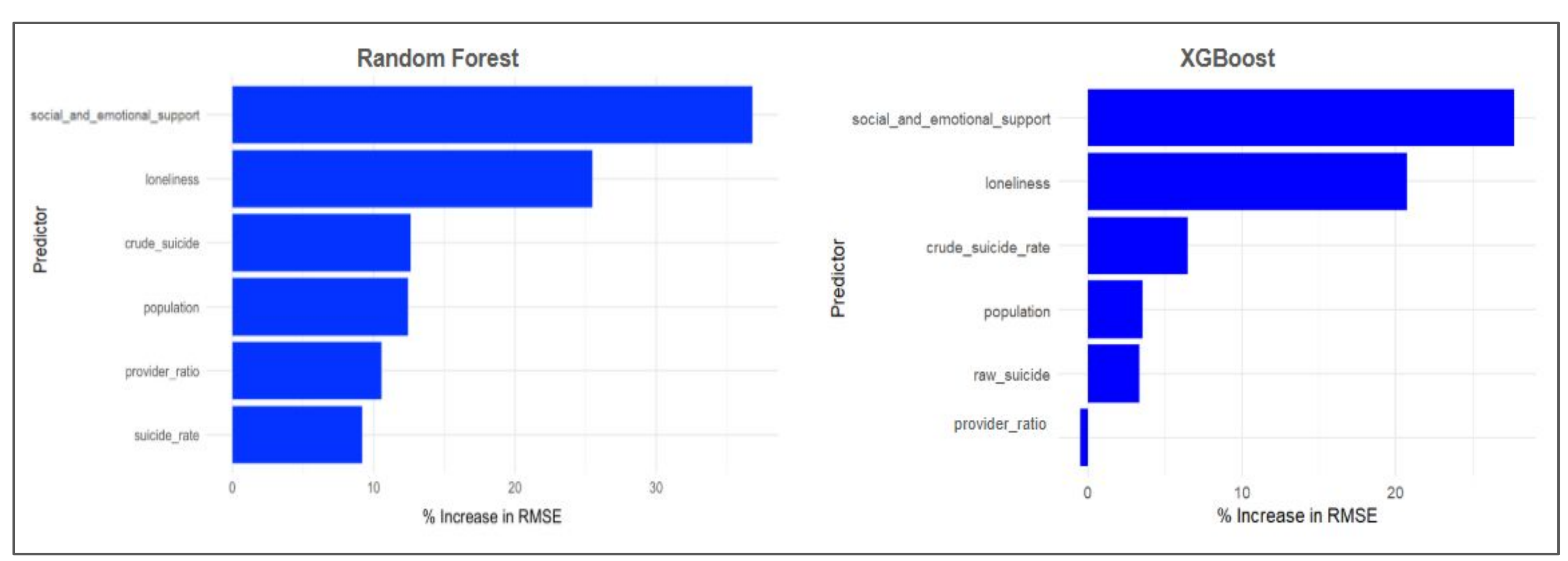

Figure 6: Feature importance of Random Forest and XGBoost with % increase in RMSE

- Both models weighed social and emotional support and feelings of loneliness heavily which supports the results we got from the Quasi-Poisson regression.
- The population to mental health provider ratio in the XGBoost has a negative % increase in RMSE, meaning it decreased the accuracy of the model.

Discussions

Conclusions:

- Mental health provider density does not significantly affect the number of poor mental health days.
- XGBoost is a better predictive model than Random Forest.
- Feelings of loneliness and social/emotional support are key predictors of poor mental health days.

Limitations:

- Analysis based on aggregated county-level data, so variation on an individual/personal level is unknown.
- Dataset lacked variables on broader social determinants and mental health provider metrics.
- Could not analyze the role of insurance rates in mental health outcomes.

Future work:

- Add behavioral and geographical factors, include quality of mental health care metrics, and consider using individuals-level data.