

# Enhancing Patient-Centered Outcomes Research (PCOR): Creating a National Small-Area Social Determinants of Health (SDOH) Data Platform

## Final Report

### 1. Overview and Objectives

This section should describe the purpose of the project and list the key objectives as well as the key deliverables, based on the original statement of work.

As described in the original statement of work, the goal of the project is to develop a consolidated set of national standardized databases on valid and reliable SDOH factors at the small-area and other geographic levels, building on existing databases developed by federal agencies.

The overall objectives of the project are to:

1. Conduct an environmental scan to identify a comprehensive set of U.S. Department of Health and Human Services' (HHS) and other federal datasets with existing or analyzable small-area level and other geographic level data on SDOH variables.
2. Design and create a publicly available data platform of valid and reliable standardized set of SDOH data sources at various geographic areas.
3. Coordinate and expand the data collection efforts on SDOH across HHS.
4. Use the new data platform to conduct a minimum of two patient-centered outcomes research (PCOR) studies.
5. Disseminate the SDOH data platform to end-users across the federal government, PCOR researchers, and health services researchers.
6. Establish a sustainability and growth plan for the SDOH data platform.

### 2. Background - Problems Addressed

This section should describe the problems addressed by the project in terms of the OS PCORTF functionalities and how it will advance the field. Again this should be based on the original statement of work. This section should clearly describe the research questions that can now be answered as a result of the work as well.

The United States spent more than \$3.3 trillion on healthcare in 2016, yet America lags behind other developed countries in terms of health outcomes such as infant mortality and life expectancy. Concern about whether the constantly changing health care system is providing value, and whether there are data-driven solutions to provide more effective health and human services and improve the health of all Americans. In particular, there is a high demand for a data platform that integrates information on SDOH, health service utilization and systems of care. Understanding the context in which the health care system operates is critical to being able to inform and find solutions to pressing health care problems, and evaluate the impact of those solutions.

Researchers, private organizations, federal, state and local investigators spend substantial resources combining multiple datasets to create one-off data files for their analyses or reinventing the collection of SDOH information, frequently lacking standardized metrics or estimates at the small-area level. There is no complete source of longitudinal information with uniformly-formatted community-level data on SDOH readily available for health services research.

The goal of the project is to develop a consolidated set of national standardized databases on valid and reliable SDOH factors at the small-area level, building on existing databases

developed by Federal agencies. Data elements to be included will span the SDOH landscape and include measures of income, employment, food, housing, environment, economics, education, safety, transportation, justice system, market structure, health status, and health care access and utilization. The SDOH data platform will be constructed longitudinally with a planned 10-year lookback. Public use files and supporting documentation in a standardized structured format will be developed and made available to the public.

By expanding data capacity, this project will enable research that examines the value of interventions that can prevent the disease before it occurs or progresses and take into account the community in which patients live, the effectiveness of interventions tailored to the whole person, and the utility of system-focused interventions that can pay for patient health outcomes with consideration to the community in which the system operates. These types of PCOR studies will ultimately inform the value transformation of the healthcare system, enabling policymakers, providers and payers to make better decisions for the patients and the healthcare system as a whole.

There are a range of questions that can be answered using the SDOH Database. Examples include:

- Hospital billing data (HCUP): Are differences in county and/or zip code characteristics associated with differences in COVID-19 hospitalizations and outcomes?
- Claims data (Medicare, Medicaid claims): To what extent do changes in health care utilization for chronic illnesses during the COVID-19 pandemic differ by SDOH characteristics of patients' county/ zip code of residence?
- Individual-level survey data (MEPS-HC): Do individuals living in areas with a shortage of primary care providers have difficulty finding a Usual Source of Care?
- Health care system/EHR data (Individual practices or systems): How does knowing information on the community in which a patient lives help clinicians improve care? Does this differ depending on the availability of patient-level social needs data?

### **3. Methodology**

This section should provide a brief overview of what was done (in lay person's language).

To begin the project, AHRQ conducted an environmental scan to identify data sources by SDOH domain, geographic level, and year. Using that information, AHRQ developed a beta version of the SDOH Database, which contained a set of variables from secondary data sources covering all SDOH domains, for years 2009 through 2018. AHRQ improved the beta version based on extensive research and end-user feedback to identify additional sources of SDOH variables at the county, ZIP code, and census tract levels. The updated SDOH Database, for the years 2009 through 2020, is available on [AHRQ's website](#).

The SDOH Database includes harmonized variable names and structures across different data sources to make the database coherent and user-friendly. In choosing data elements, we expanded our initial environmental scan from 2019 to include additional, relevant SDOH variables. We were guided by research questions in the domains of relevance to end users, actionability (e.g., "Do the data elements inform the expansion of the research infrastructure regarding SDOH and health outcomes?"), and accessibility ("Are the data sources available at a county level or below to enable analysis at a community level?"). We held a series of end-user feedback sessions for additional guidance.

