



ASPE
ASSISTANT SECRETARY FOR
PLANNING AND EVALUATION

OFFICE OF
HEALTH POLICY



OFFICE OF THE SECRETARY
**PATIENT-CENTERED OUTCOMES
RESEARCH TRUST FUND**

Building the Data Capacity for Patient-Centered Outcomes Research:

The 2021 Annual Report Executive Summary

Office of Health Policy
Assistant Secretary for Planning and Evaluation
U.S. Department of Health and Human Services

May 2022

This report was prepared under Contract No. HHSP233201500048I/75P00121F37020 between the Department of Health and Human Services' Office of the Assistant Secretary for Planning and Evaluation (ASPE) Office of Health Policy and NORC at the University of Chicago.

Suggested citation: Dullabh P, Dhopeswarkar R, Hovey L, Desai P, Jimenez F, Callaham M, Kane R, Heaney-Huls K, Lumsden S, and Smith S. Building the Data Capacity for Patient-Centered Outcomes Research: The 2021 Annual Report. Washington, DC: Office of the Assistant Secretary for Planning and Evaluation, U.S. Department of Health and Human Services. May 20, 2022.

Acknowledgements

We would like to thank the following individuals for their contributions to this report:

OS-PCORTF Project Leads

Kevin Abbott (NIH/NIDDK)	Shin Kim (CDC)
Robert Ball (FDA)	Ray King (CDC)
Elise Berliner (AHRQ)	Emily Madden (ASPE)
Steve Bernstein (AHRQ)	Danica Marinac-Dabic (FDA)
Arlene Bierman (AHRQ)	Maria Michaels (CDC)
Carlos Blanco (NIH)	Lisa Mirel (CDC)
Kate Brett (CDC)	James Mork (NIH/NLM)
Valeria Butler (ACF)	Jenna Norton (NIH/NIDDK)
Shanna Cox (CDC)	Greg Pappas (FDA)
Carol DeFrances (CDC)	Emma Plourde (ASPE)
Sarah Dutcher (FDA)	Matthew Rahn (ONC)
Stephanie Garcia (ONC)	Lok Wong Samson (ASPE)
Ken Gersing (NIH)	Daniel Schwartz (ASPE)
Robin Ghertner (ASPE)	Steven Schwartz (CDC)
Aly Goodman (CDC)	Heather Stone (FDA)
Violanda Grigorescu (CDC)	Marissa Sucusky (CDC)
Janey Hsiao (AHRQ)	Brian Ward (CDC)
Stefan Jaeger (NIH/NLM)	Adam Wong (ONC)
Patricia Keenan (AHRQ)	

ASPE Office of Health Policy, Office of Healthcare Quality and Outcomes

Lucas Arbulu
Sara Couture
Emily Evans
Aldren Gonzales
Violanda Grigorescu
Madjid Karimi
Euny Lee
Susan Lumsden
Scott Smith
Marcos Trevino
Sara Wei

Table of Contents

I. Introduction	1
II. OS-PCORTF Portfolio Overview	1
OS-PCORTF Reauthorization and New Strategic Plan	1
III. Portfolio Contributions to Key HHS Priorities	3
Addressing COVID-19 through Data, Research, and Infrastructure Investments	3
Improving Maternal Health Research and Outcomes	6
Data and Data Infrastructure to Combat the Opioid Crisis	8
Advancing Equity and Inclusion through Improved Collection and Use of SDOH and Other Non-Healthcare Data to Enhance Person-Centered Care	10
Summary of Contributions to National Health Priorities	11
IV. Completed Projects' Major Accomplishments	11
Capstone for the Outcome Measures Harmonization Project (AHRQ)	12
Enhancing Data Resources for Researching Patterns of Mortality in Patient Centered Outcomes Research: Project 1 - Adding Cause-Specific Mortality to the National Hospital Care Survey by Linking to the National Death Index (CDC/NCHS)	13
Enhancing Data Resources for Researching Patterns of Mortality in Patient Centered Outcomes Research: Project 2 – Pilot Linkage of National Death Index+ to Commercially and Publicly Insured Populations (FDA)	13
V. Conclusion	14
VI. References	15

