

County Health Rankings & Roadmaps Working Paper

Updates to Methods and Tools for Practical Application to Improve Health and Equity

March 2024



**County Health
Rankings & Roadmaps**



**University of Wisconsin
Population Health Institute**
UNIVERSITY OF WISCONSIN
SCHOOL OF MEDICINE AND PUBLIC HEALTH

Table of Contents

Introduction	3
1. Refine CHR&R approach to county-level summary measures of population health	4
1A. New and improved methods: Cluster analysis and data-informed county comparisons for summary measures of population health.....	5
1B. How a data-informed approach to comparing county health can support action	6
2. Respond to data source infrastructure changes and implications for population counts	7
2A. Background on presentation standards for age and racialized groups.....	7
2B. Proposed methods changes affecting a subset of CHR&R measures	8
Appendices	10
References	15

Introduction

This working paper summarizes two aspects of research and development (R&D) for the County Health Rankings & Roadmaps (CHR&R) program related to updated methods and tools for practical application in 2024:

1. Refine CHR&R approach to county-level summary measures of population health.
2. Respond to data source infrastructure changes and implications for population counts.

Since 2010, CHR&R has provided an easy-to-use snapshot of county health based on a model of population health, including:

- Health Outcomes—based on an equal weighting of Length of Life and Quality of Life, and
- Health Factors—based on weighted scores of four factors: Health Behaviors, Clinical Care, Social and Economic Factors, and the Physical Environment.

CHR&R has ranked the health of nearly all U.S. counties using summary measures of the Health Outcomes and Health Factors, weighted according to their relative contribution.

CHR&R continually engages in R&D related to what we measure and what key concepts we elevate, why and how. We are evolving our suite of tools and resources through innovation focused on **social solidarity¹ⁱ and the structural determinants of healthⁱⁱ**.

¹ Social solidarity is the shared responsibility and carrying of costs to assist others. It recognizes interdependence between groups of people. Social solidarity is an action that reflects the values of reciprocity, cooperation, and shared responsibility by contributing wealth, time resources, and changes for the benefit of others.

1. Refine CHR&R approach to county-level summary measures of population health

Summary measures (also referred to as composite indicators or indices) combine individual measures with the intent to capture relevant, multi-part aspects of a concept, such as length and quality of life. CHR&R, like numerous data platforms, distills vast amounts of data into a small subset of measures that allow comparisons of dimensions such as health outcomes or socioeconomic environments between states, counties and/or demographically similar communities.

Summary measures of population health are important to assess, prioritize and improve the health of communities, and therefore must be generated with attention to responsible use of data and transparency in methods.

To calculate summary measures, CHR&R first standardizes the component measures to a single scale (a z-score), and then aggregates the component measures using assigned weights.

- For Health Outcomes, CHR&R standardizes and aggregates five measures into two outcome areas: Length of Life and Quality of Life. These outcome areas are aggregated using equally assigned weights to calculate a z-score reflecting a Health Outcome summary measure for each county.
- For Health Factors, CHR&R standardizes and aggregates nearly 30 measures into 13 health factor areas using assigned weights (see Appendix A), which are further aggregated into four health factor components (i.e., Health Behaviors, Clinical Care, Social & Economic Factors, and Physical Environment). These health factor components z-scores are then aggregated into a Health Factor summary measure for each county.

Until 2024, CHR&R sorted the county z-score for Health Factors and Health Outcomes on a spectrum of healthiest to least healthy within states to provide each county with an ordinal rank. For example, each county in a state receives a unique rank from #1 (healthiest) to #X (total ranked counties in a state, least healthy). The ordinal rank has provided a simple measure of the relative health of a county within a state. However, ordinal ranks do not necessarily reflect practical or even statistically significant differences between counties and, without an understanding of the underlying data, can be misinterpreted as hard truths. Random variation in underlying county values, or “statistical noise,” sometimes makes it difficult to make meaningful distinctions between counties that are in the middle of the distribution. Thus, community action driven primarily by ordinal ranks, especially for counties in the middle quartiles, is at a greater risk of being misinformed. In addition, the creation of within-state rankings has precluded the comparison of counties across the U.S., a feature that would be useful for counties on the border lines of states, for instance, or counties with shared historic, geographic, demographic, or other similarities across the country. Below, we outline an improved method for summary measures of population health at the county level. The new method supports data-informed comparisons across the U.S. and a focus on meaningful differences that can better support action.

