CREATING A CULTURE OF HEALTH IN APPALACHIA

Disparities and Bright Spots

IDENTIFYING BRIGHT SPOTS IN APPALACHIAN HEALTH: STATISTICAL ANALYSIS

The second report in a series exploring health issues in Appalachia

Photo: Brian Stansberry

Prepared by:

APPALACHIAN REGIONAL COMMISSION
Washington, D.C.

PDA, INC.
Raleigh, N.C.

THE CECIL G. SHEPS CENTER FOR HEALTH SERVICES RESEARCH
The University of North Carolina at Chapel Hill, Chapel Hill, N.C.

July 2018

ACKNOWLEDGMENTS

Robert Wood Johnson Foundation

The Robert Wood Johnson Foundation (RWJF) is the nation's largest philanthropy dedicated to health. For more than 40 years, RWJF has supported research and programs targeting some of the nation's most pressing health issues.

RWJF provided funding for this research project as part of its Culture of Health Initiative. David M. Krol, MD, MPH, FAAP, Senior Program Officer, provided the impetus for the *Creating a Culture of Health in Appalachia: Disparities and Bright Spots* project and solicited the Appalachian Regional Commission and the Foundation for a Healthy Kentucky as partners. Hilary Heishman, MPH, Senior Program Officer, provided final guidance.

Appalachian Regional Commission

The Appalachian Regional Commission (ARC) provided funding, leadership, and project management for the project. ARC is an economic development agency of the federal government and 13 state governments focusing on 420 counties across the Appalachian Region. ARC's mission is to innovate, partner, and invest to build community capacity and strengthen economic growth in Appalachia to help the Region achieve socioeconomic parity with the nation.

Foundation for a Healthy Kentucky

The Foundation for a Healthy Kentucky served as the grantee and fiscal agent for the project. Funded by an endowment, the mission of the nonpartisan Foundation for a Healthy Kentucky is to address the unmet health needs of Kentuckians by developing and influencing policy, improving access to care, reducing health risks and disparities, and promoting health equity. Since the Foundation opened its doors in 2001, it has invested more than $27 million in health policy research, advocacy, and demonstration project grants across the Commonwealth.

Principal Investigators

To implement the research, the Appalachian Regional Commission and the Foundation for a Healthy Kentucky named two Principal Investigators: Julie L. Marshall, PhD, Senior Economist, Division of Planning and Research for the Appalachian Regional Commission, and Gabriela Alcalde, DrPH, Vice President, Policy & Program for the Foundation for Healthy Kentucky. Mary Jo Shircliffe, MBA, Vice President, Operations and Administration, assumed the role for the Foundation for a Healthy Kentucky to complete the project.

Authors

PDA, Inc. in Raleigh, North Carolina prepared this report in collaboration with The Cecil G. Sheps Center at the University of North Carolina at Chapel Hill for the Foundation for Healthy Kentucky and the Appalachian Regional Commission. The principal authors of the report include:

G. Mark Holmes, PhD, Center Director, Cecil G. Sheps Center for Health Services Research, University of North Carolina

Nancy M. Lane, MA, President, PDA, Inc.

William Holding, MHA, Senior Consultant Health Economist, PDA, Inc.

Randy Randolph, MRP, Applications Analyst Programmer, Cecil G. Sheps Center for Health Services Research, University of North Carolina

Jonathan Rodgers, MHA, Consultant and Research Associate, PDA, Inc.

Pam Silberman, JD, DrPH, Associate Director for Policy Analysis, Cecil G. Sheps Center for Health Services Research, University of North Carolina

Lisa Villamil, MAE, Assistant Professor, UNC Chapel Hill School of Media and Journalism

Thomas A. Arcury, PhD, Program Director, Wake Forest Baptist Medical Center Clinical and Translational Science Institute

Kelly Ivey, BS, Project Manager, PDA, Inc.

Daniel Goolsby, BS, Data Analyst, PDA, Inc.

Ashli Keyser, Project Assistant, PDA, Inc.

J & J Editorial, LLC, Editors

In addition, a number of research staff from the Sheps Center contributed research to the report, including Lindsay McCall, Julie Perry, George Pink, Steve Rutledge, Alex Schulte, Kristie Thompson, Rebecca Whitaker, and Dunc Williams.

Advisors and Contributors

The project benefited from insight from an Advisory Committee assembled by ARC, RWJF, and the Foundation for a Healthy Kentucky: Rui Li, PhD, Health Economist, Division of Reproductive Health, Centers for Disease Control and Prevention; Rozelia Harris, MBA, Director of Mississippi Office of Rural Health and Primary Care, Mississippi State Department of Health; Amy Martin, DrPh, MSPH, College of Dental Medicine, Medical University of South Carolina; Wayne Myers, MD; and Kadie Peters, Vice President of Community Impact for Health, United Way of Central Alabama.

Others who assisted with review and feedback include: Logan Thomas, ARC Research Analyst; Eric Stockton, ARC Health Program Manager; Keith Witt, ARC Geographic Information Specialist; Kostas Skordas, ARC Director, Planning and Research; Karen Entress, ARC Communications staff; Diane Smith, ARC Communications staff.

TABLE OF CONTENTS

GLOSSARY OF TERMS

Appalachian Region — The Appalachian Region is defined in the federal legislation from which the Appalachian Regional Commission derives its authority. The Region covers 205,000 square miles, and 420 counties in 13 states. It stretches more than 1,000 miles from Mississippi to New York, and is home to more than 25 million people.

Appalachian Regional Commission — The Appalachian Regional Commission (ARC) is an economic development agency of the federal government and 13 state governments focusing on 420 counties across the Appalachian Region. ARC's mission is to innovate, partner, and invest to build community capacity and strengthen economic growth in Appalachia to help the Region achieve socioeconomic parity with the nation.

ARC Economic Index — ARC uses an index-based classification system to compare each county in the nation with national averages on three economic indicators: three-year average unemployment rates, per capita market income, and poverty rates. Based on that comparison, each Appalachian county is classified within one of five economic status designations—distressed, at-risk, transitional, competitive, or attainment.

Bright Spot — For the purposes of this research, a "Bright Spot" is a county identified through statistical methods that demonstrates better-than-expected health outcomes, given its characteristics and resources. Counties are ranked by overall magnitude of "brightness." Bright Spot counties include only those ranking in the top decile in overall "brightness."

Centers for Disease Control and Prevention — Centers for Disease Control and Prevention (CDC) is the leading national public health protection agency in the United States. CDC administers various data collection programs vital for health researchers. These include CDC's WONDER database, which includes detailed mortality information, the Behavioral and Risk Factor Surveillance System (BRFSS), and other disease prevalence estimates.

County Health Rankings — The *County Health Rankings & Roadmaps* program is a collaboration between the Robert Wood Johnson Foundation and the University of Wisconsin Population Health Institute. The goals of the program are to build awareness of the multiple factors that influence health; provide a reliable, sustainable source of local data to communities to help them identify opportunities to improve their health; engage and activate local leaders from many sectors creating sustainable community change; and connect and empower community leaders working to improve health.

Culture of Health — This is a Robert Wood Johnson Foundation initiative aimed at strengthening the complex social factors that enable all persons to live the healthiest life possible.

Economic Distress	Distressed counties are the most economically depressed counties and rank in the worst 10 percent of the nation's counties. In fiscal year 2017, 84 Appalachian counties qualify for distressed county status on the basis of low per capita income and high rates of poverty and unemployment.
Euclidean Distance	Euclidean Distance is a measure of the degree of dissimilarity between two units, calculated as the square root of the summed squared distances. In a two dimensional framework, it is analogous to a hypotenuse on a right triangle. See Appendix C for further explanation.
Metro Counties	Counties that fall within Metropolitan Statistical Areas (MSA), as defined by the U.S. Census Bureau and the federal Office of Management and Budget (OMB) are considered "Metro" for the purposes of this report. The delineations used in this report are based on the most recent delineations as of the creation of this report, which are July 2015 OMB designations.
Nonmetro Counties	"Nonmetro" counties are counties that are not included in an MSA according to the 2015 OMB delineations.
Propensity Score	A propensity score is the probability that a unit with certain characteristics will be assigned to the treatment group (as opposed to the control group). The scores can be used to reduce or eliminate selection bias in observational studies by balancing covariates (the characteristics of participants) between treated and control groups. In this report, the propensity score matching measured the probability that a given county was "Appalachian-like," given its unique set of drivers.
Residual	In regression analysis, the residual is the difference between the observed (actual) value for the dependent variable and the value predicted by the regression model (expected value).
R-Square	In regression analysis, R-square measures the percentage of variation in a dependent variable that can be explained by variations in all of the independent variables. In this report, the dependent variables are health outcomes and the independent variables are health drivers.
Standardized	See Z-Score.
Subregion	ARC divides Appalachia into five subregions: Northern, North Central, Central, South Central, and Southern. These subregions may be referred to as Northern Appalachia, North Central Appalachia, etc. Counties within each subregion share similar characteristics, such as topography, demographics, and economics.
Years of Potential Life Lost (YPLL)	Years of potential life lost (YPLL) is a measure of premature mortality. It differs from other death rates as it puts more weight on deaths that occur at younger ages. The earlier the death, the greater years of potential life lost. The method for calculating YPLL in this report matches the methodology used in the *County Health Rankings* calculations. The measure is Years of Potential Life Lost per 100,000 population. Its calculation involves a threshold year—in this report, it is age 75.

Z-Score A z-score (aka, a standard score) indicates how many standard deviations an element is from the mean. A z-score can be calculated from the following formula: $z = (X - \mu) / \sigma$ where z is the z-score, X is the value of the element, μ is the population mean, and σ is the standard deviation. In a normal distribution, 68 percent of the sample lies within one standard deviation of the mean. A variable is "standardized" when it is converted to a z-score. For example, the average height of an adult female is 63.8 inches, and the standard deviation is approximately 2.8 inches, so a woman who is 6 feet tall (72 inches) would represent a z-score of (72 – 63.8) / 2.8 or 2.9, almost three standard deviations from the mean.

ABBREVIATIONS

ARC	Appalachian Regional Commission
BRFSS	Behavioral Risk Factor Surveillance System
CDC	The Centers for Disease Control and Prevention
CMMI	Center for Medicare and Medicaid Innovation
CMS	Centers for Medicare and Medicaid Services
COPD	Chronic Obstructive Pulmonary Disease
FFS	Fee for Service
HCC	Hierarchical Case Condition Score for Medicare
IHI	Institute for Healthcare Improvement
LBW	Low Birth Weight Births
O/E	Observed divided by Expected value
OMB	United State Office of Management and Budget
PCP	Primary Care Physician
RWJF	Robert Wood Johnson Foundation
Rx	Prescription
UIC	Urban Influence Code
YPLL	Years of Potential Life Lost before age 75

LIST OF FIGURES

Executive Summary

Introduction

Data and Methods

Results

LIST OF TABLES

Executive Summary

Data and Methods

Results

Field Study Site Selection

Executive Summary

Identifying Bright Spots in Appalachian Health: Statistical Analysis Executive Summary

This report, *Identifying Bright Spots in Appalachian Health: Statistical Analysis,* was produced through the "Creating a Culture of Health in Appalachia: Disparities and Bright Spots" research initiative funded by the Robert Wood Johnson Foundation (RWJF) and the Appalachian Regional Commission (ARC), and administered by the Foundation for a Healthy Kentucky. To date, this multi-part health research project has produced the following three reports:

1. *Health Disparities in Appalachia* (August 2017) measures population health in the Appalachian Region and documents disparities between Appalachia and the nation as a whole, as well as disparities within the Region.

2. *Identifying Bright Spots in Appalachian Health: Statistical Analysis* (July 2018) describes the results of the regression analysis used to assess how each of the Appalachian Region's 420 counties scored on 19 different health indicators, and then identifies counties with better-than-expected outcomes, given their characteristics and resource levels. Through this process, 42 Appalachian counties were classified as Bright Spot counties.

3. *Exploring Bright Spots in Appalachian Health: Case Studies* (July 2018) presents in-depth studies of 10 of the 42 Bright Spot counties identified through the statistical analysis. This report explores local perceptions of practices that may be associated with better-than-expected health outcomes, and summarizes promising strategies that may be replicable in other communities.

As described above, this report is the second in the series and is the quantitative companion to the third report, *Exploring Bright Spots in Appalachian Health: Case Studies.*

The reports offer a basis for understanding and addressing health in the Appalachian Region and for identifying factors that support a culture of health in Appalachian communities. They also explore activities, programs, and policies that may encourage better-than-expected health outcomes, many of which may be replicable in other communities.

The fourth and final report in the series, expected to be published in late 2018, will provide recommendations for practical strategies and activities that build on the findings of the first three reports.

ABOUT THE APPALACHIAN REGION

The current boundary of the Appalachian Region includes all of West Virginia and parts of 12 other states: Alabama, Georgia, Kentucky, Maryland, Mississippi, New York, North Carolina, Ohio, Pennsylvania, South Carolina, Tennessee, and Virginia (see Figure 1). The Region covers 205,000 square miles and 420 counties, and is home to more than 25 million Americans. Forty-two percent of the Region's population is rural, compared with 20 percent of the national population.

Figure 1: Map of the Appalachian Region

Source: Appalachian Regional Commission

BRIGHT SPOTS ANALYSIS

Overview

The primary objective of this report is to use regression analysis to identify Bright Spots, or counties in the Appalachian Region that have better-than-expected health outcomes given their characteristics and resource levels—that is, the socioeconomics, demographics, behaviors, health care facilities, and other factors that influence health outcomes. The second objective is to develop a systematic way to match the counties identified as Bright Spots with other counties in the Region to facilitate the exchange of ideas and lessons learned. The third objective is to identify the factors that appear to have the greatest impact on health outcomes—although this statistical analysis will provide *suggestions* rather than definitive *conclusions*.

The use of regression models and residual-based analyses is common in health research, though typically the focus of these research efforts is on how much the factors included in the model *explain* the outcome. In the Bright Spots analysis, however, the *unexplained* portions are of the greatest interest—that is, which counties are the most "unexpectedly" healthy given their characteristics and resources.

The model in this report bears some similarity to another research approach used to identify positive health outcomes that occur despite difficult circumstances. That approach, *positive deviance*, identifies the individuals, groups, and organizations affecting change at the local level, as opposed to macro-level policies at the state and national levels. The underlying principle of positive deviance is that by identifying individuals and groups that are overcoming challenges affecting a large number of people in a given community, researchers can identify specific, simple best practices that can be shared with other communities. Many of the best practices uncovered via the positive deviance approach originate from within the community and are implementable despite resource constraints.

Although the Bright Spots model does not fit wholly under the umbrella of positive deviance, this approach provided the motivation for our framework and the foundation for exploring counties through in-depth, field-based case studies.

Methods

The general approach in this analysis assumes we can broadly measure health in a community, compare actual outcomes to expected outcomes, and determine whether a community exceeds expectations.

We first identified 19 county-level *outcome* measures that capture the overall health of a community (see Table 1). Some examples of these measures include the infant mortality rate, cancer mortality rate, percentage of adults who are obese, prevalence of diabetes, and prevalence of depression among Medicare beneficiaries. The selected outcomes represent both physical and behavioral health, as well as diagnosed and perceived health.

We then identified 29 county-level *drivers* known to affect individual and community health (see Table 2). The drivers were organized into broad categories, such as social determinants, health behaviors, and access to health care services. Examples of drivers include median income, percentage of adults with some college education, percentage of the population under age 65 who are uninsured, number of primary care physicians per 100,000 population, and percentage of adults who smoke.

A multivariate regression analysis determined the relationship between the 19 health outcome measures and the 29 driver measures, producing one expected value for each of the 19 outcome measures *for each Appalachian county*. The expected outcomes were then compared to the actual, observed outcomes for each county to identify counties that performed better than expected. In most counties, some of the 19 outcomes were better than expected and some were worse than expected. Each outcome residual was then standardized into a z-score to allow comparison across all outcome measures. We reversed signs on the outcomes so that positive z-scores indicated "good health."

By using the average *degree* to which a county's observed health outcomes exceeded expected values, the Bright Spots model identified counties that either did very well on a few outcomes or exceeded expectations—perhaps only marginally—across many outcomes.

Because access to resources differs between metropolitan and nonmetropolitan areas, the statistical analysis was applied independently to two geographic groups: metropolitan and nonmetropolitan counties. Inclusion in a U.S. Census Metropolitan Statistical Area defined a county as metro. However, because the metro and nonmetro datasets are distinct, average residuals of the two groups are *not* comparable.

A county whose average of all 19 standardized outcome residuals scored in the top decile in either the metropolitan or nonmetropolitan group was classified as a Bright Spot.

The variation in county location and economic status lends support to the study design—we did not aim to identify healthy counties with high levels of resources and the sorts of characteristics that support positive health outcomes, but rather counties encompassing a wide range of resource levels and characteristics that all managed to find a way to be healthier than expected. Bright Spots are places that exceed expectations, *regardless of the values of the driver*s. This is a strength of the approach, one that allowed us to focus on the positive aspects of communities *relative to their own characteristics and resource levels.*

Table 1: Outcome Measures

Category	Measure
Mortality	Years of potential life lost per 100,000
	Stroke mortality per 100,000
	All cancer mortality per 100,000
	Unintentional injury mortality per 100,000
	COPD mortality per 100,000
	Heart disease mortality per 100,000
Mental Health	Average mentally unhealthy days per person per month
	Suicide mortality per 100,000
	Percentage Medicare beneficiaries with depression
Child Health	Percentage of live births with low birth weight (<2500g)
	Infant mortality per 1,000 births
Chronic Disease	Percentage adults with diabetes
	Medicare heart disease hospitalizations per 1,000
	Average Hierarchical Condition Category (HCC) risk score per Medicare beneficiary [a]
	Percentage adults with obesity (BMI>30)
	Average physically unhealthy days per person per month
Substance Abuse	Percentage residents drinking excessively
	Poisoning mortality per 100,000
	Opioid prescriptions as percentage of Part D claims

Notes: a. Unless noted, information on each measure is included in the Disparities report in this series (Marshall, et al., 2017).

Table 2: Driver Measures

Category	Measure
Child Health	Teenage births per 1,000
Environment	Full-service restaurants per 1,000 population [a]
	Percentage with access to exercise opportunities [a]
	Air pollution (average daily particulate matter, $PM_{2.5}$) [a]
	Grocery stores per 1,000 population
	Students per teacher (primary and secondary school)
	Average travel time to work in minutes
Health Behaviors	Percentage of adults currently smoking
	Percentage of adults not physically active
	Chlamydia incidence per 100,000
Health Care System and Utilization	Primary care physicians per 100,000 population
	Dentists per 100,000 population
	Specialty physicians per 100,000 population
	Mental health providers per 100,000 population
	Percentage of physicians that e-prescribe
	Percentage under 65 who are uninsured
Quality	Percentage of Medicare diabetics with HbA1c testing
	Percentage of Medicare women with recent mammogram
Social Determinants	Percentage of total population in paid Social Assistance jobs [a]
	Income inequality ratio [a]
	Percentage eligible enrolled in SNAP (Food Assistance) [a]
	Percentage of households with no car and low access to grocery stores [a]
	Percentage of households spending >30% of income on housing [a]
	ARC Economic Index
	Social association rate per 10,000 population
	Percentage receiving disability benefits (OASDI and/or SSI)
	Percentage of adults with some college education
	Percentage of households with income below poverty line
	Median household income

The Bright Spot Counties

Appalachian counties with an average standardized residual score in the top decile (10 percent) in either the metropolitan or the nonmetropolitan groups were identified as Bright Spots. Thus, the model identified 42 Appalachian Bright Spot counties: 15 in the metropolitan group and 27 in the nonmetropolitan group. The 42 counties in the top decile represent *the best of the better than expected*. In fact, scores for 202 counties were better than expected; the Bright Spots are simply those counties with scores in the top ten percent in their respective groups.

Bright Spot counties are located in all five Appalachian subregions and represent the diversity of communities across the Appalachian Region (see Figure 2).

Figure 2: Map of the Bright Spot Counties in Appalachia

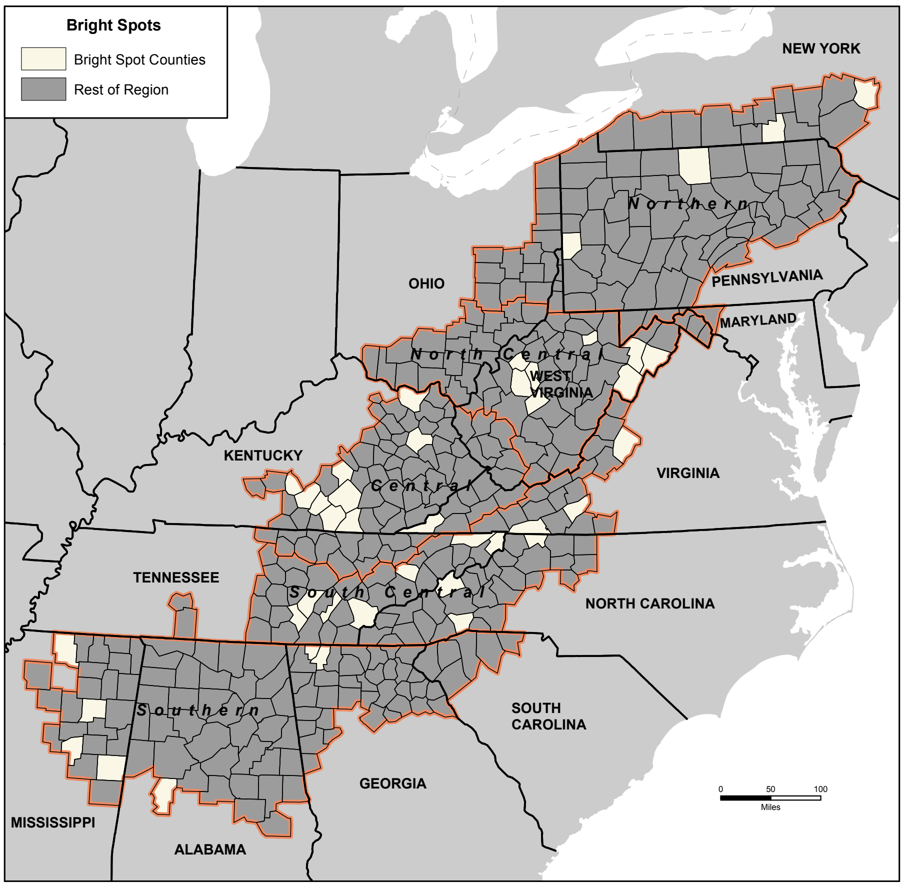


Table 3 lists the metropolitan Bright Spot counties, their average standardized residual score, and the outcome with the highest residual, which reflects the greatest *over performance* relative to available resources. Table 4 shows the same information for nonmetropolitan Bright Spot counties.

The higher the residual score, the more a county outperformed its expectations. The standardized residual scores represent standard deviations. For example, outcomes in a county with an average residual score of 0.47 were, on average, 0.47 standard deviations above the expected outcomes.

Table 3: Metropolitan Appalachian Bright Spot Counties, Ranked by Average Outcome Residual

Rank	County	State	Average Standardized Residual Score [a]	Highest Individual Residual [b]	
1	Wirt	West Virginia	0.47	Injury mortality	1.58
2	Clay	West Virginia	0.40	Heart disease mortality	1.51
3	Henderson	North Carolina	0.35	% obese adults	0.98
4	Hale	Alabama	0.35	Depression prevalence	1.10
5	Sequatchie	Tennessee	0.31	Poisoning mortality	1.22
6	Floyd	Virginia	0.30	COPD mortality	1.08
7	Sullivan	Tennessee	0.30	Poisoning mortality	1.23
8	Marshall	Mississippi	0.30	% opioid Rx claims	1.58
9	Madison	North Carolina	0.29	% obese adults	1.26
10	Whitfield	Georgia	0.29	Depression prevalence	0.97
11	Tioga	New York	0.27	Stroke mortality	0.87
12	Schoharie	New York	0.25	Average HCC risk score	0.83
13	Beaver	Pennsylvania	0.25	Average HCC risk score	1.00
14	Jefferson	Tennessee	0.24	Average HCC risk score	1.06
15	Catoosa	Georgia	0.24	Stroke mortality	0.90

Notes: a. Average residual score for the regression analysis involving 152 Appalachian metro counties
b. Highest of the 19 standardized residual outcome scores for each county.

Table 4: Nonmetropolitan Appalachian Bright Spot Counties, Ranked by Average Outcome Residual

Rank	County	State	Average Standardized Residual Score [a]	Highest Individual Residual [b]	
1	Wayne	Kentucky	0.72	Stroke mortality	1.79
2	Noxubee	Mississippi	0.58	COPD mortality	2.19
3	Calhoun	West Virginia	0.58	Injury mortality	2.02
4	Grant	West Virginia	0.49	Cancer mortality	1.88
5	McCreary	Kentucky	0.45	Poisoning mortality	1.94
6	Potter	Pennsylvania	0.45	Heart disease mortality	1.44
7	Taylor	West Virginia	0.42	Heart disease hospitalizations	1.20
8	Rockbridge	Virginia	0.41	Heart disease hospitalizations	1.37
9	Pulaski	Kentucky	0.40	Poisoning mortality	1.64
10	Green	Kentucky	0.40	YPLL	1.38
11	Lee	Virginia	0.40	Poisoning mortality	2.29
12	Russell	Kentucky	0.40	Heart disease hospitalizations	1.68
13	Bledsoe	Tennessee	0.39	Cancer mortality	1.88
14	Grayson	Virginia	0.39	Injury mortality	1.83
15	Hardy	West Virginia	0.38	% opioid Rx claims	1.21
16	Johnson	Tennessee	0.38	Poisoning mortality	1.52
17	Lincoln	Kentucky	0.37	% obese adults	1.37
18	Meigs	Tennessee	0.36	% opioid Rx claims	2.17
19	Pendleton	West Virginia	0.36	Poisoning mortality	1.48
20	Choctaw	Mississippi	0.35	Cancer mortality	1.69
21	Adair	Kentucky	0.35	Injury mortality	1.57
22	Lewis	Kentucky	0.34	Depression prevalence	1.78
23	Roane	West Virginia	0.33	Heart disease hospitalizations	1.35
24	Monroe	Tennessee	0.32	COPD mortality	1.18
25	Alleghany	North Carolina	0.31	YPLL	1.18
26	Chickasaw	Mississippi	0.31	Stroke mortality	1.61
27	Morgan	Kentucky	0.28	Injury mortality	0.92

Notes: a. Average residual score for the regression analysis involving 268 Appalachian nonmetro counties
b. Highest of the 19 standardized residual outcome scores for each county

KEY FINDINGS AND OBSERVATIONS

Bright Spot Patterns and Clusters

The Bright Spots are not distributed evenly among the Appalachian states—Kentucky and Mississippi have proportionately more Bright Spot counties than other states.