I. Introduction

Through the administration of the Office of the Secretary's Patient-Centered Outcomes Research Trust Fund (OS-PCORTF), the Office of the Assistant Secretary for Planning and Evaluation (ASPE) coordinates across agencies to improve the collection, linkage, and analysis of data for patient-centered outcomes research. The Fiscal Year 2021 (FY 2021) OS-PCORTF Annual Report highlights the accomplishments of 34 multi-agency projects to build data capacity and support the mission of the Department of Health and Human Services (HHS) to advance science and improve knowledge about the comparative effectiveness of health care interventions. In FY 2021, the report focuses on projects aligned with four national health priorities—COVID-19, maternal health, health equity and social determinants of health (SDOH), and the opioid crisis. The report also highlights major accomplishments from projects that completed their work in 2021 and provides links to the publicly-available products they have developed to support patient-centered outcomes research (PCOR).ⁱ

II. OS-PCORTF Portfolio Overview

Data are essential to research that improves the health and well-being of all Americans by increasing our understanding about the outcomes and effectiveness of health care treatments and interventions. HHS and its agencies routinely collect, link, and analyze data that can be used to generate new scientific knowledge as well as important information about Federal programs and the populations these programs serve. As a consumer, producer, and regulator of key national health data, HHS is uniquely positioned to coordinate its programs to build national data capacity in support of the mission, statutory authorities, and annual priorities of each HHS agency and the Department as a whole.

Through the administration of the OS-PCORTF, ASPE coordinates across agencies to improve the collection, linkage, and analysis of data for patient-centered outcomes research. PCOR studies focus on understanding the comparative safety and effectiveness of treatments, services, and other health care interventions to address questions identified as important to patients, caregivers, clinicians, and policymakers. Conducting PCOR studies require timely access to relevant, high-quality data and the use of rigorous and appropriate research methods.

As part of its work to build data capacity for patient-centered outcomes research, the OS-PCORTF provides critical support for addressing HHS priorities, including targeting funding toward specific gaps in its annual solicitation. Leveraging the shared interest in building data capacity for patient-centered outcomes research and addressing critical national health priorities, the OS-PCORTF brings together the expertise of HHS agency leaders, informaticians, technologists, and researchers to identify priorities, share expertise and resources, and collaborate on projects.

OS-PCORTF Reauthorization and New Strategic Plan

On December 20, 2019, Public Law 116-94 reauthorized the OS-PCORTF through 2029.¹ The reauthorization extended the Trust Fund's work to include two new research priorities—intellectual and developmental disabilities (ID/DD) and maternal mortality. The reauthorization also calls for a broader assessment of the types of outcomes considered in patient-centered outcomes research, including potential burdens and economic impacts of health care interventions. To address the Trust Fund's

ⁱ To align with other HHS agencies, we refer to patient-centered outcomes research without using the acronym except in its adjectival form (e.g., PCOR studies, PCOR data infrastructure).

expanded scope and to leverage recent advances in analytic methods and the increasing volume of data from within and outside of health care, ASPE is updating its ten-year strategic plan for the OS-PCORTF.

The new [OS-PCORTF Strategic Plan for 2021-2029](#) will set forth a long-term plan for PCOR data capacity. The previous decade of OS-PCORTF investments focused on five functionalities needed to build PCOR data infrastructure. ASPE intends to build upon this infrastructure to ensure responsiveness to emergent health challenges and evolving HHS priorities over the next decade.

The agency's new strategic plan is the result of an extensive deliberation process among internal and external experts:

- **Initial Listening Session (2020): [Challenges and Improvements for PCOR Data Infrastructure: Results from a Stakeholder Prioritization Activity](#).** On behalf of ASPE, NORC gathered perspectives on challenges and potential improvements for PCOR data infrastructure from stakeholders with a wide range of occupational backgrounds including policy, health care delivery, research, and informatics.
- **Research Data Network Report (2021): [Patient Centered Outcomes Research Trends and Opportunities: Scan and Interviews with Key Informants Report](#).** On behalf of ASPE, The MITRE Corporation conducted an environmental scan of 15 research networks and 8 key informant interviews with representatives from networks involved in PCOR activities. Respondents were asked about the challenges they face and potential opportunities to enhance access, quality, and scope of data for patient-centered outcomes research.
- **U.S. Department of Health and Human Services Stakeholder Engagement Report (2021):** On behalf of ASPE, The MITRE Corporation, conducted interviews and analyzed responses from the interviews. The HHS stakeholder engagement and interviews were designed in partnership with a group of agency representatives referred to as ASPE's Strategic Planning Engagement Council. The main goals of the interviews were to understand (1) key agency priorities related to patient-centered outcomes research, (2) existing work that can be leveraged to improve data capacity, (3) gaps and opportunities, (4) research trends and legislative and policy drivers in the environment, and (5) how the current process for creating and funding OS-PCORTF projects could be improved within the context of an evolving health care system and new priorities for patient-centered outcomes research. From March to May 2021, The MITRE Corporation conducted 32 interviews with 62 participants, including HHS agency leaders, leaders of OS-PCORTF projects, and agency data experts.
- **Study Group and Public Workshops (2021): [Building Data Capacity for Patient-Centered Outcomes Research: An Agenda for 2021 to 2030](#).**
 - A **National Academy of Sciences, Engineering and Medicine (NASEM) Study Group** was convened to identify critical needs for building PCOR data capacity and generating new evidence to inform health care decisions. The study investigated opportunities to address: 1) data user needs over the next decade; 2) data standards, methods, and policy; and 3) creating a comprehensive ecosystem for patient-centered outcomes research.
 - **Three public workshops** (virtual) were conducted and recorded in May and June 2021. [NASEM has published](#) the reports from the three workshops and the final study report.

