

# Artificially Intelligent Social Risk Adjustment: Development and Pilot Testing in Ohio

Lisa M. Lines, Marque C. Long, Jamie L. Humphrey,  
Crystal T. Nguyen, Suzannah Scanlon, Olivia K. G.  
Berzin, Matthew C. Brown, and Anupa Bir



RTI Press publication RR-0047-2209

RTI International is an independent, nonprofit research organization dedicated to improving the human condition. The RTI Press mission is to disseminate information about RTI research, analytic tools, and technical expertise to a national and international audience. RTI Press publications are peer-reviewed by at least two independent substantive experts and one or more Press editors.

### Suggested Citation

Lines, L. M., Long, M. C., Humphrey, J. L., Nguyen, C. T., Scanlon, S., Berzin, O. K. G., Brown, M. C., and Bir, A. (2022). *Artificially Intelligent Social Risk Adjustment: Development and Pilot Testing in Ohio*. RTI Press Publication No. RR-0047-2209. Research Triangle Park, NC: RTI Press. <https://doi.org/10.3768/rtipress.2022.rr.0047.2209>

This publication is part of the RTI Press Research Report series.

RTI International  
3040 East Cornwallis Road  
PO Box 12194  
Research Triangle Park, NC  
27709-2194 USA

Tel: +1.919.541.6000  
E-mail: [rtipress@rti.org](mailto:rtipress@rti.org)  
Website: [www.rti.org](http://www.rti.org)

©2022 RTI International. RTI International is a trade name of Research Triangle Institute. RTI and the RTI logo are U.S. registered trademarks of Research Triangle Institute.



This work is distributed under the terms of a Creative Commons Attribution-NonCommercial-NoDerivatives 4.0 license (CC BY-NC-ND), a copy of which is available at <https://creativecommons.org/licenses/by-nc-nd/4.0>

<https://doi.org/10.3768/rtipress.2022.rr.0047.2209>

[www.rti.org/rtipress](http://www.rti.org/rtipress)

# Contents

About the Authors	i
Acknowledgments	ii
<b>Abstract</b>	ii
<b>Introduction</b>	1
Overview	1
Background	1
Aims	3
<b>Methods</b>	3
Conceptual Framework	3
Addressing Bias	4
Population and Sample	5
Outcome Measures	5
Predictor Data Inputs for the RF Model	5
Comparison Measures	5
Analytic Approach: Machine Learning	5
<b>Results</b>	8
Characteristics of LSI Scores: Accuracy and Variable Importance	8
Descriptive Characteristics: Ohio	8
Covariate Balance Between the Ohio CPC and Comparison Groups	10
Multivariate Modeling Results	10
<b>Discussion</b>	12
Limitations	14
Implications for Policy and Practice	14
<b>References</b>	15

## About the Authors

**Lisa M. Lines**, PhD, MPH, Senior Health Services Researcher, Center for Health Innovation, Research, and Policy, RTI International and Assistant Professor, Quantitative and Population Health Research, UMass Chan Medical School

**Marque C. Long**, MPH, Public Health Analyst, RTI International

**Jamie L. Humphrey**, PhD, MPH, Health Geographer, Center for Advanced Methods Development, RTI International and Research Associate, Drexel University, Dornsife School of Public Health

**Crystal T. Nguyen**, PhD, MS, Biostatistician and Research Public Health Analyst, Center for Advanced Methods Development, RTI International

**Suzannah Scanlon**, MS, Research Economist, Payment Model Design and Implementation, RTI International

**Olivia K. G. Berzin**, MPH, Research Associate, Specialty Care for Seriously Ill Populations, RTI International

**Matthew C. Brown**, MS, Data Scientist, Center for Advanced Methods Development, RTI International

**Anupa Bir**, SD, MPH, Vice President, Center for Health Innovation, Research, and Policy, RTI International

## RTI Press Associate Editor

Cynthia Bland

## Acknowledgments

Our team has included more than 20 contributors since 2019. The co-authors gratefully acknowledge the input and contributions of the following additional current and former RTI staff (in alphabetical order): Monica Borges, Elena Bravo-Taylor, Anusha Chaudhry, Josh Clemson, Denise Clayton, Erin Erikson, Sara Freeman, Mark Hatem, Ariella Hirsch, Marc Horvath, Yevgeniya Kaganova, Krish Maypole, Marisa Morrison, Rasika Ramanan, Emily Schneider, Clyde Schwab, Kara Sokol, and Joe Wasserman.

Portions of this analysis were previously published in blog form at <https://www.themedicalcareblog.com/artificially-intelligent-social-risk-adjustment/> and presented at Health Information and Management Systems Society, August 2021, and the Academy Health National Health Policy Conference and Health Datapalooza, February 2021.

## Abstract

Prominent voices have called for a better way to measure, predict, and adjust for social factors in healthcare and population health. Local area characteristics are sometimes framed as a proxy for patient characteristics, but they are often independently associated with health outcomes. We have developed an “artificially intelligent” approach to risk adjustment for local social determinants of health (SDoH) using random forest models to understand life expectancy at the Census tract level. Our Local Social Inequity score draws on more than 150 neighborhood-level variables across 10 SDoH domains. As piloted in Ohio, the score explains 73 percent of the variation in life expectancy by Census tract, with a mean squared error of 4.47 years. Accurate multidimensional, cross-sector, small-area social risk scores could be useful in understanding the impact of healthcare innovations, payment models, and SDoH interventions in communities at higher risk for serious illnesses and diseases; identifying neighborhoods and areas at highest risk of poor outcomes for better targeting of interventions and resources; and accounting for factors outside of providers’ control for more fair and equitable performance/quality measurement and reimbursement.

---

## Introduction

### Overview

Numerous research studies have shown that social, economic, environmental, political, and other non-medical factors play a large role in shaping health status.<sup>1</sup> These factors have been found to account for 80 to 90 percent of a person's health outcomes.<sup>2</sup> Social determinants of health (SDoH) are “the conditions in the environments where people are born, live, learn, work, play, worship, and age that affect a wide range of health, functioning, and quality-of-life outcomes and risks.”<sup>3</sup> Implementing programs, policies, and investments that address differences in SDoH could help to improve the health of populations at high risk for serious illnesses and worse health outcomes.<sup>4</sup>

SDoH include both risk and protective factors. Income, for example, is a SDoH; associated measures and concepts can be both risk factors (e.g., poverty) and protective factors (e.g., wealth). Risk factors including poverty, racism, and housing insecurity are just a few of the nonmedical factors associated with poor health outcomes.<sup>5,6</sup> Differences in SDoH and risk factors play a large role in shaping health disparities in the United States.<sup>1,3</sup> Disparities in health and SDoH are the metrics we use to measure progress toward achieving health equity and reducing health inequity,<sup>7</sup> which is defined as the systematic differences in the opportunities groups have to achieve optimal health, leading to unfair and avoidable differences in health outcomes.<sup>4</sup>

For decades, health and medical practitioners, researchers, and policymakers have called for a better way to measure, predict, and adjust for SDoH in healthcare and population health.<sup>8,9</sup> The goal of this work is to answer that call in a way that supports the achievement of health equity: attainment of the highest level of health for all people, where everyone has a fair and just opportunity to attain their optimal health regardless of race, ethnicity, disability, sexual orientation, gender identity, socioeconomic status, geography, preferred language, or other factors that affect access to care and health outcomes.<sup>10</sup>

We approached this research through the lens of risk adjustment for SDoH, also known as social

risk adjustment. This manuscript outlines a novel approach to developing a health equity-focused index to expand risk adjustment models and demonstrates the utility of the measure in an analysis of a patient-centered medical home program in Ohio. We find that accurate multidimensional, cross-sector, small-area social risk scores could be useful in understanding the impact of healthcare innovations, payment models, and interventions on social drivers of health on high-risk communities; identifying neighborhoods and areas at highest risk of poor outcomes for better targeting of interventions and resources; and accounting for factors outside of providers' control for more fair and equitable performance/quality measurement and reimbursement.

### Background

Despite the enormous impact of SDoH on patient outcomes, current health system performance and quality measures generally do not consider SDoH, nor are SDoH generally included in most risk adjustment formulas.<sup>11,12</sup> Risk adjustment is a “statistical process that takes into account the underlying health status and health spending of the enrollees in an insurance plan when looking at their health care outcomes or health care costs.”<sup>13</sup> Risk adjustment is used in quality measurement and in calculating reimbursements to healthcare providers.<sup>14</sup> Risk adjustment modifies payments to providers to account for differences in expected healthcare expenditures because of factors such as individual age, sex, and diagnoses.<sup>15</sup> This ensures that providers treating patients with higher risk (e.g., older patients) are paid more to treat such patients compared with those treating lower-risk patients. Risk adjustment is common in value-based payment (VBP) models, which provide payments to providers based on achievement of quality goals or cost savings as an important way to reduce incentives to care for only young, healthy, inexpensive-to-treat patients.