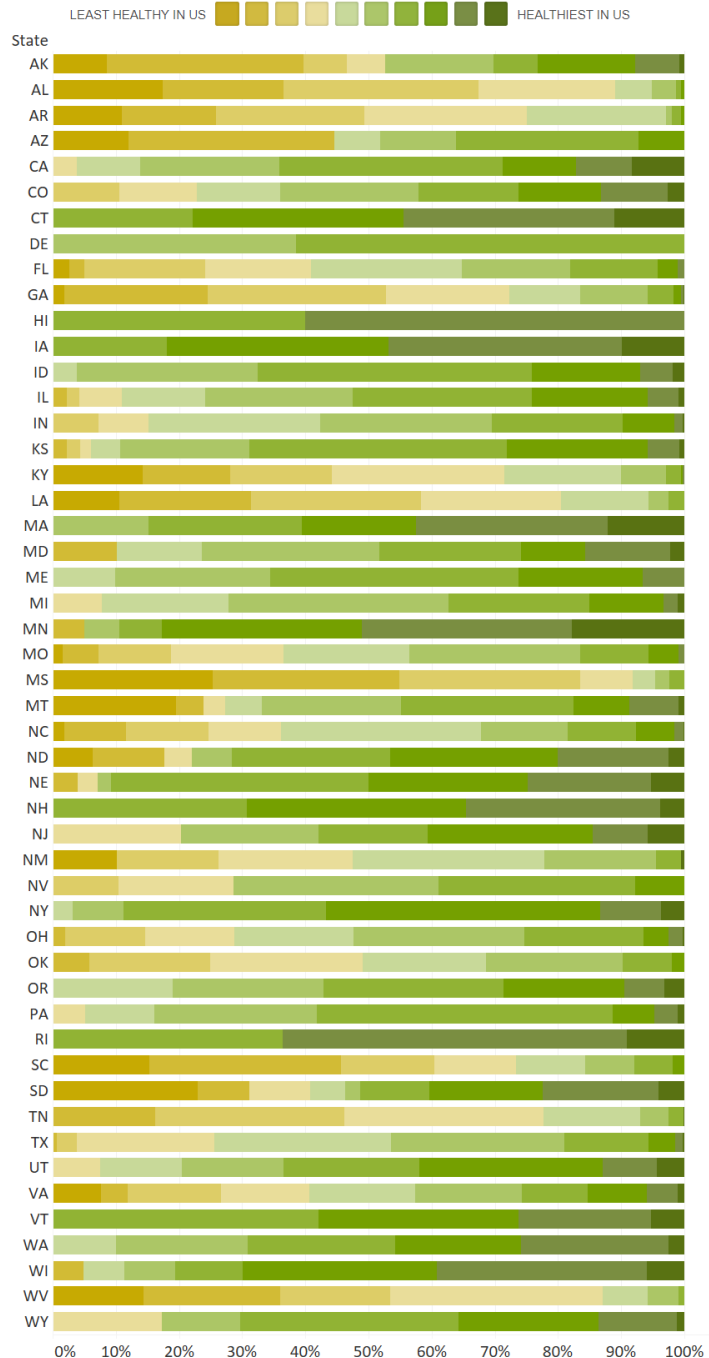
1A. New and improved methods: Cluster analysis and data-informed county comparisons for summary measures of population health

Cluster analysis, a way of creating data-informed groupings, is a collection of methods that can be used to determine which counties are and are not meaningfully different from one another – and how much they differ. Data-informed comparisons can focus attention on meaningful differences and where to prioritize action.