This multi-stakeholder consensus process resulted in the creation of four priority goals that reflect ASPE's role in building PCOR data infrastructure: 1) data capacity for national health priorities; 2) data standards and linkages for longitudinal research; 3) technology solutions to advance research; and 4) person-centeredness, inclusion, and equity. ASPE intends to finalize and release the strategic plan in the summer of 2022. Once released, the new [OS-PCORTF Strategic Plan for 2021-2029](#) will set forth a long-term plan for increasing PCOR data capacity through partnerships with federal agencies and the broader PCOR community.

III. Portfolio Contributions to Key HHS Priorities

ASPE's investments in data infrastructure are intended to fill high-priority gaps in PCOR data capacity, targeting opportunities that are within its legislative purview, aligned with its strategic goals, and responsive to urgent national health priorities. Projects funded by the OS-PCORTF focus on impacting researchers' ability to conduct new PCOR studies that assess the relative benefits and harms of interventions in real-world settings and across populations. Examples of such impact include addressing new questions, producing higher quality data, improving research efficiency.



During the past year, the portfolio has increased its investments in four key areas:

- COVID-19
- maternal health research and outcomes
- the opioid crisis
- health equity and SDOH

Below, we highlight OS-PCORTF projects that are addressing these areas, including both existing projects and newly funded FY 2021 projects.

Addressing COVID-19 through Data, Research, and Infrastructure Investments

The COVID-19 pandemic exposed multiple weaknesses in the health system's ability to conduct syndromic surveillance, efficient testing and lab reporting, and other critical public health functions. This prompted ASPE to increase the OS-PCORTF portfolio's efforts to improve and modernize the health data infrastructure needed for effective pandemic response.²

The COVID-19 Collaborative. As part of the project funding, ASPE formed a COVID-19 Learning Collaborative so that project teams can support each other's research activities and establish new cross-agency connections or enhance the existing ones. The COVID-19 Collaborative functions as a learning network in which these projects share updates, feedback, and knowledge among fellow researchers. They meet monthly and have formed communities of practice (CoP) around two areas of need: SDOH and data linkages.

- The SDOH CoP developed a [scoping review](#) to assess research on SDOH and COVID-19 risks and outcomes. The CoP is interested in identifying SDOH that are associated with COVID-19 infection and adverse clinical outcomes, how risks and outcomes are currently being defined and measured, and if standardized approaches are being used to do so.
- The Data Linkages CoP is focused on supporting the project “**Building Infrastructure and Evidence for COVID-19 Related Research, Using Integrated Data from the National Center for Health Statistics (NCHS) Data Linkage Program**” (CDC, funded in 2021). The linkage methodologies and resulting linked data can be used to expand data infrastructure capacity to support new PCOR studies. The CoP is using the project work as a basis for developing a data linkages paper focused on applying the project’s lessons to enhance data infrastructure capacity and evidence building for COVID-19 related research.

COVID-19 Individual Project Contributions. The seven individual projects address gaps in infrastructure by aggregating and linking data from diverse sources to inform research and patient care, improving data sharing and interoperability, and developing mechanisms to collect, track, and study patient outcomes in the short- and long-term.

