

Full Proposal Narrative

Identifying Information

Project Title: Building a Federal Social Risk Assessment Capacity to Support Clinical, Community, and Resource Responses (BaSiC SUCCoR)

Application I.D.: 109840

Applicant Name: Robert L. Phillips, Jr MD MSPH

Legal Name of Applicant Organization: ABFM Foundation, Inc

PROBLEM STATEMENT

The Department of Health and Human Services has struggled for seven years to respond to a charge from Congress to develop a process for adjusting clinical payments for social risk. Health care clinicians and public health have similarly struggled to find a reliable and low-burden process for identifying their patients and populations at increased health risk due to social factors. An aligned solution is needed so that clinical capacity, public health, and payment policy support, identify and address social needs that impact health and wellbeing. Area-based socioeconomic indices are used in other developed countries to direct social service and health care resources with the goal of improving health equity. These indices derive from empiric assessment of social determinants of health as contributors to health inequities. In the U.S., a handful of similarly developed “deprivation” indices are highly associated with a range of health outcomes and are increasingly used in federally funded research to understand health disparities and guide interventions but are not yet in broad use for clinical or public health applications, or for payment policy. A primary barrier to the use of these measures is that there hasn’t been a clear and transparent process for developing these measures in conjunction with stakeholders, and the equitable use of these measures across key subgroups of the population has not been validated. A partnership with the U.S. Census Bureau could help to produce one or more small area measures of social deprivation with clinical/public health/policy utility, low-burden implementation, and secure an ongoing data steward capacity for regular updating and validation of the measure or measures.

PROJECT

Objectives and Activities

The American Board of Family Medicine Foundation (ABFM), the Stanford Center for Population Health Sciences (PHS), and the U.S. Census Bureau are collaborating to use highly sensitive data to identify the area-based socioeconomic measure or measures that have the least variability across diverse population groups for prediction of health outcomes, for use in a range of public policy applications.¹ While many such measures currently exist, our project addresses a tremendous current barrier for implementation – lack of validation data using individual-based measures. The ABFM and Stanford PHS have more than two decades of experience with

research and the development of area-based measures, and in testing their predictive value in assessing social risks and their relationship to health outcomes.²⁻¹¹ The ABFM has an ongoing collaboration with the Centers for Medicare & Medicaid Services (CMS) and other health agencies to develop a policy proposal for adjusting payments using small, area-based deprivation indices.⁷ The ABFM has collaborated with large Federally Qualified Health Center research networks to test their utility in practice and developed an electronic health record (EHR)- and registry-linked tool to help practices and health systems use these indices in a low-burden manner that fosters Community Oriented Primary Care and Community Health Needs Assessments.^{6,12-17}

Our approach offers a low-burden, reliable, transparent way to identify the best area-based socioeconomic measure to adjust payments and to identify patients and populations with increased social risk who are likely to have related social needs. The focus is on validating the measure that has the highest public health, policy, and clinical utility. Importantly, our project is focused on not just a one-time recommendation but developing a process for creating a federal data steward for renewably testing, improving, maintaining, and sharing area-based measures for public use as a function of the Federal Data Strategy and/or recent Executive Orders to improve population health data. As part of this approach, we will transform our validation outcome data into a common data model to allow stakeholders to replicate our validation approach.

Our objective is to produce one or more Social Determinants of Health (SDOH) indices that are the least biased for predicting: 1) quality measures (PRIME Registry electronic health record data); 2) area prevalence of conditions, quality, utilization, costs (claims data); and 3) mortality (state-collected death data or derivatives). Our approach to assessing bias is based on a similar degree of prediction between individual and area-based measures for key subgroups of the population based on race, ethnicity, age, geography, rurality, and gender.

This research is needed to improve existing area-based measures.^{2,18-20} It will produce empirically based and theoretically informed recommendations for the area-based measure that best captures social inequalities in health and health-related outcomes in an unbiased way across subgroups of the population. It will help clinicians by reducing burden and increasing pre-test probability for guiding assessment of social need at patient visits.⁷ It will help health systems and public health departments by enabling a population-based approach to identifying patients likely to have social needs and working with community-based organizations to address them.²¹ In a like manner, it could enhance capacity for community health needs assessments required of non-profit hospitals, accredited public health departments, and federally qualified

health centers. It will help payers by offering a reliable mechanism for adjusting payments, monitoring, and understanding health equity, and reducing risk-score gaming.^{7,22,23} It will help bridge primary care and public health by identifying neighborhoods with social disadvantage and for which they have shared responsibility.²¹

