Healthy Places, Healthy Lives
Methodology

3/3/2020

Introduction

For many Americans the neighborhood in which they live has a significant impact on their opportunity for optimal health. However, most measures of community characteristics and health outcomes are made at the county or metro level. These often hide vulnerable populations, diluting their impact in mapping and analysis.

Healthy Places, Healthy Lives bridges this gap by combining two robust, Census Tract-level, measures into a single visualization. For health outcomes, the visualization uses LEB estimates created by the United States Small-Area Life Expectancy Estimates Project (USALEEP). For neighborhood characteristics the visualization incorporates the Health Opportunity Index (HOI), a Census Tract-level index comprising 13 Social Determinant of Health (SDOH) indicators, created by the Virginia Department of Health, Office of Health Equity.

The project includes two major facets: 1) a statistical examination of state and regional relationships among the HOI indicators and the USALEEP LEB estimates, and, 2) an interactive visualization that allows users to explore these relationships and potential strategies for improving regional life expectancy.

Partners at the Virginia Commonwealth University Department of Biostatistics performed Weighted-Quantile Summation regression (WQS) at both the state and regional level. WQS regression identified the level of correlation and assigned weights to each of the HOI indicators. These weights rank the relative influence of each indicator on life expectancy and allow estimates that changes in these indicators may have on life expectancy.

The visualization identifies the five HOI indicators expected to have the most impact on life expectancy regionally. Using interactive sliders, users can explore the expected effect changes in these indicators may have on LEB in their communities. The visualization includes a list of organizations and evidence-based

practices to address each SDOH factor, while an interactive map allows users to identify the census tracts where interventions are needed most.

Life Expectancy at Birth Estimates

The United States Small-Area Life Expectancy Estimates Project (USALEEP) is the first public health outcome measure available nationwide at the census tract-level measuring Life Expectancy at Birth (LEB) for nearly every census tract in the country. A joint effort of The Robert Wood Johnson Foundation (RWJF), National Association for Public Health Statistics and Information Systems (NAPHSIS), and the National Center for Health Statistics (NCHS) at the Centers for Disease Control (CDC), USALEEP data provide unparalleled insights into community health and demonstrate that not everyone has the same opportunity to be healthy where they live.

There were 207 Census Tracts in with missing LEB measures in the Virginia USA LEEP data file. For this project, missing estimates were imputed using k-nearest neighbor clustering using *KNNImpute* function. "k" was selected as the square root of the sample size (N = 1875) rounded to the nearest integer. The raw and imputed LEB measures are summarized and it was observed that imputation did not meaningfully change the mean, median or standard deviation.

The Health Opportunity Index

The Health Opportunity Index (HOI) is a composite measure of the Social Determinants of Health (the social, economic, educational, demographic, and environmental factors that relate to a community's well-being and the health status of a population) developed at the Census Tract level. It is comprised of 13 indicators that reflect a broad array of community concerns relating to (1) Air Quality, (2) Population Churning, (3) Population Weighted Density, (4) Walkability, (5) Affordability, (6) Education, (7) Food Accessibility, (8) Material Deprivation, (9) Employment Access, (10) Income Inequality, (11) Job Participation, (12) Access to Care, (13) Spatial Segregation. These indicators are weighted, grouped into four profiles using Principal Component Analysis (PCA). The HOI is the weighted sum of the indicators. Census Tracts are scaled into quintiles corresponding to "Very Low Opportunity", "Low Opportunity", "Moderate Opportunity", "High Opportunity" and "Very High opportunity" levels. Each level represents the opportunity to lead a long and healthy life, with higher opportunity corresponding with increased chance for healthy outcomes.

To provide assurance that the HOI was truly sensitive to major shifts in health status throughout the state, we used the most general definitions of good health, i.e., "life expectancy at birth" (LE) data to determine if the HOI demonstrated systematic public health differences between areas

When the Life Expectancy was modeled against the composite index (HOI), they explained over 80 percent for life expectancy when space was taken into account (Geographically Weighted Regression). Individual indicators contributed differently to the overall prediction coefficient weights depending on location within the state. This indicates that space and place are vital to understanding the distribution of life chances

Weighted Quantile Sum Regression

To estimate the relationship of the HOI indicators with the USA LEEP data, the project used Weighted Quantile Sum (WQS) regression.[1,2] The WQS regression method is able to deal with collinearity issues associated with highly correlated data, which violates assumptions of other methods. HOI indicator values were converted into z-scores by subtracting the indicator's mean and dividing by that indicator's standard deviation, while indicators negatively associated with LEB were also multiplied by -1. These choices implied that an increasing numerical value for any indicator represented an increasingly beneficial value and lead to a positive HOI framing, with larger values indicative of less disadvantaged census tracts. Weighted quantile summation regression (WQS) was used to estimate indicator weights by regressing the 13 indicators against LEB, where each indicator is weighted to maximize the association between the overall HOI scores and LEB. These indicator weights sum to one, so that each weight is interpretable as a percentage of association with LEB explained by the corresponding indicator. The R computational software was used, specifically the WQS package for determining indicator weights for calculating the HOI.

