All-cause Mortality Rates and Primary Care Physician Supply in US Counties 2021

Chaohua Li

2024-04-05

# 1. Abstract

The all-cause mortality rate is an important indicator of general population health, and has been increasing in the US for the past ten years. Personal level access to primary care is highly associated with mortality, morbidity, and healthcare costs. There is a need to study how primary care physician supply on the geographic levels is associated with all-cause mortality rates in the US. In the project, I conducted a cross-sectional study on the county level and examine the potential association between age-adjusted all-cause mortality rate and primary care physician supply for US counties. Multilevel linear models were used to assess the association, and multiple potential confounders were accounted for in the models. Log-transformed PCP supply was negatively associated with all-cause mortality rate (beta=-19.9, 95% CI=-25.0–14.7, p-value<.001), after adjustment for some socioeconomic measures.

# 2. Introduction

The all-cause mortality rate serves as important indicator of general population health, reflecting the cumulative impact of social, economic, and healthcare factors on mortality outcomes (DeSalvo, Bloser, Reynolds, He, & Muntner, 2006). In the United States, the age-adjusted all-cause mortality rate exhibited a steady decline from 1935 until 2014, signifying improvements in public health, healthcare delivery, and medical technology (Aizer, Lleras-Muney, & Stabile, 2005). However, a notable increasing trend after 2015 has been observed (Woolf, Wolf, & Rivara, 2023).

Primary care, characterized by the provision of first-contact, person-focused, ongoing and comprehensive care, plays a pivotal role in the healthcare system. It is instrumental in preventing illnesses and death, managing chronic diseases, and coordinating care (Starfield, Shi, & Macinko, 2005). Personal level access to primary care is highly associated with mortality, morbidity, and healthcare costs (Peart, Lewis, Brown, & Russell, 2018). Despite its importance, the distribution of primary care physicians (PCPs) varies significantly across the United States, with many areas experiencing pronounced shortages, especially for rural and underserved communities.

Emerging research suggests a link between PCP supply and health outcomes, including mortality rates (Starfield et al., 2005). However, the relationship between PCP supply and all-cause mortality rates at the geographic level remains inadequately explored. This gap in the literature highlights the need of comprehensive studies to understand how variations in PCP supply might influence mortality across US counties, accounting for socioeconomic disparities.

The hypothesis of this project is that primary care physician supply is negatively associated with age-adjusted all-cause mortality rate among US counties. Since counties can be considered nested within states or regions, multilevel linear models were used.

# 3. Methods

Three data sources were used for this study. (1) Multiple Cause of Death File of 2021 from CDC WONDER (https://wonder.cdc.gov/mcd.html). County-level age-adjusted all-cause mortality rates (deaths/100k people) were obtained from this source. (2) Area Health Resource File of 2021 from U.S. Department of Health & Human Services (https://data.hrsa.gov/data/download). County-level primary care physician supply (physicians/100k people) and rural/urban categorization of the county were obtained from this file. (3) Estimates of 2021 from 2017-2021 5-year American Community Survey (https://www.nhgis.org/). States, regions, and some county-level socioeconomic measures as potential confounders for modeling were obtained from this source. These measures include %Living under poverty line, %Without a high school diploma, %Hispanic, %NH-Black, %Unemployed, %Without health insurance. Data from the three sources were merged by using the Federal Information Processing Standards (FIPS) code which served as the identifiers for counties.

The outcome was county-level age-adjusted all-cause mortality rate (n/100k), and the main predictor was county-level PCP supply (n/100k). The socioeconomic measures were consider as covariates. Since counties can be considered nested within states, multilevel linear models with random intercepts for states were used to assess the association.

Descriptive statistics were calculated and reported for the overall sample and subgroups by metro status. Bivariate association between the outcome with the main predictor and each of the covariates were assessed using bivariate multilevel linear models. Parameter coefficients, 95% confidence intervals of coefficients and p-values were reported.

To improve generalizability of results and prevent overfitting, the data was randomly split into a train set and a test set with 4/1 ratio. Cross-validation was used to assess the model performance of models with and without covariates. The final model (model with all covariates) was fit using the train set. Parameter coefficients, 95% confidence intervals of coefficients and p-values were reported. The final model fit was applied to the test set and made predictions. Predicted values vs observed values using the train set and test set were plotted to assess the prediction accuracy of the final model. All P-values were 2-sided and a P < 0.05 was considered statistically significant. RStudio was used to perform all analyses.