On the other hand, the model did not identify any Bright Spot counties in Ohio, a state with 32 Appalachian counties. The other two states with no identified Bright Spot counties, South Carolina and Maryland, have only a few Appalachian counties: six and three, respectively. The absence of Bright Spots in these two states may be the result of small sample sizes, whereas the Ohio result suggests a pattern of lower-than-expected outcomes.

Several Bright Spot counties appear in geographic clusters, suggesting that factors leading to better-than-expected health may prevail across broad, multicounty areas. Clustering suggests the presence of some common factor that has improved the health of the cluster. The unit of analysis, the county, may be a proxy for a larger "community." These communities may be in the service area of a particularly effective program, health care provider, or other resource. Alternatively, other factors, such as environment, local culture, and tradition, may also support a culture of health.

Correlation of Specific Outcomes with Overall Health

Our approach allows us to broadly measure health in a community and determine whether that community exceeds expectations. We developed the average standardized outcome residual for this purpose, as it captures the degree to which a county's outcomes exceeded expectations. However, it is important to keep in mind that the average standardized residual does not represent the entire composition of a county's health status. For individual outcomes, even among counties identified as Bright Spots, there were still lower-than-expected values. These results suggest that community health cannot be painted with one broad brushstroke; rather, it is more accurately represented as a multidimensional combination of many different aspects of health.

One key aspect was to model the actual value of outcomes, rather than incorporate scales or indices. With this approach, we were able to find certain *individual* outcome measures that were more highly correlated with *overall* good health outcomes. Three of the 19 health outcome measures were consistently better than expected in the Bright Spot counties:

- **Premature mortality;**
- **Unintentional injury mortality; and,**
- **Poisoning mortality.**

Premature mortality (YPLL) had the highest correlation with the average standardized residual, supporting its use as a comprehensive measure of community health. Further, outcomes such as injury mortality and poisoning mortality were highly correlated with average standardized outcome residuals in the top-performing counties. Bright Spot counties—those in the top decile of average outcome residuals—tended to have better-than-expected poisoning mortality rates.

Unintentional injury was the fourth-leading cause of death in the United States in 2014, and includes deaths due to car accidents, falls, and poisoning. Poisoning mortality includes deaths due to overdose.

Among the ten counties with the *lowest* average standardized residuals in both the metropolitan and nonmetropolitan groups—20 counties altogether—only one county performed better than expected on poisoning mortality; many others had much higher poisoning mortality rates than expected. This suggests that poisoning mortality—and by extension, substance abuse—may have an important link to overall health for all counties.

Seven High-Impact Drivers

The results of this analysis suggest that the following seven drivers predicted the most variation in the 19 health outcomes (the direction generally associated with better health is shown in parentheses):

- **Median income (higher);**
- **ARC Economic Index value (lower);**
- **Poverty rate (lower);**
- **Percentage of adults that smoke (lower);**
- **Percentage of adults that are physically inactive (lower);**
- **Percentage of the population receiving disability payments (lower); and,**
- **Teen birth rates (lower).**

These seven drivers were better predictors of health outcomes in the Bright Spot counties than drivers describing the supply of health resources, such as the supply of primary care physicians or the supply of specialty physicians.

These findings suggest that focusing on improvements in these seven drivers may lead to the greatest overall impact on health in a community.

Notably, a county's teen birth rate emerged as a key driver of community health across most of the outcomes. Teen pregnancy serves as a marker for economic opportunity in the community, captures "risky behavior" among teenagers, including unprotected sex, which is often associated with substance use (Salas-Wright, Vaughn, Ugalde, & Todic, 2015), and can have long-lasting effects on young parents. The teen birth rate serves as a marker for life course outcomes: daughters of teenage mothers are more likely to become teenage mothers themselves (Albert, 2002). Practical strategies aimed at limiting teen birth rates, such as comprehensive, medically accurate sex education courses and other public health interventions, may have long-lasting, positive effects.

Matching Bright Spot Counties to Other Counties

Because the purpose of the research was to identify Bright Spots, explore best practices—or aspects of local culture in those communities that may be associated with better-than-expected health outcomes—and ultimately share those features with other communities, we calculated a measure to determine the similarity between each Appalachian county and each Bright Spot county.

Based on demographics, resources, and community characteristics, we used Euclidean distance analysis to create a propensity score, reflecting the similarity between the Bright Spot counties and other Appalachian counties. Table 13 in Appendix C shows the closest match for each county in Appalachia to one of the Bright Spot counties selected for case studies. Table 14 in Appendix C then shows each Appalachian county's best match out of the 42 Bright Spots identified through the statistical analysis.

Opportunity to Live Healthy

Research shows that positive health behaviors consistently have large, statistically significant relationships to good health outcomes (National Institutes of Health, 2015). Results in this report support and amplify this finding. In this study, the drivers that described behaviors, such as the percentage of adults who smoke, the percentage of adults who are physically inactive, and the teen birth rate, were more highly correlated with good health outcomes than drivers quantifying the supply of health resources. Our findings suggest that traditional public health initiatives should accompany efforts to develop community health infrastructure. For example, funding for community health workers trained to communicate chronic disease prevention behaviors might reach deeper into community values and have a greater impact on population health than the supply of additional providers alone.

Overall, this study supports an emerging body of literature that attests to the association between positive population health outcomes and a community's social, economic, and environmental factors.

NEXT STEPS: CASE STUDIES

From the 42 Bright Spots, we selected ten counties for in-depth, field-based investigations. The ten case study counties represent the diversity of communities in the Appalachian Region. Table 5 shows the ten case study counties, which include two counties in each of the five Appalachian subregions, an even distribution between metro and nonmetro counties, and three of ARC's five economic status classifications.

Table 5: Characteristics of Selected Case Study Sites

County	State	Subregion	Metro / Nonmetro	Average Outcome Residual[a]	2014 Population	Economic Status[b]
Wirt	WV	North Central	Metro	0.47	5,810	At-Risk
Hale	AL	Southern	Metro	0.35	15,393	Distressed
Sequatchie	TN	South Central	Metro	0.31	14,431	Transitional
Tioga	NY	Northern	Metro	0.27	50,464	Transitional
Madison	NC	South Central	Metro	0.29	20,951	At-Risk
Wayne	KY	Central	Nonmetro	0.72	20,728	Distressed
Noxubee	MS	Southern	Nonmetro	0.58	11,240	Distressed
Grant	WV	North Central	Nonmetro	0.49	11,829	Transitional
McCreary	KY	Central	Nonmetro	0.45	18,073	Distressed
Potter	PA	Northern	Nonmetro	0.45	17,451	Transitional

Sources: see Table 1 in Appendix B and Tables 3 and 5 in Appendix C

a. *Average outcome residuals are not comparable between metro and nonmetro groups*
b. *ARC economic designation, fiscal year 2017*

The findings from the case studies are discussed in the third report in this series, *Exploring Bright Spots in Appalachian Health: Case Studies*.

Introduction

ABOUT THE PROJECT

Culture of Health

Creating a Culture of Health in Appalachia: Disparities and Bright Spots is an innovative research initiative sponsored by the Robert Wood Johnson Foundation (RWJF) and the Appalachian Regional Commission (ARC) and administered by the Foundation for a Healthy Kentucky. This multi-part health research project has produced the three reports: *Health Disparities in Appalachia* measures population health and documents health disparities in the Appalachian Region; *Identifying Bright Spots in Appalachian Health: Statistical Analysis* establishes a framework for identifying Appalachian "Bright Spots," or communities with better-than-expected health outcomes, given their characteristics and resources; and *Exploring Bright Spots in Appalachian Health: Case Studies* which presents in-depth case studies and explores local perceptions of practices that may be associated with better-than-expected health outcomes, and summarizes promising strategies that may be replicable in other communities.

The Robert Wood Johnson Foundation's vision for a national Culture of Health—enabling all in our diverse society to lead healthier lives—is based on ten underlying principles:

1. Good health flourishes across geographic, demographic, and social sectors.
2. Attaining the best health possible is valued by our entire society.
3. Individuals and families have the means and the opportunity to make choices.
4. Businesses, government, individuals, and organizations work together to build healthy communities.
5. No one is excluded.
6. Everyone has access to affordable, quality health care.
7. Health care is efficient and equitable.
8. The economy is less burdened by excessive and unwarranted health care spending.
9. Keeping everyone as healthy as possible guides public and private decision making.
10. Americans understand that we are all in this together.

According to the Robert Wood Johnson Foundation, building a Culture of Health means creating a society that gives every person an equal opportunity to live the healthiest life they can—whatever their ethnic, geographic, racial, socioeconomic, or physical circumstances happen to be. A Culture of Health recognizes that health and well-being are greatly influenced by where we live, how we work, the safety of our surroundings, and the strength and connectivity of our families and communities—and not just by what happens in the doctor's office. The Culture of Health is operationalized through an action framework that is organized around the following four Action Areas:

1. Making health a shared value
2. Fostering cross-sector collaboration
3. Creating healthier, more equitable communities
4. Strengthening integration of health services and systems

Activity across these four action areas will, with time, lead to outcomes of improved population health, well-being, and equity.

The ten principles and the Culture of Health Action Framework served as the foundation for the first report in the series, *Health Disparities in Appalachia,* which provided a comprehensive picture of health in the Appalachian Region, focusing on how the Region compares to the United States as a whole and how parts of the Region compare to one another. Building on the findings of the first report, this second report develops a model for identifying Bright Spot counties, or areas where health outcomes are better than expected given the communities' sociodemographic and behavioral profiles as well as their levels of access to health care.

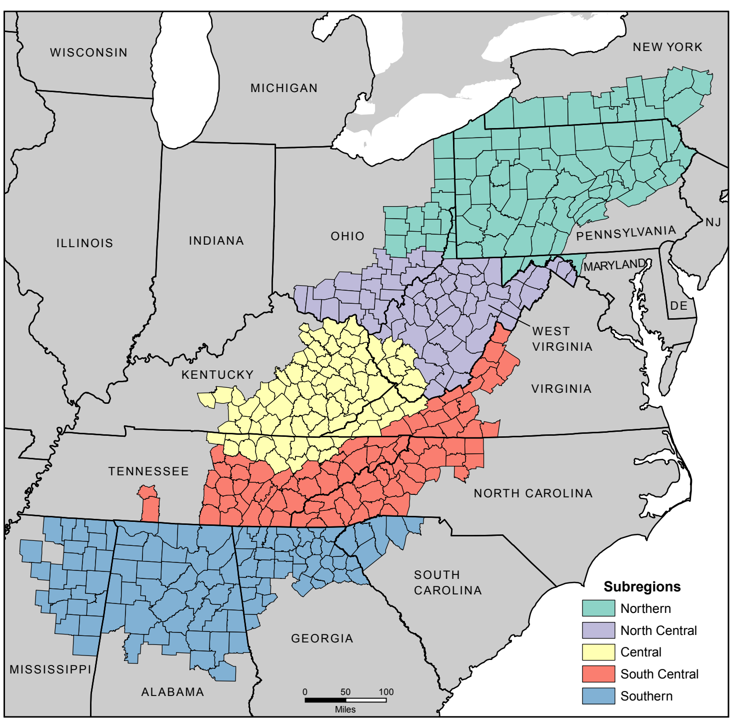
THE APPALACHIAN REGION

Geographic Subregions

The current boundary of the Appalachian Region includes all of West Virginia and parts of 12 other states: Alabama, Georgia, Kentucky, Maryland, Mississippi, New York, North Carolina, Ohio, Pennsylvania, South Carolina, Tennessee, and Virginia. The Region is home to more than 25 million people and covers 420 counties and almost 205,000 square miles.

The Appalachian subregions are contiguous regions with relatively similar characteristics (topography, demographics, and economics) within Appalachia (see Figure 3). Originally consisting of three subregions, ARC revised the classification system in 2009 and now divides the Region into five subregions. These smaller areas, the boundaries of which are based on recent economic and transportation data, allow for greater analytical detail.

Figure 3: Appalachian Subregions



Data source: Appalachian Regional Commission, November 2009

Rurality in Appalachia

We separate counties into metropolitan and nonmetropolitan groups by using the 2015 U.S. Office of Management and Budget (OMB) definition of a Metropolitan Statistical Area (MSA). This separation recognizes that metropolitan and nonmetropolitan counties can be quite different in terms of resources and overall population size, and that these differences can affect the degree to which health drivers affect health outcomes. The OMB metropolitan delineation is broad; some metropolitan counties (e.g., "bedroom counties") classify as such because of their high levels of commuting to core urban areas. Otherwise, they may resemble nonmetropolitan areas in both population size and density. However, to the extent that metropolitan status captures integration with a metropolitan center, the chosen delineation is appropriate for this model.

County Economic Status in Appalachia

ARC classifies counties based on economic status. The following information is based on ARC's report, "County Economic Status in Appalachia, FY 2017." Figure 4 shows Appalachian counties by economic status for fiscal year 2017.

The Appalachian Regional Commission uses an index-based county economic classification system to identify and monitor the economic status of Appalachian counties. The system involves the creation of a national index of county economic status through a comparison of each county's averages for three economic indicators—three-year average unemployment rate, per capita market income, and poverty rate—with national averages. The resulting values are summed and averaged to create a composite index value for each county. Each county in the nation receives a rank based on its composite index value, with higher values indicating higher levels of distress.

Each Appalachian county is classified into one of five economic status designations, based on its position in the national ranking.

<u>Distressed</u>
Distressed counties are the most economically depressed. They rank in the worst 10 percent of the nation's counties.

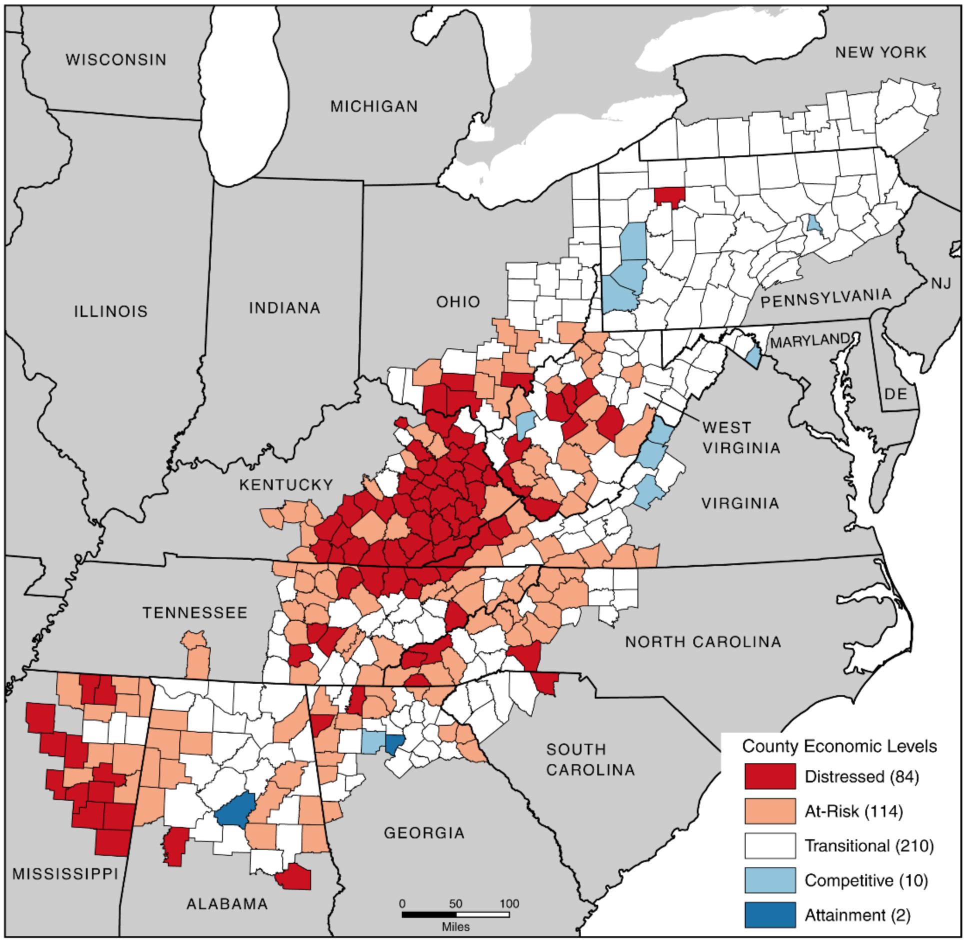
<u>At-Risk</u>
At-Risk counties are those at risk of becoming economically distressed. They rank between the worst 10 percent and 25 percent of the nation's counties.

<u>Transitional</u>
Transitional counties are those transitioning between strong and weak economies. They make up the largest economic status designation. Transitional counties rank between the worst 25 percent and the best 25 percent of the nation's counties.

<u>Competitive</u>
Competitive counties are those that are able to compete in the national economy but are not in the highest 10 percent of the nation's counties. Counties ranking between the best 10 percent and 25 percent of the nation's counties are classified as competitive.

<u>Attainment</u>
Attainment counties are the economically strongest. Counties ranking in the best 10 percent of the nation's counties are classified as attainment.

Figure 4: County Economic Status in Appalachia, FY 2017



Created by the Appalachian Regional Commission, March 2016
Data Sources:
Unemployment data: U.S. Bureau of Labor Statistics, LAUS, 2012–2014
Income data: U.S. Bureau of Economic Analysis, REIS, 2014
Poverty data: U.S. Census Bureau, American Community Survey, 2010–2014

Effective October 1, 2016
through September 30, 2017

REPORT OBJECTIVES AND METHODS

The primary objective of the Bright Spots analysis is to identify, through regression analysis, specific counties among the 420 Appalachian counties that have better-than-expected health outcomes given their characteristics and resource levels—that is, the sociodemographics, behaviors, health care facilities, and other characteristics of the communities that influence health outcomes (see Study Design and Methods for the detailed list). From this set of counties, the research team then selected ten locations for field-based case studies.

The second objective is to develop a systematic way to pair Bright Spot counties with other counties in the Region to facilitate the exchange of ideas and lessons learned.

The third objective is to identify the factors that appear to have the greatest impact on health outcomes. However, this statistical analysis provided *suggestions* rather than definitive *conclusions*.

Residual-Based Statistical Models and Positive Deviance Research

The use of regression models and residual-based analyses is common among health researchers, though typically the focus of these research efforts is on how much the factors included in the model *explain* the outcome. In the Bright Spots analysis, however, the *unexplained* portions are of the greatest interest—that is, which counties are the most "unexpectedly" healthy given their resources. Here, we review recent studies that used residual techniques to identify overperformers in health-related outcomes across the United States.

In 2011, the Institute for Healthcare Improvement (IHI) published a report entitled "Counties of Interest: Achieving Better- or Worse-Than-Expected Health Outcomes" (IHI, 2011). The IHI report evaluated County Health Rankings data to find counties with high residuals in a two-variable linear regression model. Though the model used in the Bright Spots report uses many more variables to create residuals, the goal is similar to that of the IHI report, which identified 17 counties with better-than-expected health outcomes and 11 counties with worse-than-expected health outcomes. The IHI study also included interviews and site visits with several of the counties. As such, the IHI paper served as a model for our analysis and allowed us to compare results and insights.

The observed over expected (O/E) ratio for a health outcome is a common metric used in health outcomes research and is similar to the residual-based approach used in this study. Best and Cowper used O/E ratios to determine the impact of expert care in Veterans Administration hospitals (Best & Cowper, 1994). Another example of research using the O/E approach is a study by Feudtner et al. evaluating the differences in mortality rates in children's hospitals across the United States (Feudtner, et al., 2011).

The model in this report bears some similarity to another research approach aimed at identifying positive health outcomes that occur despite difficult circumstances. That approach, *positive deviance*, identifies the individuals, groups, and organizations affecting change at the local level, as opposed to macro-level policies at the state and national levels. Although the Bright Spots model does not fit wholly under the umbrella of positive deviance, this approach provided the motivation for our framework and the foundation for exploring counties through in-depth, field-based case studies.

Perhaps the seminal publication on the topic is *The Power of Positive Deviance*, which details findings from years of community-level health research and observation across the globe (Pascale, Sternin, & Sternin, 2010). The underlying principle is that by identifying individuals and groups that are overcoming challenges affecting a large number of people in a given community, researchers can identify specific, simple best practices that can be shared with the rest of the community. Many of the best practices uncovered via the positive deviance approach originate from within the community and are implementable despite resource constraints.

Other researchers have combined the positive deviance approach with statistical analysis in order to identify positive health outcomes in communities or individuals with limited resources. One report with a similar framework to the Bright Spots analysis used the positive deviance framework, along with residual analysis to identify local health departments in three states that have better-than-expected outcomes in maternal and child health (Klaiman, Pantazis, Chainani, & Bekemeier, 2016), while a study conducted in France identified individuals who were more effective than others at consuming healthy food with a limited budget (Marty, et al., 2015).

In this report, we applied a positive deviance approach to the Appalachian Region in order to identify counties where health outcomes are exceeding expectations and to form a foundation to explore some of these counties through in-depth, field-based research.

CONCEPTUAL FRAMEWORK

The purpose of the Bright Spots statistical analysis was to identify three elements that are useful in identifying practices and strategies that improve health in Appalachia. These elements are:

1. "Bright Spots," or Appalachian counties with better-than-expected health outcomes given their characteristics and resources (e.g., socioeconomics, health system infrastructure, and behaviors);
2. A way to match Bright Spot counties with similar Appalachian counties in order to to facilitate the exchange of ideas and lessons learned; and
3. Individual health drivers and outcomes that are consistently associated with Appalachian Bright Spot counties.

The statistical model comprises two components: *drivers,* or determinants of health, and *outcomes,* or the measured results in individual and community health status. We interpret characteristics and resources broadly, and include drivers that reflect the health care system, the environment, sociodemographics, and the economy.

We start with county-level metrics for drivers and outcomes, most of which are described in the first report in this series, *Health Disparities in Appalachia* (Marshall, et al., 2017). We used multivariate regression analysis to predict expected values for 19 outcomes for each of the 420 Appalachian counties. Next, we measured the difference between the expected and the actual outcome variables for each county and described it in terms of a statistical residual. Counties whose average residual for all outcomes ranked in the top decile in either the metropolitan or nonmetropolitan group were identified as Bright Spots.

The third report in this series, *Exploring Bright Spots in Appalachian Health: Case Studies*, explores ten of the Bright Spot counties identified through the statistical analysis for possible factors, identifiable only by visiting the community, which may contribute to the better-than-expected outcomes.

To assist in facilitating the exchange of ideas and lessons learned between counties, we provide a metric for matching Bright Spot counties with each Appalachian county based on their relative similarities. Applying Euclidean distance method, we determine the similarity between each driver measure (e.g., median income, the percentage of adults with some college, number of primary care physicians) for each Bright Spot county and the same drivers for all other Appalachian counties. The resulting Euclidean distance calculation illustrates the similarity between each Appalachian county and each Bright Spot county based the drivers in the model. The best matches are listed in Table 13 and Table 14 in Appendix C.

We also examined the relationship between overall performance on the health outcome measures and the individual drivers to determine which county-level characteristics were most consistently associated with better-than-expected health outcomes.

The next chapter outlines the details of our approach.

Data and Methods

Objective 1: Identify Bright Spot Counties

Objective 2: Match Bright Spot Counties to Other Counties

Objective 3: Identify Health Drivers Contributing to Positive Outcomes

The previous chapter provided an overview of the study design. This chapter provides details on the study's approach.