The Centers for Medicare & Medicaid Services (CMS) Innovation Center estimates that by 2030, 100 percent of people with Original Medicare (Parts A & B) will be cared for under a VBP model.<sup>16</sup> This reinforces the imperative that risk adjustment models accurately account for underlying individual- and community-

level risk when predicting patient expenditures. In the zero-sum world of pay-for-performance and VBP models, omitting SDoH from risk calculations can lead to several unintended consequences. First, practices with patients who are at a lower risk of serious illnesses may receive rewards that reflect patient mix rather than quality, while practices with patients who are at a higher risk of serious illnesses may lose out on resources for improving quality.<sup>17,18</sup> Second, providers may object to being held accountable for factors outside of their control.<sup>19</sup> Third, payers have incentives to enroll members with lower risk wherever possible, potentially leading to cream-skimming and cherry-picking.<sup>20</sup> Finally, lack of good data on SDoH can lead to interventions targeting populations with a lower risk of serious illnesses, attenuating potential benefits from those interventions.<sup>21</sup>

Previous work on social risk adjustment has employed analyses at the county or ZIP code levels or has approached measurement or risk adjustment for SDoH with factor analysis, principal component analysis, or simple weighted formulas.<sup>22,23</sup> Several US small-area-based indices of social risk and socioeconomic deprivation are available, including the Area Deprivation Index (ADI), Social Deprivation Index (SDI), and the Social Vulnerability Index (SVI). These indices consider relatively few social factors and rely heavily on US Census data. Some of these indices are currently used in VBP models to increase equity in payments based on the area-level social risk factors of the population served. For example, the CMS Accountable Care Organization (ACO) Realizing Equity, Access, and Community Health (REACH) model uses the ADI to adjust ACO benchmarks based on the number of beneficiaries in the highest and lowest deprivation areas.<sup>24</sup>

The ADI, originally developed by Singh et al. more than 20 years ago,<sup>23</sup> has been used in numerous studies to understand the association between neighborhood-level characteristics and health outcomes, such as hospital readmissions and expenditures.<sup>25,26</sup> The ADI relies upon 17 socioeconomic indicators organized in four domains (income, education, employment, and housing quality), all of which come from the American

Community Survey (ACS) 5-year sample at the Census block group level.<sup>23</sup>

The SDI, developed by Butler et al.<sup>27</sup> and maintained by the Robert Graham Center of the American Association of Family Physicians, is based on seven variables collected in the ACS at the Census tract level: (1) poverty rate, (2) adults without high school diploma, (3) percent single parent household, (4) percentage living in rented housing unit, (5) percentage living in overcrowded housing unit, (6) percentage of households without a car, and (7) unemployment.

The SVI (available at census tract and county levels), developed by the Centers for Disease Control and Prevention (CDC) and the Agency for Toxic Substances and Disease Registry, incorporates 15 Census variables, including poverty, lack of access to transportation, and crowded housing, within four domains: (1) socioeconomic status, (2) household composition, (3) race/ethnicity/language, and (4) housing/transportation. The SVI was specifically designed to understand social vulnerability in the context of natural disasters and emergency management.<sup>22</sup>

The ADI (available at 9-digit ZIP code and Census block levels)<sup>23</sup> and SDI (available at county, census tract, aggregated ZIP Code Tabulation Area, and Primary Care Service Area levels)<sup>27</sup> were developed using factor analysis, a method with poor replicability over different space and time periods.<sup>28</sup> In addition, the ADI is roughly 80 percent correlated with owner-occupied property values, which can be misleading. For example, the ADI rates neighborhoods with low homeownership rates as excellent when those few owned homes are very expensive, despite economic circumstances being difficult for most residents.

In this pilot study, we sought to develop a comprehensive neighborhood-level index to enable risk adjustment models to correct for inadequate adjustment for social risk factors. Although individual-level SDoH data are best for predicting individual-level outcomes, collecting this information can be difficult.<sup>29</sup> Area-level measures are associated with health outcomes and can be useful either in conjunction with individual-level data or on their own.<sup>12,27,30-35</sup>

We have developed an “artificially intelligent” approach to risk adjustment for SDoH using random forests (RFs) to understand life expectancy (LE) at the Census tract (CT) level. In the United States, LE inequality persists along racial, sex, and urban/rural lines.<sup>35</sup> Although LE has increased from a national average of 69.7 years in 1950 to 78.8 years in 2015, the averages vary dramatically across small geographic areas.<sup>36,37</sup> For example, within Franklin County, Ohio, CT-level LE ranges from 60.0 to 87.6 years.<sup>38</sup>

Artificial intelligence (AI) is “a branch of computer science that explores the ability of computers to imitate aspects of intelligent human behavior, such as problem-solving, reasoning and recognition.”<sup>39,40</sup> Machine learning (ML), an application of AI, is a popular tool in public health and health services research.<sup>41,42</sup> Researchers have used these tools in four different ways: (1) to predict an outcome  $Y$  based on a set of covariates  $X$  without ascribing any causal value to these covariates, (2) to identify bias in data and decision-making, (3) to enhance the evaluation of how treatment affects an outcome, and (4) to synthesize new data.<sup>41</sup> For this paper, we focus on its pure prediction application via RFs. An RF model is a supervised ML model that creates multiple regression trees and aggregates their results with the purpose of forming a generalization about the relationship of covariates to the outcome of interest.<sup>43</sup> We used RF models to understand which SDoH variables explain the variation in LE at the CT level and to generate a composite SDoH measure, the Local Social Inequity (LSI) score, which is a ranked percentile score based on the RF-predicted outcome measure. Section 2 of the online Supplement provides additional background information about RFs (supplementary material can be found online at <https://doi.org/10.5281/zenodo.7026658>).

## Aims

The aims of this research were threefold. First, we created a health equity–focused conceptual framework for risk adjustment of the SDoH. The framework provides context and rationale for the inclusion of SDoH variables and their links to our population health outcome, LE. We used LE at birth as our outcome measure because it is frequently used to measure the overall health status of a

population.<sup>44</sup> Second, we built upon our conceptual model to develop a health equity–focused index that comprehensively measures neighborhood-level social inequity. We used supervised ML and publicly available data to develop a new SDoH composite index that is robust, comprehensive, cross-disciplinary, spatially refined, and highly predictive. Finally, we pilot-tested the composite index in a patient-centered medical home program in Ohio and estimated associations between the index and individual-level core outcome measures for the evaluation: total expenditures, inpatient admission, hospital readmissions, and emergency department (ED) visit rates.

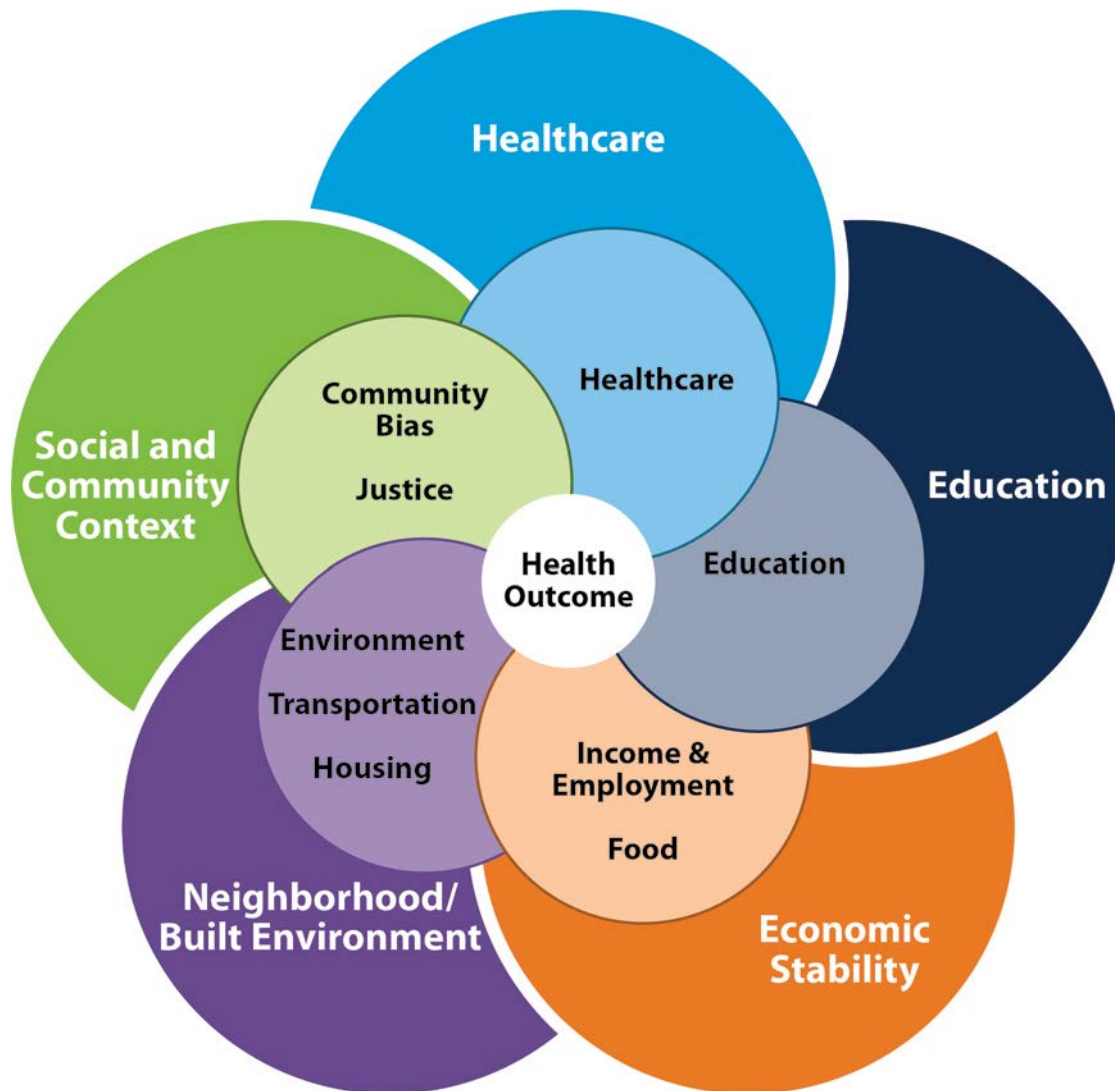
---

## Methods

### Conceptual Framework

We started conceptualizing SDoH for this study by reviewing the literature on existing conceptual frameworks.<sup>3,45,46</sup> We ultimately started with and then adapted the five domains in CDC’s Healthy People framework: healthcare access and quality; education access and quality; social and community context; economic stability; and neighborhood and built environment. We expanded these five domains into 10 domains, as follows, based on data and measure availability and potential associations with LE. First, we kept the CDC’s domains of (1) healthcare and (2) education. We expanded the social/community context domain to include three separate areas: community health, well-being, and healthy behaviors; bias, stress, and trauma; and justice, crime, and incarceration. We split the CDC’s economic stability domain into two: (1) food security and access to healthy food and (2) poverty, inequality, and employment. Finally, we expanded the neighborhood/built environment domain into three domains: (1) housing adequacy, crowding, and structural health; (2) environmental quality; and (3) transportation access, infrastructure, and safety. Figure 1 shows the domains at a glance and emphasizes their interconnectedness. The online Supplement Exhibit 1 provides the full list of measures from which the RF drew predictor variables in the main model and sensitivity analyses and rationales for inclusion.