In 2024, CHR&R will apply a cluster analysis approach to summary measures of population health, namely the composite Health Outcomes and Health Factors z-score values for each county. Specifically, CHR&R will calculate z-scores using a national distribution (rather than within-state) and use K-means clusteringⁱⁱⁱ to partition n observations into $k=10$ clusters, identifying the optimal grouping of the counties for each possible cluster. Clusters are determined by creating 10 random centroids of the data and then assigning each data point to the nearest centroid. The centroid of each cluster is then moved to the average of the data in the cluster and the process is repeated until no data points change groups. See Figure 1 and Appendix B for data-informed clusters by state and the geographic distribution across the U.S.

CHR&R will impose a cap of 10 clusters, k , based on analyses to assess the potential loss of information in limiting clusters (using the Wasserstein metric,

Figure 1. Distribution of Health Outcome National Z-score Clusters by State



also known as Earth Mover's Distance, a measure of the distance between two probability distributions) and to support ease of communication.

CHR&R will apply a cluster analysis to all counties nationally to generate the updated data-informed approach to comparing counties. In 2024, each county will be grouped within data-informed clusters (e.g., cluster 1-10) based on their z-score rather than sorting z-scores and applying an ordinal rank.

1B. How a data-informed approach to comparing county health can support action

CHR&R's updated approach to comparing county health (cluster analysis) provides additional context for summary measures of population health and a data-informed understanding of the health status of a county that can be compared within several settings. For example, by switching from 50 state-specific distributions of z-scores to a single national distribution, this approach enables comparison of a county with similar counties of a state, region, or nationally, based on the data-informed grouping.

Data-informed comparisons are an enhancement of ordinal rank methods in several ways that can support community-led health improvement efforts. Specifically, the updated approach to comparing counties more fully considers the underlying spread and imprecision in the data and better identifies the relative position of counties within a setting. This approach can be applied to several types of summary measures, including measures of health disparities (forthcoming), reveal counties that are similar or not meaningfully different from one another, and support identification of peer-health counties across geographies and over time. This approach can also support comparisons within settings that extend outside of state jurisdiction where structural determinants influence the health of populations, such as regions with types of policy-relevant classification (e.g., rural and urban), disinvestment or development, environmental resources or disaster, or sovereignty (e.g., Tribal nations).

Counties that are similar in Health Outcomes or Health Factors within and across states may be more motivated to work together to advocate for structural interventions that can advance health and equity under the new approach. Unlike with ordinal ranks, a county does not have to outperform other counties to see an improvement because data-informed groupings are not constrained to a certain number of counties in each group. Rather than communicating solely a frame of competition based on ordinal rankings within states, the updated approach to comparing counties may encourage collaboration and solidarity, leading to resource allocation according to need.

2. Respond to data source infrastructure changes and implications for population counts

In 2022, the National Center for Health Statistics (NCHS) – the primary source of CHR&R’s vital statistics – made changes to the way data are reported by age and race. In response, CHR&R will adapt its methods for vitality measures calculated using specific age groups and those disaggregated by race.

2A. Background on presentation standards for age and racialized groups leading to an evolved approach

CHR&R currently offers data disaggregated by race for 21 Health Outcome and Health Factor measures; 13 of these measures are calculated using NCHS data (Appendix C). CHR&R presents these disaggregated data with the intent to communicate inequitable exposures to less healthy community conditions, which would otherwise be hidden at the overall county level.

Where possible, CHR&R follows the Office of Management and Budget (OMB) standards for presentation of racialized population groups. As OMB standards and data collection methods have evolved, the practice of race bridging has been used to preserve comparable categories representing racialized groups across data sources. Race bridging has been used to introduce compatibility between multi-race and single-race data collection systems such that race-specific statistics can be compared over time even as data collection systems evolve and diverge.² While race bridging methods can maintain compatible categories over time, these methods have the disadvantage of clouding self-identification through mathematical reassignment of multi-race identities to multiple single-race categorizations and disproportionately impact racialized groups with smaller populations like those classified as American Indian and Alaska Native and Asian or Pacific Islander^{iv}.