# 4. Results

## 4.1 Descriptive Statistics

We included 3079 US counties in this study, 1162 of which were defined as metropolitan counties (Table 1). The mean age-adjusted all-cause mortality rates were 1380/100k for all counties, 1190/100k for metropolitan counties, and 1490/100k for the non-metro counties. Metro and non-metro counties had similar mean percentages of Hispanic population and unemployment rate. On average, metro counties had higher PCP supply(60.1/100k vs 46.2/100k) and percentages of NH-Black population (10.8% vs 7.8%). Average percentages of population without high school diploma (10.5% vs 13.0%), living under federal poverty line (12.6% vs 15.6%), and without health insurance (8.5% vs 10.3%) were lower among metro counties compared to non-metro counties.

According to the scatter plot between mortality rate and log-transformed PCP supply, there was a linear association between the two, and the association did not considerably differ by metro status of the county (Figure 1).

|  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
| Table 1: Descriptive statistics of US counties in 2021.   |  | Metro | Non-metro | Overall | | --- | --- | --- | --- | |  | (N=1162) | (N=1917) | (N=3079) | | Mortality rate (n/100k) |  |  |  | | Mean (SD) | 1190 (332) | 1490 (343) | 1380 (368) | | Median [Min, Max] | 1170 [411, 3590] | 1490 [288, 2910] | 1380 [288, 3590] | | PCP supply (n/100k) |  |  |  | | Mean (SD) | 60.1 (40.2) | 46.2 (34.3) | 51.5 (37.2) | | Median [Min, Max] | 54.3 [0, 581] | 41.2 [0, 508] | 45.5 [0, 581] | | % Hispanic |  |  |  | | Mean (SD) | 10.5 (13.0) | 9.30 (14.5) | 9.76 (13.9) | | Median [Min, Max] | 5.80 [0, 95.5] | 3.82 [0, 96.3] | 4.51 [0, 96.3] | | % NH-Black |  |  |  | | Mean (SD) | 10.8 (13.4) | 7.80 (14.8) | 8.95 (14.4) | | Median [Min, Max] | 5.50 [0, 79.7] | 1.26 [0, 87.0] | 2.21 [0, 87.0] | | % No high school diploma |  |  |  | | Mean (SD) | 10.5 (4.84) | 13.0 (6.14) | 12.1 (5.81) | | Median [Min, Max] | 9.67 [2.12, 50.3] | 11.9 [1.38, 47.9] | 10.8 [1.38, 50.3] | | % Poverty |  |  |  | | Mean (SD) | 12.6 (4.97) | 15.6 (6.39) | 14.5 (6.08) | | Median [Min, Max] | 12.0 [1.80, 38.5] | 14.5 [3.54, 59.0] | 13.5 [1.80, 59.0] | | % Unemployed |  |  |  | | Mean (SD) | 5.15 (1.93) | 5.35 (2.92) | 5.27 (2.59) | | Median [Min, Max] | 4.90 [0.699, 26.2] | 4.93 [0, 32.4] | 4.92 [0, 32.4] | | % No health insurance |  |  |  | | Mean (SD) | 8.48 (4.15) | 10.3 (5.37) | 9.62 (5.02) | | Median [Min, Max] | 7.81 [1.06, 33.2] | 9.13 [1.17, 44.9] | 8.51 [1.06, 44.9] |   **Note:** aSD, standard deviation; PCP, primary care physician. |

|  |
| --- |
| Figure 1: Association between age-adjusted all-cause mortality rate and Log (PCP supply) by metropolitan status among US counties, 2011 |