OBJECTIVE 1: IDENTIFY BRIGHT SPOT COUNTIES

Residual-Based Analysis

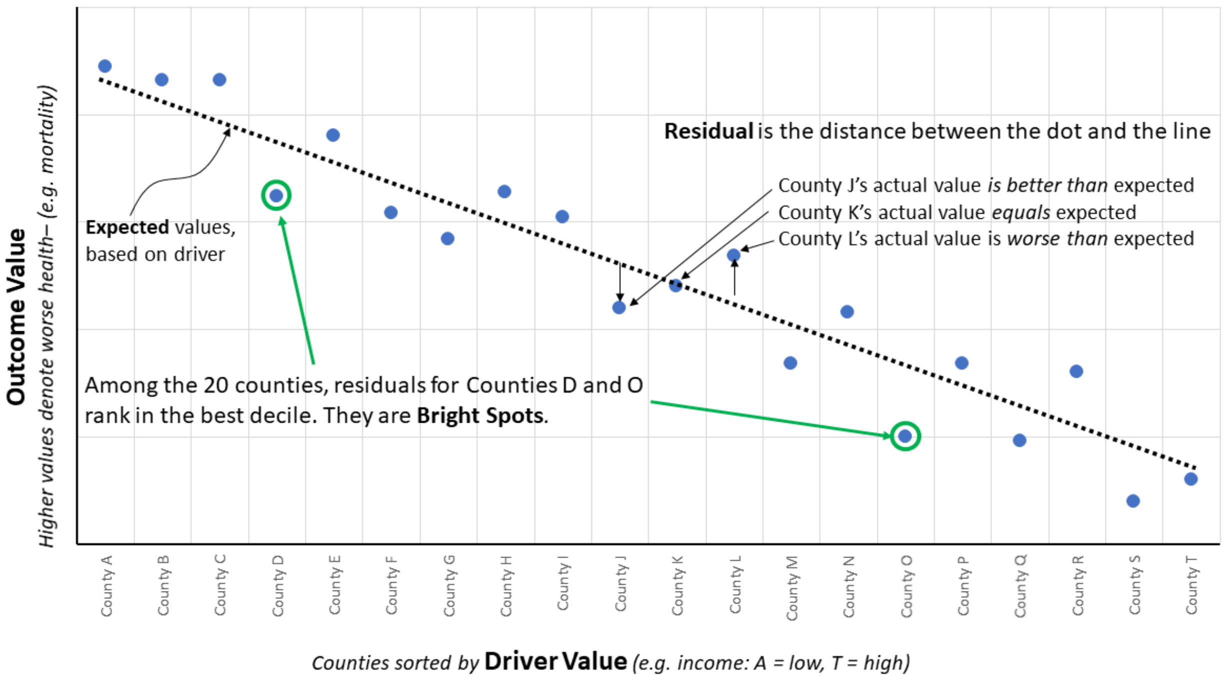
The Bright Spots analysis is similar in approach to the residual models used in previous work ((Topmiller, 2016; IHI, 2011); see discussion in Introduction), with a few major departures. First, the Bright Spots model involved a larger, multifaceted array of 29 health drivers. This larger selection of variables permitted a better representation of the variations in the health driver profiles of each county. Second, the Bright Spots model used 19 outcome measures, each with its own regression equation. This enabled the model to identify counties that broadly exceeded predictions and avoided the narrower view of health presented in other studies that used single outcome measures (e.g., mortality alone). By looking at multiple health outcomes, the model provided meaningful, nuanced conclusions about overall population health. Some communities, for example, may perform well on mental health measures, while others may perform well on measures related to child health. Our goal was to identify counties that performed well across a wide variety of measures.

Quantitative Model

Figure 5 illustrates the theoretical model for the statistical analysis. The figure sorts 20 hypothetical counties (labeled A–T) in ascending order of a single health *driver* (e.g., income). For illustration purposes, the driver values on the horizontal, or x-axis, differ by equal increments. County A has the lowest health driver value and County T has the highest. Dots in the figure correspond to the vertical y-axis and denote actual health *outcome* values for each x-axis county (for simplicity purposes, we assume that the value of the driver is related linearly to the county; that is, the difference in driver values between consecutive counties is identical).

The dotted line represents the *expected* values of the health outcome as the *actual* values of the health driver increases. In the figure below, the relationship is negative; the outcome values decrease as the value of the driver increases (moving from left to right on the horizontal axis). For example, the relationship between infant mortality and income is *negative*—that is, infant mortality rates decrease as population income increases. The relationship between infant mortality and the adult smoking rate is an example of a *positive* relationship—the value of the outcome increases as the value of the driver increases.

Figure 5: Conceptual Model for Bright Spots Statistical Analysis



The health outcomes in this report have differing scales; for example, years of potential life lost (YPLL) is measured in the thousands, while the number of mentally unhealthy days per person per month has an average in the single digits. To make the outcome values (and their associated residuals) comparable, we first standardized the outcomes into "z-scores"—a standard method measuring the number of standard deviations the value differs from the average value. For example, a z-score of 0 means the county's value is the average while a z-score of -1 refers to a value that is one standard deviation below the average. The residual represents the difference between the *actual* and the *expected* value. In Figure 5, we show three types of residuals: better than, equal to, and worse than expected.

In a final step, the model rank ordered the average of the standardized residuals for all 19 outcomes and classified the top decile of the rank-ordered counties as Bright Spots. In Figure 5, 2 of the 20 counties (10 percent) represent the top decile for a single outcome measure. The top decile was the threshold applied to the sum of the standardized residuals; those with an average residual in the top 10 percent of the 420 Appalachian counties (27 nonmetropolitan counties and 15 metropolitan counties) were classified as Bright Spot counties.

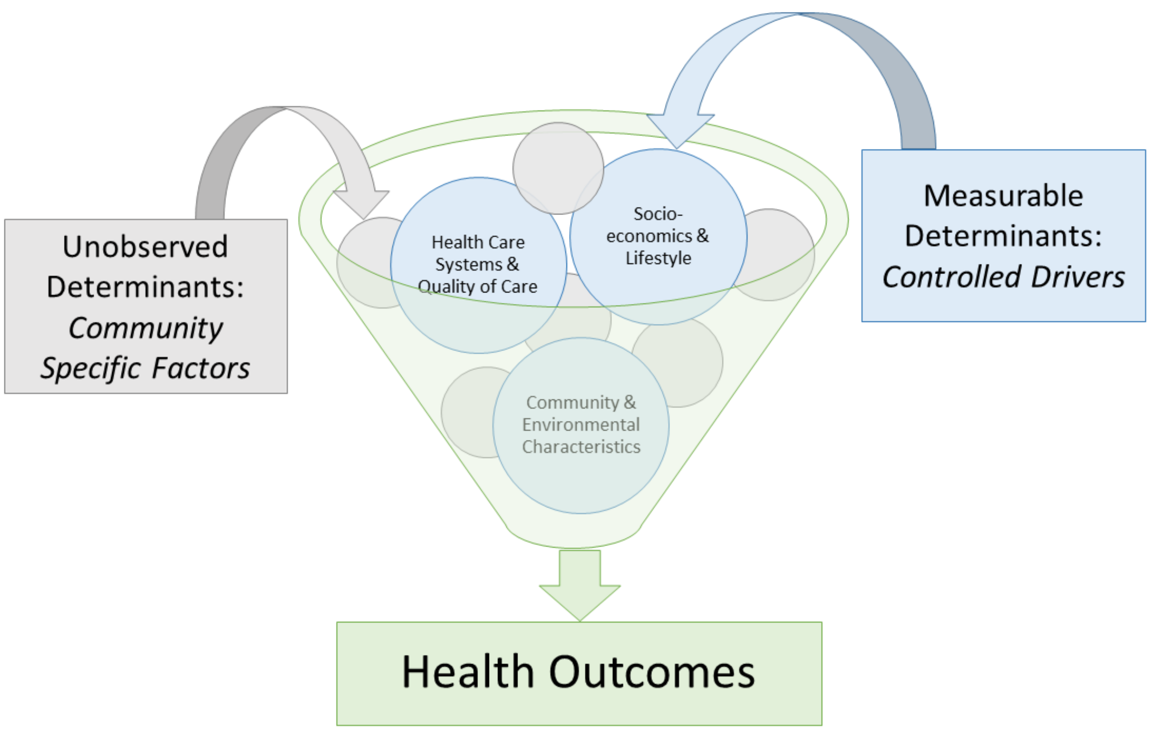
To improve the strength of the analysis, we expanded the sample size from the 420 Appalachian counties to all counties in the United States. The method of expansion took into account the uniqueness of the Appalachian Region and weighted non-Appalachian counties based on their similarity to counties in Appalachia, so that counties outside of the Region with a similar profile to Appalachia more heavily influenced the regression model than counties dissimilar to those in the Region. Appalachian counties had the highest weight.

Establishing a Framework for Selecting Measures

The Bright Spots analysis involves a social-ecological model that describes health as the output of individual, societal, and cultural factors (McLeroy, Bibeau, Steckler, & Glanz, 1988). Building this as a statistical model required outcome and driver metrics. Although certain health drivers are well defined, measurable, and lend themselves to statistical analysis, many other factors that affect health outcomes do not have well-established, standardized metrics; the assessment of these measures requires alternative means.

Figure 6 illustrates the framework. The large circles represent "Measurable Determinants," those health driver measures used in the regression model (see Figure 7 below for all 29 measures). The smaller circles represent "Unobserved Determinants," those community-specific factors that may influence health outcomes but operate outside of the statistical model. Examples include unique institutions, individual community members and leaders, nongovernmental safety net organizations such as federally qualified health centers, traditions, and any other components of local culture. Although these community-specific factors could explain why any given county outperforms its expected health outcomes, the regression model cannot take them into account because of the difficulty or cost of measuring these factors.

Figure 6: Social-Ecological Framework for the Bright Spots Model of Health



The drivers in the Bright Spots statistical model include measures known to influence health at both a community level (e.g., income, education, housing expenditures) and at an individual level (e.g., smoking, excessive drinking, risky sexual behavior). The design of the regression model allowed for *identification* of the high achievers. However, the model could not determine the specific, local factors not included that may have led to the "unexpectedly high" health values. To determine how much any given high-achiever county's community-specific factors contribute to its actual outcomes requires another level of research. The companion to this report, *Exploring Bright Spots in Appalachian Health: Case Studies* (Lane, et al., 2017), investigated community-specific factors that may have contributed to the better-than-expected outcomes in ten of the Bright Spot communities.

Relationship to the Culture of Health Action Areas

In selecting the health drivers, we considered the RWJF Culture of Health's four Action Areas. Although some of the Action Areas straightforwardly connect to national, county-level data, most indicators associated with the Action Areas were not uniformly available as county-level measures. Consequently, although the Bright Spots regression model included some measures that related directly to the Action Areas, in the aggregate, the model also included a number of variables related to the *concepts* the Action Areas represent.

Research Hypotheses

The statistical analysis was built on the following five hypotheses:

1. Place matters. Just as particular socioeconomic factors (e.g., low median income) can negatively influence a county's health status, other positive factors—many of which may be difficult to systematically quantify (e.g., "leadership")—can be strong enough to allow some counties in Appalachia to achieve better-than-expected health outcomes despite poor performance in health drivers.
2. A county could be a Bright Spot regardless of its absolute health or socioeconomic status.
3. A county that outperforms expectations across a broad array of health outcome measures may have qualities worthy of study and emulation by similar counties and, thus, should be classified as a Bright Spot.
4. In Bright Spot counties, one might expect to find elements of the RWJF Culture of Health Action Areas.
5. Metropolitan counties have access to more resources and should outperform nonmetropolitan counties, even after controlling for social determinants. As such, the analysis should separate metropolitan and nonmetropolitan counties.

Outcome Measure Selection Criteria

A central task of building the model was selecting a group of outcome measures that collectively represented the overall population health of a county. We recognize that in a group of communities evaluated across multiple health measures, some may perform exceptionally on a few measures but poorly on others. Although it is difficult to accurately characterize any community's health with only a few measures, it is more likely that one performing well on multiple measures is healthier than one performing well on just a few. Thus, the model was built on the simple underlying hypothesis that a wide range of variables would likely capture a county's overall health. The guiding principle for the selection of individual outcome measures was that, taken together, the outcome measures would identify counties where systemic characteristics, culture, and programs work together to improve overall health. Under this construct, assessing health across a wide range of outcomes increases the likelihood that the county is high performing; evaluating performance on only one or two outcomes may yield false identification of high performers if the outcomes were "randomly" low one year.

To qualify for inclusion in the model, outcome measures had to meet the following four criteria:

1. Available to the public (including those for which permission must be obtained);
2. Calculated at the county level and available for all counties in the United States;[1]
3. Relevant to the overall concept of population health; and,
4. Fit one of the model's five selected dimensions of health: mortality, mental health, child health, chronic disease, or substance abuse.

Although we initially sought to include oral health measures, national county-level data for dental health largely do not exist. We added substance abuse as its own category because of the national epidemic of

[1]Some intracounty smoothing is required in counties with small sample sizes for certain measures.

drug abuse occurring throughout much of the United States, with a particularly high concentration in Appalachia (Reynolds, 2016).

During the initial writing and data collection stages, we chose the most recent time period available for each measure. In many cases, we incorporated data from several years to increase the validity of the measurement and reduce the possibility of data suppression because of a low sample size.

From an initial pool of approximately 50 potential measures, we selected 19 that seemed to best meet the selection criteria. Table 6 lists the Bright Spots model's outcome measures grouped by category for ease of understanding. The domains described in the *Health Disparities in Appalachia* report provided inspiration for the categories used in this report, but the two are not identical. The slight difference stems from the distinction in the Bright Spots model between drivers and outcomes, a distinction not emphasized in the *Disparities* report.

Table 6: Outcome Measures by Category

Category	Measure
Mortality	Years of potential life lost per 100,000 population
	Stroke mortality per 100,000 population
	All cancer mortality per 100,000 population
	Unintentional injury mortality per 100,000 population
	Chronic obstructive pulmonary disease (COPD) mortality per 100,000 population
	Heart disease mortality per 100,000 population
Mental health	Average mentally unhealthy days per person per month
	Suicide mortality per 100,000 population
	Percentage of Medicare free-for-service (FFS) beneficiaries with depression
Child health	Percentage of live births with low birth weight (<2,500 g)
	Infant mortality per 1,000 births
Chronic disease	Percentage of adults with diabetes
	Medicare heart disease hospitalizations per 1,000 Medicare FFS beneficiaries
	Average Medicare beneficiary health care conditions risk score
	Percentage of adults with obesity (BMI >30)
	Average physically unhealthy days per person per month
Substance abuse	Percentage of residents drinking excessively
	Poisoning mortality per 100,000 population
	Opioid prescriptions as percentage of Part D Medicare claims

Appendix B provides data sources and year(s) for the selected driver and outcome measures. The *Disparities* report also includes a discussion related to most of the measures listed above. The Bright Spots analysis includes one additional outcome measure not found in the *Disparities* report: the average Medicare beneficiary hierarchical condition category (HCC) risk score, which calculates the increased representation of chronic disease outcomes (Center for Medicare and Medicaid Services, 2016).

Driver Selection Criteria

The regression model requires a sufficient array of health drivers, also referred to as determinants, to adequately distinguish communities from one another based on county-level characteristics. These drivers are the variables used to calculate a county's "expected" health outcomes. Each of the drivers selected has a documented relationship with at least one of the 19 outcomes included in the model. Not every driver has a strong statistical relationship with every outcome, but many drivers do relate to multiple outcomes.

Like the outcome measures, all drivers had to be available to the public, calculated at the county level, and available for all counties in the United States. The drivers include eight measures not found in the *Disparities* report, increasing the representation of environmental and social determinants. The eight additional measures include the following:

1. Full service restaurants per 1,000 population (Papas, Alberg, Ewing, & Helzlsouer, 2007);
2. Percentage of individuals in a county who live reasonably close to a location of physical activity ((Catlin, 2015); (County Health Rankings, 2016));
3. Air pollution or average daily particulate matter per M (Gan, FitzGerald, Carlsten, Sadatsafavi, & Brauer, 2013);
4. Percentage of the total population employed in social assistance jobs (Bradley & Taylor, 2016);
5. Income inequality ratio, comparing household income at the 80th percentile to household income at the 20th percentile;
6. Percentage of eligible recipients enrolled in the Supplemental Nutrition Assistance Program (SNAP) food assistance (Kreider, Pepper, & Gunderson, 2012);
7. Percentage of housing units with no car and low access to grocery stores (Treuhaft & Karpyn, 2010; Bell, Mora, Hagan, Rubin, & Karpyn, 2013); and,
8. Percentage of households spending 30 percent or more of their income on housing (Maqbool, Viverios, & Ault, 2015).

The 29 health drivers are shown in Table 7.

Table 7: Driver Measures by Category

Category	Measure	Culture of Health Action Area where applicable
Child health	Teenage births per 1,000 females ages 15–19	Making Health a Shared Value
Environment	Full-service restaurants per 1,000 population	
	Percentage with access to exercise opportunities	Fostering Cross-sector Collaboration to Improve Well-Being
	Air pollution (average daily particulate matter, $PM_{2.5}$)	Creating Healthier, More Equitable Communities
	Grocery stores per 1,000 population	Creating Healthier, More Equitable Communities
	Students per teacher (primary and secondary school)	
	Average travel time to work in minutes	
Health behaviors	Percentage of adults currently smoking	
	Percentage of adults not physically active	Making Health a Shared Value
	Chlamydia incidence per 100,000	Making Health a Shared Value
Health care system and utilization	Primary care physicians per 100,000 population	
	Dentists per 100,000 population	
	Specialty physicians per 100,000 population	
	Mental health providers per 100,000 population	
	Percentage of physicians that e-prescribe	Fostering Cross-sector Collaboration to Improve Well-Being
	Percentage under 65 who are uninsured	Strengthening Integration of Health Services and Systems
Quality	Percentage of Medicare diabetics with HbA1c testing	
	Percentage of Medicare women with recent mammogram	Fostering Cross-sector Collaboration to Improve Well-Being
Social determinants	Percentage of total population in social assistance jobs	Fostering Cross-sector Collaboration to Improve Well-Being
	Income inequality ratio	
	Percentage eligible enrolled in SNAP (food assistance)	
	Percentage of households with no car and low access to grocery stores	
	Percentage of households spending >30% of income on housing	
	ARC Economic Index	
	Social association rate per 10,000 population	Fostering Cross-sector Collaboration to Improve Well-Being
	Percentage receiving disability benefits (OASDI and/or SSI)	
	Percentage of adults with some college education	
	Percentage of households with income below poverty line	
	Median household income	

Calculation of Expected Outcomes

With the measures selected, the next task was to create a multivariate regression equation to calculate expected values for each outcome measure for each county. This formed a set of analytical equations. For example, the analytical equation for one of the outcomes, YPLL, is the following:

$$\textit{Expected YPLL} = \beta_0 + \beta_1 \times \textit{Driver 1} + \beta_2 \times \textit{Driver 2} + \beta_3 \times \textit{Driver 3} + \beta_4 \times \textit{Driver 4} \ldots + \beta_{29} \times \textit{Driver 29}$$

where β_0 equals the regression constant and β_1 through β_{29} equal the regression coefficients for each driver for that outcome.

To estimate the coefficients in the analytic equations, we applied regression analyses to each of the 19 outcomes with the set of 29 drivers serving as independent variables. We used Stata (version 14.1; College Station, Texas) to run the analysis and produced 19 different regression equations, one for each of the 19 outcomes (see below for additional detail). Each equation varies in its predictive power, with R-square values ranging from 0.279 to 0.834. Table 4 and Table 6 in Appendix C contain the regression results. Once the coefficients were determined, we calculated expected values for each outcome in each county by substituting the county's actual values for the driver measures in the equations.

Stratification of Study Sample by Rurality

Building on this basic framework, we enhanced the analysis to account for the relative rurality of the counties. As is the case for any region or state, nonmetropolitan Appalachia is quite different from metropolitan Appalachia. Appalachian counties vary in levels of isolation from large urban centers, and because resources—particularly health care resources—tend to concentrate in urban areas, distance from urban centers is likely to play a large role in the health of the community.

To stratify the regression equations for rurality, we separated counties into metropolitan and nonmetropolitan groups by using the 2015 U.S. Office of Management and Budget (OMB) definition of a Metropolitan Statistical Area (MSA). This separation recognized that metropolitan and nonmetropolitan counties can be quite different in terms of resources and overall population size, and that these differences can affect the degree to which health drivers affect health outcomes. The OMB metropolitan delineation is broad; some metropolitan counties (e.g., "bedroom counties") classify as such because of their high levels of commuting to core urban areas. Otherwise, they may resemble nonmetropolitan areas in both population size and density. However, to the extent that metropolitan status captures integration with a metropolitan center, the chosen delineation is appropriate for this model.

As a result of this stratification, the 19 outcome regression equations became 38, with a metropolitan and a nonmetropolitan version for each outcome measure.

Enhancing the Sample with Non-Appalachian Counties

We expanded the sample size beyond the 420 Appalachian counties to generate more precision. Generally, a larger study sample yields more statistical power. However, we recognize that Appalachian counties differ substantially from other parts of the United States; in statistical terms, this is an internal validity/external validity tradeoff. Therefore, we explored different options for defining the study sample. These options are outlined in Table 8.

Table 8: Study Sample Options

Option	Included counties	Description of sample
1	Appalachian only	420 Appalachian counties
2	United States	3,113 counties in the United States
3	Appalachian-like	420 Appalachian counties plus those with *similar characteristics* to Appalachian counties
4	Appalachia and proximate	420 Appalachian counties plus those *geographically close* to Appalachian counties

Options 1 and 2 have limitations. Option 1, which maximizes internal validity, suffers from low sample size. Its limited power to detect the effect of observable factors would identify Bright Spots with lower precision; we would find Bright Spots that are not worthy of inspection. The increased size of the national sample in Option 2 would have lower internal validity. Los Angeles County and the New York City boroughs are quite different from counties in the Appalachian Region and, therefore, are less relevant to the analysis.

Options 3 and 4 balance the tradeoff of increased sample size and lower internal validity. Option 3 limits the sample to counties that are empirically similar to Appalachian counties. For example, the northern lower peninsula of Michigan shares many characteristics (e.g., median income, percentage receiving disability benefits, and the supply of primary care physician) with the Appalachian Region. Option 4 takes a similar approach with distance by using a geographically weighted regression approach (Holmes & Ricketts, 2007). This approach expands the sample based on proximity to the Appalachian Region. For example, Ohio might be included, but Oregon would not.

Ultimately, we implemented Option 3. The next section describes the technical method for identifying Appalachian-like counties.

Finding Appalachian-Like Counties and Weighting the Sample

Weighting Approach

To determine the similarity of non-Appalachian counties to Appalachian counties, we used an approach similar to propensity score weighting. We derived a propensity score and used it to assign weights to each non-Appalachian county in the United States. These weights determined the extent to which non-Appalachian counties affected the regression analyses.

Propensity Score Analysis

Propensity scoring analysis derives from Rosenbaum and Rubin's 1983 work to create a better method to assign control cases in clinical research (Rosenbaum & Rubin, 1983). Historically, research methods were aimed at evaluating the effect of a given treatment on a condition or outcome focused on matching control cases (e.g., subjects given a placebo) with treated cases (e.g., subjects given a new drug). A common problem with this method is an inability to generate enough control cases, which need to have similar characteristics to treated cases (e.g., same gender, ethnicity, age, risk factors). To address this, Rosenbaum and Rubin's work created a method to identify cases that were likely to have had treatment based on a set of characteristics. Their method represents this likelihood as a probability called the *propensity score.*

While clinical research provided the inspiration for Rosenbaum and Rubin, propensity scoring analysis can apply to any situation with a continuous outcome and a dichotomous treatment. We created a probability for the dichotomous "treatment" of whether a county is likely to be in Appalachia. After creating propensity weights, researchers can then match control cases to treatment cases. The Methodological and Technical Notes section in Appendix E contains additional background on this approach.

The following section outlines the detailed approach for this adjustment. Readers not interested in the technical details may skip the section and just recognize that the concept is to put more weight on counties "similar to" Appalachia in the regression and put very little weight (including, potentially, zero weight) on those "different" from Appalachia. For example, the northern lower peninsula of Michigan and parts of Arkansas are demographically and economically similar to Appalachia, so counties from these regions were more important to the Bright Spots analysis sample. Southern California is not similar to Appalachia, so counties from this region were less important to the Bright Spots analysis sample.

Propensity Scoring in This Model

Basic Approach

Unlike a traditional propensity score analysis, the goal of our model was not to estimate the effect of a treatment but, rather, to identify non-Appalachian counties that are "similar" to Appalachian counties. Thus, we used the logic of the common support—find a unidimensional index on which we could match counties, i.e., the propensity score. Whereas the textbook method's goal is to "balance" a sample—ensure that the distribution of the confounders is similar in the matched control and treatment sample—the goal of our method was to boost sample size beyond the 420 Appalachian counties while weighting counties similar to those of Appalachia counties more than counties "less similar" to those of Appalachia.

The Bright Spots model followed the textbook case with differences as noted:

1. Run a logistic regression of "county is in the Appalachian Region" regressed on the drivers;
2. Generate propensity scores, the predicted probability of being an Appalachian county given the driver values; and
3. Calculate Appalachian weights: 1.0 if a county is in the Appalachian Region and 0.0 if perfectly dissimilar to an Appalachian county

The following section illustrates the sequence of steps that produced the two separate sets of propensity scores, one for metropolitan counties, and one for nonmetropolitan counties.

Propensity Scoring Steps

1. Run a logistic regression of "county is in the Appalachian Region" regressed on the drivers

Table 9 and Table 10 show the output of a logistic regression model that defines the probability of being an Appalachian county as a function of the drivers in the Bright Spots model. The chi-square statistic, which measures whether the drivers were statistically important, and R-square, which measures the predictive ability, showed that the drivers were quite good at assessing whether a county was in the Appalachian Region. Table 9 summarizes the metropolitan logistic regression results, and Table 10 shows the nonmetropolitan results.