A good example that our project will be addressing is the reticence of most state Medicaid programs to follow Massachusetts's lead in using an area-based index to adjust managed care payments, even as nearly 40 states have made collection of social determinants data a requirement.^{20,24,25} Collection of individual patient social needs broadly has proven to be a burden, inconsistent, and at-risk and unstable for missing the most disadvantaged patients, and frustrating for clinicians if they don't have aligned resources for addressing needs, thus motivating our focus on area-based measures.^{26–28}

Deliverables and Outcomes

This project will have immediate impacts for producing quantitative evidence for which area-based socioeconomic measures most consistently predict outcomes of interest across both pre-specified, theoretically informed groups.

Our primary deliverable will be a detailed report, suitable for circulation and dissemination across a wide variety of stakeholder groups, describing in detail the results of the exhaustive validation approaches of each of the assessed area-based socioeconomic measures. The report will focus on the rationale and data behind our recommended measure. We will also include all of the scripts required to replicate our analyses with data. In order to allow replication of our findings with private data, we will convert our data into a common data model, to allow interested users to run scripts on our data without actually have to access the data. This will be achieved through OMOP transformation of the EHR data.

The second deliverable will be to facilitate a federal data steward for renewably testing, improving, maintaining, and sharing area-based measures for public use. This process is designed to be consistent with the Federal Data Strategy and recent Executive Orders to improve population health data. Having a federal steward for managing these social deprivation/risk measures ensures that they are created from the most reliable and precise data and are most up to date. It also reduces the risk of having a variety of different, possibly proprietary, processes, that are not aligned.

Research Method

The project will utilize five high quality sources of data that will enable and empower our team to apply a variety of models to test the predicting power of SDOH indices on outcomes, including

both traditional parametric models, and non-parametric machine learning models First, all area-based socioeconomic measures will be created from public use U.S. Census data, in order to have a firmly established sample that is of the highest quality possible for the assessment of small area socioeconomic measures that are collected without sampling bias. We will assess individual and census group-level data for their ability to improve index prediction capacity. Second, we will use individual-level U.S. Census data accessed within the Stanford and University of Kentucky Census Federal Statistical Research Data Centers (FSRDCs). This will enable us to compare individual socioeconomic characteristics to the area-based measures in terms of prediction and bias with health and wellbeing outcomes. We have previously worked in the FSRDC and have approval for the current project (see letter of support). Third, we will utilize EHR data that has already been collected as part of the PRIME Registry by the ABFM. The registry focuses on data from nearly 800 small primary care practices, and the data is managed by the Stanford PHS, as part of a collaboration with the ABFM. This PRIME Registry data is the largest outpatient clinical registry open to all primary care clinicians in the United States. There were several important considerations in our choice of health outcome data to use for our validation. First, of critical importance were the geographic spread of the population across the U.S., in order to increase the validity of our approach across all U.S. populations. Secondly, we needed to use data that was large enough to be adequately powered to assess heterogeneity across subpopulations based on race, ethnicity, age, census region, urban/rural designations, and gender. Finally, we wanted to select a source of data that was among the most commonly collected and appropriate to health care organizations and governmental agencies. We refer to the research version of this data as the American Family Cohort (AFC). Fourth, we will use Medicare and commercial claims data that afford the smallest geographical assessment of health care costs and utilization. Fifth, we will use National Death Index or public use derivatives to assess relationships between index constructs and mortality.

We propose to accomplish this collaborative research project through achieving the following goals over the two years of the proposed project. In the first year, we will create the proposed range of area-based socioeconomic measures from publicly available U.S. Census data. Our analyses of measures are based initially on the most commonly used area deprivation indexes currently in use in the U.S.. First, the Area Deprivation Index (ADI), composed of income, education, employment, housing quality. Second, the CDC Social Vulnerability Index (SVI), composed of socioeconomic status (below poverty, unemployed, income, no high school diploma); Household composition & disability (aged 65 or older, aged 17 or younger, older than

age 5 with a disability, single parent); Minority status (minority, speak English “less than well”); Housing type & transportation (multi-unit structures, mobile homes, crowding, no vehicle, group quarters). Thirdly, the social deprivation index (SDI), composed of percent living in poverty, percent with less than 12 years of education, percent single-parent household, percent living in rented housing unit, percent living in overcrowded housing unit, percent of households without a car, and percent non-employed adults under 65 years of age. While some of these measures are used at only one level of aggregation, in order to do the most comprehensive comparison possible, we will create these measures at four levels of geography: county, Zip code tabulation area, census tract, and census block group. In addition to the multicomponent indexes listed in the table below, we will also examine single item education and poverty measures. Specifically, percent of the population with a high school Diploma, percent of the population living below poverty, and percent of the population living below 185% of poverty (Federal definition).