#### 4. Accomplishments by Final Deliverables

Based on the objectives specified in the Overview, this section should provide the “story of accomplishments”. Most importantly, it should provide a description of each of the end products, including target audience, how it is to be used, and location of products (how target audience will access them).

Accomplishments for each of the 6 objectives listed in the overview are described below.

##### **1) Conduct an environmental scan to identify a comprehensive set of U.S. Department of Health and Human Services’ (HHS) and other federal datasets with existing or analyzable small-area level and other geographic level data on SDOH variables.**

The environmental scan served as the starting point to inform the development of AHRQ’s SDOH beta data files publicly released in December 2020. AHRQ publicly released the environmental scan as a resource for analysts interested in identifying SDOH-related data sources. The scan aims to identify as many SDOH data sources as possible as of July 2020 at the ZIP Code, county, and State level for the domains of social context, economic context, education, physical infrastructure, and healthcare context. The scan is organized in an Excel spreadsheet to maintain the filter functionality of each column in the scan so that analysts can sort by variables such as SDOH domain or level of geography. The environmental scan is available on AHRQ’s website.

##### **2) Design and create a publicly available data platform of valid and reliable standardized set of SDOH data sources at various geographic areas.**

AHRQ first developed a beta version of the SDOH Database, which contained a set of variables from secondary data sources covering all SDOH domains, for years 2009 through 2018, posted in December 2020. AHRQ then improved the beta version based on extensive research and end-user feedback to identify additional sources of SDOH variables at the county, ZIP code, and census tract levels. The updated version was posted on AHRQ’s website in July 2022.

The SDOH Database is derived from 43 publicly available data sources and one restricted data source, including over 1,366 unique variables across years 2009 to 2020. Variables span the SDOH domains of social context, economic context, education, physical infrastructure, and healthcare context. It makes community-level SDOH analysis easier in the following ways:

- It uniquely contains variables at three small area levels - county, ZIP code, and census tract - allowing for analysis of SDOH factors at the community level.
- Researchers can use the database to analyze the relationships among SDOH factors and health outcomes as a stand-alone source as well as linked to other data resources such as the Medical Expenditure Panel Survey (MEPS) and the Healthcare Cost Utilization Project (HCUP).
- It contains variables spanning five key SDOH domains and 31 different topic areas within those domains; it is expansive to accommodate different end-users such as PCOR researchers; healthcare delivery systems; and local, state, tribal, and federal programs.
- It is publicly available with documentation, including codebooks by year, and guidance for end-users.

With over 1,366 unique variables, users can analyze a wide range of topics using the SDOH Database. Topic examples span from immigration, internet connectivity, socioeconomic disadvantage indices, environmental data, and distance to health care providers.

Among data sources newly added are the [Community Resilience Estimates](#) variables from the U.S. Census Bureau which is a metric for understanding how vulnerable every neighborhood is in the U.S. to disasters (including COVID-19) and the [Centers for Disease Control and Prevention \(CDC\) and Office of Minority Health's Social Vulnerability Index](#) which supports the identification of racial and ethnic health disparities. There are variables regarding weather and the environment from the Environmental Protection Agency, the National Oceanic and Atmospheric Administration, and Washington University in St. Louis. Other examples of timely data include variables relevant to the opioid crisis, disparities in healthcare quality and outcomes, and 2020 data that can provide context for understanding trends during the early stages of the COVID-19 public health emergency.

### **3) Coordinate and expand the data collection efforts on SDOH across HHS.**

This project took place during a time of rapid expansion of activity and data collection regarding SDOH in HHS. This project was designed prior to the onset of the COVID epidemic, responding to data needs to respond to questions about how value-based purchasing affects health care systems and communities with more vulnerable people. With the onset of the COVID epidemic, efforts to measure SDOH information have grown rapidly across multiple settings, including among individual health care providers, multiple types of CMS data, and in federal surveys. To support information sharing, AHRQ has been actively participating in a series of ASPE-led cross-HHS workgroups including the HHS SDOH workgroup, the HHS SDOH Data subgroup, and the HHS SDOH Small Leadership team. These meetings have served the purpose of coordinating and sharing information across HHS about ongoing and new data collection and development efforts. In addition, these meetings provided a source of ongoing federal expert input throughout the SDOH Database development.

In addition to information from these ongoing HHS SDOH meetings, AHRQ held extensive one-on-one meetings and received feedback from stakeholders interested in the SDOH Database and conducted structured “end user” interviews to inform Database development. The meetings are summarized in the “SDOH Database meeting summary” and “SDOH-RD End-User Report to AHRQ” documents. In addition, we developed an internal synthesis report, “Proposed Use Cases for Community-Level SDOH Data,” reflecting information from external input, published studies, and reports, to further inform database development and dissemination.