Table 9: Logistic Regression Model for Metropolitan Counties

Regression statistics	
Number of observations	1,146
LR chi-square (29)	486.33
Probability > chi-square	0.000
Log likelihood	−205.3416
Pseudo R-square	0.5422

Driver	Coef.	Std. Err.	z	*P*	95% confidence interval	
Social association rate	−0.02	0.04	−0.52	0.60	−0.11	0.06
Percent in social assistance jobs	37.88	35.00	1.08	0.28	−30.73	106.48
Income inequality ratio	0.78	0.36	2.14	0.03	0.07	1.50
Percent enrolled in SNAP	0.04	0.02	1.88	0.06	0.00	0.09
Grocery stores per 1,000	−7.60	2.36	−3.22	0.00	−12.22	−2.97
Restaurants per 1,000	1.00	0.60	1.66	0.10	−0.18	2.17
Percent with no car and low access to grocery stores	0.03	0.01	2.44	0.02	0.01	0.05
Access to exercise	0.05	0.01	4.57	0.00	0.03	0.07
Percent spending >30% income on housing	−0.11	0.04	−2.58	0.01	−0.19	−0.03
Doctors who e-prescribe	0.00	0.01	0.34	0.74	−0.01	0.01
Percent adults who smoke	−0.12	0.07	−1.62	0.10	−0.26	0.02
Percent adults physically inactive	0.02	0.05	0.33	0.74	−0.08	0.11
Chlamydia incidence	0.00	0.00	−4.22	0.00	−0.01	0.00
Diabetes A1C testing	0.01	0.05	0.26	0.80	−0.09	0.11
Breast cancer screening	0.06	0.03	2.22	0.03	0.01	0.11
Percent receiving disability	0.76	0.13	5.76	0.00	0.50	1.02
Teen birth rate	−0.06	0.02	−3.42	0.00	−0.10	−0.03
Student-teacher ratio	−0.34	0.08	−4.30	0.00	−0.50	−0.19
Percent with some college	−0.07	0.03	−2.67	0.01	−0.12	−0.02
Average daily air pollution	1.72	0.22	7.70	0.00	1.28	2.15
Primary care physician ratio	0.02	0.01	2.03	0.04	0.00	0.03
Dentist ratio	0.00	0.01	0.22	0.83	−0.02	0.02
Physician specialist ratio	0.00	0.00	0.22	0.83	0.00	0.01
Mental health provider ratio	0.00	0.00	−1.86	0.06	−0.01	0.00
Percent households below poverty	−0.08	0.08	−0.97	0.33	−0.24	0.08
ARC Economic Index	0.01	0.01	0.40	0.69	−0.02	0.03
Median income	0.00	0.00	−2.37	0.02	0.00	0.00
Average travel time to work	0.13	0.04	2.96	0.00	0.04	0.22
Percent uninsured under 65	0.03	0.05	0.59	0.56	−0.07	0.13
Constant	−25.01	6.52	−3.83	0.00	−37.79	−12.22

Table 10: Logistic Regression Model for Nonmetropolitan Counties

Regression statistics	
Number of observations	1,965
LR chi-square (29)	938.89
Probability > chi-square	0.000
Log likelihood	−313.310
Pseudo R-square	0.5997

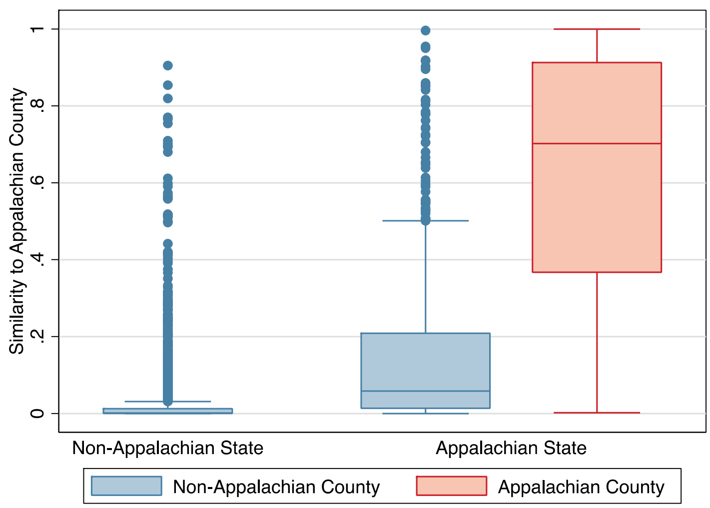
Driver	Coef.	Std. Err.	z	*P*	95% confidence interval	
Social association rate	−0.06	0.03	−2.17	0.03	−0.12	−0.01
Percent in social assistance jobs	−2.24	19.85	−0.11	0.91	−41.16	36.67
Income inequality ratio	0.33	0.24	1.38	0.17	−0.14	0.79
Percent enrolled in SNAP	−0.05	0.02	−2.58	0.01	−0.09	−0.01
Grocery stores per 1,000	−2.88	1.12	−2.57	0.01	−5.07	−0.68
Restaurants per 1,000	−0.46	0.44	−1.03	0.30	−1.33	0.41
Percent with no car and low access to grocery stores	−0.04	0.01	−3.57	0.00	−0.06	−0.02
Access to exercise	0.03	0.01	5.26	0.00	0.02	0.05
Percent spending >30% income on housing	−0.03	0.03	−0.88	0.38	−0.08	0.03
Doctors who e-prescribe	0.00	0.00	−0.16	0.88	−0.01	0.01
Percent adults who smoke	−0.04	0.06	−0.75	0.45	−0.15	0.07
Percent adults physically inactive	0.05	0.03	1.60	0.11	−0.01	0.12
Chlamydia incidence	0.00	0.00	−5.34	0.00	−0.01	0.00
Diabetes A1C testing	0.05	0.02	1.96	0.05	0.00	0.09
Breast cancer screening	−0.01	0.02	−0.43	0.66	−0.04	0.03
Percent receiving disability	0.39	0.07	5.39	0.00	0.24	0.53
Teen birth rate	−0.06	0.01	−5.45	0.00	−0.08	−0.04
Student-teacher ratio	−0.21	0.05	−4.45	0.00	−0.30	−0.11
Percent with some college	−0.04	0.02	−2.45	0.01	−0.08	−0.01
Average daily air pollution	1.53	0.16	9.75	0.00	1.22	1.84
Primary care physician ratio	0.00	0.01	−0.42	0.68	−0.01	0.01
Dentist ratio	0.01	0.01	0.85	0.40	−0.01	0.02
Physician specialist ratio	0.01	0.00	3.80	0.00	0.01	0.02
Mental health provider ratio	0.00	0.00	0.13	0.90	0.00	0.00
Percent households below poverty	−0.11	0.05	−2.30	0.02	−0.21	−0.02
ARC Economic Index	0.00	0.01	0.43	0.66	−0.01	0.02
Median income	0.00	0.00	−3.78	0.00	0.00	0.00
Average travel time to work	0.07	0.03	2.29	0.02	0.01	0.13
Percent uninsured under 65	0.03	0.04	0.80	0.43	−0.05	0.11
Constant	−8.99	4.76	−1.89	0.06	−18.32	0.34

The regressions identified which measures best predicted whether a county was in the Appalachian Region. For illustration, we discuss the first two variables shown in Table 9 that are statistically significant: the income inequality ratio and the number of grocery stores per 1,000 population. Both variables were associated with being in the Appalachian Region in the metropolitan groups ($p < 0.05$). That is, metropolitan counties with higher income inequality or with fewer grocery stores per 1,000 residents were more likely to be in Appalachia. The relative effect of these factors varies by rurality; for example, the social association rate is predictive of being an Appalachian county among nonmetropolitan counties but not among metropolitan counties.

2. Develop Appalachian propensity scores

From the regression estimates, we calculated the probability that a county was "Appalachian-like." Figure 7 shows probability distributions in box-plot format for non-Appalachian states, non-Appalachian counties in Appalachian states, and Appalachian counties. The y-axis in Figure 7 shows the probability that a given county is similar to an Appalachian county. Figure 7 combines propensity score data for all counties, both metropolitan and nonmetropolitan.

Figure 7: Probability of Similarity between a U.S. County and a Typical Appalachian County

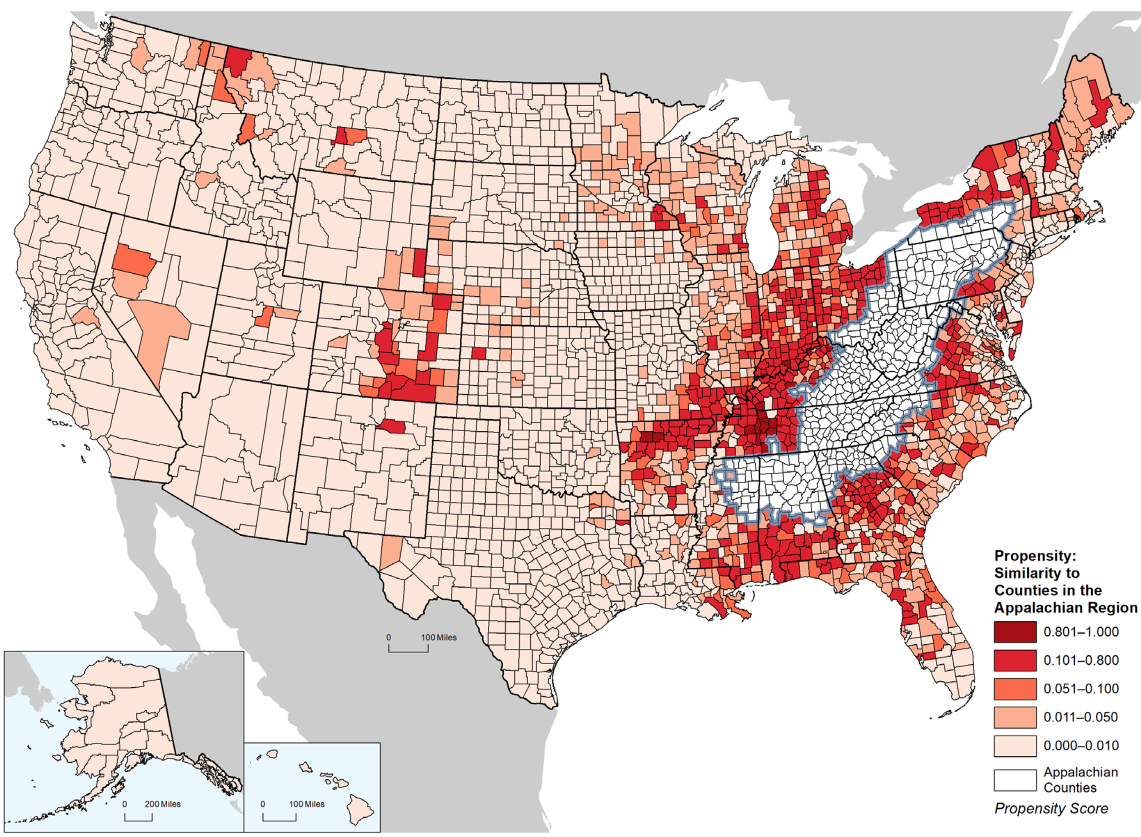


These results show diversity even among Appalachian counties. The Appalachian box (far right) shows that the characteristics of half of the Appalachian counties have at least a 70 percent probability of being an Appalachian county (the horizontal line within the box denotes the median county). That is, if provided county characteristic values for such a county, there is a 70 percent chance that county is in Appalachia. Only a few Appalachian counties are unlikely to be Appalachian. Sullivan County, Pennsylvania, and Marshall County, Mississippi, are the most atypical nonmetropolitan and metropolitan Appalachian counties, respectively. However, overall, counties in the Appalachian Region are highly similar to the typical Appalachian county. Non-Appalachian counties in Appalachian states (e.g., northwestern Ohio, eastern North Carolina, and western Kentucky) are less similar to those in the Region but have many counties with similar profiles to those in the Region. For example, Carroll, Decatur, and Wayne Counties in Tennessee have many similarities with the typical Appalachian nonmetropolitan county. Likewise, Twiggs County, Georgia, and Hickman County, Tennessee, are similar to metropolitan Appalachian counties. In Figure 7 the high similarity scores, or blue dots, at the top of the box plot represent these counties and other non-Appalachian counties very similar to Appalachian counties.

Though a few counties outside the Appalachian states share similarities with those in the Region (e.g., Searcy and Newton Counties, Arkansas), most counties outside the Region have low similarity to counties in the Region. Figure 7 shows that the bulk of the non-Appalachian state "box" has values near zero.[2] The results show what we expect: non-Appalachian counties generally look very different from Appalachian counties. This shows that the model is differentiating between Appalachian and non-Appalachian counties rather well and, at the same time, is identifying counties that are similar to Appalachian counties.

Figure 8 maps the results and includes both metropolitan and nonmetropolitan counties; for counties outside Appalachia, the darker the color, the more similarity to Appalachian counties. Most of the counties that are similar to Appalachian counties are located close to the Region, with some exceptions, such as northeastern Arkansas and eastern Colorado. A few counties close to the Region are rather dissimilar. For example, Franklin County, Ohio (Columbus), lies just outside the Appalachian Region but has very low similarity (light color) because of its more educated, higher-income population and the presence of more health care providers than the typical metropolitan Appalachian county. Appalachian counties are marked white on this map.

Figure 8: National Map of Appalachian Probability Scores



[2] Seventy-two percent of counties in non-Appalachian states have a propensity score ("probability" of a county with that profile being Appalachian) of 0.01 or lower.

3. Create Appalachian weights

To boost the sample size and give priority to Appalachian counties, the final step converted propensity scores to weights by simply defining the weight as the propensity score for all counties outside Appalachia. All Appalachian counties had a weight defined to 1.0. Table 13 summarizes the distribution of weights. Specifically, the 268 nonmetropolitan Appalachian counties had a weight of 1.0. The remaining 1,697 nonmetropolitan counties had an average propensity score, or weight, of 0.056.[3] The 152 metropolitan Appalachian counties had a weight of 1.0. The remaining 994 metropolitan counties had an average propensity score, or weight, of 0.063. This meant that, in aggregate, the weights from the 1,697 nonmetropolitan non-Appalachian counties equaled the weight of about 95 Appalachian counties, and the weights from the 994 metropolitan non-Appalachian counties equaled the weight of about 63 Appalachian counties. In other words, we effectively boosted the sample size but acknowledged the fact that these extra counties differ from Appalachian counties.

Table 11: Average Weights for All Counties, Stratified by Rurality

Rurality	County type	Frequency	Average weight	Sum of weights
Nonmetro	Non-Appalachian county	1,697	0.056	95.03
	Appalachian county	268	1.000	268
Metro	Non-Appalachian county	994	0.063	62.62
	Appalachian county	152	1.000	152

Table 7 in Appendix C, contains the propensity scores and Appalachian weights for all counties in the United States.

Identification of Bright Spots

To identify Bright Spots, we used the *residual*s from each regression equation. As illustrated in Figure 5, residuals are the difference between the observed outcomes and the expected outcomes. In this case, they identified counties where outcomes were better than expected based on the driver measures.

To make residuals comparable across outcomes, we standardized the raw outcomes for each county into *z*-scores. *Z*-scores measure the number of standard deviations a value lies from the mean of a given data set (here, outcomes). We reversed the *z*-score signs on the outcomes such that positive *z*-scores indicated “good health.” Because all measures in the outcome set had a “lower is better” orientation, all measures were reverse coded in this manner. Each of the 38 regression equations (19 outcomes each for metropolitan and nonmetropolitan county sets) had different results, with counties that had high residuals (better-than-expected) and counties that had low residuals (worse-than-expected).

By using our concept of overall health, we sought counties that outperformed expectations across multiple outcomes. Because *z*-scores, or standardized residuals, represented each outcome residual, we could average the standardized residuals across all 19 outcomes and generate a single average outcome residual for every county. This average standardized residual was the core measure of “brightness.”

[3] Two counties, Petersburg, Alaska, and Broomfield, Colorado, are not included in the analysis because of suppressed data.

We then defined *Bright Spots* as those counties where the average standardized residual was in the top decile within each stratum (metropolitan or nonmetropolitan). There was no natural break point in the distribution of average standardized residuals in either metropolitan or nonmetropolitan sets. Other national health care outcome comparisons, such as Hospital Compare (Center for Medicare and Medicaid Services, 2016), use the top decile to identify the best performers. With no clear demarcation points in the score rankings, we followed the CMS pattern and classified the counties in the top decile in both the metropolitan and nonmetropolitan groups as Bright Spots. The Results chapter of this report lists all 42 Bright Spot counties—27 nonmetropolitan and 15 metropolitan.

OBJECTIVE 2: MATCH BRIGHT SPOT COUNTIES TO OTHER COUNTIES

Because the purpose of the research was to identify Bright Spots, explore best practices or aspects of local culture in those communities that may be associated with better-than-expected health outcomes, and ultimately, share those features with other communities, we calculated a measure to determine the similarity between each Appalachian county and each Bright Spot county. We matched Bright Spot counties to other counties in Appalachia by finding the closest match between Bright Spot counties and all other Appalachian counties based on their driver values.

To determine the degree of similarity between each Appalachian county and a Bright Spot county, we calculated the Euclidean distance—the square root of the sum of squared distances between the standardized values for each county in the United States and the Bright Spot counties. (This is not the typical use of "distance," or how far apart the counties were *geographically*; rather, we used this measure to determine how "far" apart counties were based on *demographics* and the other drivers). We applied the calculation separately to metropolitan and nonmetropolitan counties.

Previous researchers have used this general method to describe the similarity of larger regions as a predictor for how physicians migrate. (Holmes & Fraher, In press). In a two-variable situation, this can be conceptualized as the "distance" between points in a (x, y) plane. For example, if the (standardized) smoking rate values for two counties were 0 and 1.5, and the supply of dentists for the same two counties were −1 and 2, then the Euclidean distance between the two counties would be calculated as follows:

$$\text{EUCLIDEAN DISTANCE} = \sqrt{(0-1.5)^2 + (-1-2)^2}$$

$$= \sqrt{2.25 + 9}$$

$$\text{EUCLIDEAN DISTANCE} = \sqrt{(11.25)}$$

$$\text{EUCLIDEAN DISTANCE} = 3.35$$

Smaller distance numbers denoted "closer" and, thus, more similar communities. The full results are listed in Table 13 and Table 14 in Appendix C. For example, among the 15 metropolitan Bright Spot counties, Bibb County, Alabama, was most similar to Wirt, West Virginia.

OBJECTIVE 3: IDENTIFY HEALTH DRIVERS ASSOCIATED WITH POSITIVE OUTCOMES

To determine which health drivers had the greatest impact on the 19 outcomes, we conducted additional univariate regression analyses that showed the relationship between each of the 29 drivers and the 19 outcomes. To determine the univariate regression equations, we continued the stratification of metropolitan and nonmetropolitan counties. To motivate the approach, first consider a basic univariate regression equation for the relationship between the social association rate and YPLL:

$$YPLL_i = \alpha + SOCASSOC_i\beta + \epsilon_i$$

where *i* denotes county. We are interested in how the social association rate is associated with *all* of the outcomes, so we "stack" the outcomes, allowing outcome-specific coefficients and constants:

$$OUTCOME_{ij} = \alpha_j + SOCASSOC_i\beta_j + \epsilon_{ij}$$

where *j* denotes the specific outcome (e.g., YPLL, cancer mortality, etc.) Each nonmetropolitan model has 1,900 counties × 19 outcomes or roughly 36,000 observations. This approach has one model for each driver (29) for each metropolitan status (2) for a total of 58 different models (29 × 2 = 58). We specify the outcome variables in the standardized (z-score) form and allow for clustering at the county level. As in the multivariate models, we weighted regressions by using the propensity weights and stratified by rurality. To test the statistical significance of the driver, we tested the null hypothesis that all β are zero. To assess predictive power, we calculated the partial R-square resulting from the addition of the driver in explaining the variation in the outcomes: the difference in R-square between the model above and an "indicator only" model specifying α for each outcome but not including the social association variable.

Results

SUMMARY

Appalachian counties with an average standardized residual score in the top decile (10 percent) of either the metropolitan or the nonmetropolitan group were classified as Bright Spots. The model identified a total of 42 Bright Spot counties: 15 in the metropolitan group and 27 in the nonmetropolitan group (see Table 13 and Table 14). The 42 counties in the top decile represent *the best of the better than expected*.

The Bright Spot counties are located in all five Appalachian subregions, include both economically distressed and non-distressed counties, and represent the diversity of communities across the Appalachian Region.

The variation in county location and economic status lends support to the study design—we did not aim to identify healthy counties with high levels of resources and the sorts of characteristics that support positive health outcomes, but rather counties encompassing a wide range of resource levels and characteristics that all managed to find a way to be healthier than expected. Bright Spots are places that exceed expectations, *regardless of the values of the driver*s. This is a strength of the approach, one that allowed us to focus on the positive aspects of communities *relative to their own characteristics and resource levels.*

By using the average degree to which a county's observed health outcomes exceeded expected values, the Bright Spots model identified the counties that either did very well on a few measures or exceeded expectations—perhaps only marginally—across many health outcomes. Nonetheless, 3 of the 19 health outcome measures were strongly correlated with better-than-expected overall health in the Bright Spot counties:

- Premature mortality (Years of potential life lost);
- Unintentional injury mortality; and,
- Poisoning mortality.

The second objective of this study was to match Bright Spot counties to other Appalachian counties to facilitate the exchange of ideas and lessons learned. Table 13 and Table 14 in Appendix C contain these matches. This matching allows community leaders throughout Appalachia to quickly identify the Bright Spot that is most similar to their circumstances, allowing a starting point in designing specific strategies to improve the health of their community.

The model also provides a foundation for the study's third objective, which is to identify drivers that appear to have the biggest impact on health outcomes. The nature of the empirical approach used to identify the Bright Spots made it challenging to draw explicit conclusions about specific drivers. For example, results from the multivariate regression were not sufficiently robust to support a statement such as, "median income is the most important factor in determining health outcomes." The results of this analysis suggest that the following seven drivers (direction generally associated with improved health shown in parentheses) had the most significant impact on the collective set of the 19 health outcomes:

- Median income (higher);
- ARC Economic Index (lower);
- Poverty rate (lower);
- Percentage of adults that smoke (lower);
- Percentage of adults that are physically inactive (lower);
- Percentage of the population receiving disability payments (lower); and,
- Teen birth rates (lower).

OBJECTIVE 1: IDENTIFY BRIGHT SPOT COUNTIES

Regression Results

The analysis involved separate regression equations for metropolitan and nonmetropolitan Appalachian counties. Each set of equations includes 19 different regressions, one for each health outcome (dependent variable), and each regression had the same 29 health drivers (independent variables). Thus, each individual regression run produced 29 coefficients (one for each driver), plus an intercept, resulting in 570 overall regression parameters (30 × 19 = 570).

In multivariate regression, the coefficients measure the estimated impact of an independent variable on the dependent variable, holding constant ("controlling for") the other independent variables. When independent variables themselves are highly correlated—for example, the poverty rate and median income—this is known as multicollinearity; in the presence of multicollinearity the individual coefficient estimates may yield unexpected results. In some analyses, multicollinearity might limit interpretations and require additional mitigation. However, because the focus of this model was on the *aggregate prediction*, rather than the predicted effect from an individual driver, multicollinearity was of less concern.

Appendix C contains the parameters for the 19 regression equations for both the metropolitan and nonmetropolitan counties. The unstandardized outcomes provide the bases for the regression parameters presented in Table 4 and Table 6 in Appendix C. The standardized residuals for each county/outcome combination are reported in Table 3 and Table 5 in Appendix C. The standardized residuals are presented as *z*-scores, created by standardizing the outcome data sets in terms of standard deviations from the mean and then reversing the signs for all measures, so positive *z*-scores indicate "good health." Thus, in most of the discussion that follows, a positive residual means "healthier than expected" even when referring to a measure of "poor health." For example, unless otherwise stated, a positive residual for mortality indicates mortality rates that are lower than expected.

Model Strength for Each Outcome

The model's ability to predict each of the individual dependent variables differed for a number of reasons:

- Several outcome measures are synthetic estimates developed from similar variables. For example, many County Health Rankings measures are derived from Behavioral Risk Factor Surveillance System (BRFSS) state sample data that adjusted to the county level by using a synthetic process that "smoothed" on the basis of county characteristics. Thus, to the degree that the variables used to "smooth" and derive the county estimates correlated with the model's drivers, the model may appear to capture variation in outcomes when it may, instead, be capturing more variation in the *synthetic* outcome, which is itself a function of some of the drivers. An example of this is the BRFSS measure "physically unhealthy days."
- Some outcome variables are inherently less statistically precise. For example, heart disease mortality has less annual variation than mortality from a condition that occurs less frequently (e.g., poisoning). Likewise, the difference in denominators for the different outcome measures will lead to different levels of precision. For example, infant mortality rates are based on the number of deliveries, which, even when aggregated over several years, will be based on a much smaller number than the denominators for measures like heart disease hospitalizations or the cancer mortality rate. Values for outcome measures with smaller denominators will tend to vary more.

- Some outcome measures simply vary less nationally. For example, although cancer incidence rates vary with differences in the prevalence of risk factors such as smoking, the variance in national cancer mortality is less than half as much as the variance in heart disease mortality. Cancer mortality rates vary across counties by about ±15 percent, while heart disease rates vary by ±25 percent (Marshall, et al., 2017).

Table 12 includes the R-square values for all 19 outcome equations for both metropolitan and nonmetropolitan groups. R-square values represent the proportion of variation in the outcome explained by the model (that is, the driver variables). Higher values of R-square explain more variation in the outcome; an R-square of 1.0 means the model explains all of the variability.

The highest R-square value in both data sets was for physically unhealthy days at 0.87 and 0.89 for metropolitan and nonmetropolitan counties, respectively. These high R-square values indicate that the variables included in the model are strong predictors for physically unhealthy days, which is not surprising given the synthetic nature of that measure. Conversely, the model was the least predictive for Medicare Part D opioid prescriptions at 0.28 and 0.20 for metropolitan and nonmetropolitan, respectively. In general, the models explained more variation in mortality outcomes than other outcomes.

The model does not fully explain the variance in any individual outcome; most health outcomes are the result of cumulative factors, including local characteristics. In the aggregate, the model confirms that social determinants are clear drivers of health outcomes, especially mortality outcomes, but the fact that the R-square values for the non-synthetic outcomes are below 0.80 means that factors that are *not* included in this model explain 20 percent or more of the variation in every outcome measured except physically unhealthy days.

Table 12: R-square Values for Each Outcome in Both Metropolitan and Nonmetropolitan Models

Outcome	Metro	Nonmetro
Physically unhealthy days[a]	0.87	0.89
Percent drinking excessively[a]	0.75	0.79
Mentally unhealthy days[a]	0.79	0.74
YPLL	0.78	0.67
Cancer mortality	0.70	0.59
% low birth weight	0.58	0.59
Poisoning mortality	0.42	0.56
COPD mortality	0.56	0.55
Injury mortality	0.62	0.54
% adults with obesity	0.59	0.54
Heart disease mortality	0.52	0.53
Diabetes prevalence	0.66	0.53
Heart disease hospitalizations (Medicare)	0.38	0.42
Infant mortality	0.43	0.37
Depression prevalence (Medicare)	0.43	0.33
Average HCC risk score (Medicare)	0.38	0.31
Stroke mortality	0.37	0.23
Suicide mortality	0.35	0.21
% opioid Rx claims (Medicare)	0.28	0.20

[a] *These measures are derived from BRFSS survey data and synthesized for counties with small sample sizes. As explained in the text, these values may have a high correlation to many of the drivers, and, therefore, it is not surprising that they have a high R-square value in this model.*

Expected Values and Residuals

Each linear regression for each individual outcome generated an expected value for each county. The focus of the Bright Spots analysis is the *residual*—the difference between the actual (or observed) value and the expected (or predicted) value.