We will evaluate the area socioeconomic measures (as compared to individual measures) based on associations with healthcare utilization and health outcomes, as described in Table 1 below. These measures are already available from AFC data. We have previously demonstrated that quality measures are associated with a better SDI score.⁴

Table 1: The Preventive Care, Quality Measures and Outcomes examined; AFC

NQF#	CMS#	Measure Title
18	CMS165	Controlling High Blood Pressure
59	CMS122	Diabetes: Hemoglobin A1c (HbA1c) Poor Control (> 9%)
N/A	N/A	Provider Level Continuity Measure
2372	CMS125	Breast Cancer Screening
32	CMS124	Cervical Cancer Screening
34	CMS130	Colorectal Cancer Screening
38	CMS177	Childhood Immunization Status
0041/0041e	CMS147	Preventive Care and Screening: Influenza Immunization
N/A	CMS127	Pneumococcal Vaccination Status for Older Adults
55	CMS131	Diabetes: Eye Exam
0081/0081e	CMS135	Heart Failure (HF): Angiotensin-Converting Enzyme (ACE) Inhibitor or Angiotensin Receptor Blocker (ARB) Therapy for Left Ventricular Systolic Dysfunction (LVSD)
N/A	CMS347	Statin Therapy for the Prevention and Treatment of Cardiovascular Disease
1392	N/A	Well Child Visits in the first 15 months of life

1516	N/A	Well Child Visits in 3rd, 4th, 5th, and 6th yrs of life
46	N/A	Screening for Osteoporosis for Women Aged 65-85 Years of Age

We will also explore the use Medicare and Optum claims data and state All Payer Claims data to assess claims-based quality measures (Table 2), total costs of care, and hospitalization rates.

Table 2: The Preventive Care, Quality Measures and Outcomes examined; Medicare

NQF#	CMS#	Measure Title
2512	0032	Cervical Cancer Screening (CCS-AD)
2372	4005	Breast Cancer Screening
0057	2838	Comprehensive Diabetes Care: Hemoglobin A1c (HbA1c) Testing (HA1C-AD)
0055	4020	Diabetes Care Eye Exam
0062	4021	Diabetes Care Kidney Disease Monitoring
0041	439	Preventive Care and Screening: Influenza Immunization

We will appeal to Census and CDC to match National Death Index (NDI) data with Census and American Community Survey respondents in order to assess age at death for the three indexes and as outcome for index construct from individual SDOH elements. Failing access to the NDI we will use county-level, 5-year mortality rates (2014-2018) obtained from HDPulse.

In the second year of the project, we will examine the prediction of all previous outcomes examined by population subgroup. The first approach will be an *a priori* specified stratified models to examine whether individual-level versus area-level prediction differs among different population subgroups. We will first evaluate the correlation structure of the area-based socioeconomic measures within strata of race, ethnicity, age, census division, urban/rural designations and gender will form a basis for a descriptive understanding of the measures across populations. The purpose of this analysis is to understand how similar the indexes are to each other in the population overall and within subgroups of the population. We will conduct a counterfactual simulation using the area-based socioeconomic factors measures to adjust primary care payment. The more sufficient reimbursement on primary care for targeted patients who are facing tough social determinants of health will lead to better outcomes and higher value of the healthcare delivery system. The counterfactual simulation will quantify the value of such intervention to inform CMS payment policy.

Through these research methods, we will produce empirically based and theoretically informed recommendations for the area-based measure that best captures social inequalities in health

and health-related outcomes in an unbiased way across subgroups of the population. The recommendations can be used by a broad range of constituencies including health care systems, physicians, payers, and health departments.