The WQS regression revealed several details about the relationship between social determinants of health and life expectancy. The Health Opportunity Index has a strong, positive correlation with the USALEEP LEB estimates, explaining 76 percent of the variation in LEB. At the state level, housing and transportation affordability (weight = 26%) and education (weight = 37%) were allotted the largest

[1] 2015. Czarnota, et al. "Assessment of Weighted Quantile Sum Regression for Modeling Chemical Mixtures and Cancer Risk'. *Cancer Informatics*. Vol 14 (S2). Pp. 159-171.

[2] 2015. Carrico, et al. "Characterization of a Weighted Quantile Sum Regression for Highly Correlated Data in a Risk Analysis Setting." *Journal of Agricultural, Biological, and Environmental Statistics.* Vol 20(1). Pp. 100-120.

proportion of the association between the HOI and LEB. Other indicators with greater than 5 percent weight were population churning (7%), material deprivation (10%), and job participation (7%).

The positive correlation was generally strong at the regional level as well, including at the Local Health District level. In Virginia, Local Health Districts vary in scale from a single independent city to several counties. The WQS regression found statistically significant, positive correlations in 29 of Virginia's 35 Local Health Districts. The remaining districts were all rural and combined with other districts in the Healthy Places, Healthy Lives (HPHL) regions used in the final analysis. For a full list of weights, see the appendix. The full results from the WQS analysis are available here.

Visualization Calculations

The average LEB for each Healthy Places, Healthy Lives region was calculated by summing up all the predicted LEB values of all census tracts that are within the HPHL region and dividing the sum by the number of census tracts.

Moving any of the five slider bars on the right to a desired percentage changes the indicator value in question by the amount shown. A new z-score is calculated for that indicator for each census tract, using the unadjusted mean and standard deviation. The predicted LEB for each Census Tract is then calculated by multiplying the transformed z-score by the model weights and summed. The model slope, specific for each region, is then multiplied by the aggregated sum and the result added to the intercept and the residual.

Appendix

Indicator Weights per Healthy Places, Healthy Lives Region													
	Air Quality	Population Churning	Population Density	Walkability	Affordability	Education	Food Accessibility	Material Deprivation	Employment Access	Income Inequality	Job Participation	Access to Care	Segregation
Alleghany-Roanoke Alleghany Roanoke	0.05	0.16	0.06	0.03	0.08	0.21	0.08	0.17	0.03	0.03	0.06	0.03	0.01
Central Shenandoah	0.05	0.14	0.00	0.00	0.09	0.24	0.12	0.03	0.00	0.06	0.19	0.00	0.07
Central Virginia	0.09	0.04	0.01	0.01	0.17	0.40	0.12	0.02	0.01	0.05	0.06	0.01	0.01
Chesapeake Eastern Shore Three Rivers	0.06	0.05	0.09	0.02	0.22	0.04	0.09	0.01	0.01	0.01	0.01	0.11	0.27
Greater Richmond Chesterfield Henrico Richmond City	0.12	0.01	0.01	0.02	0.10	0.33	0.03	0.17	0.09	0.04	0.03	0.02	0.03
Hampton Roads Chesapeake Hampton Norfolk Portsmouth Virginia Beach	0.06	0.01	0.01	0.01	0.17	0.41	0.01	0.19	0.04	0.02	0.03	0.03	0.02

Lord Fairfax	0.04	0.04	0.01	0.03	0.26	0.37	0.05	0.09	0.03	0.01	0.06	0.00	0.00
New River	0.02	0.01	0.09	0.04	0.13	0.30	0.03	0.13	0.06	0.11	0.07	0.01	0.00
Northern Virginia Alexandria Arlington Fairfax Loudoun	0.01	0.17	0.01	0.05	0.28	0.17	0.12	0.06	0.01	0.01	0.01	0.09	0.02
Peninsula	0.04	0.11	0.01	0.00	0.16	0.37	0.04	0.15	0.00	0.05	0.03	0.00	0.04
Piedmont-Southside Piedmont Southside	0.26	0.03	0.02	0.04	0.01	0.07	0.02	0.12	0.06	0.28	0.01	0.01	0.07
Pittsylvania-Danville	0.06	0.28	0.08	0.14	0.11	0.03	0.07	0.07	0.01	0.05	0.01	0.03	0.04
Prince William	0.01	0.20	0.03	0.13	0.03	0.15	0.14	0.05	0.03	0.14	0.01	0.02	0.07
Southwest Virginia Cumberland-Plateau Lenowisco Mount Rogers	0.13	0.04	0.05	0.00	0.21	0.17	0.01	0.15	0.17	0.09	0.03	0.01	0.09
Thomas Jefferson	0.05	0.04	0.01	0.05	0.19	0.20	0.03	0.14	0.03	0.05	0.01	0.08	0.12
Upper Rappahannock River Rappahannock Rappahannock Rapidan	0.06	0.01	0.04	0.11	0.04	0.37	0.04	0.02	0.03	0.01	0.01	0.14	0.12

West Piedmont	0.08	0.05	0.12	0.03	0.08	0.23	0.08	0.05	0.00	0.05	0.01	0.19	0.03
Western Tidewater	0.13	0.09	0.00	0.01	0.03	0.16	0.00	0.26	0.01	0.06	0.02	0.02	0.21