**Figure 1. Conceptual Framework: Social Determinants of Health**

### Addressing Bias

We addressed potential biases in our framework in several ways. First, we “tuned” our model to predict LE, because it is a high-level population health outcome that is commonly used in population health research. Second, we involved people in the development process from a variety of backgrounds and disciplines, bringing expertise and insights from multiple viewpoints—including people with direct experience with the social risks (e.g., food insecurity and poverty) we sought to measure. The team included staff with expertise in health services research, social epidemiology, risk adjustment, health economics, public health and health policy, health

geography, AI, data science, and visual displays of quantitative information. Third, we explicitly called out bias, stress, and trauma in its own domain, and we included measures of racial residential segregation and inequality in the model. We included measures that capture racism and bias, rather than race and ethnicity, as underlying causes of health inequity.

### Population and Sample

Ohio’s population in 2010–2015 was approximately 11.5 million people, residing in 2,952 CTs. The area-level predictor variables for the LSI scores were drawn from this population. We linked the scores to individual-level data on a population of Medicaid



enrollees in Ohio ( $n = 2,427,829$ ) for whom we had individual characteristics (including age, sex, and ZIP code) and Medicaid claims data for the period spanning 2015 to 2019. This analysis was undertaken as part of the State Innovation Models (SIM) Initiative, Round 2 evaluation sponsored by CMS's Innovation Center.

## Outcome Measures

Our outcome measure for the ecological analysis was CT-level LE, as estimated from 2010 to 2015 by the CDC's US Small-area Life Expectancy Estimates Project (USALEEP). The USALEEP provides estimates of LE at birth—the average number of years a person can expect to live—in each CT. Mean LE in Ohio in 2010–2015 was 76.6 years (range: 60–89.2). For the person-level analysis, we examined four outcome measures: total Medicaid expenditures, hospitalizations, readmissions, and ED visit rates.

## Predictor Data Inputs for the RF Model

We used theoretically driven, publicly available information to explain variation in LE. Most variables were available at the CT level; variables at other geographies (e.g., block group, ZIP code tabulation areas [ZCTAs], county), were rescaled to CTs either using population weights or direct assignment. Because our LE outcome was based on 2010–2015 estimates, we used data from contemporaneous years whenever possible (e.g., we included 2010–2014 ACS estimates). We assumed that predictors from pre-2010 and post-2015 were similar to contemporaneous conditions; we included pre-2010 and post-2015 predictors only when contemporaneous estimates were not available. When necessary, we averaged across estimates from multiple years to provide a single point estimate for each CT. In some cases, we normalized variables, created rates per 100k population, or performed other simple transformations to make the variable more standardized or reliable. Because the LE variable was based on the 2010 CT vintage, we performed population-based conversions to convert non-2010 vintage variables to be 2010 vintage. Along with the ACS, data were also obtained from PolicyMap, the Robert Wood Johnson Foundation, and numerous state and federal government agencies (17 data

sources in total). Before inclusion in modeling, the variables were all standardized (z-scored) to prevent scale bias. In total, 155 variables were available as predictors in the RF models. The full list of available predictor variables is provided in Exhibit 1 of the online Supplement.

## Comparison Measures

We benchmarked our scores against three previously described comparison measures: the ADI, SDI, and SVI. We used population weighting to scale the ADI from block group to CT.

## Analytic Approach: Machine Learning

We trained an RF model (see Supplement Section 2 for more information) to predict LE estimates from the 10 domains of SDoH data described earlier using the randomForestSRC package in R.<sup>48</sup> CTs missing their LE value from the USALEEP were omitted from the model. Variables with missing data in more than 30 percent of CTs were categorized into deciles with an additional category labeled “missing.” Variables missing in fewer than 30 percent of CTs were imputed using a multiple imputation method that uses RFs to simultaneously estimate the missing predictor values based on the other predictors.<sup>49,50</sup>

A stepwise approach was taken for variable importance: we iteratively fit the model and removed variables with a variable importance (VI) below 0.025. These predictors were likely to be noise variables, and fitting the model without them improved its overall accuracy. Additionally, we selected the following tuning parameters: a seed of 2021 was used for replicability; our forest consisted of 1,000 trees; one-third of predictors were sampled at each node for potential splitting; and nodes had to have at least 15 in-bag data points to continue splits. The model-predicted outcome, or  $\hat{Y}$ , was then used to create CT-level percentile rank scores that represent relative LSI in an Ohio neighborhood.

## Analytic Approach: Evaluating the Ohio CT-level Scores

Of the 2,952 CTs in Ohio, LSI scores were not estimated for 181 tracts (6.1 percent) that were missing LE estimates from the USALEEP (with roughly 511,900 residents, or 4.4 percent of

Ohio's population). These CTs also had a high degree of missingness of feature variables, making model prediction for these CTs difficult. Instead, we interpolated the LSI score by averaging the scores of neighboring areal units using the *spdep* package.<sup>51</sup> Neighbors were defined using a 1st order queens contiguity matrix; all neighboring units were equally weighted, ignoring neighbors with missing LE. If all 1st order neighbors also had missing LE, then 2nd order neighbors were used. Recognizing that areal interpolation may induce spatial dependency, we included latitude and longitude of each CT as predictors in the RF model to address this potential issue.

To transform LSI scores from the CT to the ZIP code level, the *zcta* package<sup>52</sup> was used to first transform the CT-level risk scores to  $n = 1,197$  ZCTAs, which are generalized representations of US ZIP codes using population weighting. Weights were derived from the 2010 decennial Census 100 percent population counts at the block level. ZCTA-level scores were then cross-walked to ZIP codes using the UDS Mapper ZIP Code to ZCTA Crosswalk<sup>53</sup> and merged with the person-level data for the SIM2 analysis.

### Validation

We validated the LSI scores in several ways. First, we compared the percent variance explained by the LSI scores and the model-predicted LE estimates with the variance explained by our three benchmark composite measures: the ADI, SDI, and SVI. This involved simple post-hoc linear regressions. Second, we ran models within all other states and across the United States to compare rankings of variable importance and understand their stability. We also tested the stability of the VI rankings by undertaking several alternative modeling approaches, including lasso regression, stepwise regression, and principal components analysis followed by RF estimation. We provide further information about validation analyses in Section 3 of the Supplement.

### Analytic Approach: Pilot Test in Ohio

The Ohio SIM Initiative<sup>54</sup> began in 2015 with the goals of enrolling 80 to 90 percent of residents in a VBP model and covering 50 percent of the state's medical spending within 5 years. Ohio aimed to

achieve these goals through two key strategies, one of which was a Medicaid patient-centered medical home model known as Ohio Comprehensive Primary Care (Ohio CPC). To be eligible for the Ohio CPC model before January 2019, primary care practices were required to have at least 500 attributed Medicaid beneficiaries. Practices also agreed to meet activity requirements—such as extended patient hours—associated with the provision of patient-centered primary care.

The federally funded evaluation of Ohio CPC included three cohorts of practices, each joining the model between January 2017 and January 2018. The timeframe for the evaluation was January 1, 2014, through December 31, 2018. The intervention group included 1,116,460 Medicaid beneficiaries who were attributed to Ohio CPC practices, and the comparison group included 1,888,767 similar Ohio Medicaid beneficiaries attributed to practices that did not participate in Ohio CPC. The Ohio Department of Medicaid attributed Medicaid beneficiaries to Ohio CPC practices on a quarterly basis. Medicaid beneficiaries were assigned to the intervention group if they spent more months during the entire analysis period attributed to Ohio CPC practices rather than non-Ohio CPC practices. Of the approximately 3 million Medicaid beneficiaries—including both children and adults—who were ever in the evaluation sample, 53 percent of these beneficiaries were included in the sample between 2014 and 2018.

At the beneficiary level, the intervention start date was defined according to when a beneficiary was first attributed to an Ohio CPC practice. The first possible date for attribution to an Ohio CPC practice was January 2017, when the first cohort of practices joined the model.

The evaluation assessed the effects of Ohio CPC on four core outcomes, including annualized total spending per beneficiary per month (PBPM), the probability of an inpatient admission, the probability of a readmission within 30 days after a live hospital discharge, and total annual ED visits. We estimated total spending PBPM, inpatient admissions, and ED visits at the beneficiary year level for both children and adults. We evaluated readmissions at the inpatient discharge level for individuals who were

aged 18 years or older. We employed a difference-in-differences (D-in-D) quasi-experimental design using an unbalanced longitudinal panel and inverse propensity score weights to adjust for intervention and comparison group differences.

The analysis used an ordinary least squares (OLS) model to obtain D-in-D estimates for total spending PBPM, a logistic regression model to obtain D-in-D estimates for the inpatient admission and readmissions outcomes, and a negative binomial model to obtain D-in-D estimates for ED visits. The estimated probability of any inpatient admission and the ED visit count were multiplied by 1,000 to obtain an approximate rate per 1,000 member years. The estimated probability of a readmission was multiplied by 1,000 to obtain an approximate rate per 1,000 live discharges per year.

Because some Medicaid beneficiaries are not enrolled in Medicaid for a full year, we annualized the total spending and ED visit outcomes. We divided each beneficiary's total annual spending by an eligibility fraction: the number of months in a year that an individual was enrolled in Medicaid divided by 12. We then converted annualized total spending to a PBPM value by dividing it by 12. For each beneficiary in each year, we divided ED visit counts by that beneficiary's annual eligibility fraction and rounded the annualized count of ED visits to the nearest integer.