Communications

The goal of our communications strategy is to involve stakeholders as part of the process for building the measure in order to facilitate the broad implementation of the measure at the completion of the project. We will conduct an initial set of focus groups with key stakeholders, including the consultants listed below, in terms of our process and deliverables. Our primary output for communications will be the report recommending a particular measure and presenting detailed justification for that recommendation. We will produce a full, detailed report, including code for replication, as well as a more accessible short version of the report for broader dissemination. We will also produce two peer-reviewed publications detailing the findings of our project and describing their applicability for clinical, public health, and policy use. We will also work with RWJF on related briefs or blogs that summarize project objectives and findings.

Staffing

ABFM and The Center for Professionalism & Value in Health Care

Robert L. Phillips, Jr, MD, MSPH is Founding Executive Director, The Center for Professionalism & Value in Health Care. He will serve as principal investigator (15% FTE). He is currently PI for a related effort to help federal stakeholders develop a policy for adjusting clinical payments for social risk (supported by the Commonwealth Fund, Arnold Ventures, the Samuelli Foundation, and 3M). Dr. Phillips is also PI on current projects with the Centers for Disease Control and Prevention and the Office of the National Coordinator for Health Information Technology. Dr. Phillips recently served as co-chair of the Population Health Subcommittee and as a member of the executive committee of the National Committee for Vital and Health Statistics (advisory committee staffed by CDC's National Center for Health Statistics). He recently co-chaired the Committee on Implementing High-Quality Primary Care for the National Academies of Science, Engineering, and Medicine. He served as a Fulbright Specialist to the Netherlands in 2012 and New Zealand in 2016. Dr. Phillips was elected to the National Academy of Medicine in 2010. Dr. Phillips' research and policy experience led him to his selection by the Secretary of Health and Human Services to serve on a Federal Negotiated Rule Making Committee for the redesignation of shortage and underservice areas. **Andrew W. Bazemore, MD, MPH** will assist in coordinating research efforts between the ABFM research team and Stanford (10% FTE). He is the Senior Vice President of Research and Policy for the ABFM and co-director of The Center for Professionalism & Value in Health Care. Dr. Bazemore previously served as the Director of the Robert Graham Center for Policy Studies in Family

Medicine in Washington, DC. He developed HealthLandscape, an innovative data engagement platform entirely funded by grants and contracts, including the nearly \$1.5 million/year federal Uniform Data System (UDS) Mapper contract that guides funding for all the nation's Federally Qualified Centers. He was elected as a member of the National Academies of Medicine in 2016. Drs. Phillips and Bazemore will lead efforts with Census and CDC to reach agreement on data and index stewardship. **Zhou Yang, PhD** is the Senior Health Economist at the ABFM and is leading its research portfolio for Artificial Intelligence/Machine Learning (15% FTE). She will work with Research Analyst, **Zachary Morgan, MS** (10% FTE) in the University of Kentucky Federal Statistical Research Data Center (FSRDC) to apply this methodology to testing relationships between health care quality and outcomes and SDOH constructs. The ABFM research team will also assist with analytic design, output review, and index selection. **Mikel Severson** is The Center for Professionalism & Value in Health Care Administrator, and she will assist with project logistics coordination (2.5% FTE). An ABFM Research Coordinator (2.5% in Year 1, and 5% in Year 2 FTE) will assist with project coordination.

Stanford Center for Population Health Sciences

David Rehkopf, ScD, MPH is an Associate Professor of Epidemiology and Population Health, Medicine, and Sociology and Director of the Stanford Center for Population Health Sciences (10% FTE). He was a Robert Wood Johnson Foundation Health and Society Scholar (Cohort 4). He has two decades of experience in developing and applying socioeconomic measures in research. For the past two years he has worked with the team at the ABFM to transform the ABFM PRIME Registry data to data for research. **Ayin Vala, MS** (5% FTE) is the lead data manager and subject matter expert on the AFC data. He was a fellow at Google, Chief Data Scientist at Foundation for Precision Medicine, and Head of Data Science at April health. He has expertise in cloud computing, machine learning, and natural language processing of healthcare data. **Isabella Chu, MPH** (15% FTE) will assist with regulatory and contractual processes for data acquisition and use and in the documentation and preparation of data for research use. **Postdoctoral scholar** (75% FTE Year 1, 100% FTE in Year 2). The scholar will be under the direct supervision of Dr. Rehkopf and will work at the FSRDC on the analysis of data to complete the goals of the project.