For example, a portion of the linear regression equation for YPLL in metropolitan areas by using the 29 variables is:

$$\text{Expected YPLL} = 41.722 \times [\text{social association rate}] + 58.322 \times [\text{income inequality}] - 11{,}000 \times [\%\ \text{social assistance jobs}] + 0.418 \times [\text{SNAP benefits}] - 0.00011 \times [\text{grocery stores}] + 70.216 \times [\text{restaurants}] - 10.139 \times [\text{no car}] \ldots$$

Tables 4 and 6 in Appendix C contain the non-standardized regression results for the expected outcomes. For illustration, consider the example above. It shows that an increase of one unit in the social association rate would lead to an increase of 41.722 in expected YPLL. Meanwhile, an increase of 1 percent in income inequality would lead to an increase of 58.322 in YPLL. Note that the social association rate runs counter to expectations and serves as an example of the multicollinearity issue discussed above.

The focus, however, is not on the coefficients of the individual drivers, but rather on the *overall prediction of better-than-expected outcomes*. For example, as illustrated in Figure 9 below, the largest outlier (highest residual) for YPLL in nonmetropolitan counties was for Green County, Kentucky, shown as a solid red circle. Green County's observed average YPLL was 7,080. By using the regression formula above, we found that Green County's expected YPLL was 10,264, which is 3,184 years higher than the observed value for the county. Figure 9 illustrates the observed and expected YPLL values for all 420 Appalachian counties.

Figure 9: Plot of Actual and Predicted YPLL, All Appalachian Counties

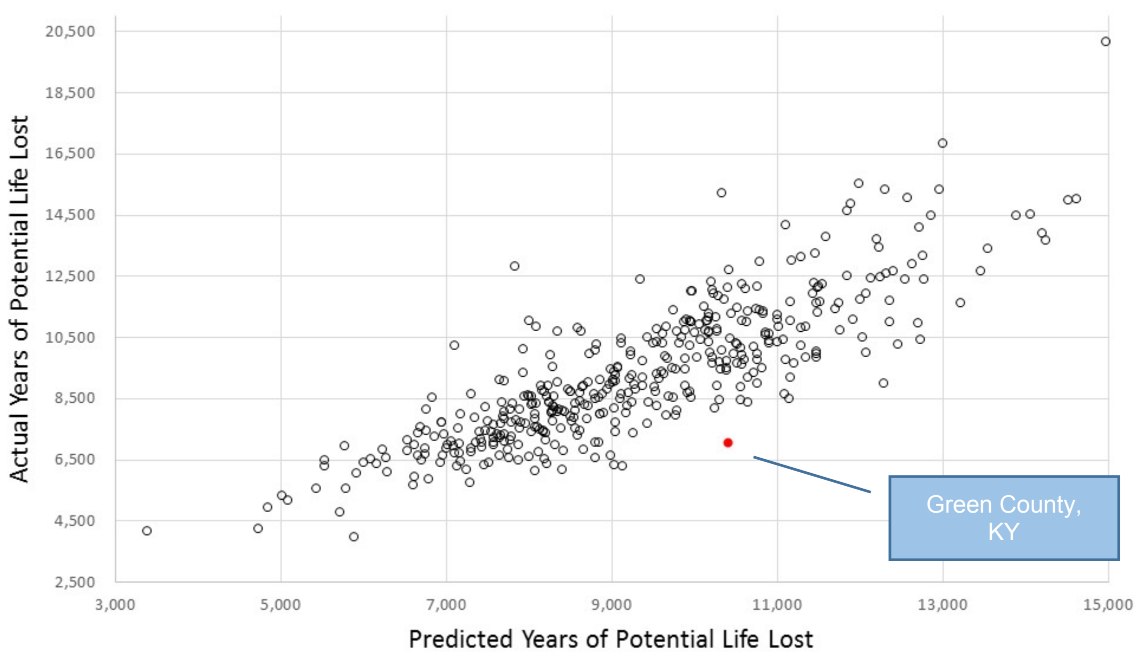


Figure 10 plots the residuals for premature mortality, controlling for median income. It shows that after adjusting for median income, median income has little impact on the residual (the portion unexplained by the drivers). This is exactly as expected; after adjusting for income, there should be no relationship between income and the portion unexplained by income. In other words, there are counties with high and low incomes that have better-than-expected YPLL, and there are counties with high and low incomes with worse-than-expected YPLL. The solid red points on the map are counties in the top decile for YPLL residuals, the 10 percent in which outcomes are highest relative to expectations. The dark horizontal line represents the zero residual, and the dark vertical line represents median Appalachian income.

Figure 10: Residual Plot of Nonmetropolitan Premature Mortality (YPLL), by Median County Income

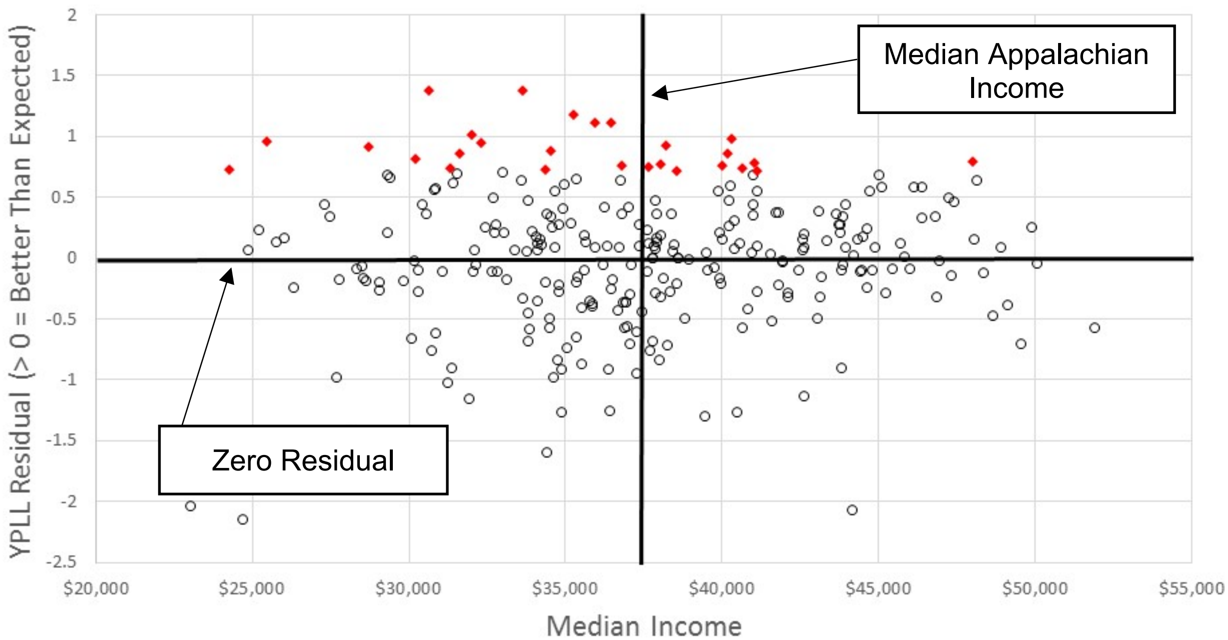


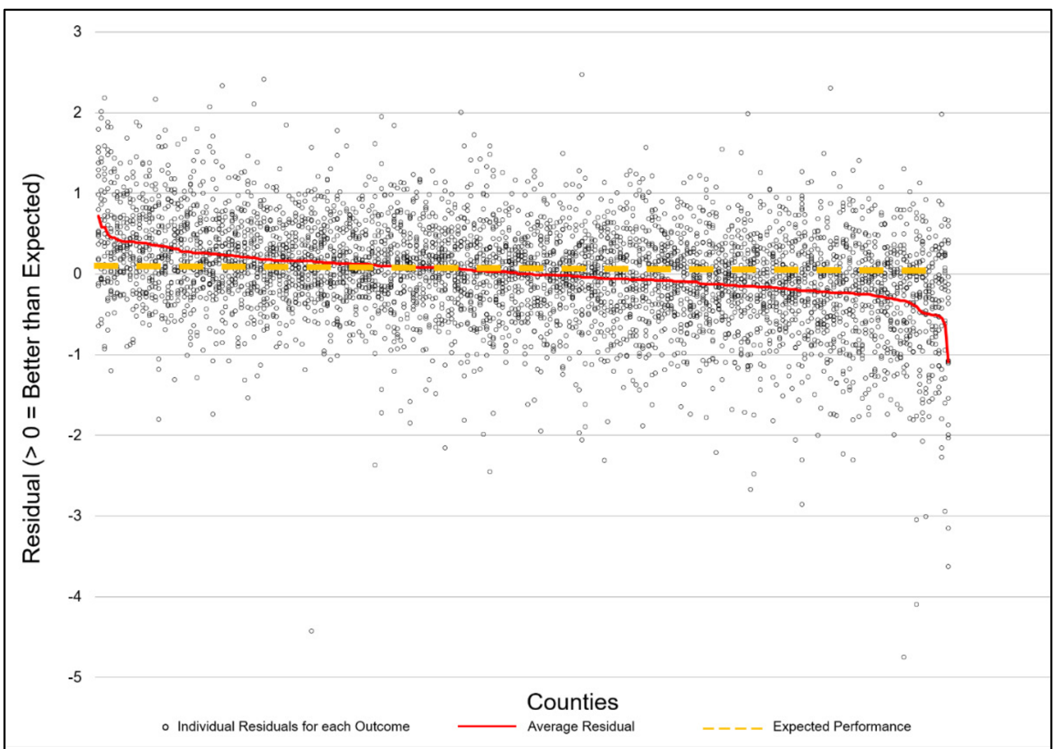
Figure 10 underscores a key principle of the approach: Bright Spots are places that exceed expectations, *regardless of the values of the drivers*. In Figure 10, there are low-income, middle-income, and high-income communities that exceed expectations. This is a strength of the approach, one that allowed us to focus on the positive aspects of communities *relative to their own characteristics and resource levels.* This approach lets us identify Bright Spot counties throughout the Region, despite the variation in characteristics and resources throughout the 420 Appalachian counties.

Moreover, because the outcomes were converted to a common metric (*z*-score), the residuals are comparable across outcomes. Tables 3 and 5 in Appendix C contain the standardized residuals for each county-outcome combination for both the metropolitan and nonmetropolitan data sets. Because the outcomes have been standardized, the interpretation of the residual is in standard deviation units. For example, a standardized residual of 0.5 means the county outcome value was half of a standard deviation better than expected, on average.

Average Residuals

The standardized residuals for all outcomes in the model were averaged to create an average residual. The average residual represents "overall brightness." The average residual for each county in the metropolitan and nonmetropolitan groups are shown in Table 3 and Table 5 in Appendix C, respectively. The higher the average residual, the healthier, on average, a county is *relative to expectations*. The average residuals provide a relative ranking of all counties in the metropolitan and nonmetropolitan groups. The red line in Figure 11 shows the average residuals for all outcomes for every nonmetropolitan county in Appalachia. The black circles are the individual residuals for each county-outcome combination. The vertical axis represents the range of outcome residuals. The counties are represented along the horizontal axis, sorted by the county with the highest average residual on the left (Wayne County, Kentucky) and the county with the lowest average residual on the right. The distribution looks similar to the scatterplot for the metropolitan counties in Figure 12.

Figure 11: Scatterplot of All Residuals for Nonmetropolitan Counties



The red lines in Figures 10 and 11 represent an average. High-performing counties perform better on average across all measures but underperform on some measures. Similarly, low-performing counties overperform on some measures.

Figure 12: Scatterplot of All Residuals for Metropolitan Counties

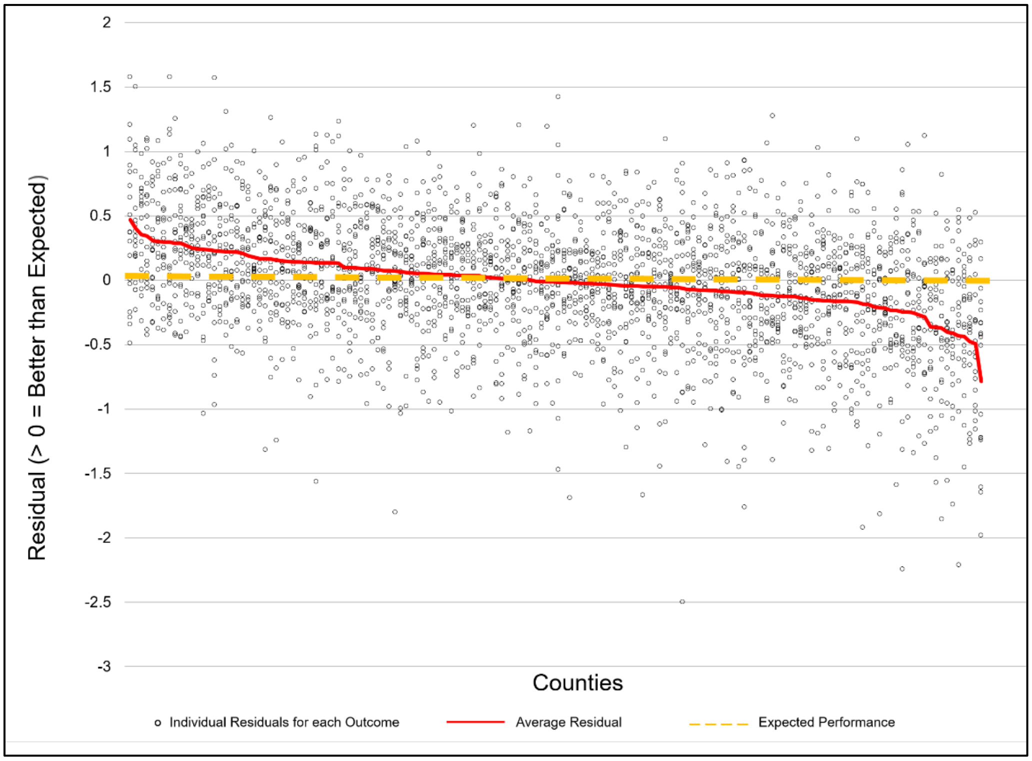


Figure 11 and Figure 12 illustrate that counties with a high average residual tend to outperform expectations on more of the individual outcomes. Thus, the range of black circles tends to shift toward the top of the chart for the counties with a high average residual and toward the bottom for counties with a low average residual. Because the metro and nonmetro datasets are distinct, average residuals of the two groups are *not* comparable.

In this model, a single county has 19 different individual outcome residuals (the black circles in the figures above). Even counties with a high average residual do not have top-ranked residuals for every outcome. A county with a very high YPLL residual and, thus, a much better-than-expected YPLL may exhibit worse-than-expected performance on the obesity measure and, therefore, rank low for that measure. For example, Wayne County, Kentucky (on the far left in Figure 11), has the highest average residual among nonmetropolitan counties. Wayne shows much better-than-expected outcomes in stroke mortality (residual = 1.8), injury mortality (residual = 1.2), depression (residual = 1.6), and YPLL (residual = 1.4). However, Wayne underperforms on some measures, such as percent of opioid prescription claims (residual = -0.6) and physically unhealthy days (residual = -0.1).

Even the county with the lowest average residual performed better than expected on some outcomes. The range of residuals across outcomes for the lowest performing county, shown in the far right of the figures above, overlaps significantly with the range of the residuals of the highest performing county on the far left. We note this to emphasize that the model uncovered counties that appear to perform *better generally*. High-performing counties do not perform better than other counties on all outcomes. As such, the model is not one from which we can make absolute statements; Bright Spot counties were not better than expected on all measures. Similarly, Bright Spot counties were not better on all outcomes than the lower ranked counties. But, taken in aggregate, the Bright Spot counties exceeded expected outcomes to a *greater degree* than other counties.

Bright Spot Counties

To recap, Bright Spots are the counties whose average standardized residual ranked in the top decile among either metropolitan or nonmetropolitan Appalachian counties. Table 13 and Table 14 rank the Bright Spot counties by their average standardized residual. The tables also show the county's outcome measure with the highest individual residual. For example, because residuals for an individual measure can be interpreted as the number of standard deviations from the expected value, Wirt County, West Virginia's injury mortality outcome value is 1.51 standard deviations better than expected (lower because lower mortality means better health).

As anticipated, the expected results differed between the metropolitan and nonmetropolitan counties. Scores for metropolitan counties have less variation, and the top metropolitan scores are lower than the top nonmetropolitan scores. However, the variation reflects the difference in the comparison groups, not an absolute difference in health. That is, we expected nonmetropolitan residuals to have higher variation (higher maximum values) for multiple reasons. The most straightforward reason is that "bigger surprises" are possible in smaller counties. For example, even with multiyear data, one bad car accident could spike YPLL in a less populous county but would be less consequential in a more populous, metropolitan county.

Table 13: Metropolitan Bright Spot Counties Ranked by Average Standardized Residual

Rank	County	State	Average residual score[a]	Highest individual outcome residual score[b]	
1	Wirt	West Virginia	0.47	Injury mortality	1.58
2	Clay	West Virginia	0.40	Heart disease mortality	1.51
3	Henderson	North Carolina	0.35	% adults with obesity	0.98
4	Hale	Alabama	0.35	Depression prevalence	1.10
5	Sequatchie	Tennessee	0.31	Poisoning mortality	1.22
6	Floyd	Virginia	0.30	COPD mortality	1.08
7	Sullivan	Tennessee	0.30	Poisoning mortality	1.23
8	Marshall	Mississippi	0.30	% opioid Rx claims	1.58
9	Madison	North Carolina	0.29	% adults with obesity	1.26
10	Whitfield	Georgia	0.29	Depression prevalence	0.97
11	Tioga	New York	0.27	Stroke mortality	0.87
12	Schoharie	New York	0.25	Average HCC risk score	0.83
13	Beaver	Pennsylvania	0.25	Average HCC risk score	1.00
14	Jefferson	Tennessee	0.24	Average HCC risk score	1.06
15	Catoosa	Georgia	0.24	Stroke mortality	0.90

[a] *Average residual score for the regression involving 152 metropolitan counties.*
[b] *Highest individual outcome residual score for this county and the associated measure.*

Table 14: Nonmetropolitan Bright Spot Counties Ranked by Average Standardized Residual

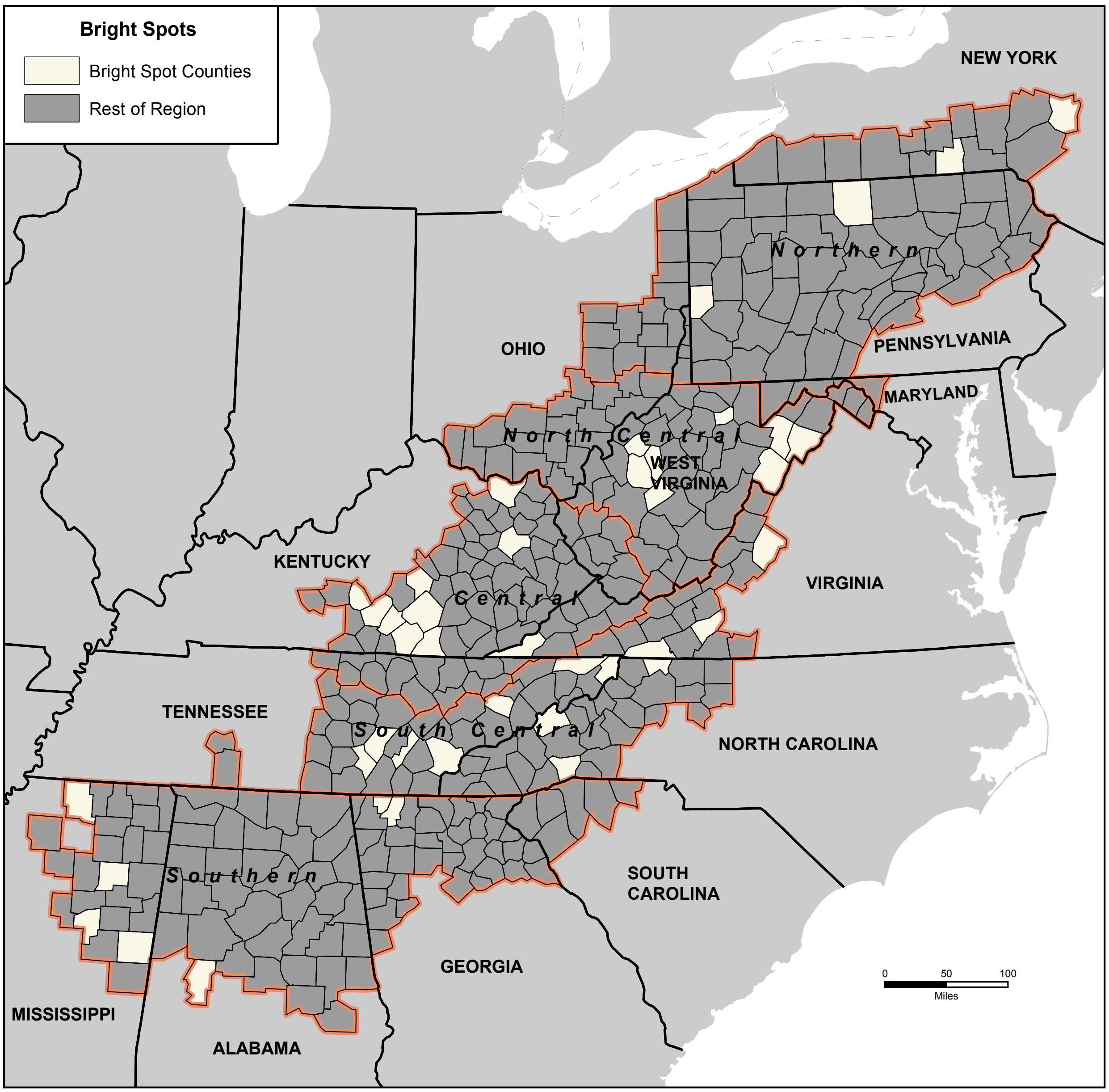
Rank	County	State	Average residual score[a]	Highest individual outcome residual score[b]	
1	Wayne	Kentucky	0.72	Stroke mortality	1.79
2	Noxubee	Mississippi	0.58	COPD mortality	2.19
3	Calhoun	West Virginia	0.58	Injury mortality	2.02
4	Grant	West Virginia	0.49	Cancer mortality	1.88
5	McCreary	Kentucky	0.45	Poisoning mortality	1.94
6	Potter	Pennsylvania	0.45	Heart disease mortality	1.44
7	Taylor	West Virginia	0.42	Heart disease hospitalizations	1.20
8	Rockbridge	Virginia	0.41	Heart disease hospitalizations	1.37
9	Pulaski	Kentucky	0.40	Poisoning mortality	1.64
10	Green	Kentucky	0.40	YPLL	1.38
11	Lee	Virginia	0.40	Poisoning mortality	2.29
12	Russell	Kentucky	0.40	Heart disease hospitalizations	1.68
13	Bledsoe	Tennessee	0.39	Cancer mortality	1.88
14	Grayson	Virginia	0.39	Injury mortality	1.83
15	Hardy	West Virginia	0.38	% opioid Rx claims	1.21
16	Johnson	Tennessee	0.38	Poisoning mortality	1.52
17	Lincoln	Kentucky	0.37	% adults with obesity	1.37
18	Meigs	Tennessee	0.36	% opioid Rx claims	2.17
19	Pendleton	West Virginia	0.36	Poisoning mortality	1.48
20	Choctaw	Mississippi	0.35	Cancer mortality	1.69
21	Adair	Kentucky	0.35	Injury mortality	1.57
22	Lewis	Kentucky	0.34	Depression prevalence	1.78
23	Roane	West Virginia	0.33	Heart disease hospitalizations	1.35
24	Monroe	Tennessee	0.32	COPD mortality	1.18
25	Alleghany	North Carolina	0.31	YPLL	1.18
26	Chickasaw	Mississippi	0.31	Stroke mortality	1.61
27	Morgan	Kentucky	0.28	Injury mortality	0.92

[a] *Average residual score for the regression involving 268 nonmetropolitan counties.*
[b] *Highest individual outcome residual score for this county and the associated measure.*

The strength of this research approach is its capacity to identify counties with better-than-expected health. However, determining what causes a county to exceed expectations requires further investigation.

Figure 13 shows the geographic distribution of Bright Spot counties. One notable result is the degree to which the Bright Spot counties are distributed throughout the Region. The other notable result is the clear geographic clustering of many of the Bright Spot counties.