For each outcome, we ran two sets of models, each with a different set of covariates. One model adjusted for person-level variables (gender, age, Medicaid enrollment because of disability, race, a count of total months enrolled in Medicaid during the measurement year, the logged Chronic Illness and Disability Payment System score, residence in a metropolitan statistical area, and changes to attribution status across calendar quarters) and a full set of county-level variables (supply of short-term acute care hospital beds, percentage of county residents who are uninsured, percentage of county residents living in poverty, median age of the county population, and a count of Federally Qualified Health Centers [FQHCs] per 1,000 people). The second model adjusted for person-level variables, a limited set of county-level variables (supply of short-term acute care hospital beds and a count of

FQHCs per 1,000 people) and the ZIP-level LSI score. We constructed these two separate models to determine whether LSI scores could substitute for area-level covariates without affecting the direction, magnitude, and significance of D-in-D estimates. (The second model retained the hospital bed supply and FQHC count covariates because those variables were not included in the pilot version of the LSI scores for Ohio.)

We derived person-level covariates from Medicaid claims and enrollment data. We created an indicator for residence in a metropolitan statistical area by matching ZIP codes of residence in enrollment data to data from the Area Health Resources Files. The covariate for changes to attribution status indicated whether beneficiaries were attributed to both Ohio CPC practices and non-CPC practices while enrolled in Ohio Medicaid. We included this covariate in the model because a state-funded evaluation of Ohio CPC indicated that in 2017 and 2018, the practice to which a beneficiary was attributed changed between one quarter and the next for approximately 30 percent of beneficiaries in Ohio CPC for at least two successive quarters.<sup>55</sup> All county-level covariates were derived from the Area Health Resources Files.

To improve balance between the intervention and comparison observations on observable covariates, we created propensity scores for the comparison sample at the beneficiary year level for total spending, inpatient admissions, and ED visits and at the inpatient discharge year level for readmissions. We used the propensity scores to calculate inverse probability weights ( $\text{propensity score}/(1 - \text{propensity score})$ ), which we applied to observations in the comparison group. We set the inverse probability weight to 1 for all observations in the intervention group. Propensity score weights were capped to be between 0.05 and 20 because extreme weights may give undue influence to a small set of observations, leading to greater sensitivity in the impact results. Observations in regression models for beneficiary year-level outcomes (spending, inpatient admissions, and ED visits) were weighted by an analytic weight, which we created by multiplying the annual propensity score weight by each beneficiary's annual eligibility fraction (number of months enrolled in

Medicaid in a year divided by 12). In the inpatient discharge–level model for readmissions, observations were weighted by a propensity score weight. In all regression models, standard errors were clustered at the county level to account for the correlation in outcomes within geographic regions.

All outcome models assumed that Ohio CPC and comparison group outcome trends were parallel during the baseline period. We produced annual D-in-D estimates for each intervention year, overall D-in-D estimates, and adjusted means for both versions of our D-in-D models (the one with LSI scores and the one without). We calculated the annual D-in-D estimates from non-linear models as an average marginal effect. We created overall D-in-D estimates as a weighted average of the annual D-in-D estimates. For each outcome, we produced regression-adjusted means for the intervention and comparison groups during the baseline period, in each intervention year, and for the overall intervention period. Adjusted means during the baseline period represented the average of the outcome during the middle baseline year (e.g., the second year in a 3-year baseline period) after controlling for model covariates. Adjusted means during the intervention period represented the average of the outcome for each specific intervention year after controlling for model covariates. The overall adjusted mean is a weighted average of the year-specific adjusted means calculated for each intervention year. The weights used in this calculation were chosen to account for variation in size of the intervention (and comparison) group across years. The weights were calculated as the number of intervention (or comparison) beneficiaries observed within each year divided by the total number of intervention (or comparison) beneficiary years observed during the intervention period. The total weighted number of observations for all models except the readmission outcome was 8,229,002; the weighted number of observations for the readmission outcome was 696,087. These numbers include all person year (or discharge year) observations for both the Ohio CPC and comparison group.

## Results

### Characteristics of LSI Scores: Accuracy and Variable Importance

As piloted in Ohio, the LSI score had a mean squared error of 4.47 and explained 73 percent of the variance in LE at the CT level. In comparison, the SDI explains 50 percent, the SVI explains 58 percent, and the ADI explains 63 percent of variation in tract-level LE in Ohio. Complete results from the validation analyses are provided in Section 3 of the Supplement.

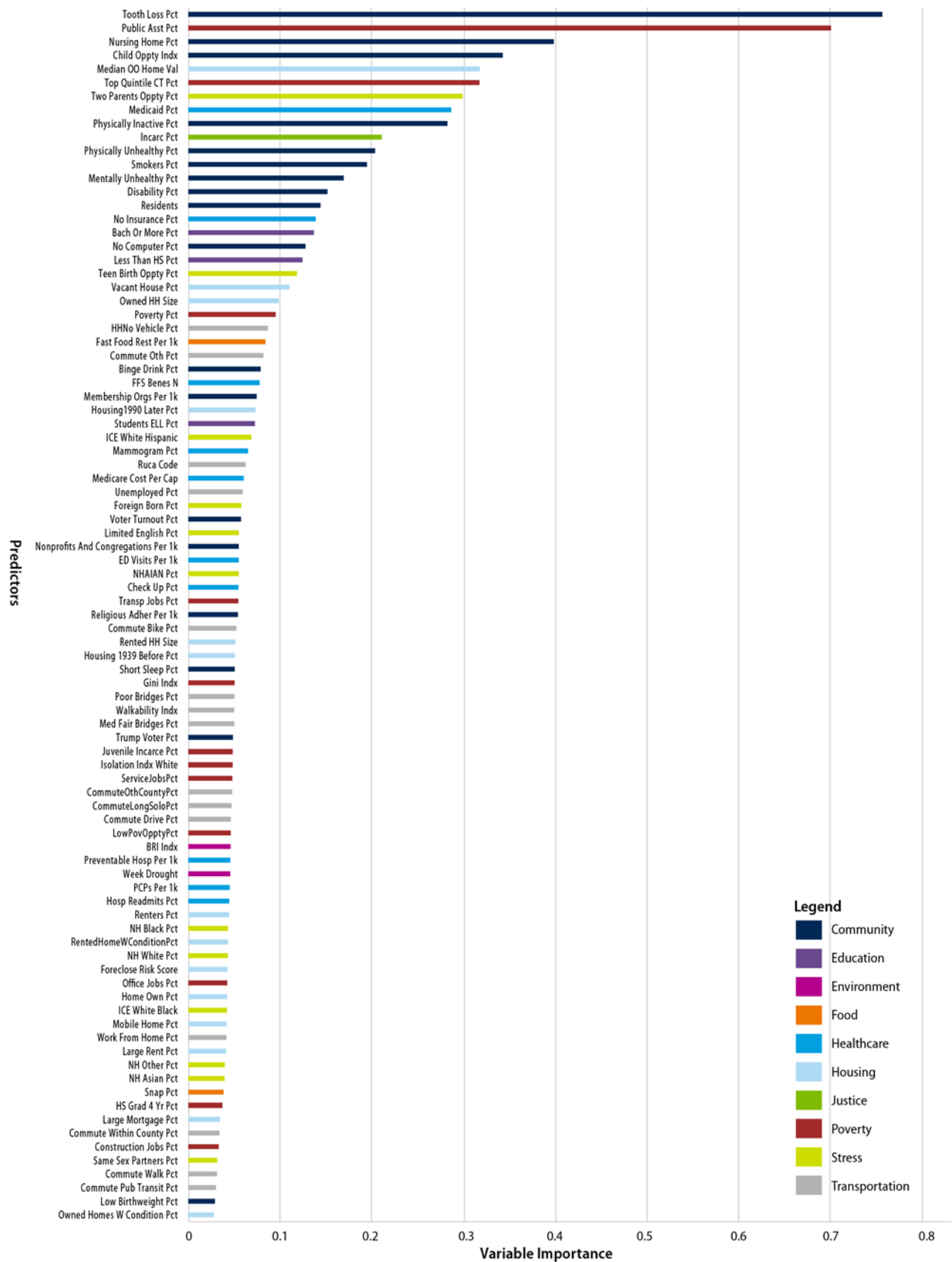
Figure 2 shows the 89 variables retained in the final model and how they ranked in terms of variable importance; Supplement Exhibit 1 provides the complete list of candidate measures. The most important individual predictors—in terms of VI— included percentage with tooth loss; percentage receiving public assistance, including food and heating assistance; percentage of the population in nursing homes; the Child Opportunity Index; and owner-occupied property values.

### Descriptive Characteristics: Ohio

**Table 1** provides comparative statistics for Ohio— overall and in the highest- and lowest-risk deciles of the LSI score. We show both the characteristics of the outcome measure and of the top 10 predictors in terms of variable importance.

The values for many predictors vary starkly between the highest-decile and lowest-decile areas. In the highest-decile neighborhoods, nearly half of residents were receiving public assistance (e.g., the Supplemental Nutrition Assistance Program, or SNAP), compared with only 3 percent of those in the lowest-decile neighborhoods. The percentage of residents receiving public assistance is a poverty measure, but it also reflects state policy: the average SNAP benefit in Ohio in 2015 was \$126 per month.<sup>57</sup> Similarly, property values are more than just a measure of relative affluence—they also reflect the legacy of racism, segregation, and historical redlining.<sup>58</sup> In the lowest-decile neighborhoods, property values were nearly 4.5 times higher than in the highest-decile neighborhoods.