## 4.2 Basic statistical analysis

Without adjustment of other measures, log-transformed PCP supply was negatively associated with mortality rate (beta=-37.3, 95% CI=-43.0- -31.5, p-value<.001), non-metro counties had significantly higher mortality rates (beta=310.5, 95% CI=287.2-333.8, p-value<.001) (Table 2); %Hispanic was negatively associated with mortality rates (beta=-9.8, 95% CI=-11.0 - -8.7, p-value<.001), %NH-Black was not significantly associated with mortality rate (beta=0.3, 95% CI=-0.9-1.5, p-value=0.657) (Table 3); both %No high school diploma (beta=16.7, 95% CI=14.2 -19.1, p-value<.001) and %Poverty (beta=23.5, 95% CI=21.4–25.6, p-value<.001) were positively associated with mortality rates (Table 4); %Unemployed (beta=32.7, 95% CI=27.7 -37.7, p-value<.001) and %No health insurance (beta=13.0, 95% CI=9.7 -16.4, p-value<.001) were associated with higher mortality rates (Table 5).

|  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
| Table 2: Results of unadjusted multilevel linear regression.   | Predictors | Estimates | CI | p | Estimates | CI | p | | --- | --- | --- | --- | --- | --- | --- | | (Intercept) | 1444.6 | 1389.1 – 1500.0 | <0.001 | 1129.1 | 1076.0 – 1182.3 | <0.001 | | log(pcp\_100k + 0.01) | -37.3 | -43.0 – -31.5 | <0.001 |  |  |  | | Metropolitan status:Non-metro |  |  |  | 310.5 | 287.2 – 333.8 | <0.001 | | N | 51 state | 51 state | 51 state | 51 state | 51 state | 51 state | | Observations | 3079 |  |  | 3079 |  |  | | Marginal R2 / Conditional R2 | 0.042 / 0.260 |  |  | 0.156 / 0.371 |  |  | | AIC | 44497.021 |  |  | 44036.552 |  |  |   **Note:** ap, p-values; CI, 95% confidence interval. |

|  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
| Table 3: Results of unadjusted multilevel linear regression.   | Predictors | Estimates | CI | p | Estimates | CI | p | | --- | --- | --- | --- | --- | --- | --- | | (Intercept) | 1407.0 | 1350.3 – 1463.6 | <0.001 | 1311.5 | 1257.4 – 1365.7 | <0.001 | | % Hispanic | -9.8 | -11.0 – -8.7 | <0.001 |  |  |  | | % NH-Black |  |  |  | 0.3 | -0.9 – 1.5 | 0.657 | | N | 51 state | 51 state | 51 state | 51 state | 51 state | 51 state | | Observations | 3079 |  |  | 3079 |  |  | | Marginal R2 / Conditional R2 | 0.119 / 0.352 |  |  | 0.000 / 0.230 |  |  | | AIC | 44402.708 |  |  | 44658.061 |  |  |   **Note:** ap, p-values; CI, 95% confidence interval. |

|  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
| Table 4: Results of unadjusted multilevel linear regression.   | Predictors | Estimates | CI | p | Estimates | CI | p | | --- | --- | --- | --- | --- | --- | --- | | (Intercept) | 1133.5 | 1080.8 – 1186.1 | <0.001 | 993.7 | 943.8 – 1043.5 | <0.001 | | % No high school diploma | 16.7 | 14.2 – 19.1 | <0.001 |  |  |  | | % Poverty |  |  |  | 23.5 | 21.4 – 25.6 | <0.001 | | N | 51 state | 51 state | 51 state | 51 state | 51 state | 51 state | | Observations | 3079 |  |  | 3079 |  |  | | Marginal R2 / Conditional R2 | 0.067 / 0.236 |  |  | 0.150 / 0.285 |  |  | | AIC | 44483.556 |  |  | 44215.238 |  |  |   **Note:** ap, p-values; CI, 95% confidence interval. |

|  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
| Table 5: Results of unadjusted multilevel linear regression.   | Predictors | Estimates | CI | p | Estimates | CI | p | | --- | --- | --- | --- | --- | --- | --- | | (Intercept) | 1137.5 | 1078.6 – 1196.4 | <0.001 | 1200.6 | 1141.8 – 1259.3 | <0.001 | | % Unemployed | 32.7 | 27.7 – 37.7 | <0.001 |  |  |  | | % No health insurance |  |  |  | 13.0 | 9.7 – 16.4 | <0.001 | | N | 51 state | 51 state | 51 state | 51 state | 51 state | 51 state | | Observations | 3079 |  |  | 3079 |  |  | | Marginal R2 / Conditional R2 | 0.049 / 0.271 |  |  | 0.030 / 0.240 |  |  | | AIC | 44494.841 |  |  | 44598.291 |  |  |   **Note:** ap, p-values; CI, 95% confidence interval. |