U.S. Census Bureau (Consultant)

Victoria Udelova, PhD is a research economist at the Center for Economic Studies at the U.S. Census Bureau (see letter of support). She leads the Census Bureau's initiative to combine health records with Census data to expand research possibilities, particularly around social determinants of health. She will contribute to analytic design and analyses, output review, index selection, and discussions about index stewardship. She also enables data use authorization for

the project for Census and Census-managed data. The agreement between Census, Stanford, and the ABFM is a formal project of the Enhancing Health Data (EHealth) Program which has data sharing and cost-sharing parameters already established by agreement. The proposed costs reflect the terms of this agreement and the U.S. Census Bureau is able to receive external funding by agreement.

Additional Project Consultants

We will rely heavily on input from three outside project consultants, who have agreed to support our project *pro bono* (see relevant letters of support). **Donald M Berkwick, MD** is a former Administrator of the Centers for Medicare & Medicaid Services and has substantial experience in working to identify patients with social needs and to ensure sufficient payment to address those needs. He will contribute his experience to helping with the development and testing of the deprivation indexes, as well as helping with the transition to implementation by the Center for Medicare Services. **Richard Gilfillan, MD** was the first director of the Center for Medicare and Medicaid Innovation (CMMI), where he worked with payers and providers to develop innovative models for improving patient care and reducing costs, and was recently CEO of Trinity Health, which focused on the mission of serving communities, not just patients. In addition to reviewing our findings, he will offer key translation advice on how different federal agencies and payers might use the tool to shape payment and practice policies. **Xingyou Zhang, PhD** is a researcher at the Bureau of Labor Statistics who previously served as the Chief of the U.S. Census Bureau Small Area Methods Branch. Prior to his work at the Census, he spent seven years with the U.S. Centers for Disease Control and Prevention as a geographer and statistician, contributing substantially to the 500 Cities Project. He has also led work with various foundations and organizations to validate small area measures. He will meet regularly with the study investigators as an outside consultant contributing his substantial experience on both the statistical and implementation side of creating small area measures.

Timeline

We propose a 2-year project timeline beginning January 1, 2022 and concluding December 31, 2023. In the first year of the project, we will complete the linkage of health records and outcomes to the individual-level census data, we will create the area-based deprivation measures, and we will test the prediction of area and individual deprivation measures with the health outcomes. In the second year of the project, we will assess heterogeneity of association between area and individual deprivation measures and health outcomes by key social and demographic characteristics. We will also complete the reports and articles for communicating our recommendations.

STRATEGY

The proposed project directly addresses the primary barriers that exist for the implementation of area-based measures for cost adjustment in the United States. These barriers have been identified through ongoing discussions with the multiple stakeholders involved in the implementation and use of area-based measures for cost adjustment. Our strategy is thus to use our data access and partnerships to directly address these barriers to implementation, with a focus on transparency of the validation process, community stakeholder input on the process, and a process centered on health equity. We define success as meeting our specific deliverables for this proposal, which are both technical and translational in content. We define success as meaningful engagement with stakeholders throughout the process, and the production of data and theory-based recommendations for the use of an area-based socioeconomic measure. Our specific, measurable objectives are: 1) linkage of AFC data to the U.S. Census, 2) validation of the area versus individual measures, 3) report on biases of measures by key social groups, 4) documentation of the process of stakeholder engagement, 5) production of recommendations for the use of the area-based measure, including scripts and data for replication. The current project is focused on the Foundation's goals of building a culture of health. For this project, we view a more equitable, transparent, and agreed upon area-based socioeconomic measure for payment adjustment as fundamental to this goal. We expect the project to be sustained after the grant period by developing an approach to continue to validate and assess bias in the measure over time.

RISK AND CHALLENGES

We do not foresee any financial risks or reputational risks as part of the project but do identify several implementation risks based on our past experience. One potential challenge is the delay in availability of Census Data. While possible, we have sought to address this issue by completing the agreement with the Census and establishing the protocol for data transfer and linkage. In addition, this is a high-priority project for the Census, so it will receive their full attention. A second related issue is the potential for delay in permissions to use claims data within the FSRDC. In this case, if there are unexpected delays, we can work to analyze the outcomes in the AFC data, and begin writing code and preparing for the claims data in order to complete this analysis in the second year of the project rather than the first, as anticipated. This would then not be expected to delay the production of the report and publications by the end of year 2. A more substantive challenge is what our process would be if multiple measures were determined to be equally equitable and predictive of health outcomes, or very similar. In this case, we would fully present this in our report but then seek input from stakeholders and advisors as to which measure is the easiest to understand, which we expect to be the measure with the fewest measures combined.