**Figure 2. Variable Importance Rankings in the Pilot Ohio Model**



**Table 1. Characteristics of the Ohio population overall and in the highest- and lowest-risk tracts and correlation of top 10 predictors with LE**

	Correlation with Life Expectancy	Statewide Mean	Highest-risk LSI Score Decile Mean	Lowest-risk LSI Score Decile Mean
LE at birth, years, 2010–2015	[Outcome]	76.6	70.3	81.6
<b>Top predictors by variable importance</b>				
Tooth loss prevalence rate, %, 2018	–78%	18%	35%	9%
Public assistance received in past year, % of households, 2014	–77%	18%	46%	3%
Owner-occupied home value, median, \$, 2010–2014	69%	\$127,013	\$58,851	\$251,585
Nursing home residents, %, 2010	–13%	0.71%	0.72%	0.56%
Children claimed by two people in the year they were born, %, 1978–1985 <sup>a</sup>	68%	71%	41%	87%
Child Opportunity Index, mean, 2010, 2015	76%	2.44	1.01	3.74
Probability of reaching the top 20% income bracket (among children born in the same year), %, 1978–2015 <sup>a</sup>	75%	18%	4%	36%
Physical inactivity prevalence rate, %, 2018	–76%	28%	40%	19%
Medicaid enrollees, %, 2010–2014	–76%	20%	45%	5%
Adult incarceration rate, %, 2010	–5%	0.41%	0.73%	0.00%

Notes: Analysis includes 2,952 CTs in Ohio with an LSI score. Higher LSI scores indicate lower LE and higher social inequity.

<sup>a</sup> Opportunity Atlas measures drawn from an analysis of data spanning 1978–2015.<sup>56</sup>

Figure 3 provides maps of the tract-level outcome, LE at birth (on left), and LSI scores (on right) in Ohio. The maps identify, in bright yellow, areas in urban cores and in the southern Appalachian areas of Ohio where community-level social inequities are more prevalent.

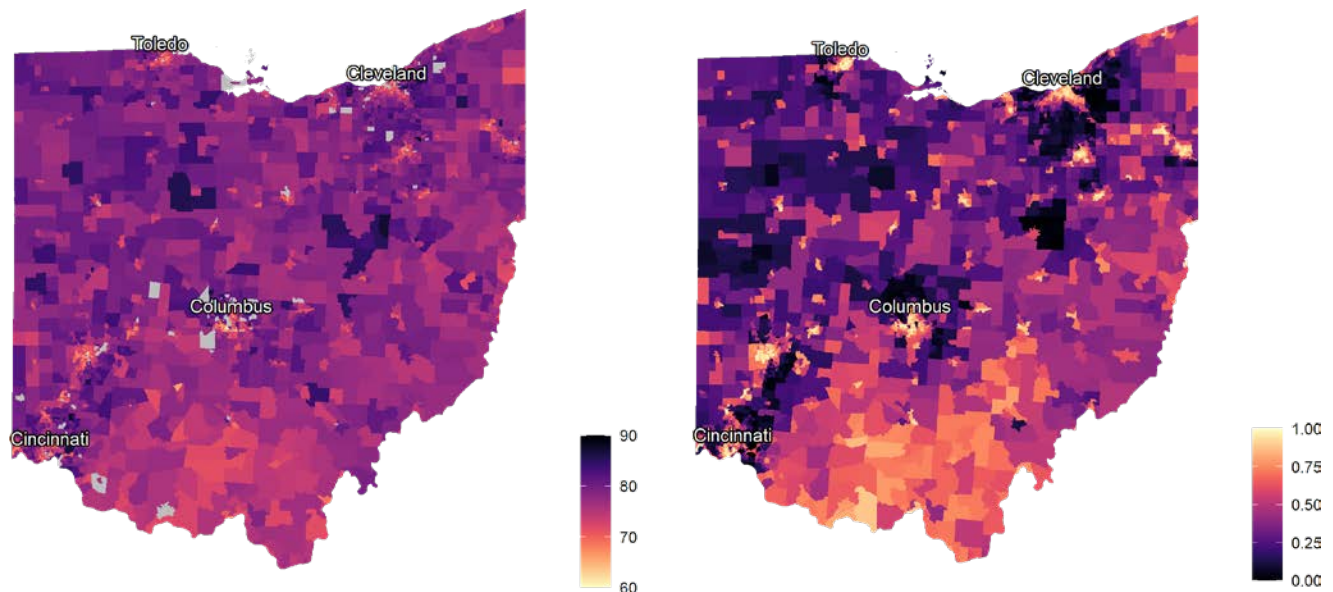
### Covariate Balance Between the Ohio CPC and Comparison Groups

Table 2 shows the covariate balance between the Ohio CPC and comparison groups for the study sample in the last baseline year for the overall study sample (2016). The analysis estimated propensity scores in each analysis year by using logistic regressions in which the dependent variable was an indicator of inclusion in the Ohio CPC group. Although the analysis calculated propensity scores in each analysis year, means and standardized differences were similar across years, so covariate balance results are presented for the last baseline year only. Before propensity score weighting, standardized differences were greater than 0.10 for some individual- and county-level characteristics. After propensity score weighting, standardized differences were all below the 0.10 threshold, indicating an acceptable level of covariate balance.

### Multivariate Modeling Results

Table 3 shows the estimates of the Ohio CPC model on total PBPM, inpatient admissions, ED visits, and readmissions for Medicaid beneficiaries attributed to Ohio CPC practices relative to comparison beneficiaries in a D-in-D model that uses the LSI score as a covariate. The “standard” column in Table 3 also shows D-in-D estimates for the same outcomes from a model that uses county-level covariates from the Area Health Resources Files. A negative value for the regression-adjusted D-in-D corresponds to a greater decrease or a smaller increase in an outcome after the implementation of Ohio CPC relative to the comparison group. A positive value corresponds to a greater increase or a smaller decrease in an outcome in the Ohio CPC group relative to the comparison group after Ohio CPC implementation. The relative difference is the D-in-D estimate as a percentage of the Ohio CPC baseline period adjusted mean.



**Figure 3. Maps of life expectancy (left) and local social inequity (right) by Ohio Census tract (n = 2,949)**

Source: Authors' analysis of life expectancy (LE) and social determinants data in Ohio Census tracts. Brighter yellow indicates shorter LE and higher local social inequity (LSI) scores; dark purple indicates longer LE and lower LSI scores.

The following were results from the model that included the LSI score:

1. Total spending per beneficiary per month (PBPM) increased for both Medicaid beneficiaries attributed to the Ohio CPC and the comparison group but increased by \$8.37 less (\$9.31 less in the standard model) for the Ohio CPC group during the first 2 years of Ohio CPC implementation ( $P < .001$ ).
2. Inpatient admissions decreased for both Medicaid beneficiaries attributed to Ohio CPC and the comparison group but decreased by 11.99 more admissions per 1,000 beneficiaries (11.85 more admissions per 1,000 beneficiaries in the standard model) for the Ohio CPC group during the first 2 years of Ohio CPC implementation ( $P < 0.001$ ).
3. ED visits decreased for the Ohio CPC group and increased for comparison beneficiaries, leading to a relative decrease of 24.80 visits per 1,000 beneficiaries (17.79 visits per 1,000 beneficiaries in the standard model) for Medicaid beneficiaries attributed to Ohio CPC practices ( $P < 0.001$ ).
4. Readmissions within 30 days of discharge increased for both Medicaid beneficiaries attributed to Ohio CPC and the comparison group but increased by 5.14 fewer readmissions

per 1,000 discharges (5.63 fewer readmissions per 1,000 discharges in the standard model) for the Ohio CPC group during the first 2 years of Ohio CPC implementation ( $P = 0.02$ ).

We compared the model that excluded the LSI score but included additional area-level covariates with the model that included the LSI score. The impact estimates remained significant for all outcomes, and including the single LSI score variable in lieu of several area-level covariates resulted in significant differences by LSI in Ohio CPC's impacts on total spending, inpatient admissions, ED visits, and readmissions. Because a D-in-D model with the LSI score was also more parsimonious than the model with area-level variables and produces more accurate estimates (in terms of narrower Confidence Intervals), there are likely further opportunities to apply the LSI score to impact evaluations.

## Discussion

Our LSI score, piloted in Ohio, explains 73 percent of the 29-year variance in LE at birth, with a mean squared error of 4.47. After developing the scores for Ohio, we validated them in several ways, including by evaluating their performance after linking with individual data in the SIM2 evaluation and by



**Table 2. Covariate balance between the Ohio Comprehensive Primary Care (OH CPC) and comparison groups in the last baseline year (2016)**

Variable	Unweighted mean or percentage, OH CPC	Unweighted mean or percentage, comparison group	Unweighted standardized difference	Weighted mean or percentage, comparison group	Weighted standardized difference
<b>Individual-level measures</b>					
Percentage of people who are female	54.3	53.3	0.02	54.4	0.001
Age in years	22.8	25	0.12	22.5	0.02
Percentage of people who are disabled	5.7	5.6	0.004	5.7	0.002
Percentage of people whose race is non-white	50.6	42.0	0.17	51.1	0.01
Total months of enrollment during year	11.0	10.8	0.07	11	0.02
CDPS risk score, logged <sup>a</sup>	-0.1	-0.1	0.05	-0.1	0.01
Percentage of people who had changes to attribution status across calendar quarters <sup>b</sup>	39.4	22.5	0.37	41.0	0.03
<b>ZCTA-level measures</b>					
LSI scores	0.70	0.70	0.04	0.70	0.0002
<b>County-level measures</b>					
Percentage of people residing in a Metropolitan Statistical Area	84.3	78.1	0.16	84.7	0.01
Hospital beds per 1,000 people (2015)	3.8	3.4	0.19	3.8	0.01
FQHCs per 1,000 people (2016)	< 0.001	< 0.001	0.10	< 0.001	0.004

Notes: CDPS = Chronic Illness and Disability Payment System; FQHC = Federally Qualified Health Center; ICD = International Classification of Diseases; LSI = local social inequity; ZCTA = ZIP code tabulation areas; OH CPC = Ohio Comprehensive Primary Care.

<sup>a</sup> CDPS score is a risk adjustment score calculated from ICD-9 and ICD-10 diagnosis codes included on hospital and outpatient claims, with larger CDPS scores corresponding to a larger number of comorbidities or a more severe set of comorbidities.