Through 2023, CHR&R had, alongside many other data systems, presented data disaggregated among the four minimum categories specified in the 1977 OMB standards for measures constructed from NCHS data. NCHS continued to use the older standards as states individually updated their reporting to meet the 1997 standards according to different timetables. This was made possible via methods collaboratively developed by NCHS and the U.S. Census Bureau that bridged the 31 race categories introduced by the 1997 OMB standards (and used in the decennial Census since 2000) to the four categories widely adopted as a result of the 1977 OMB standards. With all 50 states finally reporting to the same standard in 2017³, NCHS no longer had the need to produce custom bridged-race population estimates and has since switched to using the U.S. Census Bureau’s publicly available, annual, postcensal population estimates.

In 2024, CHR&R will likewise adjust methods and shift to census data as the primary source of population estimates. The discontinuation of the NCHS race bridging methods and shift to census data

² National Vital Statistics System. Bridged Race Categories. https://www.cdc.gov/nchs/nvss/bridged_race.htm

³ West Virginia was the last state to report multiple-race data to NCHS in September 2017. https://www.cdc.gov/nchs/data/nvsr/nvsr68/nvsr68_09-508.pdf

population estimates introduces two key issues for the 13 CHR&R measures of mortality and natality which previously used NCHS bridged-race population estimates:

1. Changes to race categories: The census population estimates follow the 1997 OMB standards and there is currently no guidance for the construction of bridged-race population estimates compatible with the 2020 census methods of collection and coding for race and ethnicity data. This means that CHR&R (and other data systems) must adopt the race categories specified in the 1997 OMB standards or pursue reconstruction of race bridging methods compatible with the 2020 census data collection and coding methods.
2. Loss of data for granular age groups: The census population estimates are presented in 5-year age categories in contrast to the single-year and 10-year age categories that were formerly available through NCHS data. The loss of the single-year categories has implications for measures that are age-adjusted and measures that utilize specific, granular population estimates in their calculation. Appendix C indicates CHR&R measures implicated and includes Premature Death (Years of Potential Life Lost), Life Expectancy, Premature Age-Adjusted Mortality, Child Mortality, Suicides, and Infant Mortality. With the shift to census population estimates, these measure calculations will require new methods that draw from available data.

2B. Proposed methods changes affecting a subset of CHR&R measures

Categorization of race data

The data infrastructure changes impacting the race categories available in population estimations offers CHR&R an opportunity to reevaluate our approach to data disaggregation and refocus our efforts to increase the visibility of structural racism and its effects. These changes also provide an opportunity to modernize our approach to the categorization of racialized population groups to better align with the identities that individuals and communities have self-assigned.

In 2024, CHR&R will shift from the current four minimum categories for race representative of the 1977 OMB standards toward the 31 race categories introduced by the 1997 OMB standards. As first step towards the ultimate goal of presenting as many of the 31 categories as data availability will allow, in 2024 CHR&R will provide the following six categories in county data snapshots: White; Black or African American; American Indian or Alaska Native; Asian; Native Hawaiian or Other Pacific Islander; and “Two or more races”. We recognize that the practice of aggregating those reporting more than one race into a joint category for the sake of simplicity and preservation of small numbers can both aid and hinder efforts to advance racial equity – simultaneously maintaining visibility for those who would otherwise not be captured among the single-race categories while creating a category too heterogeneous to hold meaning for interpretation of the group’s health experience. CHR&R will present a “Two or more races” category to support comparability with external data presentations and between jurisdictions while continuing to explore the capability of providing a flexible set of categories for representation of racialized people in county and state snapshots.

Table 1: CHR&R Categories for Presentation of Data for Racialized Population Groups

2023	2024
American Indian or Alaska Native	American Indian or Alaska Native
Asian	Asian
Black	Black
Hispanic	Hispanic
White	Native Hawaiian or Other Pacific Islander
	Two or more races
	White

Age group categorization

The loss of single-year and 10-year age group data previously available through the NCHS population estimates implicates two classes of affected measures: age-adjusted measures and those that require population estimates for specific age groups not represented in the default 5-year census categories, such as Infant Mortality. To maintain as much consistency as possible with previous methods, CHR&R will combine available data from NCHS and the census to recreate the age groupings necessary to calculate CHR&R measures and adjust the definition of measures where this is not possible.