Figure 13: Map of the Bright Spot Counties in Appalachia



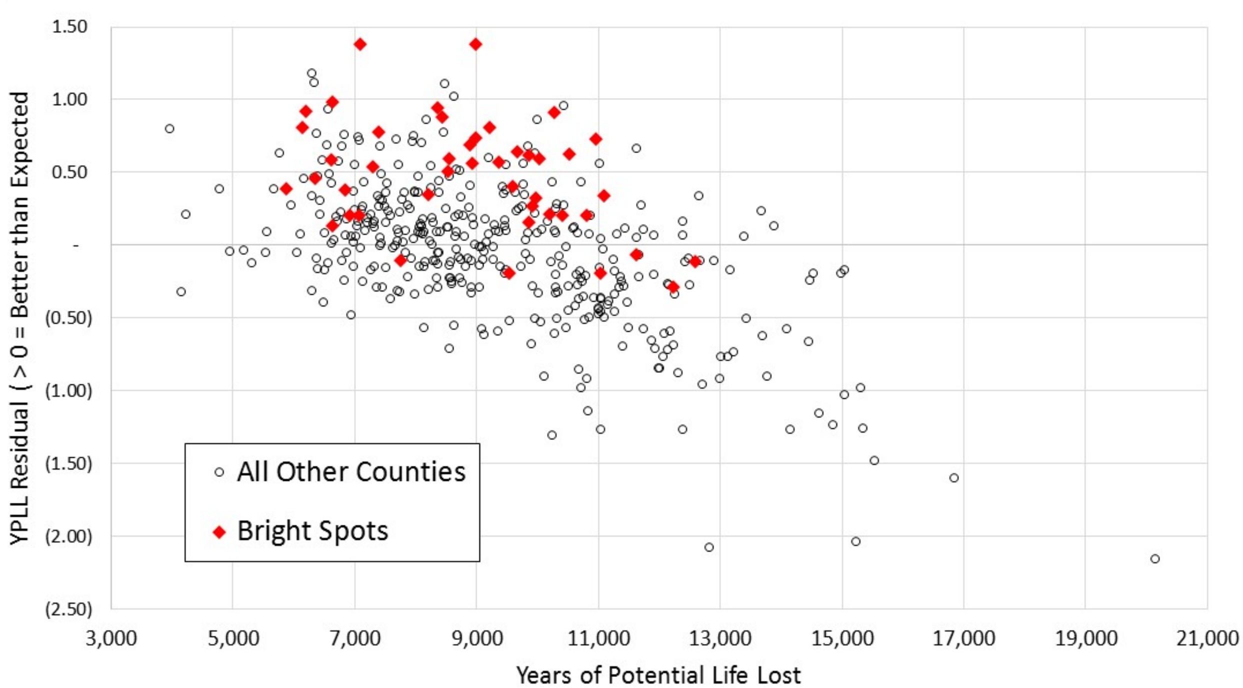
What Makes a Bright Spot "Bright?"

The Bright Spots predictive model controlled for the 29 driver measures to the extent possible. However, the model was limited in that it excluded other factors that drive health outcomes. For example, the model did not include preventive oral health care; coordination and effectiveness of local health systems; strong local health department funding; or aspects of local tradition or culture that are difficult to measure. These other factors may hold answers to questions about *why* a given county's health outcomes were better than expected.

Nonetheless, the relationship of outcome residuals to each other provided some insight into what makes a Bright Spot "bright." The model used 19 outcome measures to identify the Bright Spots and gave each measure the same weight in the average residual score. For example, cancer mortality carried the same weight as Medicare Part D opioid prescription claims. One way to understand the model is to evaluate how the Bright Spots, the top decile overall, performed on the individual outcomes.

Figure 14 shows the residuals for the Bright Spots compared with all other counties for the premature mortality measure, YPLL. The chart shows that the Bright Spot counties do not have uniformly better-than-expected residuals for YPLL. Instead, the Bright Spot counties performed better *as a group* than all other Appalachian counties on the YPLL measure. Figure 14 shows that some Bright Spot counties had slightly worse-than-expected YPLL, but most had better-than-expected YPLL (i.e., the solid red diamonds on the chart are better, on average, than the open circles).

Figure 14: Observed YPLL Versus YPLL Residual for All Appalachian Counties



For the period 2011–2013, YPLL in most Appalachian counties was worse than the national average (8,291 per 100,000 compared to 6,658 per 100,000). However, regardless of the observed YPLL values for the Bright Spots, the YPLL *residuals*—that is, the degree to which observed YPLL deviated from the expected—were, on average, better than the YPLL residuals for other counties.

Outcome measures that had a strong residual correlation with identification as a Bright Spot were:

- Premature mortality (YPLL);
- Unintentional injury mortality; and
- Poisoning mortality.

The individual outcome residuals that track closely with identification as a Bright Spot may be more likely to influence overall health. To illustrate this, Table 15 contains the correlations between the residuals of each individual outcome and the overall average residual. A correlation coefficient of 1.0 represents perfect correlation and 0 represents no correlation. None of the correlation coefficients were close to 1.0, but for the three outcomes listed above, the coefficients were over 0.5 in both the metropolitan and nonmetropolitan groups. This implies that the counties that tended to do well on one of these measures also tended to fare better than expected on many other health outcome measures. It's possible that if a community has a culture that not only discourages drug use but also dedicates resources to recovery and rehabilitation efforts, that same culture may also impact health outcomes in a positive way.

If a specific community has a culture that not only discourages substance abuse, but also dedicates resources to recovery and rehabilitation, that same culture may also impact other health outcomes in a positive way.

Table 15: Correlation between Average Residual and Individual Residuals for Each Outcome

Outcome	Correlation coefficient	
	Metro	Nonmetro
YPLL	**0.63**	**0.71**
Injury mortality	**0.59**	**0.65**
Poisoning mortality	**0.59**	**0.58**
COPD mortality	0.49	0.42
Average HCC risk score[a]	0.49	0.40
Heart disease mortality	0.45	0.35
Suicide mortality	0.44	0.28
Depression prevalence	0.43	0.32
Heart disease hospitalizations	0.38	0.41
Cancer mortality	0.37	0.42
Stroke mortality	0.37	0.36
% opioid Medicare Part D Rx claims	0.35	0.30
% adults with obesity	0.34	0.17
Infant mortality	0.26	0.21
Physically unhealthy days	0.24	0.13
Diabetes prevalence	0.24	0.17
Mentally unhealthy days	0.23	0.22
% births with low birth weight	0.18	0.24
% drinking excessively	0.12	0.11

[a] *Medicare risk score of the complexity of the disease among hospitalized patients: hierarchical condition categories (HCC).*
[b] *Bold text indicates a correlation greater than 0.5.*

Table 15 is sorted by the nonmetropolitan correlations. For nonmetropolitan counties, the top four correlations are mortality measures. The fifth highest is the average Medicare hierarchical condition categories (HCC) risk score measure. This suggests that HCC, a measure of a Medicare patient's risk of a complex high-cost disease not typically measured as an outcome in public health studies, is an important measure to track when studying counties and communities that have better-than-expected health.

Distribution of Bright Spot Counties among Appalachian States

The Bright Spots are not distributed evenly among the Appalachian states—Kentucky and Mississippi have proportionately more Bright Spot counties than other states.

On the other hand, the model did not identify any Bright Spot counties in Ohio, a state with 32 Appalachian counties. The other two states with no identified Bright Spot counties, South Carolina and Maryland, have only a few Appalachian counties: six and three, respectively. The absence of Bright Spots in these two states may be the result of small sample sizes, whereas the Ohio result suggests a pattern of lower-than-expected outcomes.

Several Bright Spot counties appear in geographic clusters, suggesting that factors leading to better-than-expected health may prevail across broad, multicounty areas. Clustering suggests the presence of some common factor that has improved the health of the cluster. The unit of analysis, the county, may be a proxy for a larger "community." These communities may be in the service area of a particularly effective program, health care provider, or other resource. Alternatively, other factors, such as environment, local culture, and tradition, may also support a culture of health.

Table 16: Appalachian States Ranked by Percentage of Bright Spots

State	Appalachian counties				
	State	Bright Spots			
		Total	Metro	Non-metro	Total
Mississippi	24	4	1	3	17%
Kentucky	54	9	0	9	17%
Virginia	25	4	1	3	16%
West Virginia	55	8	2	6	15%
New York	14	2	2	0	14%
Tennessee	52	7	3	4	13%
North Carolina	29	3	2	1	10%
Georgia	37	2	2	0	5%
Pennsylvania	52	2	1	1	4%
Alabama	37	1	1	0	3%
Maryland	3	0	0	0	0%
Ohio	32	0	0	0	0%
South Carolina	6	0	0	0	0%
TOTAL	420	42			10%
Total metro	152		15		
Total nonmetro	268			27	

Bright Spots Relative to Other Appalachian Counties

Drivers

On average, Bright Spot counties tended to have slightly lower income and slightly less favorable health drivers than other Appalachian counties. However, for all drivers, the range of raw values for the Bright Spots overlapped substantially with the range for the rest of the Region.

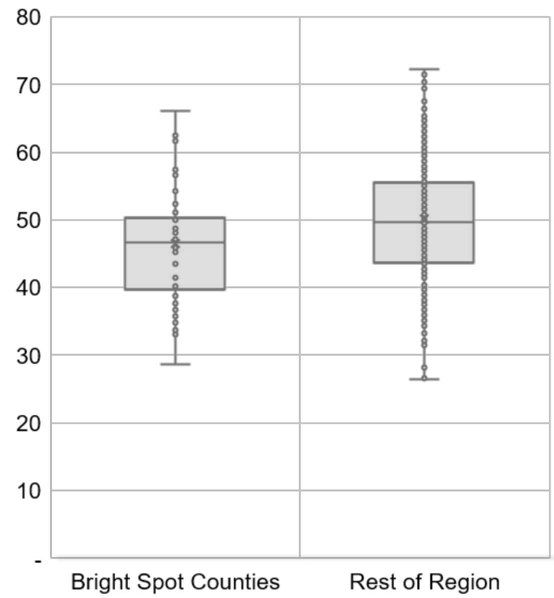
Box plots in Appendix D compare the actual health driver values for Bright Spot counties and the counties in the rest of the Region. The box plots confirm that, in this model, a Bright Spot is not necessarily extraordinary *because* of its quantifiable health drivers. Rather, a Bright Spot is extraordinary *in spite of* the values of its drivers.

The box plots show the ranges of values for Bright Spots and the rest of the Region, excluding outliers. Each box shows the middle 50 percent of values; the line is the median value (see the Introduction in the *Disparities* report for how to read a box plot) (Marshall, et al., 2017).

As examples, Figure 15 and Figure 16 show box plots for two of the drivers: the percentage of adults with some college education, and the percentage of adults who smoke. These plots illustrate the trend for most drivers. The median Bright Spot county is slightly less advantaged, and the ranges for Bright Spots and the rest of the Region overlap substantially.

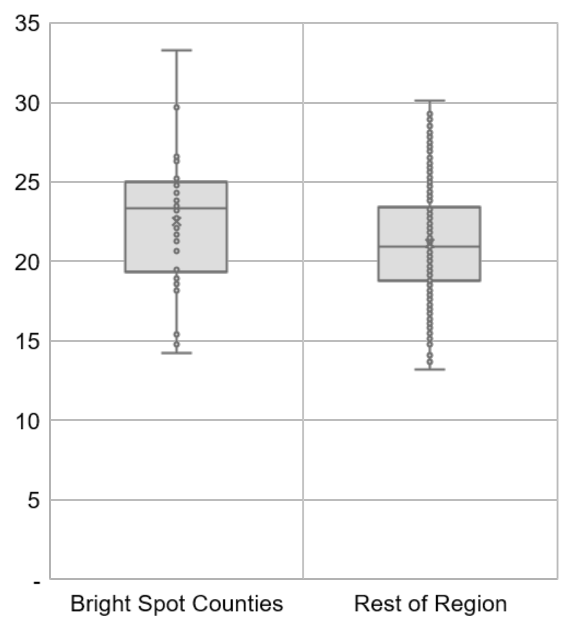
> *A Bright Spot is not necessarily extraordinary because of its quantifiable health drivers. Rather, a Bright Spot is extraordinary in spite of the values of its drivers.*

Figure 15: Percentage of Adults with Some College Education



Outliers excluded.[4]

Figure 16: Percentage of Adults that Smoke



Outliers excluded.

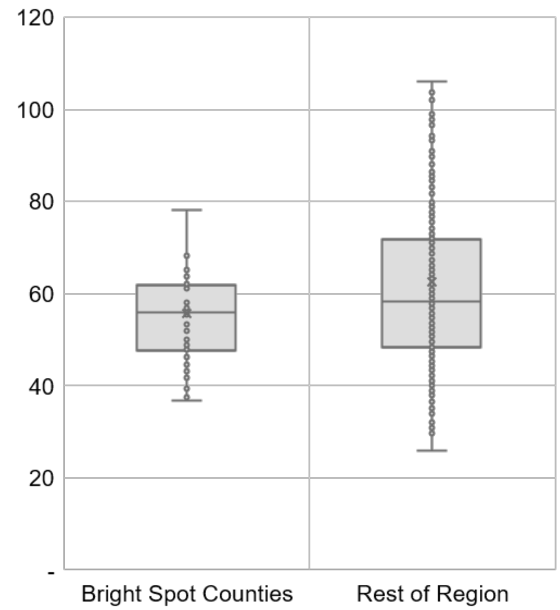
[4] In the box plots here, "Outliers excluded" indicates that, for simplicity, unusually high or low values are not shown on the graph. See page 31 of the *Disparities* report for more detail.

Outcomes

As expected, with regard to outcomes, Bright Spot counties performed better than the rest of the Region. This was true both for absolute values and values relative to expectations. By definition, Bright Spots have more favorable health outcomes than expected; therefore, on an absolute basis, we should expect the group to outperform the rest of the Region by at least a small margin. Ranges for Bright Spot outcomes were also tighter.

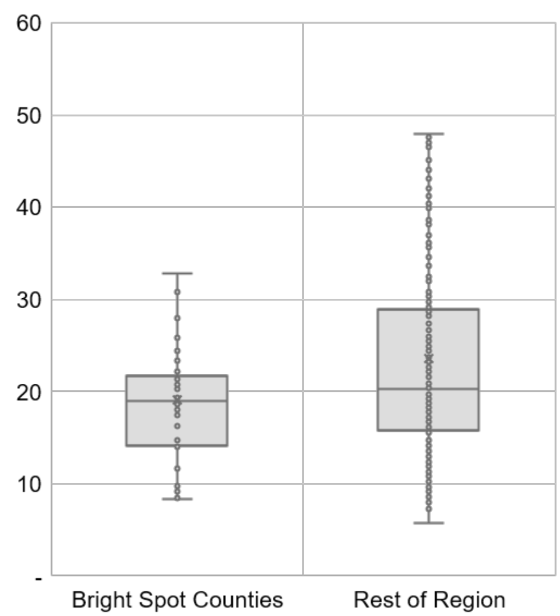
Figure 17 and Figure 18 show how the Bright Spots compared with the rest of Appalachia on two outcomes related to substance abuse: unintentional injury mortality and poisoning mortality. These figures contain the absolute input data; therefore, low values represent better health. While the Bright Spot counties had lower injury and poisoning mortality rates on average than the rest of the Region, the plots show that the ranges between the two groups overlapped. Appendix D contains the plots for all of the health outcomes.

Figure 17: Unintentional Injury Mortality per 100,000 Population



Outliers excluded

Figure 18: Poisoning Mortality per 100,000 Population



Outliers excluded

OBJECTIVE 2: MATCH BRIGHT SPOT COUNTIES TO OTHER COUNTIES

To facilitate the transference of lessons learned from the Bright Spot communities, we matched each Appalachian county with a Bright Spot county by using Euclidean distance analysis to create a score to assess similarity. This technique allowed us to gauge each county's similarity to the Bright Spot counties on the basis of resources and community characteristics. For example, Northumberland County, Pennsylvania, which was not one of the 42 Bright Spots, closely matched Potter County, Pennsylvania, a county identified as a Bright Spot. These counties have very similar economic indicators, such as median income; similar health behaviors, such as the percentage of adults who smoke; and other similar social determinants, such as the percentage of residents receiving disability benefits. Table 13 and Table 14 in Appendix C contain the data, including the closest Bright Spot match for each county in Appalachia as well as the closest of the Bright Spot case study matches for each county. The smaller the score, the closer the match.

Table 17: Top Ten Matches between Bright Spot Counties and Other Appalachian Counties

County	State	Best Bright Spot match	Score
Northumberland	Pennsylvania	Potter, Pennsylvania	9.43
Marion	Tennessee	Jefferson, Tennessee	9.84
Warren	Tennessee	Monroe, Tennessee	9.96
Elliott	Kentucky	Lewis, Kentucky	10.13
Claiborne	Tennessee	Roane, West Virginia	10.31
Jackson	West Virginia	Hardy, West Virginia	10.35
Smith	Tennessee	Sequatchie, Tennessee	10.67
Hart	Kentucky	Monroe, Tennessee	10.76
Lawrence	Kentucky	Roane, West Virginia	10.98
Prentiss	Mississippi	Choctaw, Mississippi	11.15

OBJECTIVE 3: IDENTIFY HEALTH DRIVERS ASSOCIATED WITH POSITIVE OUTCOMES

To examine the relationship between drivers and outcomes, we ran a series of univariate regression analyses, repeating the general model approach (stratification, propensity weighting); however, instead of using all 29 drivers in the same model, we estimated the effects of one driver on all the outcomes jointly. As outlined in the Study Design chapter, this was accomplished by a "stacked" regression with full interaction between the driver and indicators for the specific outcomes, clustering on county. We calculated the following three summary measures for each model:

- the partial R-square resulting from the addition of the driver to the model;
- an F-statistic testing whether the coefficient was zero for all 19 outcomes; and,
- the number of outcomes for which the driver was positively or negatively statistically significant.

Table 18 below contains key results. Tables 9 and 10 in Appendix C, contain all output from this univariate analysis. For ease of reading Table 18, partial R-square values above 0.30 are shaded green—identifying the drivers that predict the most variation in the outcomes. In addition, F-statistics above 40 are shaded green. These high F-statistics indicate that the driver has a strong statistical relationship between the driver and at least one of the outcomes. With only two exceptions—the e-prescription rate in nonmetropolitan counties and grocery store availability in metropolitan counties (marked by daggers)—there was strong evidence ($p < 0.001$) that the driver was associated with at least one outcome.

Drivers explaining the largest average proportion of variation across the outcomes are

- *percentage of adults that smoke*
- *percentage adults physically inactive*
- *percentage receiving disability*
- *teen birth rate*
- *median income*
- *ARC Economic Index*
- *poverty rate*

Seven drivers had the most significant impact on the collective set of outcomes. The highest partial R-squares across the outcomes were for the percentage of adults that smoke, the percentage of adults that were physically inactive, the percentage receiving disability, the teen birth rate, median income, the value of ARC's economic index, and the poverty rate. Physical inactivity and the disability rate were the best predictors for metropolitan areas, while the disability rate and the percentage of adults who smoke were the best predictors for nonmetropolitan counties.

The number of outcomes to which drivers significantly correlated in a single direction are also shaded green if the number was 16 or more. Outcomes were modeled as "good health," so "+" means increases in the driver are associated with better health. For example, in nonmetropolitan counties, more social associations lead to better health in 17 of the 19 outcomes. A higher percentage of adults who smoke leads to worse health ("-" indicates that a higher smoking rate decreases the rate of positive health outcomes) for 18 of the 19 outcomes. The relationships between individual outcomes and drivers in this model were generally as expected. Unsurprisingly, the economic conditions in the community matter—in both metropolitan and nonmetropolitan counties, lower median income, higher values of the ARC Economic Index, and more households below the poverty line were associated with worse health outcomes. Although the specific metropolitan and nonmetropolitan results differ slightly (e.g., a driver may be statistically significant in 15 outcomes for one group and 16 in the other), they generally track together. The most notable difference is the social association rate; higher values of this variable are highly predictive in nonmetropolitan areas but not in metropolitan areas. Other notable differences are for housing, student-teacher ratios, air pollution, the supply of mental health providers, and travel time to work.

The teen birth rate in the county emerged as a key driver of community health across most of the outcomes. Teen pregnancy serves as a marker for economic opportunity in the community, captures “risky behavior” among teenagers, including unprotected sex, which is often associated with substance use (Salas-Wright, Vaughn, Ugalde, & Todic, 2015), and can have long-lasting effects on young parents. The teen birth rate serves as a marker for life course outcomes: daughters of teenage mothers are more likely to become teenage mothers themselves (Albert, 2002). Policies aimed at limiting teen birth rates, such as comprehensive, medically accurate sex education courses and other public health interventions may have long-lasting effects.

Numerous studies show that socioeconomic conditions are some of the strongest predictors of community health. This is true at both the individual and community level—lower-resourced individuals are generally less healthy and areas with worse economic conditions are less healthy (Catlin & Williams Van Dijk, 2015).

Like changing health behavior, improving the economic conditions of a community takes time. As demonstrated in the *Health Disparities in Appalachia* report (Marshall, et al., 2017), measured health status in the Region improved over the last two decades. Although the Region has not yet reached parity with the nation, results of this study suggest that aggregate health improvement will accompany economic improvement.

Table 18: Univariate Regression Results

Driver	Nonmetropolitan				Metropolitan			
	Part. R2[5]	F Stat[6]	Outcomes significant at $p < 5\%$[7]		Part. R2	F stat	Outcomes significant at $p < 5\%$	
Social association rate	0.09	15.5	17+	1−	0.01	3.9	1+	2−
Percentage in social assistance jobs	0.01	2.8	2+	1−	0.03	6.6	10+	3−
Income inequality	0.08	11.1	2+	15−	0.03	16.2	2+	8−
Percentage enrolled in SNAP	0.06	31.3	3+	15−	0.05	21.2	1+	14−
Grocery stores per 1,000	0.02	2.6	1+	9−	0.01	1.4†	0+	5−
Restaurants per 1,000	0.08	13.7	16+	1−	0.08	7.2	15+	1−
Percentage with no car and low access to grocery stores	0.02	6.1	8+	1−	0.04	5.9	12+	2−
Access to exercise	0.03	6.5	13+	1−	0.10	18.2	15+	2−
Percentage spending >30% income on housing	0.01	8.3	5+	1−	0.04	7.3	13+	1−
Physicians who e-prescribe	0.01	1.4†	9+	1−	0.02	4.6	7+	1−
Percentage of adults who smoke	0.22	177.8	1+	18−	0.23	75.6	1+	18−
Percentage of adults physically inactive	0.14	52.2	1+	18−	0.25	54.0	1+	18−
Chlamydia incidence	0.05	23.8	6+	9−	0.04	14.8	2+	10−
Diabetes HbA1c testing	0.01	3.2	11+	2−	0.03	4.9	13+	1−
Breast cancer screening	0.11	15.2	17+	1−	0.07	8.0	15+	1−
Percentage receiving disability	0.24	26.3	1+	18−	0.25	48.0	1+	18−
Teen birth rate	0.17	37.1	1+	18−	0.22	50.0	1+	17−
Student-teacher ratio	0.02	2.9	1+	10−	0.01	4.8	1+	1−
Percentage with some college	0.06	9.4	16+	1−	0.15	23.7	15+	1−
Average daily air pollution	0.02	18.5	3+	9−	0.03	10.1	2+	14−
Supply of primary care physicians	0.01	2.9	10+	1−	0.03	13.1	12+	3−
Supply of dentists	0.01	2.8	11+	1−	0.05	15.0	12+	3−
Supply of specialty physicians	0.01	4.3	11+	1−	0.02	20.6	8+	3−
Supply of mental health providers	0.01	7.0	1+	0	0.04	6.4	13+	3−
Percentage households below poverty	0.17	79.1	1+	17−	0.14	30.5	1+	16−
ARC Economic Index	0.16	49.7	1+	17−	0.17	31.0	1+	18−
Median income	0.19	89.3	18+	1−	0.18	32.0	18+	1−
Average travel time to work	0.02	4.2	1+	12−	0.02	12.7	4+	6−
Percentage uninsured under 65	0.03	15.5	4+	12−	0.05	12.8	2+	12−

[5] R-square above 0.15 predicts most variation.

[6] F above 40 indicates strongest evidence there is a relationship between driver and at least one outcome; †Not statistically significant at $p < 0.001$.

[7] Significance: tabulation of whether the variable is statistically significant at 5% p (+ = positive and significant; − = negative and significant); green shaded cells have over 15 statistically significant outcome relationships in either direction.

It is also instructive to examine how much variation some of the drivers explained for the outcomes for which we expected little association. For example, the percentage of persons with diabetes receiving appropriate HbA1c testing was statistically significant for a majority of the outcomes, although the hypothesized relationship between HbA1c testing and, for example, injury mortality, is weak at best. It is more likely that the *drivers* also correlate; for example, counties with better HbA1c testing likely had higher quality of care in other domains. This underscores the use of the approach used for prediction; the results for the individual drivers are less important than the overall predictive power.

Opportunity to Live Healthy

Other research shows that positive health behaviors consistently have large, statistically significant relationships to good health outcomes (National Institutes of Health, 2015). Results in this report support and amplify this finding. In this study, the drivers that described behaviors, such as the percentage of adults who smoke, the percentage of adults who are physically inactive, and the teen birth rate, were more highly correlated with good health outcomes than drivers quantifying the supply of health resources. Our findings suggest that traditional public health initiatives should accompany efforts to develop community health infrastructure. For example, funding for community health workers trained to communicate chronic disease prevention behaviors might reach deeper into community values and have a greater impact on population health than the supply of additional providers alone.

Overall, this study supports an emerging body of literature that attests to the association between positive population health outcomes and a community's social, economic, and environmental factors.

Field Study Site Selection

Criteria for Field Study Site Selection

Better-than-Expected Outcomes in Case Study Counties

Matching Case Study Sites to All Other Counties

The next step in this research project involves qualitative case studies of ten Bright Spot counties to explore firsthand how the practices and activities in these counties may support better-than-expected health outcomes.

CRITERIA FOR FIELD STUDY SITE SELECTION

Three criteria determined whether a county was eligible for case study selection. The Bright Spot statistical analysis identified the prospective counties. We wanted to ensure that the case studies represented the diversity of Appalachian communities, so we selected an equal number of metropolitan and nonmetropolitan communities. Finally, we selected at least one case study from each of the five Appalachian subregions to maximize representativeness.

Top Decile of Average Standardized Residual Scores

The statistical model identified counties that—on average across the 19 outcomes—scored better than expected. Counties that had average residual scores in the top decile in either the metro or nonmetro group were classified as Bright Spots. The average standardized residual scores ranged from +0.72 to −1.09, where 1.0 represents one standard deviation from expected.

Metropolitan and Nonmetropolitan Groups

The design of the Bright Spots model controlled for the impact of urban resource advantages by separating metropolitan and nonmetropolitan counties. Hence, for balance, the case studies include an equal number of cases from the metropolitan and nonmetropolitan groups. The differences between urban and rural life mean that the estimated relationship between drivers and outcomes may differ between those groups. For example, the results of Table 18 show that the social association rate has a stronger relationship with outcomes in nonmetropolitan areas compared to metropolitan areas. Likewise, as outlined in the Results chapter, county size affects the precision of the model, so lower-populated counties may have less precise results. Since population correlates with metropolitan status, stratifying on metropolitan status may be more informative.