## 4.3 Full analysis

The train set containted 2463 counties from all the 51 states. After adjusting for socioeconomic measures, the effect of log-transformed PCP supply on mortality rate was attenuated but still significant (beta=-19.9, 95% CI=-25.0–14.7, p-value<.001). Non-metro counties had significantly higher mortality rates (beta=177.8, 95% CI=153.4-202.3, p-value<.001); %Hispanic (beta=-13.4, 95% CI=-14.6 - -12.2, p-value<.001) and %NH-Black (beta=-4.2, 95% CI=-5.2–3.1, p-value=0.657) were significantly associated with mortality rate; both %No high school diploma (beta=15.0, 95% CI=12.0 -18.1, p-value<.001), %Poverty (beta=15.0, 95% CI=12.4–17.7, p-value<.001) and %Unemployed (beta=8.7, 95% CI=3.5–13.9, p-value=0.001) were positively associated with mortality rates; %No health insurance (beta=-4.3, 95% CI=-7.8 –0.8, p-value=0.016) was negatively associated with mortality rates (Table6). In Figure 2, the predicted/observed values for the test data illustrated similar distribution with the train data, indicating absence of considerable overfitting of the final model.

|  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
| Table 6: Results of multivariate multilevel linear regression.   | Predictors | Estimates | CI | p | | --- | --- | --- | --- | | (Intercept) | 1057.6 | 999.2 – 1116.1 | <0.001 | | PCP supply | -19.9 | -25.0 – -14.7 | <0.001 | | Metropolitan status:Non-metro | 177.8 | 153.4 – 202.3 | <0.001 | | % Hispanic | -13.4 | -14.6 – -12.2 | <0.001 | | % NH-Black | -4.2 | -5.2 – -3.1 | <0.001 | | % No high school diploma | 15.0 | 12.0 – 18.1 | <0.001 | | % Poverty | 15.0 | 12.4 – 17.7 | <0.001 | | % Unemployed | 8.7 | 3.5 – 13.9 | 0.001 | | % No health insurance | -4.3 | -7.8 – -0.8 | 0.016 | | N state | 51 | 51 | 51 | | Observations | 2463 |  |  | | Marginal R2 / Conditional R2 | 0.422 / 0.559 |  |  | | AIC | 34403.777 |  |  |   **Note:** ap, p-values; CI, 95% confidence interval. |

|  |
| --- |
| Figure 2: Predicted vs Observed Values by Train/Test Data |

# 5. Discussion

## 5.1 Summary and Interpretation

## 5.2 Strengths and Limitations

## 5.3 Conclusions

# 6. References

Aizer, A., Lleras-Muney, A., & Stabile, M. (2005). Access to Care, Provider Choice, and the Infant Health Gradient. *The American Economic Review*, *95*(2), 248–252. <https://doi.org/10.1257/000282805774670248>

DeSalvo, K. B., Bloser, N., Reynolds, K., He, J., & Muntner, P. (2006). Mortality prediction with a single general self-rated health question. A meta-analysis. *Journal of General Internal Medicine*, *21*(3), 267–275. <https://doi.org/10.1111/j.1525-1497.2005.00291.x>

Peart, A., Lewis, V., Brown, T., & Russell, G. (2018). Patient navigators facilitating access to primary care: A scoping review. *BMJ Open*, *8*(3), e019252. <https://doi.org/10.1136/bmjopen-2017-019252>

Starfield, B., Shi, L., & Macinko, J. (2005). Contribution of primary care to health systems and health. *The Milbank Quarterly*, *83*(3), 457–502. <https://doi.org/10.1111/j.1468-0009.2005.00409.x>

Woolf, S. H., Wolf, E. R., & Rivara, F. P. (2023). The New Crisis of Increasing All-Cause Mortality in US Children and Adolescents. *JAMA*, *329*(12), 975. <https://doi.org/10.1001/jama.2023.3517>