Most 5-year age categories from the census can be combined to create the age categories CHR&R uses for age-adjustment with a few key exceptions in the youngest age groups. For instance, when CHR&R calculates Premature Death, the youngest three age groups used are <1 year, 1-14 years, and 15-24 years. The census, on the other hand, has the 5-year age categories <5 years, 5-9 years, 10-14 years, 15-19 years, and 20-24 years. To recreate the required <1 year and 1-14 years age groups, CHR&R will approximate the size of the infant population in a given year using the number of live births from NCHS birth data. This approach to approximating the infant population aligns with CDC WONDER, NCHS’s web-based system for disseminating public health data. This number will then be subtracted from the census’ <5 years age group to create a custom 1-4 years age group, which will be added to the census’ 5-9 years and 10-14 years age groups resulting in the necessary <1 year and 1-14 years groupings. The single-year age group of <1 year made available by this method will also be used to calculate the CHR&R Infant Mortality measure (deaths before one year of age per 1,000 live births) and other CHR&R measures that require an approximation of the infant population.

The census age categorization will also require an adjustment to the definition and calculation of the CHR&R measure, Child Mortality. Through 2023, CHR&R defined Child Mortality as the number of deaths among residents under age 18 per 100,000 population. The number of residents under age 18 was readily obtainable through a combination of single-year age groups from the NCHS population estimates. The 5-year census age groups necessitate an adjustment to the definition and calculation of Child Mortality. In 2024, Child Mortality will be defined as the number of deaths among residents under age 20 per 100,000 population and calculated using the census 5-year age categories and the infant age population approximated by the number of live births.

Appendices

- **Appendix A. Select Measures of Population Health and Assigned Weights (2024)**
- **Appendix B. Geographic Distribution of the National Health Outcomes Z-score Values and the 10 National Health Outcomes Z-score Clusters**
- **Appendix C. CHR&R Measures Affected by Changes in Data Source Infrastructure**

Appendix A. Select Measures of Population Health and Assigned Weights (2024)