<sup>b</sup> The covariate for changes to attribution status indicates whether beneficiaries were attributed to both OH CPC practices and non-OH CPC practices while enrolled in OH Medicaid. OH CPC attribution was on a quarterly basis, and there was significant churn in attribution status, so this analysis assigned beneficiaries to the OH CPC and comparison groups based on each beneficiary's total length of time attributed to OH CPC versus non-OH CPC practices.

Source: Federal Evaluation Team analysis of Ohio Medicaid claims data from the OH Department of Medicaid.

benchmarking their performance against alternative, commonly used indices, which explain only 50 to 63 percent of variation in CT-level LE in Ohio. Thus, we find substantial improvement in variance explained relative to other composite measures of SDoH.

It is important to note that the measures against which we compare our performance were not developed to predict or explain LE, whereas ours was. However, we contend that if the comparison composite measures we identified were able to explain variance in LE or other population health outcomes, they could be more useful to the field. As Allik and colleagues wrote in 2020:

Comparisons of indicators and measures have shown that those with a strong foundation in theory are better able to explain variations in health... [A] combined measure should be better able to capture the unmeasured concept of deprivation than the individual indicators themselves. Empirically, this should be reflected in the composite measure having a stronger association to the outcome of interest than any of the variables on their own.<sup>28</sup>

Several prior studies have used RF models to predict LE. Meshram used RFs to model country-level LE on health outcomes and a few SDoH measures (schooling, GDP), and they compared the RF results

**Table 3. Differences in the pre–post change in total spending, inpatient admissions, readmissions, and ED visits, in Ohio CPC and the comparison group**

Outcome	Baseline period adjusted mean, OH CPC	Baseline period adjusted mean, comparison group	Intervention period adjusted mean, OH CPC	Intervention period adjusted mean, comparison group	LSI score: Regression-adjusted D-in-D (90% CI)	Standard: Regression-adjusted D-in-D (90% CI)	Relative difference (%)	p-value
<b>Total spending PBPM (\$)</b>								
Year 1	358.44	349.74	388.48	398.35	–18.63 (–23.84, –13.41)	–19.17 (–24.72, –13.63)	–5.2	< 0.001
Year 2	358.44	349.74	411.83	400.09	2.98 (–2.90, 8.86)	1.61 (–4.82, 8.05)	0.8	0.40
Overall	358.44	349.74	399.56	399.13	–8.37 (–12.28, –4.46)	–9.31 (–13.53, –5.09)	–2.3	< 0.001
<b>All-cause acute inpatient admissions per 1,000 population</b>								
Year 1	96.28	97.10	82.23	100.64	–17.60 (–21.06, –14.14)	–17.17 (–20.07, –14.28)	–18.3	< 0.001
Year 2	96.28	97.10	76.44	82.81	–5.77 (–11.78, 0.24)	–5.96 (–11.86, –0.07)	–6.0	0.11
Overall	96.28	97.10	79.48	92.66	–11.99 (–15.37, –8.60)	–11.85 (–15.04, –8.67)	–12.5	< 0.001
<b>ED visits per 1,000 population</b>								
Year 1	1,317.99	1,338.80	1,316.49	1,386.27	–48.19 (–62.50, –33.89)	–41.51 (–56.49, –26.53)	–3.7	< 0.001
Year 2	1,317.99	1,338.80	1,313.27	1,332.94	1.10 (–18.11, 20.32)	8.47 (–11.95, 28.89)	0.08	0.93
Overall	1,317.99	1,338.80	1,314.96	1,362.39	–24.80 (–36.61, –12.98)	–17.79 (–30.27, –5.30)	–1.9	< 0.001
<b>Readmissions within 30 days of discharge per 1,000 discharges</b>								
Year 1	122.47	124.53	165.48	177.88	–8.50 (–13.96, –3.03)	–8.85 (–14.00, –3.70)	–6.9	0.01
Year 2	122.47	124.53	187.70	191.70	–0.96 (–5.61, 3.68)	–1.82 (–6.36, 2.72)	–0.8	0.73
Overall	122.47	124.53	175.40	183.59	–5.14 (–8.80, –1.47)	–5.63 (–9.11, –2.15)	–4.2	0.02

Notes: LSI: Local Social Inequity; CI = Confidence Interval; D-in-D = difference-in-differences; ED = emergency department; OH CPC = Ohio Comprehensive Primary Care; PBPM = per beneficiary per month. The LSI score model fits the data slightly better than the area-level covariates model.

Source: Federal evaluation team analysis of Ohio Medicaid claims data from the Ohio Department of Medicaid.

with a decision tree model and OLS to find that their RF is the best performing model.<sup>59</sup> Schultz et. al. predicted LE using the individual items of the Frailty Index. They found that RF outperformed OLS and Elastic Net in terms of  $R^2$  and median error and that using the individual items of the Frailty Index was more accurate than the Frailty Index itself.<sup>60</sup> Makridis et al. compared a few ML models to (1) predict the well-being index composite score and (2) classify US military veterans as having high vs. low physical well-being.<sup>61</sup> They compared standard methods (OLS

and logistic regression) with ML methods (XGBoost and support vector machine).<sup>61</sup> For both model sets, they found XGBoost, a tree-based method (like RF) to perform the best, but their results “suggest that additional measures of socio-economic characteristics are required to better predict physical well-being, particularly among vulnerable groups.”<sup>61</sup>

## Limitations

Our findings should be interpreted in light of several limitations and caveats. First, we acknowledge the dangers of ecological fallacy inherent in small-area estimation. Individual-level data are needed to study individual-level outcomes. Thus, while the scores and data are useful on their own as measures of population health, they need to be linked to individual-level data to best make predictions about individuals, as we have demonstrated in the SIM2 evaluation. Similarly, bias may arise if the outcome of interest is from a different era than our outcome measure and predictor data. Changes in area-level characteristics over time are inevitable, and RF models cannot extrapolate a linear trend. We recognize the need to update the model and scores when new LE estimates and predictor data are released. Although it may be possible to generate LSI scores on an annual basis given an annual outcome measure, the current scores are limited to a single point in time and thus are not suitable for measuring progress in reducing health inequities.

One important limitation of this study is that we only estimate a predictive model, without regard to causality. Thus, the variable importance rankings should be interpreted as indicating what is driving the algorithm—not necessarily what is important to policy or practice. Certainly, some of the top predictors could potentially be amenable to interventions, but the goal of this study was not to identify actionable factors for interventions. It was to create an accurate, equity-focused composite measure of SDoH predicting a core population health outcome.

## Implications for Policy and Practice

SDoH are typically not accounted for in risk adjustment models, even though, like age, social factors influence health care outcomes and are not a consequence of the quality of healthcare delivery. When social risk factors are not adjusted for in these models, this can result in lower payments to providers who care for patients with higher social risk compared with providers who care for patients with lower social risk, when all else is equal.<sup>11,12</sup> This creates incentives to avoid caring for populations with social risk factors<sup>11</sup> which limits the ability of

VBP programs to address pervasive health inequities. Several recent publications have noted that most VBP models do not consider health inequities caused by social factors,<sup>12,62-64</sup> with the notable exception of the Medicare ACO REACH Model, which begins in 2023.<sup>65</sup> There is some evidence that certain models have actually exacerbated health inequities.<sup>66,67</sup> This past year, the CMS Innovation Center made an explicit goal of embedding health equity in all their healthcare models.<sup>16</sup> This will require comprehensive tools to measure and account for social risk factors in these payment models.

Including social factors in risk adjustment has been shown to reduce underpayments to providers that serve vulnerable populations.<sup>12</sup> There are examples of healthcare systems that adjust for social factors, including Massachusetts,<sup>68</sup> the UK, and New Zealand.<sup>69</sup> However, social risk adjustment remains controversial. Several recent articles have outlined arguments for and against social risk adjustment in value-based care.<sup>29,70-74</sup> These arguments are also reflected in a series of reports from the Department of Health and Human Services<sup>75,76</sup> and the National Academies of Sciences, Engineering, and Medicine (NASEM),<sup>11</sup> stemming from the 2014 Improving Medicare Post-Acute Care Transformation (IMPACT) Act, which required an examination of social risk and performance in VBP programs. NASEM has argued that payment adjustments can be made to avoid underpaying providers that serve socially at-risk patients by accounting for the increased resources, or estimated costs, that are required to care for more socially at-risk populations.<sup>11</sup> These resources can be used to invest in reducing those disparities while improving efficiency and quality of care.<sup>11</sup> NASEM also recommended that CT-level compositional measures of social risk factors can be used as a good proxy for individual-level data in calculations of these adjustments.<sup>11</sup> We believe that the LSI is such a measure and that it could be used both in payment adjustments and to identify populations for targeted investment.

The Department of Health and Human Services has acknowledged the need for greater payment support for improving care for patients with social risk, but unlike NASEM, they have not supported social risk

adjustment, arguing it would mask disparities.<sup>70</sup> We do not agree that social risk adjustment masks disparities. Adjusting for social factors does not remove the requirement to address those social factors; rather, it directs more funds to addressing those factors. It is also important to note that social risk adjustment is not likely to end health disparities on its own,<sup>77</sup> partly because unequal access to care may still result in underprediction of expenditures when social factors are included in risk adjustment models.<sup>9,77</sup> In addition to social risk adjustment, stratification<sup>11,75</sup> or post-adjustment<sup>29,77</sup> based on social factors can also be used to further ensure appropriate adjustment of payments for social factors.

Disparities in LE are related to numerous social factors.<sup>36,78</sup> Our approach to creating a composite SDoH index uses modern computing power to

provide a new measure that is robust, comprehensive, cross-disciplinary, spatially refined, and highly predictive of a health-equity-centered population health outcome. Considering a broader array of SDoH measures provides a more complete and higher-resolution picture of the relative risks of neighborhoods across Ohio. Accurate multidimensional, cross-sector, small-area social risk scores could be useful in understanding the impact of healthcare innovations, payment models, and interventions on social drivers of health on high-risk communities; identifying neighborhoods and areas at highest risk of poor outcomes for better targeting of interventions and resources; and accounting for factors outside of providers' control for more fair and equitable performance/quality measurement and reimbursement.