Appalachian Subregions

Representativeness across geography is also critical. Appalachia is a diverse region, with different resources, cultures, policies, and environments across its 420 counties. Thus, we selected counties for the case studies with an eye to ensure representation from all subregions. Selecting two counties from each Appalachian subregion provides, technically, the most representative Appalachian view of contributing factors. Selecting without an eye to this—i.e., picking the five “Brightest” spots in each strata—would eliminate subregions, which may have caused us to overlook positive behaviors or activities that might be specific to that subregion. Thus, we ensured representativeness across the Region by selecting one case study from each rurality-subregion strata.

Table 19 contains the list of the selected case study counties. Although the 42 Bright Spot counties tended to cluster geographically, only two of the ten case study counties—McCreary and Wayne, Kentucky—are contiguous.

The selected counties represent three of ARC's five economic status designations: Distressed, At-Risk, and Transitional. The populations in the case study counties range from 5,810 residents to 50,464 residents.

Table 19: Selected Case Study Sites

County	State	Subregion	Metro / Nonmetro	Average Standardized Residual Score	2014 Population	Economic Status [a]
Wirt	WV	North Central	Metro	0.47	5,810	At-Risk
Hale	AL	Southern	Metro	0.35	15,393	Distressed
Sequatchie	TN	South Central	Metro	0.31	14,431	Transitional
Tioga	NY	Northern	Metro	0.29	50,464	Transitional
Madison	NC	South Central	Metro	0.27	20,951	At-Risk
Wayne	KY	Central	Nonmetro	0.72	20,728	Distressed
Noxubee	MS	Southern	Nonmetro	0.58	11,240	Distressed
Grant	WV	North Central	Nonmetro	0.49	11,829	Transitional
McCreary	KY	Central	Nonmetro	0.45	18,073	Distressed
Potter	PA	Northern	Nonmetro	0.45	17,451	Transitional

Sources: See Table 1 in Appendix B and Table 3 and 5 in Appendix C

a. Economic status in fiscal year 2017

BETTER-THAN-EXPECTED OUTCOMES IN CASE STUDY COUNTIES

Outcome scores reflect the diversity in the case study counties. All were high performers on most of the 19 outcome measures, indicated by the light shaded cells in Table 20. The last column in the table shows that, on three measures—premature mortality (YPLL), injury death, and depression—all ten of the case study county outcomes were better than expected. Although all ten had better-than-expected outcomes in at least 12 of the 19 measures, Wayne, Noxubee, and Hale had the highest number of better-than-expected outcomes, as illustrated in the last row of the table.

Table 20: Better-than-Expected Outcomes in Case Study Counties

Category	Indicator	Case Study County										Total Counties Better-than-Expected per Measure *(Max = 10)*
		Wayne	Noxubee	Hale	Wirt	Sequatchie	Tioga	McCreary	Potter	Madison	Grant	
Mortality	YPLL	■	■	■	■	■	■	■	■	■	■	10
	Stroke	■	■	■	■	■	■	■	■		■	9
	Cancer	■	■	■			■	■		■	■	7
	Injury	■	■	■	■	■	■	■	■	■	■	10
	COPD	■	■	■	■	■	■	■	■		■	9
	Heart disease	■	■	■	■	■	■	■	■	■		9
Mental Health	Mentally unhealthy days	■	■	■		■				■		5
	Suicide mortality	■	■	■	■	■	■		■		■	8
	Depression	■	■	■	■	■	■	■	■	■	■	10
Child Health	% low birthweight	■	■		■		■	■	■	■		7
	Infant mortality	■	■	■	■		■	■	■	■		8
Chronic Disease	Diabetes	■					■		■	■		4
	Heart disease hospitalization	■	■	■	■	■		■	■	■	■	9
	Medicare HCC	■	■	■	■		■	■	■	■	■	9
	Obesity			■		■		■		■		4
	Physically unhealthy days		■			■			■	■		4
Substance Abuse	Percentage excessive drinking	■		■	■	■	■	■			■	7
	Poisoning mortality	■	■	■	■	■		■	■		■	8
	Medicare opioid Rx		■		■	■	■				■	5
Total Better-than-Expected Outcomes per County *(Max = 19)*		16	16	15	14	14	14	14	14	13	12	

Sources: See Tables 3 and 5 in Appendix C

MATCHING CASE STUDY SITES TO ALL OTHER COUNTIES

Table 13 in Appendix C contains the results of the Euclidean distance measurement, and provides the scores that match the field study counties to all other Appalachian counties. This table includes the closest match to the ten case study sites for the other 410 counties in Appalachia. Though not incorporated into the field research, the match tables provide a reference that may be useful for follow-up and policy analysis in future studies. Likewise, it may be a useful resource for community leaders wishing to foster a culture of health in their community—their “match” is a peer that appears to be doing something worthy of emulation by others.

APPENDICES

A. Bibliography

B. County-Level Data and Data Sources

C. Results of Statistical Analyses

D. Box Plot Comparisons: Bright Spots versus Rest of Region

E. Methodological and Technical Notes

A. BIBLIOGRAPHY

Albert, B. (2002). *With One Voice 2002: America's Adults and Teens Sound Off About Teen Pregnancy.* Washington, D.C.: The National Campaign to Prevent Teen and Unplanned Pregnancy.

Bell, M. J., Mora, M. G., Hagan, M. P., Rubin, M. P., & Karpyn, P. A. (2013). *Access to Healthy Food and Why it Matters: A Review of the Research.* The Food Trust. Oakland, CA: PolicyLink. Retrieved from http://thefoodtrust.org/uploads/media_items/access-to-healthy-food.original.pdf

Best, W., & Cowper, D. (1994, April). The ratio of observed to expected mortality as a quality of care indicator in non-surgical VA patients. *Medical Care, 32(4)*, 390-400.

Bradley, E., & Taylor, L. (2016, August 17). *How Social Spending Affects Health Outcomes.* Retrieved 2017, from Robert Wood Johnson Foundation Culture of Health Blog: http://www.rwjf.org/en/culture-of-health/2016/08/how_social_spending.html

Catlin, B. J., & Williams Van Dijk, J. (2015). *2015 County Health Rankings Key Findings Report.* University of Wisconson , Population Health Institute. Madison, Wis: County Health Rankings & Roadmaps.

Center for Medicare and Medicaid Services. (2016, December 19). *Hospital Compare*. Retrieved from Medicare.gov: https://www.medicare.gov/hospitalcompare/search.html?

County Health Rankings. (2016). *Access to exercise opportunities.* Retrieved from http://www.countyhealthrankings.org/app/west-virginia/2018/measure/factors/132/map

Feudtner, C., Berry, J. G., Parry, G., Hain, P., Morse, R. B., Slonim, A. D., . . . Hall, M. (2011). Statistical Uncertainty of Mortality Rates and Rankings for Children's Hospitals. *Pediatrics*, e966-e972.

Gan, W., FitzGerald, J., Carlsten, C., Sadatsafavi, M., & Brauer, M. (2013). Associations of Ambient Air Pollution with Chronic Obstructive Pulmonary Disease Hospitalization and Mortality. *American Journal of Respiratory and Critical Care Medicine, 187*(7), 721-727.

Holmes, G. M., & Fraher, E. (In press). Developing physician migration estimates for workforce models . *Health Services Research*.

Holmes, G. M., & Ricketts, T. C. (2007, December). Mortality and Physician Supply: Does Region Hold the Key to the Paradox? *Health Services Research*, 2233-2251. doi:10.1111/j.1475-6773.2007.00728.x

IHI. (2011). *90-day R&D Project Final Summary: Counties of Interest: Achieving Better or Worse Than Expected Health Outcomes.* Cambridge, MA: Institute for Healthcare Improvement.

Klaiman, T., Pantazis, A., Chainani, A., & Bekemeier, B. (2016). Using a positive deviance framework to identify Local Health Departments in Communities with exceptional maternal and child health outcomes: a cross sectional study. *BMC Public Health*, 602-610.

Kreider, B., Pepper, J., & Gunderson, C. a. (2012). Identifying the Effics of SNAP (Food Stamps) on Child Health Outcomes When Participation is Endogenous and Misreported. *Journal of the American Statistical Association*, 958-975. Retrieved from http://batten.virginia.edu/sites/default/files/research/attachments/JASA_KPGJ_online(1).pdf

Lane, N. M., Holmes, G. M., Arcury, T. A., Schwalbe, M. L., Randolph, R., Frank, J., . . . Keyser, A. (2017). *Exploring Bright Spots in Appalachian Health: Case Studies.* Washington, D.C.: Appalachian Regional Commission.

Maqbool, N., Viverios, J., & Ault, M. (2015). *The Impacts of Affordable Housing on Health: A Research Summary.* National Housing Conference, Housing Policy Research. Washington, D.C.: Center for Housing Policy.

Marshall, J. L., Thomas, L., Lane, N. M., Holmes, G. M., Arcury, T. A., Randolph, R., . . . Rodgers, J. (2017). *Creating a Culture of Health in Appalachia, Disparities and Bright Spots: Health Disparities in Appalachia.* Washington, D.C.: Appalachian Regional Commission.

Marty, L., Dubois, C., Gaubard, M., Maidon, A., Lesturgeon, A., Gaigi, H., & Darmon, N. (2015). Higher nutritional quality at no additional cost among low-income households: insight from food purchases of "positive deviants". *American Journal of Clinical Nutrition*, 1-9.

McLeroy, K., Bibeau, D., Steckler, A., & Glanz, K. (1988). An ecological persepctive on health promotion programs. *Health Educ, 15*(4), 351-337.

National Institutes of Health. (2015, August). Positive Emotions and Your Health , Developing a Brighter Outlook. *NIH News in Health*. Retrieved from https://newsinhealth.nih.gov/issue/aug2015/feature1

Papas, M., Alberg, A., Ewing, R., & Helzlsouer, K. (2007). The Built Environment and Obesity. *Epidemiologic Reviews*, 129-143.

Pascale, R., Sternin, J., & Sternin, M. (2010). *The Power of Positive Deviance.* Boston: Harvard Business Press.

Reynolds, C. (2016, November 30). *Appalachian Studies Final Project: Drug Abuse in Appalachia.* Retrieved from Blogs@VT Sites site: https://blogs.lt.vt.edu/appalachianiswhatshappenin/2016/11/30/final-project-drug-abuse-in-appalachia/

Robert Wood Johnson Foundation. (2017, March 15). *About a Culture of Health*. Retrieved from Building a Culture of Health: http://www.cultureofhealth.org/about/

Rosenbaum, D., & Rubin, P. (1983). The central role of the propensity score in observational studies for causal effects. *Biometrika, 70*(1), 41-55. Retrieved from https://oup.silverchair-cdn.com/oup/backfile/Content_public/Journal/biomet/70/1/10.1093/biomet/70.1.41/2/70-1-41.pdf?Expires=1488584930&Signature=LoZP4uTuVqlcq3GL5p9gWzzPdcx04rglNIsOm-w~HTDofm0JqS7wO1Anq08LM4c97FmZ1BkiOoBiW985mW1adxTdIBuvHZi6OQZE57eMJq3d5

Salas-Wright, C. P., Vaughn, M. G., Ugalde, H., & Todic, J. (2015). Substance Use and Teen Pregnancy in the United States: Evidence from the NSDUH 2002-2012. *Addict Behavior*, 218-225.

Topmiller, M. (2016, March). *Geospatial Research Brief*. Retrieved from healthlandscape.org: https://www.healthlandscape.org/Brigh_Spots_Approrpriate_Diabetes_Preventive_Care.pdf

Treuhaft, S., & Karpyn, A. (2010). *The Grocery Gap: Who has Access to Healthy Food and Why it Matters.* Oakland, CA: Policy Link, The Food Trust.

Zeiders, K., & Roosa, M. a. (2011, Marcy). Family Structure and Family Processes in Mexican American Families. *Fam Process, 50*(1), 77-91.

B. COUNTY-LEVEL DATA AND DATA SOURCES

Excel data files for this report were provided separately. This appendix includes two files:

1) County-Level Data
2) Data Sources

C. RESULTS OF STATISTICAL ANALYSES

Excel data files for this report were provided separately. This appendix includes three files:

1) Multivariate Regression Results
2) Univariate Regression Results
3) Euclidean Distance Analysis

D. BOX PLOT COMPARISONS: BRIGHT SPOTS VERSUS REST OF REGION

Driver Values Compared: Bright Spots and Remaining Appalachian Region Counties

Box plots below compare the 42 Bright Spot counties (left) with the remaining 378 Appalachian Region counties (right). Each plot compares county values from Appendix B covering the period 2007–2014. The boxes represent the middle 50 percent of values, 25th to 75th percentiles. The vertical lines show values below and above the 25th and 75th percentile, respectively. The brackets show the range of each data set.

Figure 1: Teenage Births per 1,000, 2007–2013

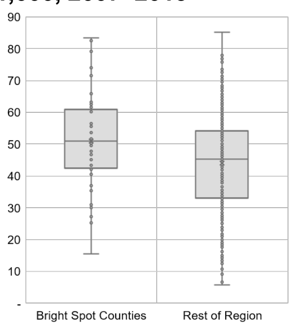


Figure 2: Full Service Restaurants per 1,000 pop, 2012

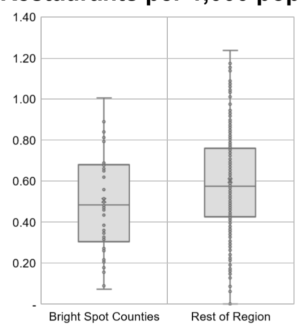


Figure 3: Access to Exercise Opportunities, 2014

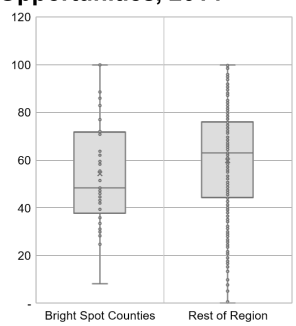


Figure 4: Avg. Daily Particulate Matter (Air Pollution), 2011

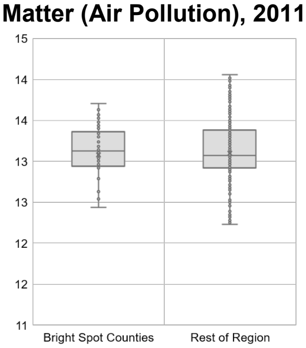


Figure 5: Grocery Stores per 1,000 pop., 2012

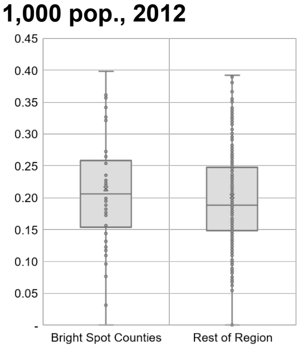


Figure 6: Student-Teacher Ratio, 2013–2014

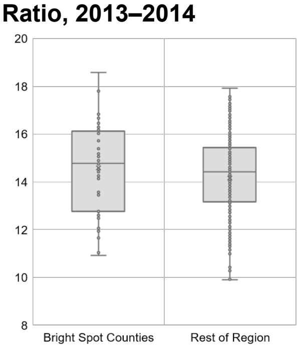


Figure 7: Avg. Travel Time to Work, 2010–2014

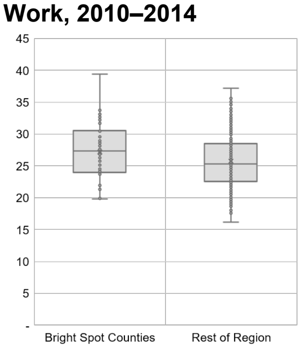


Figure 8: Percent of Adults Currently Smoking, 2014

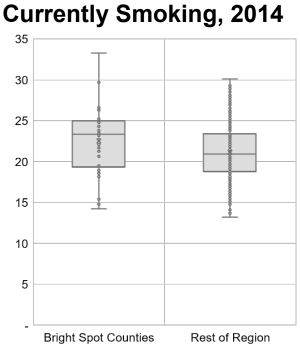


Figure 9: Percent of Adults Not Physically Active, 2012

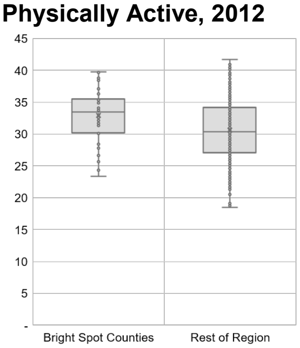


Figure 10: Chlamydia Incidence Rate per 100,000, 2013

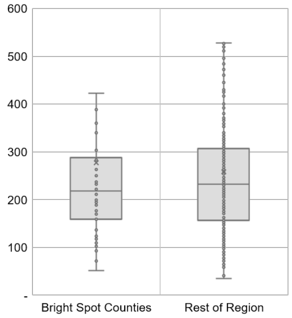


Figure 11: Primary Care Phys. per 100,000 pop., 2013

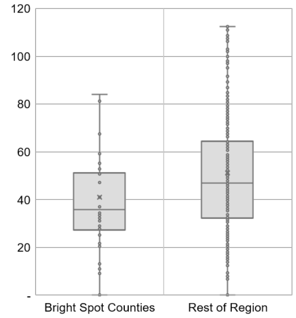


Figure 12: Dentists per 100,000 pop., 2014

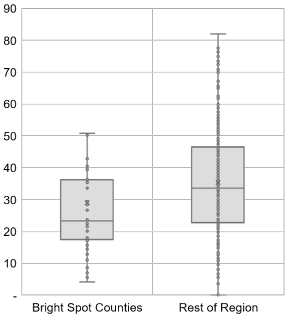


Figure 13: Specialist Physicians per 100,000 pop., 2013

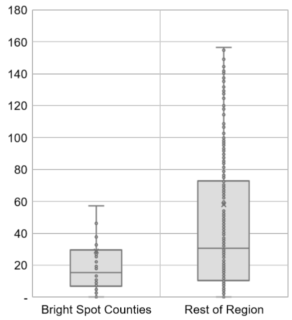


Figure 14: Mental Health Providers per 100,000 pop., 2015

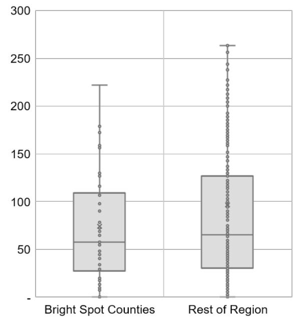


Figure 15: Percent of Doctors that Electronic Prescribe, 2014

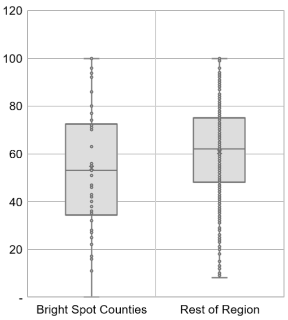


Figure 16: Rate of Persons Uninsured Under Age 65, 2013

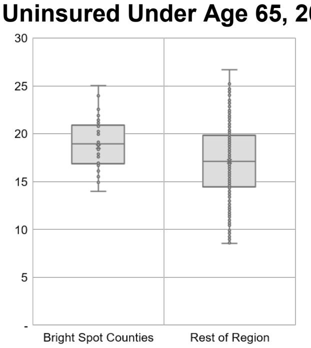


Figure 17: Percent of Persons w/ Diabetes Diagnosis w/ A1C, 2012

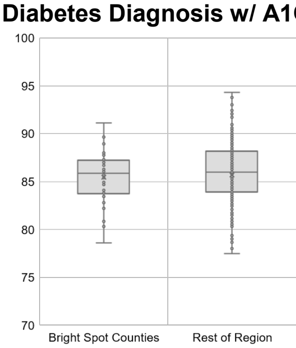


Figure 18: Pct. Medicare Women w/ Recent Mammogram, 2013

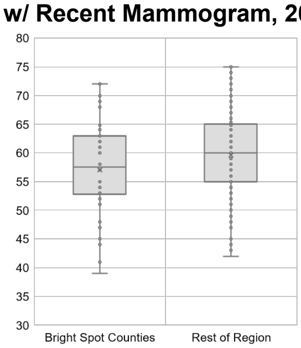


Figure 19: Percent Employed in Social Assistance Jobs, 2013

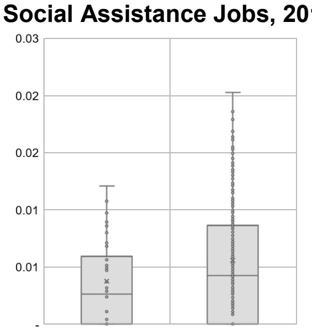


Figure 20: Income Inequality Ratio, 2010–2014

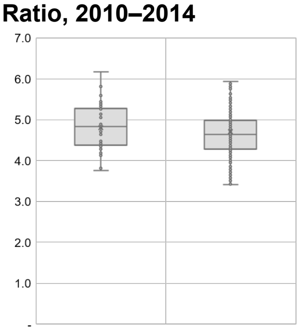


Figure 21: Percent of Eligibles Enrolled in SNAP Benefits, 2014

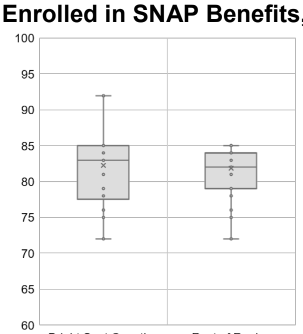


Figure 22: Percent Households w/ No Car/Low Access, 2010-2014

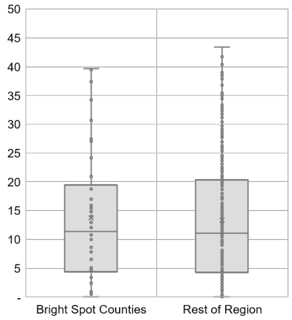


Figure 23: Pct. Spending >30% Income on Housing, 2010–2014

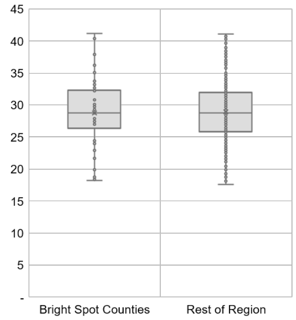


Figure 24: Economic Index, 2016

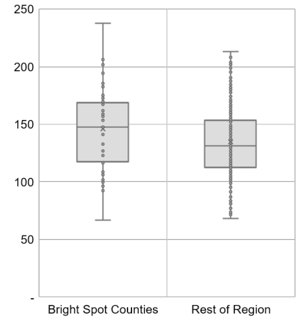


Figure 25: Social Association Rate, 2013

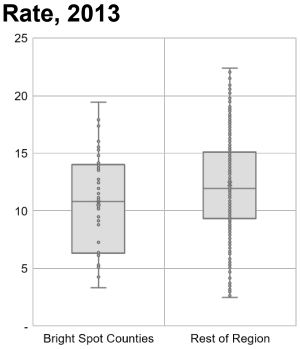


Figure 26: Percent Using Disability Benefits, 2014

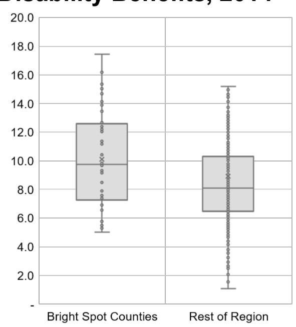


Figure 27: Percent Adults w/ at Least Some College, 2010–2014

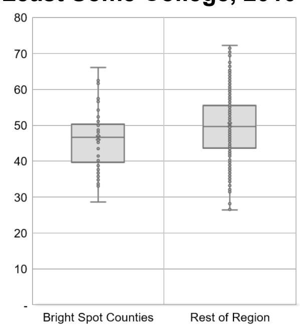


Figure 28: Percent Households with Income Below Poverty, 2014

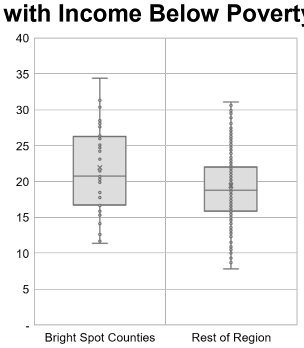
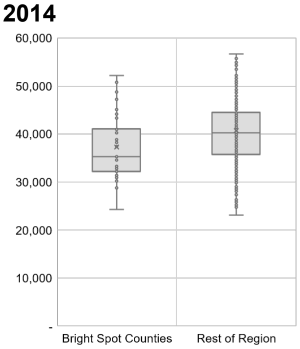


Figure 29: Median Income, 2010–2014



Outcome Values Compared: Bright Spots and Rest of Appalachian Region Counties

Box plots below compare the 42 Bright Spot counties (left) with the remaining 378 Appalachian Region counties (right). Each plot compares county values from Appendix B covering the period 2007–2014. The boxes represent the middle 50 percent of values, 25th to 75th percentiles. Mid lines are median values. Vertical lines show values below and above the 25th and 75th percentile, respectively. The brackets show the range of each data set.