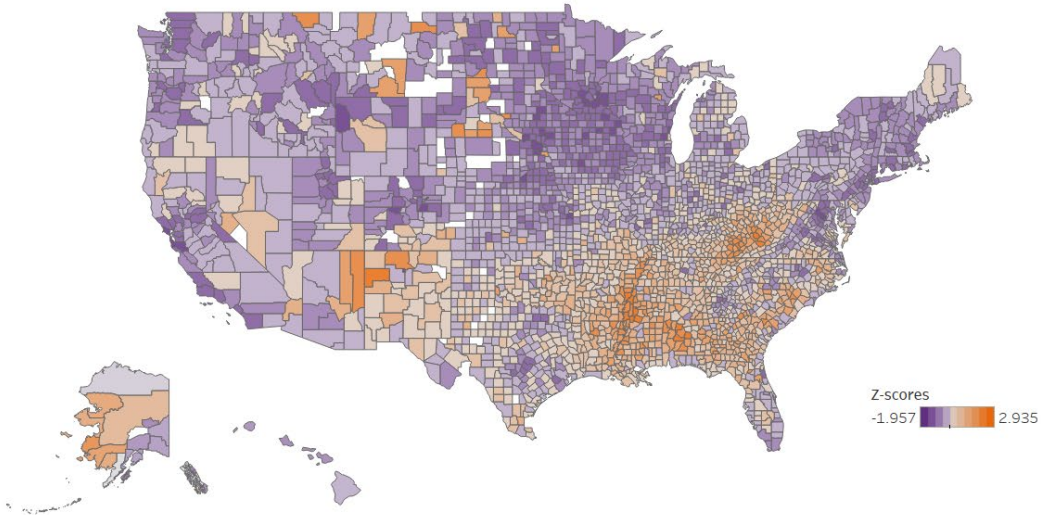
	Measure	Weight	Data Source	Years of Data
HEALTH OUTCOMES				
Length of Life	Premature Death*	50%	National Center for Health Statistics - Natality and Mortality Files; Census Population Estimates Program	2019-2021
Quality of Life	Poor or Fair Health*	10%	Behavioral Risk Factor Surveillance System	2021
	Poor Physical Health Days*	10%	Behavioral Risk Factor Surveillance System	2021
	Poor Mental Health Days*	10%	Behavioral Risk Factor Surveillance System	2021
	Low Birthweight*	20%	National Center for Health Statistics - Natality Files	2016-2022
HEALTH FACTORS				
HEALTH BEHAVIORS				
Tobacco Use	Adult Smoking*	10%	Behavioral Risk Factor Surveillance System	2021
Diet and Exercise	Adult Obesity*	5%	Behavioral Risk Factor Surveillance System	2021
	Food Environment Index	2%	USDA Food Environment Atlas; Map the Meal Gap from Feeding America	2019 & 2021
	Physical Inactivity*	2%	Behavioral Risk Factor Surveillance System	2021
	Access to Exercise Opportunities	1%	ArcGIS Business Analyst and ArcGIS Online; YMCA; US Census TIGER/Line Shapefiles	2023, 2022 & 2020
Alcohol and Drug Use	Excessive Drinking*	2.5%	Behavioral Risk Factor Surveillance System	2021
	Alcohol-Impaired Driving Deaths	2.5%	Fatality Analysis Reporting System	2017-2021
Sexual Activity	Sexually Transmitted Infections	2.5%	National Center for HIV/AIDS, Viral Hepatitis, STD, and TB Prevention	2021
	Teen Births*	2.5%	National Center for Health Statistics - Natality Files; Census Population Estimates Program	2016-2022
CLINICAL CARE				
Access to Care	Uninsured	5%	Small Area Health Insurance Estimates	2021
	Primary Care Physicians	3%	Area Health Resource File/American Medical Association	2021
	Dentists	1%	Area Health Resource File/National Provider Identifier Downloadable File	2022
	Mental Health Providers	1%	CMS, National Provider Identification	2023
Quality of Care	Preventable Hospital Stays*	5%	Mapping Medicare Disparities Tool	2021
	Mammography Screening*	2.5%	Mapping Medicare Disparities Tool	2021
	Flu Vaccinations*	2.5%	Mapping Medicare Disparities Tool	2021
SOCIAL & ECONOMIC FACTORS				
Education	High School Completion	5%	American Community Survey, 5-year estimates	2018-2022
	Some College	5%	American Community Survey, 5-year estimates	2018-2022
Employment	Unemployment	10%	Bureau of Labor Statistics	2022
Income	Children in Poverty*	7.5%	Small Area Income and Poverty Estimates; American Community Survey, 5-year estimates	2022 & 2018-2022
	Income Inequality	2.5%	American Community Survey, 5-year estimates	2018-2022
Family and Social Support	Children in Single-Parent Households	2.5%	American Community Survey, 5-year estimates	2018-2022
	Social Associations	2.5%	County Business Patterns	2021

Community Safety	Injury Deaths*	5%	National Center for Health Statistics - Mortality Files; Census Population Estimates Program	2017-2021
PHYSICAL ENVIRONMENT				
Air and Water Quality	Air Pollution - Particulate Matter	2.5%	Environmental Public Health Tracking Network	2019
	Drinking Water Violations*	2.5%	Safe Drinking Water Information System	2022
Housing and Transit	Severe Housing Problems	2%	Comprehensive Housing Affordability Strategy (CHAS) data	2016-2020
	Driving Alone to Work*	2%	American Community Survey, 5-year estimates	2018-2022
	Long Commute - Driving Alone	1%	American Community Survey, 5-year estimates	2018-2022