## References

1. Braveman P, Gottlieb L. The social determinants of health: it's time to consider the causes of the causes. *Public Health Rep* 2014;129(1 Suppl 2):19–31. <https://doi.org/10.1177/00333549141291S206>
2. O'Neill Hayes T, Delk R. Understanding the social determinants of health. Updated 2018 [cited 22 Aug 2022]. Available from: [https://www.americanactionforum.org/research/understanding-the-social-determinants-of-health/#\\_edn9](https://www.americanactionforum.org/research/understanding-the-social-determinants-of-health/#_edn9)
3. Marmot M, Friel S, Bell R, Houweling TAJ, Taylor S, WHO Commission on Social Determinants of Health World Health Organization. Closing the gap in a generation: health equity through action on the social determinants of health: Commission on Social Determinants of Health final report. *Lancet* 2008;372(9650):1661–9. [https://doi.org/10.1016/S0140-6736\(08\)61690-6](https://doi.org/10.1016/S0140-6736(08)61690-6)
4. Baciu A, Negussie Y, Geller A, Weinstein JN, National Academies of Sciences Engineering, and Medicine. The root causes of health inequity. *Communities in action: pathways to health equity*. National Academies Press; 2017. Available from: <https://www.ncbi.nlm.nih.gov/books/NBK425845/>
5. Alderwick H, Gottlieb LM. Meanings and misunderstandings: a social determinants of health lexicon for health care systems. *Milbank Q* 2019;97(2):407–19. <https://doi.org/10.1111/1468-0009.12390>
6. Muennig P, Fiscella K, Tancredi D, Franks P. The relative health burden of selected social and behavioral risk factors in the United States: implications for policy. *Am J Public Health* 2010;100(9):1758–64. <https://doi.org/10.2105/AJPH.2009.165019>
7. Braveman P. What are health disparities and health equity? We need to be clear. *Public Health Rep* 2014;129(Suppl 2):5–8. <https://doi.org/10.1177/00333549141291S203>
8. Figueroa JF, Frakt AB, Jha AK. Addressing social determinants of health: time for a polysocial risk score. *JAMA* 2020;323(16):1553–4. <https://doi.org/10.1001/jama.2020.2436>
9. Obermeyer Z, Powers B, Vogeli C, Mullainathan S. Dissecting racial bias in an algorithm used to manage the health of populations. *Science* 2019;366(6464):447–53. <https://doi.org/10.1126/science.aax2342>

10. Topchik M, Gross K, Pinette M, Brown T, Balfour B, Kein H. The rural health safety net under pressure: rural hospital vulnerability. The Chartis Group; 2020. Available from: <https://www.chartis.com/insights/rural-health-safety-net-under-pressure-rural-hospital-vulnerability>
11. National Academies of Sciences, Engineering, and Medicine. Accounting for social risk factors in medicare payment: data. Washington (DC): The National Academies Press; 2017.
12. Ash AS, Mick EO, Ellis RP, Kiefe CI, Allison JJ, Clark MA. Social determinants of health in managed care payment formulas. *JAMA Intern Med* 2017;177(10):1424–30. <https://doi.org/10.1001/jamainternmed.2017.3317>
13. HealthCare.gov. Risk adjustment. Available from: <https://www.healthcare.gov/glossary/risk-adjustment/>
14. Center for Medicare and Medicaid Services. Risk adjustment in quality measurement. 2021. Available from: <https://www.cms.gov/files/document/blueprint-risk-adjustment.pdf>
15. Irvin JA, Kondrich AA, Ko M, Rajpurkar P, Haghgoo B, Landon BE. Incorporating machine learning and social determinants of health indicators into prospective risk adjustment for health plan payments. *BMC Public Health* 2020;20(1):608. <https://doi.org/10.1186/s12889-020-08735-0>
16. Menzin J, Lines LM, Weiner DE, Neumann PJ, Nichols C, Rodriguez L, Agodoa I, Mayne T. A review of the costs and cost effectiveness of interventions in chronic kidney disease implications for policy. *Pharmacoeconomics* 2011;29(10):839–61. <https://doi.org/10.2307/2626958>
17. Guo W, DuGoff E. Including social risk factors in performance measurement: methods matter. *The Medical Care Blog* [Internet]; 2019. Available from: <https://www.themedicalcareblog.com/social-risk-factors/>
18. Lines L, Sokol S. Adjusting publicly reported performance measures for social risk factors. *The Medical Care Blog* [Internet]; 2020. Available from: <https://www.themedicalcareblog.com/adjusting-for-social-risk-factors/>
19. Lines L. The past, present, and future of risk adjustment: an interview with Arlene Ash. *The Medical Care Blog* [Internet]; 2018. Available from: <https://www.themedicalcareblog.com/risk-adjustment-interview-with-arlene-ash/>
20. Bynum J, Lewis V. Value-based payments and inaccurate risk adjustment—who is harmed? *JAMA Intern Med* 2018;178(11):1507–8. <https://doi.org/10.1001/jamainternmed.2018.4142>
21. Chen M, Tan X, Padman R. Social determinants of health in electronic health records and their impact on analysis and risk prediction: A systematic review. *J Am Med Inform Assoc* 2020;27(11):1764–73. <https://doi.org/10.1093/jamia/ocaa143>
22. Flanagan BE, Gregory EW, Hallisey EJ, Heitgerd JL, Lewis B. A social vulnerability index for disaster management. *J Homel Secur Emerg Manag* 2011;8(1). <https://doi.org/10.2202/1547-7355.1792>
23. Singh GK. Area deprivation and widening inequalities in US mortality, 1969-1998. *Am J Public Health* 2003;93(7):1137–43. <https://doi.org/10.2105/AJPH.93.7.1137>
24. Liao JM, Navathe AS. What comes next in prioritizing equity in payment? The ACO REACH Model [Internet]. *Health Affairs Forefront*; 2022. Available from: <https://www.healthaffairs.org/doi/10.1377/forefront.20220404.728371>
25. Liaw W, Krist AH, Tong ST, Sabo R, Hochheimer C, Rankin J. Living in “cold spot” communities is associated with poor health and health quality. *J Am Board Fam Med* 2018;31(3):342–50. <https://doi.org/10.3122/jabfm.2018.03.170421>
26. Raven MC, Billings JC, Goldfrank LR, Manheimer ED, Gourevitch MN. Medicaid patients at high risk for frequent hospital admission: real-time identification and remediable risks. *J Urban Health* 2009;86(2):230–41. <https://doi.org/10.1007/s11524-008-9336-1>
27. Butler DC, Petterson S, Phillips RL, Bazemore AW. Measures of social deprivation that predict health care access and need within a rural area of primary care service delivery. *Health Serv Res* 2013;48(2 Pt 1):539–59. <https://doi.org/10.1111/j.1475-6773.2012.01449.x>

28. Allik M, Leyland A, Travassos Ichihara MY, Dundas R. Creating small-area deprivation indices: a guide for stages and options. *J Epidemiol Community Health* 2020;74(1):20–5. <https://doi.org/10.1136/jech-2019-213255>
29. Phillips RL, Ostrovsky A, Bazemore AW. Adjusting Medicare payments for social risk to better support social needs [Internet]. *Health Affairs Forefront*; 2021. Available from: <https://www.healthaffairs.org/doi/10.1377/forefront.20210526.933567>
30. Chang HY, Hatef E, Ma X, Weiner JP, Kharrazi H. Impact of area deprivation index on the performance of claims-based risk-adjustment models in predicting health care costs and utilization. *Popul Health Manag* 2021 Jun;24(3):403–11.
31. Kind AJ, Jencks S, Brock J, Yu M, Bartels C, Ehlenbach W. Neighborhood socioeconomic disadvantage and 30-day rehospitalization: a retrospective cohort study. *Ann Intern Med* 2014;161(11):765–74. <https://doi.org/10.7326/M13-2946>
32. Liaw W, Krist AH, Tong ST, Sabo R, Hochheimer C, Rankin J. Living in “cold spot” communities is associated with poor health and health quality. *J Am Board Fam Med* 2018;31(3):342–50. <https://doi.org/10.3122/jabfm.2018.03.170421>
33. Link BG, Phelan J. Social conditions as fundamental causes of disease. *J Health Soc Behav* 1995;35(Spec No):80–94. <https://doi.org/10.2307/2626958>
34. Powell WR, Buckingham WR, Larson JL, Vilen L, Yu M, Salamat MS. Association of neighborhood-level disadvantage with Alzheimer disease neuropathology. *JAMA Netw Open* 2020;3(6):e207559–207559. <https://doi.org/10.1001/jamanetworkopen.2020.7559>
35. Sapra KJ, Yang W, Walczak NB, Cha SS. Identifying high-cost Medicare beneficiaries: impact of neighborhood socioeconomic disadvantage. *Popul Health Manag* 2020;23(1):12–9. <https://doi.org/10.1089/pop.2019.0016>
36. Singh GK, Daus GP, Allender M, Ramey CT, Martin EK, Perry C. Social determinants of health in the United States: addressing major health inequality trends for the nation, 1935–2016. *Int J MCH AIDS* 2017;6(2):139–64. <https://doi.org/10.21106/ijma.236>
37. Xu J, Kochanek KD, Murphy SL, Tejada-Vera B. Deaths: final data for 2014. *Natl Vital Stat Rep* 2016;65(4):1–122.
38. Tejada-Vera B, Bastian B, Arias E, Escobedo LA, Salent B. Life expectancy estimates by U.S. Census Tract, 2010–2015. Updated 2020 [cited 22 Aug 2022]. Available from: <https://www.cdc.gov/nchs/data-visualization/life-expectancy/index.html>
39. Smith MJ, Axler R, Bean S, Rudzicz F, Shaw J. Four equity considerations for the use of artificial intelligence in public health. *Bull World Health Organ* 2020;98(4):290–2. <https://doi.org/10.2471/BLT.19.237503>
40. Wellcome Trust. Ethical, social, and political challenges of artificial intelligence in health. London: Future Advocacy with the Wellcome Trust; 2019 [cited 22 Aug 2022]. Available from: <https://wellcome.org/sites/default/files/ai-in-health-ethical-social-political-challenges.pdf>
41. Kino S, Hsu YT, Shiba K, Chien YS, Mita C, Kawachi I. A scoping review on the use of machine learning in research on social determinants of health: trends and research prospects. *SSM Popul Health* 2021;15:100836. <https://doi.org/10.1016/j.ssmph.2021.100836>
42. Wiemken TL, Kelley RR. Machine learning in epidemiology and health outcomes research. *Annu Rev Public Health* 2020;41(1):21–36. <https://doi.org/10.1146/annurev-publhealth-040119-094437>
43. Kaur M, Dhalaria M, Sharma PK, Park JH. Supervised machine-learning predictive analytics for National Quality of Life scoring. *Appl Sci (Basel)* 2019;9(8):1613. <https://doi.org/10.3390/app9081613>
44. OECD. Life expectancy at birth (indicator). 2022 [cited 26 Aug 2022]. <https://doi.org/10.1787/27e0fc9d-en>
45. Kilbourne AM, Switzer G, Hyman K, Crowley-Matoka M, Fine MJ. Advancing health disparities research within the health care system: a conceptual framework. *Am J Public Health* 2006;96(12):2113–21. <https://doi.org/10.2105/AJPH.2005.077628>
46. Office of Disease Prevention and Health Promotion. Explore resources related to the social determinants of health. [Cited 22 Aug 2022]. Available from: <https://www.healthypeople.gov/2020/topics-objectives/topic/social-determinants-health/interventions-resources>
47. Ishwaran H, Kogalur UB. Fast unified random forests for survival, regression, and classification (RF-SRC). Updated 2021 [cited 22 Aug 2022]; Available from: <http://www.maths.bristol.ac.uk/R/web/packages/randomForestSRC/randomForestSRC.pdf>