### **4) Use the new data platform to conduct a minimum of two patient-centered outcomes research (PCOR) studies.**

We have conducted multiple analyses resulting from linkages to three data sources, the Health Care Cost and Utilization Project, Chronic Conditions Data Warehouse, and Medical Expenditure Panel Survey, as described below. We prioritized linkages to multiple types of data to develop understanding of how the Database can be used with a variety of sources.

## Health Care Cost and Utilization Project (HCUP)

**Manuscript:** Owens PL, Liang L, Weiss AJ, Hammonds TL, Fingar KR, Agniel D. Racial and ethnic differences in the impact of the COVID-19 pandemic on emergency department utilization related to mental illness and substance use disorders. To be submitted to the *American Journal of Public Health* by early December 2022.

**Objective.** To determine how the COVID-19 pandemic impacted emergency department (ED) utilization involving mental illness and substance use disorders (MSUD). MSUD are critical health concerns during the pandemic, and the ED is a long-standing entry point into the behavioral health treatment system.

**Methods.** Using data from the 2017–2020 Healthcare Cost and Utilization Project (HCUP), we compared racial/ethnic-specific population rates of ED visits involving MSUD for 2,253 U.S. counties before and during the pandemic overall and by county-level weekly COVID-19 burden and social vulnerability.

**Results.** ED visit rates involving MSUD were lower early in the pandemic than before the pandemic, declining further during periods of high county-level COVID-19 burden and in more socially vulnerable counties. This effect was observed for individuals with serious mental illnesses and other SUD regardless of race/ethnicity and for other mental illnesses among Black non-Hispanic individuals.

**Conclusions.** Access to adequate treatment for individuals with MSUD is always essential, and especially during the COVID-19 pandemic. Notably, MSUD treatment access during the pandemic may be particularly challenging for Black individuals and those who live in socially vulnerable communities.

**HCUP Stat Brief:** Racial and ethnic differences in emergency department visits related to substance use disorders, 2019. (co-authors: PL Owens, BJ Moore). To be posted on HCUP-US the week of December 5: <https://hcup-us.ahrq.gov/reports/statbriefs/statbriefs.jsp#>

**Summary:** This brief presents statistics on SUD-related ED visits by patient race and ethnicity using weighted estimates from the 2019 Nationwide Emergency Department Sample (NEDS). SUDs include disorders involving alcohol, cannabis, stimulants, opioids, sedatives, hallucinogens, and inhalants, as well as other substances. The number and percentage of SUD-related ED visits are presented for select patient characteristics and primary expected payer. The population rate of SUD-related ED visits is presented by race and ethnicity for select demographic characteristics, including sex, age group, rural/urban location, and the social vulnerability of the county in which the individual resides (i.e., community's ability to prevent human suffering and financial loss during a disaster).

### Findings:

- The population rate of emergency department (ED) visits related to substance use disorders (SUDs) was 28.5 ED visits per 1,000 population in 2019. The rate was highest among Black non-Hispanic individuals (48.3 ED visits per 1,000 population) and lowest for Asian/Pacific Islander non-Hispanic individuals (5.3 ED visits per 1,000 population).
- The population rate of SUD-related ED visits was approximately twice as high for men as women, regardless of an individual's race and ethnicity (37.3 vs. 20.1 ED visits per 1,000 population).

- The rate of SUD-related ED visits was higher among individuals living in the most socially vulnerable communities versus less socially vulnerable communities for all racial and ethnic groups except Asian/Pacific Islander non-Hispanic individuals.
- Population rates of ED visits for specific substances varied by race and ethnicity. Black non-Hispanic individuals had the highest rate of ED visits related to cannabis-, stimulant-, and opioid-related disorders (15.6, 11.4, and 6.5 ED visits per 1,000 population, respectively) compared with all other racial and ethnic groups. White non-Hispanic individuals had the highest rate of ED visits related to sedative-related disorders (1.1 ED visits per 1,000 population).

The following used the beta version of the SDOH files:

- HCUP Stat Brief 260: Social Determinants of Health and County Population Rates of Opioid-Related Inpatient Stays and Emergency Department Visits, 2016. <https://hcup-us.ahrq.gov/reports/statbriefs/sb260-Social-Determinants-County-Opioid-Rates-Hospital-Use-2016.pdf>
- Internal AHRQ Report on the National Nursing Home COVID-19 Action Network (*this has not been released for distribution, although there is a manuscript under development*) NORC's Report on Assessment Report on AHRQ's National Nursing Home COVID-19 Action Network. The report examines the Network's recruitment, retention, implementation, and dissemination strategies and assesses the Network's impact on the entry, spread and control of COVID-19 in nursing homes. AHRQ's SDOH database was used to examine community-level contextual factors related to recruitment, implementation and impact of the Network.