Figure 1: Suicide Incidence per 100,000 pop., 2008–2014

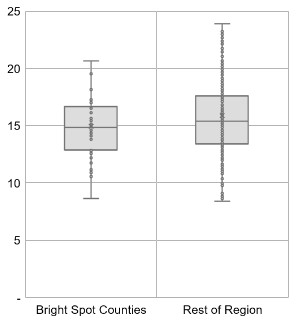


Figure 2: Medicare Beneficiaries Depression Rate, 2012

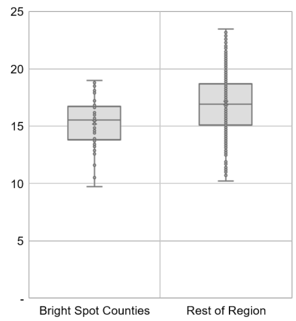


Figure 3: Percent of Residents Drinking Excessively, 2014

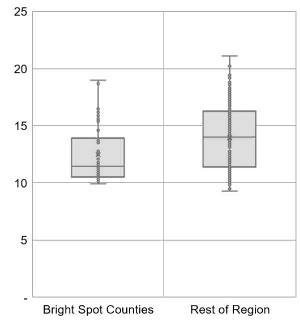


Figure 4: Poisoning Mortality per 100,000 pop., 2008–2014

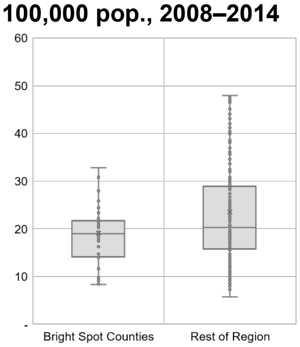


Figure 5: Opioid Prescription as a Pct. of Part D Claims, 2013

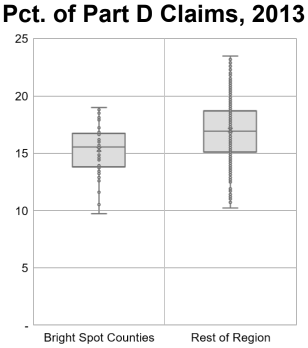


Figure 6: Heart Disease Mortality per 100,000 pop., 2008–2014

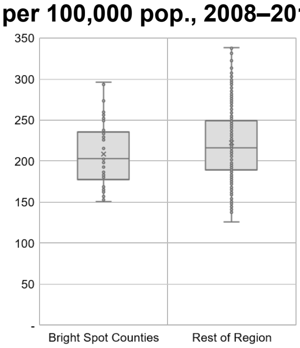


Figure 7: Years of Potential Life Lost per 100,000, 2011–2013

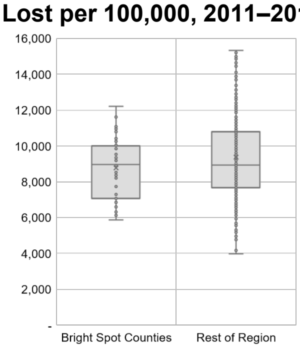


Figure 8: Cancer Mortality per 100,000 pop., 2008–2014

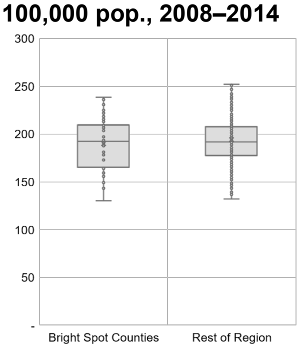


Figure 9: Injury Mortality per 100,000 pop., 2008–2004

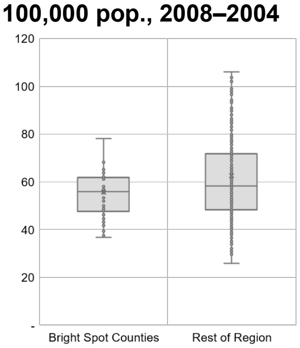


Figure 10: Stroke Mortality per 100,000 pop., 2008–2014

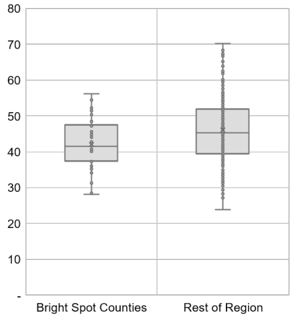


Figure 11: COPD Mortality per 100,000 pop., 2008–2014

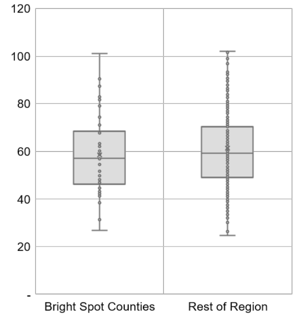


Figure 12: Physically Unhealthy Days / Month / Person, 2014

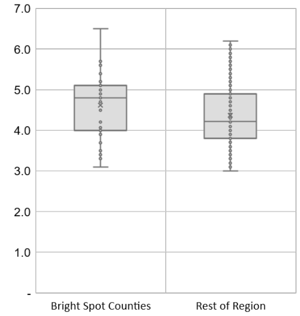


Figure 13: Mentally Unhealthy Days / Month / Person, 2014

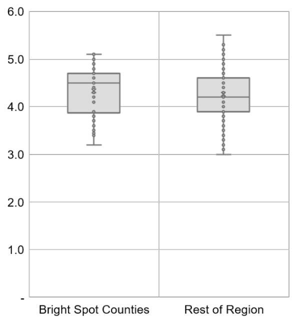


Figure 14: Average HCC Score per Medicare Beneficiary, 2013

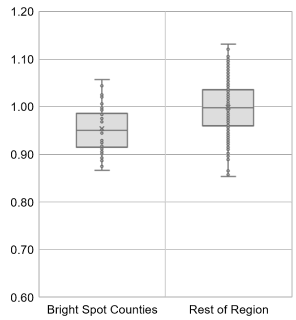


Figure 15: Percent of Adults with Diabetes, 2012

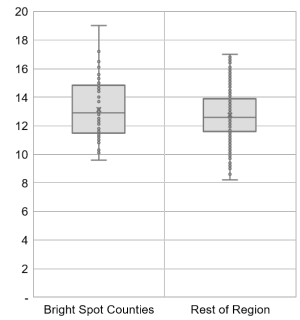


Figure 16: Percent of Adults with BMI >30, 2012

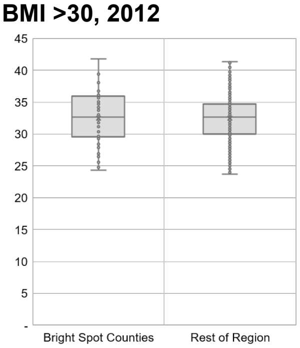


Figure 17: Heart Disease Hospitalizations

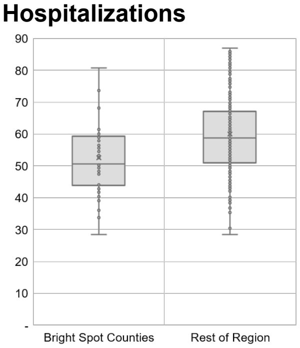


Figure 18: Low Birth Weight Births, >2,500g per 1,000 Births, 2007–2013

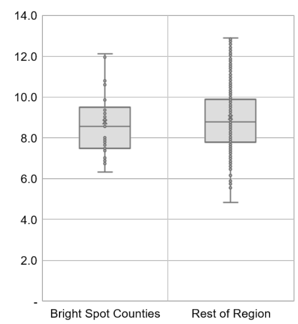
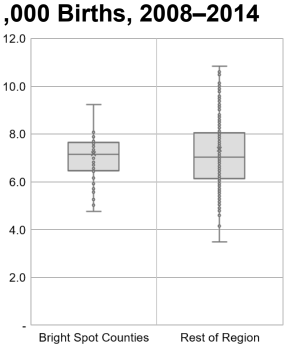


Figure 19: Infant Mortality per 1,000 Births, 2008–2014



E. METHODOLOGICAL AND TECHNICAL NOTES

Definition of a County for Purposes of Study

The data in this report come from more than 15 secondary data sources. Different institutions organize their data into a particular universe of spatial units based on their analytical needs. The county-level data files provided by these sources—or sub-county data aggregated from source files—were organized to meet the Appalachian Regional Commission's standard representation of the United States as consisting of 3,113 counties, which adheres to the Bureau of Economic Analysis's county-level delineation of the nation. However, many of the data sources disaggregate the country into 3,143 "county units" (with some slight variation around this number). The data from many of these secondary data sources thus have to be converted into the 3,113-county universe, and this is done by combining several county-level units. The most frequent example of this takes place in Virginia, where independent cities are combined with surrounding counties in order to meet the ARC/BEA organizational structure. If source data provided numerators and denominators, these values were used to compute figures such as rates and percentages for each indicator. When only computed figures were provided, a weighting variable from another source (such as the 2014 American Community Survey (ACS) population figures) was used to create a weighted average of values. In rare cases, data from source counties were distributed to more than one of the 3,113 counties. This was most often necessary for five Alaska boroughs/county-equivalents, which were recently reallocated from three county-equivalents. In these cases, values from the three county equivalents were directly assigned to the new areas based on predominate geographic overlap.

Age Adjusting Mortality Rates (All Except for Infant Mortality and YPLL)

Age adjusting mortality rates allows for comparisons among counties with different age distributions. Counties with greater numbers of elderly residents can generally expect higher mortality rates than counties with less elderly residents. Thus, a county with higher unadjusted (crude) mortality rates, which suggest poor health, may actually be relatively healthy but simply have a larger number of older residents (and thus a higher overall baseline risk of death). Using data from the Compressed Mortality Files from CDC, we compared the distributions of county populations by age cohort to the standard population distribution for the country as a whole. Using these population distributions by age cohort as a base, the mortality rates in this report are age adjusted and standardized based on the Year 2000 Standard Million Population (see Table 1). This provides the reader with the ability to accurately compare mortality across counties with different age distributions. Only infant mortality is not age adjusted in this report, because the age distribution of a population is not relevant to the measure. The YPLL indicator is obtained directly from County Health Rankings & Roadmaps, which age-adjusted YPLL prior to publishing.

The formulas to convert crude rates to weighted rates by age cohort and total population age-adjusted are:

(Deaths / pop) x 100,000 = CRUDE RATE

CRUDE RATE x (standard pop in each age cohort / 1,000,000) = WEIGHTED RATE

Sum (WEIGHTED RATES) all cohorts = AGE-ADJUSTED RATE for total population

In this example, the crude data are reported in increments of 100,000 residents.

Table 1: Year 2000 Standard Million Population for the United States

Age	2000 Standard Population Distribution
Under 1 year	13,818
1-4 years	55,317
5-9 years	72,533
10-14 years	73,032
15-19 years	72,169
20-24 years	66,478
25-34 years	135,573
35-44 years	162,613
45-54 years	134,834
55-64 years	87,247
65-74 years	66,037
75-84 years	44,841
85 years and over	15,508
All Years	1,000,000

Source: https://wonder.cdc.gov/wonder/help/cmf.html

For example, consider the crude and age-adjusted mortality rates in two states: Utah and Maine, the former of which has a relatively younger population. Table 2 shows the process of converting from crude mortality rates to age-adjusted mortality rates. While Maine has nearly double the crude rate of Utah, once the data are age adjusted, the rates become quite similar in all age cohorts. Table 3 displays the crude and age adjusted rates using two sources: CDC Wonder, an interactive web tool that allows users to calculate mortality rates for specific queries, and then also the age adjustment process used in this report. The slight differences in the mortality rates between these two sources are due to rounding and the inclusion of deaths with unknown ages.

Table 2: Comparison of Crude and Age-Adjusted Mortality Rates

Age Cohort in Years	Standard Million Population	Utah (Younger Population State)				Maine (Older Population State)			
		Deaths	Population	Crude Rate	Wtd Rate	Deaths	Population	Crude Rate	Wtd Rate
Under 1 year	13,818	4,383	840,336	522	7.2	1,345	229,045	587	8.1
1-4 years	55,317	886	3,256,680	27	1.5	227	947,922	24	1.3
5-9 years	72,533	501	3,862,337	13	0.9	170	1,291,131	13	1.0
10-14 years	73,032	600	3,669,490	16	1.2	206	1,427,186	14	1.1
15-19 years	72,169	1.972	3,698,192	53	3.8	843	1,520,896	55	4.0
20-24 years	66,478	3.047	4,021,230	76	5.0	1,166	1,314,897	89	5.9
25-34 years	135,573	6.974	6,764,886	103	14.0	2,568	2,537,252	101	13.7
35-44 years	162,613	9.417	5,545,348	170	27.6	5,060	3,157,287	160	26.1
45-54 years	134,834	16.789	4,839,035	347	46.8	12,660	3,540,842	358	48.2
55-64 years	87,247	25.609	3,575,069	716	62.5	23,653	2,888,520	819	71.4
65-74 years	66,037	36.825	2,198,826	1,675	110.6	37,485	1,868,337	2,006	132.5
75-84 years	44,841	62.962	1,308,643	4,811	215.7	61,595	1,160,897	5,306	237.9
85 years and over	15,508	72.608	475,018	15,285	237.0	70,676	462,785	15,272	236.8
All Years	1,000,000	242.573	44,055,090	550.6	734.0	217,654	22,346,997	794.0	788.0

Table 3: Comparison of Calculated Age-Adjusted Rates with CDC Wonder Reported Rates

State	Utah		Maine	
Data Type	Crude	Age-Adjusted	Crude	Age-Adjusted
Rates	550.6	734.0	974.0	788.0
Validated using CDC Wonder	550.7	734.1	974.0	788.0

Data Suppression and Smoothing

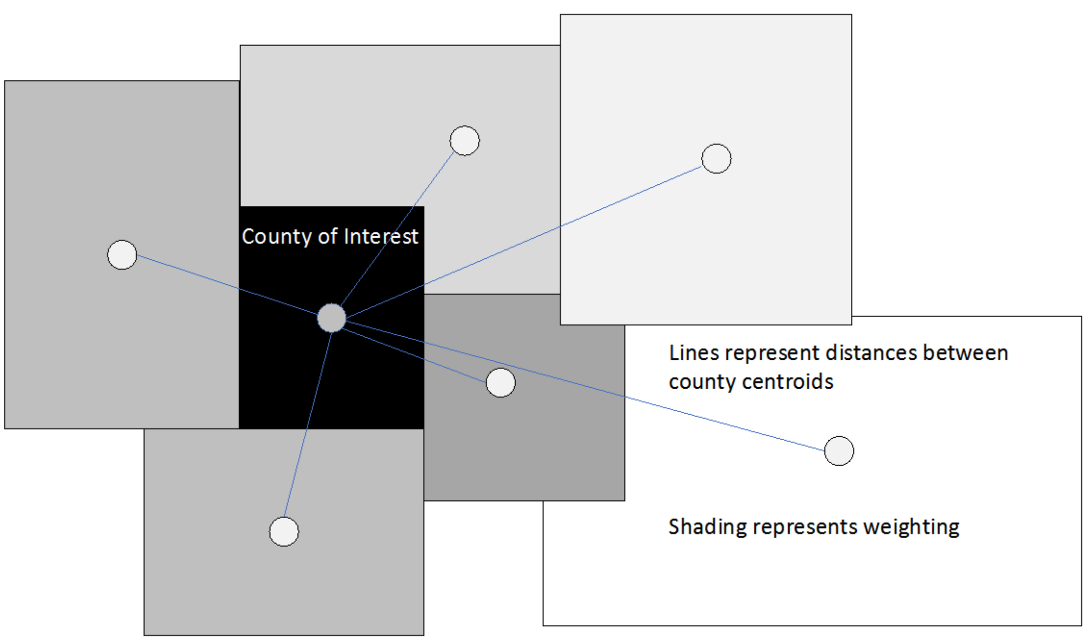
The National Center for Health Statistics of the Centers for Disease Control and Prevention prohibits reporting death counts and death rates when the unit is based on fewer than 10 cases. This results in rates that would need to be suppressed for many county-cause combinations in both the Appalachian Region and the United States. For this report the data for counties with few deaths were “augmented” by incorporating information from nearby counties. To calculate, a proportion of deaths (numerator) and population (denominator) from nearby counties were added to the base numbers from the target county. The proportion of a nearby county's information that was added to the target's decreases with increasing distance between the two. If the augmented number of deaths in the target county was fewer than 10, the death rate for that county-cause combination was suppressed in this report. This approach significantly reduces the number of unreportable area-cause death rates, while maximizing local influence on the augmented rate.

This technique is called smoothing, or spatial smoothing. It helps correct for unstable measures resulting from small population sizes. When very small population leads to very small measure numerators, such as counties with fewer than 10 poisoning deaths in a given year, rates can be smoothed to eliminate statistical instability. The below ten death criteria for spatial adjustment was applied using total deaths, but the gravity-method adjustments were applied for all individual age cohorts.

Overview of Smoothing Process

The smoothing process applies weights to counties surrounding the county with the suppressed value, the "county of interest" in Figure 1 below. The process results in an augmented value resulting from the combination of actual data from the county of interest with weighted values from nearby counties. The degree to which a nearby county contributes to the augmented value depends on the distance from the county of interest to any given nearby county. Figure 1 below demonstrates this principle. The county of interest receives a weight of 1.0 and is fully shaded. Closer counties (e.g., the small, dark gray county directly to the southeast) have darker shading and higher weights than more distant counties (e.g., the large, white county further to the southeast). For each county, the number of deaths and populations in each age category then multiplies the weights. These numbers are then aggregated across counties by age category, and ultimately an age-adjusted rate for the county of interest is calculated.

Figure 1: Example of Weighting Counties for Smoothing Based on Distance



Technical Detail

The approach is one we have used in the past (Ricketts & Holmes, 2007). First, we specified the general form of the weighting function. Here, we specified:

$$\mathrm{WEIGHT} = \exp(\lambda * \mathrm{MILES})$$

where MILES is the straight-line distance between two county centroids. (This can also be written as $e^{(\lambda * \mathrm{MILES})}$.) The parameter λ is constrained to be negative, so the weighting function is equal to one if MILES equals zero, is decreasing in MILES, and is bounded from below by zero. Larger λ (closer to zero) lead to a slower decay – that is, a λ of -0.05, for example, will place greater weight on distant counties than a λ of, say, -0.2. The tradeoff of selecting the λ is a tradeoff between faster decay (averaging over counties that are closer to the suppressed value) and using more counties (leading to more precise estimates). This approach significantly reduces the number of unreportable area-cause death rates, while maximizing local influence on the augmented rate.

We conducted a grid search for the optimal λ using the following approach. First, we randomly chose counties in the Appalachian Region to be labeled suppressed. We calculated implied rates using the above methods for a variety of potential λ. We repeated this exercise 1000 times using different λ each time and calculated the mean squared error, the average squared difference between the smoothed rate and the actual rate. This method allows us to identify the λ with the smallest mean squared error. The lambda satisfying this condition was -0.125. Thus, a county that is 10 miles away from the suppressed county would receive a weight of exp(-.125 * 10) = .29. We specified the same λ for all mortality rates.

With the λ in hand, the approach was as follows. For each suppressed county, we calculated the distance, MILES, between that county centroid and the county centroid of all other counties. We then calculated the weight associated with each county by using the weighting function.

Any county with a weight of less than 0.01 was dropped from the analysis. This approach has little practical effect and eases any rounding issues (e.g. a populous county that is very distant may still have more weight than a nearby county of moderate size). The deaths and population (numerator and denominator) were aggregated using their weights to an augmented numerator and denominator. This augmentation includes any deaths and population reported for the suppressed county as well. At this point, the suppressed county has a “smoothed” number of deaths and population, with closer counties contributing more to the value. The mortality rate can be calculated directly from these numbers.

For mortality with age-adjustment (everything except infant mortality), the augmentation occurred prior to age adjusting. Table 4 walks through a simple example. For simplicity for this example, only two ages are considered – children and adults. Five counties are displayed, along with (for each age group) the number of deaths and population. Distance from county centroid determines the weight. The final four columns calculated the weighted deaths and population. Note that county E, with a weight of less than 0.01, does not contribute to the aggregation. Age adjustment to the 2000 Standard Million population in the compressed mortality file was performed on the augmented age cohort values.

Table 4: Sample County Illustration of Augmentation for Suppressed Data

County	Deaths	Population	Age group	Distance in Miles	Weight	Children		Adults	
						Wtd deaths	Wtd pop	Wtd deaths	Wtd pop
A	2	1,000	Kids	0	1.000	2.00	1000		
A	18	10,000	Adults	0	1.000			18.00	10000.00
B	2	1,200	Kids	5	0.535	1.07	642.0		
B	24	11,500	Adults	5	0.535			12.85	6152.5
C	4	800	Kids	10	0.287	1.15	229.6		
C	37	7,900	Adults	10	0.287			10.60	2267.3
D	3	1,500	Kids	20	0.082	0.25	123.0		
D	27	16,000	Adults	20	0.082			2.22	1312.0
E	7	2,000	Kids	40	0.007	weight < .01 => 0			
E	91	18,500	Adults	40	0.007			weight < .01 => 0	
TOTAL						***4.47***	***1,994.6***	***43.7***	***19,731.8***

After the aggregated deaths and population are calculated by age category, age-adjustment occurs using the approach outlined above (see Table 4). For purposes of this example, we specify weights of 0.1 (children) and 0.9 (adults) and thus calculate an age-adjusted rate of 211.6.

Table 5: Age-Adjusted Step for Spatial Adjustment

Metric	Children	Adults
Rate	224.11	221.32
Age-adjusted weights (for example)	0.10	0.90
Weighted rate	22.41	199.19
AUGMENTED AGE-ADJUSTED RATE	***211.6***	

Reciprocal Measures

In order to report health professional supply measures consistently, we calculated the reciprocal of values pulled directly from County Health Rankings. For example, 2016 County Health Rankings Data for Bibb County, Alabama shows a PCP Ratio of 2,814:1 (persons per physician). We use the reciprocal (1/2,814), converted to primary care physicians per 100,000 people. The calculation is as follows:

$$1/2{,}814 * 100{,}000 = .000355 * 100{,}000 = 35.5$$

Similarly, the dentist ratio of 5627:1 becomes 17.8 dentists per 100,000. The only health professional supply measure that was not calculated in this manner was specialist physician per 100,000. County Health Rankings do not report this measure. We sourced the data for this measure directly from the HRSA Area Health Resources File (AHRF).

Metropolitan and Nonmetropolitan

The Office of Management and Budget (OMB) periodically publishes an official listing of Metropolitan Statistical Areas (MSAs). The OMB defines MSAs as a county or group of counties with a core urban area of over 50,000 people. Counties adjacent to the county containing the core urban area are included in the MSA if those counties have strong commuting patterns with the core. All counties not part of an MSA using the 2015 definition, including Micropolitan statistical areas, are considered "Non-metropolitan" in this report.

County-Level Estimates

Some of the data included in this report were the result of estimation techniques employed by the data owner, prior to our analysis. This often occurs when the data source is sufficiently precise at the state level, but not at the county level, and so the analysts must interpolate county-estimates using state-specific estimates and county-level characteristics. Leading examples of this include physical and mentally unhealthy days from the County Health Rankings. For those measures, the data owners use state-level data and interpolate down to county-level estimates using the known association between population characteristics and county measures of the population. Sometimes these measures are adjusted so that the county-specific estimates aggregate to the state estimate.

Imputing Missing Values for Drivers

Some data sources had a number of missing or suppressed values for certain drivers. In the absence of a solution, these counties would not contribute to the analysis. Missing values were imputed for these driver variables: smoking, teen birth rates, pollution, social assistance jobs, HbA1C testing, breast screening rates (mammograms), student/teacher ratio, STI rate, e-prescription, and exercise. Rates of "missingness" ranged from 199 (e-prescription) to five (smoking). Imputation for these values was based on housing affordability, social association rate, percent with some college, poverty rate, ARC economic index, median income, travel time, population and population squared, metropolitan status, and whether the county was in the ARC region.

Propensity Scoring

Researchers have long been interested in the effect of a "Treatment" (T) on some outcome. Traditionally, we think of a medical treatment (such as a drug) on some health outcome (such as blood pressure); but this form is general enough to apply to any situation with a continuous outcome and a dichotomous treatment.[1] In observational (i.e., non-experimental) studies, researchers have known the importance of controlling for confounders, other factors affecting the outcome and correlated to the "treatment." The most common approach is to control for these factors using a regression model. However, this approach has limitations. For example, if the sample size is small, the researchers may not want to use the degrees of freedom necessary to control for multiple confounders, or the researchers may not want to assume linearity. A common alternative, is to "match" each treated observation to a non-treated (control) observation, where matching would be done by confounders. For example, for each treated 53-year-old college-educated male living in the South, find a 53-year-old college-educated male living in the South who did NOT receive treatment and compare the outcomes. This cell-based approach requires a large sample of controls from which to pull matches, and the controls are often a limiting factor.

In their seminal 1983 work, Rosenbaum and Rubin (Rosenbaum & Rubin, 1983) proposed a method by which rather than matching on each confounding variable, the analysts could match on one uni-dimensional index that would function similarly to the cell-based matching method, but would be far easier to implement and more likely to find a control for each treated. They proposed using the "probability observations with these confounding variables would be observed to have had treatment"— the *propensity score*— as the dimension on which to match. This approach has quickly become a popular method for the reasons listed above, and this approach we chose. Although not relevant to our discussion, it is important to note that the approach addresses only *observed* confounding—any *unobserved* selection bias is not mitigated by this approach.

With the propensity score analysis completed, the analyst then turns to comparing outcomes for the treatment and control groups. There are multiple potential approaches[2] to this step, but we will describe the two most common. One approach is to *match* by taking each treated case and finding a "similar" control case. "Similar" is defined as "having a nearly identical propensity score." Common approaches include "caliper," finding controls that differ from the treatment by less than some threshold; or, nearest neighbor, finding the control with the smallest difference. The other approach is *inverse probability of treatment weighting* (IPTW), which weights the entire sample by the inverse of the probability the observation was in the category it was (i.e. 1 / Pr(T) for those in the treated category and 1/((1-Pr(T)) for the those in the control group). By using the propensity score in either of these manners, comparison of outcomes can be done using such techniques as simple as a t-test of means.

The textbook propensity score approach, then, is:

1. Run a logistic regression of T regressed on X;
2. Generate propensity scores, the predicted probability of having received treatment (T=1);
3. Match or generate IPTWs; and
4. Compare the outcomes for the treated and control groups.

[1] Of course, these assumptions can be further relaxed, but such generalization is not necessary for this discussion.

[2] An excellent resource is Austin P. An Introduction to Propensity Score Methods for Reducing the Effects of Confounding in Observational Studies. Multivariate Behav Res. 2011 May; 46(3): 399–424. 2011. doi: 10.1080/00273171.2011.568786