48. R Core Team. R: a language and environment for statistical computing. Vienna (Austria): R Foundation for Statistical Computing; 2021 [cited 22 Aug 2022]. Available from: <https://www.R-project.org/>
49. Ishwaran H. Variable importance in binary regression trees and forests. *Electronic J Stats* 2007;1(none):519–37. <https://doi.org/10.1214/07-EJS039>
50. Ishwaran H, Kogalur UB, Blackstone EH, Lauer MS. Random survival forests. *Ann Appl Stat* 2008;2(3):841–60. <https://doi.org/10.1214/08-AOAS169>
51. Bivand RS, Pebesma EJ, Gómez-Rubio V, Pebesma EJ. *Applied spatial data analysis with R*. Springer; 2008.
52. Chern J, Wasserman J. zcta: 2010 Census ZCTA Relationship File. 2021 [cited 22 Aug 2022]. Available from: <https://github.com/jjchern/zcta>
53. Tucker-Seeley RD, Li Y, Sorensen G, Subramanian SV. Lifecourse socioeconomic circumstances and multimorbidity among older adults. *BMC Public Health* 2011;11:313. <https://doi.org/10.1186/1471-2458-11-313>
54. CMS Innovation Center. Appendix H: State Innovation Model in model test states: Ohio. Updated 2021 [cited 22 Aug 2022]. Available from: <https://innovation.cms.gov/data-and-reports/2021/sim-rd2-test-final-appendix>
55. Ohio Colleges of Medicine Government Resources Center. State Innovation Model self-assessment: technical report. Updated March 2019 [cited 15 Aug 2022]. Available from: <https://grc.osu.edu/Projects/MEDTAPP/SIMSelfEvaluation>
56. Chetty R, Friedman JN, Hendren N, Jones MR, Porter SR. The opportunity atlas: Mapping the childhood roots of social mobility. National Bureau of Economic Research; 2018. <https://doi.org/10.3386/w25147>
57. Ohio Department of Job and Family Services. Public assistance monthly statistics report. 2016 Jul [cited 2022 Aug. 22]. Available from: <https://jfs.ohio.gov/pams/>
58. Krieger N, Waterman PD, Spasojevic J, Li W, Maduro G, Van Wye G. Public health monitoring of privilege and deprivation with the index of concentration at the extremes. *Am J Public Health* 2016;106(2):256–63. <https://doi.org/10.2105/AJPH.2015.302955>
59. Meshram SS. Comparative analysis of life expectancy between developed and developing countries using machine learning. 2020 IEEE Bombay Section Signature Conference (IBSSC)/IEEE; 2020:6–10. <https://doi.org/10.1109/IBSSC51096.2020.9332159>
60. Schultz MB, Kane AE, Mitchell SJ, MacArthur MR, Warner E, Vogel DS. Age and life expectancy clocks based on machine learning analysis of mouse frailty. *Nat Commun* 2020;11(1):1–12. <https://doi.org/10.1038/s41467-020-18446-0>
61. Makridis CA, Zhao DY, Bejan CA, Alterovitz G. Leveraging machine learning to characterize the role of socio-economic determinants on physical health and well-being among veterans. *Comput Biol Med* 2021;133:104354. <https://doi.org/10.1016/j.combiomed.2021.104354>
62. Werner RM, Emanuel EJ, Pham HH, Navathe AS. The future of value-based payment: a road map to 2030. 2021. Available from: <https://ldi.upenn.edu/our-work/research-updates/the-future-of-value-based-payment-a-road-map-to-2030>
63. Liao JM, Lavizzo-Mourey RJ, Navathe AS. A national goal to advance health equity through value-based payment. *JAMA* 2021;325(24):2439–40. <https://doi.org/10.1001/jama.2021.8562>
64. Sandhu S, Saunders R, McClellan M, Wong C. Health equity should be a key value in value-based payment and delivery reform. *Health Affairs Blog* [Internet]; 2020. <https://www.healthaffairs.org/doi/10.1377/forefront.20201119.836369/full>
65. Halpern MT, Holden DJ. Disparities in timeliness of care for U.S. Medicare patients diagnosed with cancer. *Curr Oncol* 2012;19(6):404–13. <https://doi.org/10.3747/co.19.1073>
66. Yasaitis LC, Pajeroski W, Polsky D, Werner RM. Physicians' participation in ACOs is lower in places with vulnerable populations than in more affluent communities. *Health Aff (Millwood)* 2016;35(8):1382–90. <https://doi.org/10.1377/hlthaff.2015.1635>
67. Liao JM, Huang Q, Ibrahim SA, Connolly J, Cousins DS, Zhu J. Between-community low-income status and inclusion in mandatory bundled payments in Medicare's comprehensive care for joint replacement model. *JAMA Netw Open* 2021;4(3):e211016. <https://doi.org/10.1001/jamanetworkopen.2021.1016>



68. Lines LM, Cohen J, Halpern MT, Smith AW, Kent EE. Care experiences among dually enrolled older adults with cancer: SEER-CAHPS, 2005-2013. *Cancer Causes Control* 2019;30(10):1137-44. <https://doi.org/10.1007/s10552-019-01218-7>
69. Huffstetler AN, Phillips RL Jr. Payment structures that support social care integration with clinical care: social deprivation indices and novel payment models. *Am J Prev Med* 2019;57(6 Suppl 1):S82-8. <https://doi.org/10.1016/j.amepre.2019.07.011>
70. Sheingold S, Zuckerman R, DeLew N, Sommers B. Health equity and value-based payment systems: moving beyond social risk adjustment. *Health Affairs Blog*; 2021. <https://www.healthaffairs.org/doi/10.1377/forefront.20210726.546811>
71. Nerenz DR, Austin JM, Deutscher D, Maddox KE, Nuccio EJ, Teigland C. Adjusting quality measures for social risk factors can promote equity in health care. *Health Aff (Millwood)* 2021;40(4):637-44. <https://doi.org/10.1377/hlthaff.2020.01764>
72. Tran LD. Social risk adjustment in health care performance measures. *JAMA Netw Open* 2020;3(6):e208020. <https://doi.org/10.1001/jamanetworkopen.2020.8020>
73. Fiscella K, Burstin HR, Nerenz DR. Quality measures and sociodemographic risk factors: to adjust or not to adjust. *JAMA* 2014;312(24):2615-6. <https://doi.org/10.1001/jama.2014.15372>
74. Joynt KE, De Lew N, Sheingold SH, Conway PH, Goodrich K, Epstein AM. Should Medicare value-based purchasing take social risk into account? *N Engl J Med* 2017;376(6):510-3. <https://doi.org/10.1056/NEJMp1616278>
75. Cubanski J, Koma W, Damico A, Neuman T. How many seniors live in poverty? Kaiser Family Foundation; 2018. <https://www.kff.org/report-section/how-many-seniors-live-in-poverty-issue-brief/>
76. Gold MR, Sofaer S, Siegelberg T. Medicare and cost-effectiveness analysis: time to ask the taxpayers. *Health Aff (Millwood)* 2007;26(5):1399-406. <https://doi.org/10.1377/hlthaff.26.5.1399>
77. Damberg CL, Elliott MN, Ewing BA. Pay-for-performance schemes that use patient and provider categories would reduce payment disparities. *Health Aff (Millwood)* 2015;34(1):134-42. <https://doi.org/10.1377/hlthaff.2014.0386>
78. Dimick J, Ruhter J, Sarrazin MV, Birkmeyer JD. Black patients more likely than whites to undergo surgery at low-quality hospitals in segregated regions. *Health Aff (Millwood)* 2013;32(6):1046-53. <https://doi.org/10.1377/hlthaff.2011.1365>

RTI International is an independent, nonprofit research institute dedicated to improving the human condition. We combine scientific rigor and technical expertise in social and laboratory sciences, engineering, and international development to deliver solutions to the critical needs of clients worldwide.

[www.rti.org/rtipress](http://www.rti.org/rtipress)

**RTI Press publication RR-0047-2209**