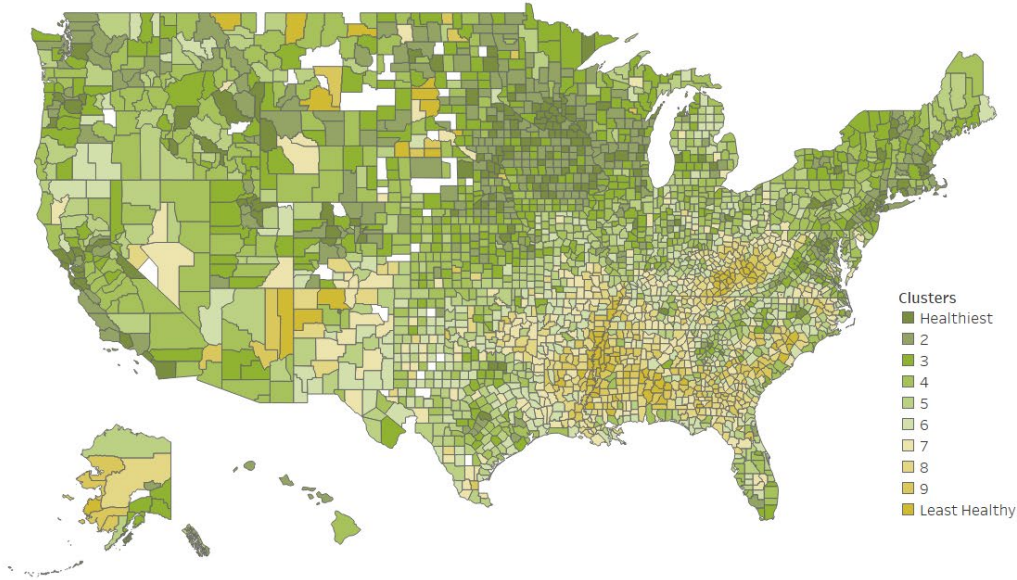
*Subgroup data available by race and ethnicity; †Data availability or recency varies by state

Appendix B. Geographic Distribution of the National Health Outcomes Z-score Values and the 10 National Health Outcomes Z-score Clusters*

Health Outcomes National Z-score



Health Outcomes National Clusters



*lower z-score value indicates better Health Outcomes, higher value indicates worse Health Outcomes

Appendix C. CHR&R Measures Affected by Changes in Data Source Infrastructure

Measure	Current Data Source	Disaggregated by racial groups	Calculated using single-year age groups
Premature Death (Years of Potential Life Lost)	NCHS – mortality files	✓	✓
Life Expectancy	NCHS – mortality files	✓	✓
Premature Age-Adjusted Mortality	NCHS – mortality files	✓	✓
Child Mortality	NCHS – mortality files	✓	✓
Infant Mortality	NCHS – mortality files	✓	✓
Drug Overdose Deaths	NCHS – mortality files	✓	
Injury Deaths	NCHS – mortality files	✓	
Homicides	NCHS – mortality files	✓	
Suicides	NCHS – mortality files	✓	✓
Firearm Fatalities	NCHS – mortality files	✓	
Motor Vehicle Crash Deaths	NCHS – mortality files	✓	
Low Birthweight	NCHS - natality files	✓	
Teen Births	NCHS - natality files	✓	

* Premature Death and Life Expectancy measure calculations account for population age structure without application of age-adjustment.

References

ⁱ Prainsack, B. & Buyx, A. (2012). Solidarity in Contemporary Bioethics: Toward a New Approach. *Bioethics*, 26(7), 343-350. doi:10.1111/j.1467-8519.2012.01987.x

ⁱⁱ Heller JC, Givens ML, Johnson SP, Kindig DA. Keeping It Political and Powerful: Defining the Structural Determinants of Health. *Milbank Q*. 2024 Feb 16. doi: 10.1111/1468-0009.12695.

ⁱⁱⁱ Pollock EA, Gagnon RE, Gennuso KP, Givens ML. Using Cluster Analysis to Inform and Communicate Ranking Uncertainty to Broad Audiences. *Population Health Metrics*. *Under Review July 2023*

^{iv} Ingram DD, Parker JD, Schenker N, Weed JA, Hamilton B, Arias E, Madans JH. United States Census 2000 population with bridged race categories. National Center for Health Statistics. *Vital Health Stat* 2(135). 2003.