### **CMS Medicare Chronic Conditions Data Warehouse (CCW)**

- We developed a series of analyses to examine the association between community-level social determinants of health and utilization and health outcomes of fee-for-service Medicare beneficiaries before and during 2020. Outcomes included telehealth visits, wellness visits, evaluation and management visits, emergency department visits, avoidable hospitalizations, premature mortality, and mortality. These results are being developed into manuscripts. In addition, we documented the steps and linkage code to share with CMS and other researchers.

### **Medical Expenditure Panel Survey (MEPS)**

- Vaccinations. We examined the association between county characteristics and receipt of an annual influenza vaccine, independent of individual demographic, socioeconomic and health characteristics. The analysis finds that living in counties with a higher percent of residents with a 4-year college degree is positively associated, and living in counties with higher social vulnerability is negatively associated, with vaccine receipt. The results suggest that community characteristics, either alone or in concert with individual characteristics, could help identify areas to target for vaccine outreach. These findings were presented as a poster at AcademyHealth in 2021 and are available on request. In addition, this methodology could be applied to future data to examine these patterns during COVID.
- Heat exposure. We examined the association between county-level temperature and likelihood of an emergency department visit. Preliminary results show an association between higher temperature and emergency department visits. We will be developing a presentation and draft manuscript in the next several months.

## **5) Disseminate the SDOH data platform to end-users across the federal government, PCOR researchers, and health services researchers.**

We employed multiple types of dissemination strategies to reach federal end-users, health services researchers, and PCOR researchers. One approach is to present at professional conferences, including organizing a Datapalooza panel on the SDOH Database in 2022, and presenting at AcademyHealth in two sessions in 2022 and one in 2020. In addition, to specifically reach HHS stakeholders, we presented twice to the OS-PCORTF seminar series. Further, to ensure that the presentation is readily accessible, we posted a presentation on AHRQ's website that provides an overview of the SDOH Database.

Another type of dissemination activity is a long-term strategy to work directly with key stakeholders on linkage resources. At CMS, we are working with two different groups to a) make SDOH Database variables available to researchers working with the Chronic Conditions Data Warehouse, and b) use SDOH Database variables in analyses focused on Medicaid long-term care services and supports. In addition, we have been actively involved in a collaborative effort between AHRQ, PCORI, and ASPE with one objective to support dissemination and use of the SDOH Database in PCORnet research and PCOR analyses generally. These efforts reflect a substantial investment of time to build relationships, explore potential collaborations, and take steps to build linkages. The aim is that, longer term, these ongoing efforts will promote awareness and use of SDOH data in support of HHS, health services, and PCOR research far beyond what would be possible absent such collaborations.

## **6) Establish a sustainability and growth plan for the SDOH data platform**

The SDOH Database project is relatively new, with data files first posted publicly at the end of December 2020 and an expanded, updated version of the SDOH Database posted publicly in July 2022. In collaboration with ASPE, we finalized a OS-PCORTF SOW for additional work to enhance the SDOH Database, to continue through 2026. Doing this work, which will include multiple years of data updates, data enhancements, evaluating the SDOH Database, and opportunities for end-user input, will continue to develop momentum and use of the SDOH Database. As part of this work, we will specify and put in place elements needed for sustainability of the SDOH database into the future.

## **5. Lessons Learned and considerations for future work**

This section should provide a thoughtful description of lessons learned and considerations for future work based on the impact of the project.

We have identified several key lessons learned and considerations for future work:

- There is a demand for timely community-level SDOH data. Since we first posted the initial SDOH Database files on AHRQ's website, in late December 2020, there have been 20,113 downloads. Since July 21, 2022 when we posted the updated database files, there have been 8,130 unique file downloads. Consistent with multiple sources of end-user feedback, timely data is a key interest. Downloads are most common for the most recent data year, 2020.
- One observation that may be relevant to future data infrastructure needs has to do with the interrelation between community and individual level SDOH / social needs data. There is substantial opportunity to establish methods for how best to utilize community-level SDOH data in conjunction with varying availability of patient-level social needs

data. Further, the ecosystem of types of individual and community level SDOH data is not well mapped in terms of available sources, level of data available (person, provider, community), data type (electronic health record, claims, administrative/program enrollment, survey, etc), and applicable privacy requirements for data used individually and in linkages.

- A further observation is that this project has taken an iterative approach to building and expanding the SDOH Database over time. We think this has benefited the project development and dissemination to take a sequential approach because it enabled a more rapid production of initial data files and provided opportunity to use those files to elicit feedback for future improvements.
- Finally, this project is inherently collaborative in multiple respects. The SDOH Database is a synthesis of data sources from multiple Executive Branch Departments as well as multiple agencies within HHS. Another important aspect of collaboration is ongoing dialogue with multiple HHS agencies to share information, receive feedback, and support development of expanded availability and use of the SDOH Database across HHS.