- **Creating a Federal COVID-19 Longitudinal Patient Outcomes Research Database Linked to Health Systems and Clinical Data** (NIH/National Center for Advancing Translational Sciences [NCATS], 2021). The new COVID-19 database will link three sources of data: 1) provider characteristics from the Agency for Healthcare Research and Quality (AHRQ) Compendium of US Health Systems; 2) clinical data in NIH’s [National COVID-19 Cohort Collaborative \(N3C\) dataset](#); and 3) claims data on Medicare fee-for-service (FFS), Medicare Advantage (MA), and Medicaid populations. The database will allow researchers to evaluate the U.S. health system’s COVID-19 response and patient care utilization patterns over time.
- **Understanding Long-term Outcomes in COVID-19 Survivors with Multiple Chronic Conditions (MCC) through e-Care Plan Development** (AHRQ and NIH/National Institute of Diabetes and Digestive and Kidney Diseases ([NIDDK], 2021). The project team has already created provider-facing and patient-facing apps to facilitate data collection and care coordination for adults with MCC. New project work will create a caregiver app and add data on COVID-19 and long COVID, with the goal of identifying risk and protective factors, natural history, and short- and long-term outcomes in COVID-19 survivors. In April 2022, NIH released a funding announcement to study the feasibility of, and develop best practices for, using interoperable health information to analyze health conditions prevalent in older adults. The announcement encouraged the use of open-source tools developed through the Multiple Chronic Conditions Electronic Care Plan project to improve interoperability of data for people living with MCCs.
- **CURE ID: Aggregating and Analyzing COVID-19 Treatment from Electronic Health Records (EHRs) and Registries** (Food and Drug Administration [FDA], 2021). The project expands the FDA’s and NCATS’ [CURE ID](#) platform by automatically extracting data from EHRs and clinical disease registries. The expanded CURE ID platform will house tens of thousands of COVID-19 case reports that can be used to identify potentially safe and effective COVID-19 treatments from among existing therapeutics.

- **Using Machine Learning Techniques to Enable Health Information Exchange (HIE) Data Sharing to Support COVID-19-focused PCOR** (Office of the National Coordinator [ONC], 2021). This project will create reliable methods to access large volumes of data from state and regional HIEs, which are underused sources of data. The project will then develop machine learning algorithms to aggregate and analyze HIE data while preserving privacy. These methods will be applicable across HIEs and support data gathering for research questions related to COVID-19 and beyond.
- **Building Infrastructure and Evidence for COVID-19 Related Research Using Integrated Data** (Center for Disease Control and Prevention [CDC]'s NCHS, 2021). The project team will utilize existing NCHS linked data sources to develop synthetic data generation methods that can be applied to future linked files that focus on the COVID-19 pandemic. The project will create synthetic linked data files that preserve privacy using data from the National Health Interview Survey (NHIS) and the National Hospital Care Survey (NHCS) that have been linked to the National Death Index (NDI), Medicare data from the Centers for Medicare & Medicaid Services (CMS), and federal housing assistance data from the Department of Housing and Urban Development (HUD). The project will produce a public dashboard that allows researchers to analyze the linked data to study, for example, the associations between SDOH and COVID-19-related health outcomes.
- **Multistate Emergency Medical Services (EMS) and Medicaid Dataset (MEMD): A Linked Dataset for PCOR** (ASPE, 2021). This project will develop a database of linked EMS data and Medicaid claims data from five states to allow researchers to study the effectiveness of emergency services. The researchers will focus on behavioral health emergencies (e.g., drug overdoses, psychotic episodes, suicidal ideation, and panic attacks), but the database they develop will support a broad range of research on emergency care, including COVID-19-related responses.
- **Data Set on Intellectual and Developmental Disabilities (ID/DD): Linking Data to Enhance Person-Centered Outcomes Research** (ASPE, 2021). This project will link state-level data sources to enable researchers to analyze relationships between various sociodemographic information, need for home and community-based services, service utilization, service expenditures, and person-centered outcomes prior to and during the COVID-19 pandemic for individuals with ID/DD.

Together, the COVID-19 projects represent significant opportunities to enhance PCOR data infrastructure ranging from exchange standards to better access to synthetic, longitudinal, and outcomes data. This infrastructure will advance goals related to COVID-19 research and will support broader PCOR goals.



- ▶ **Real-Word Data Repository for Research on Treatments.** CURE ID provides a platform for collecting data submitted by medical providers on novel uses and outcomes of existing medications to treat infectious diseases. Information on COVID-19 drugs, biologics, and vaccines from clinical trials submitted to clinicaltrials.gov is now included.
- ▶ **Accelerating Data Sharing to Respond to COVID-19.** The ONC Machine Learning project team will work with three state and regional HIEs to expand their data capacity and promote interoperability by developing repeatable, privacy-preserving machine learning algorithms that can be leveraged by the research community to address important PCOR questions.
- ▶ **Promoting Use of Interoperable Health Records in Clinical Research.** NIH's recent funding announcement encouraged the use of open source tools, including the Multiple Chronic Conditions Electronic Care Plan to study the feasibility of, and develop best practices for, using interoperable health information to analyze health conditions prevalent in older adults.

Improving Maternal Health Research and Outcomes

In the OS-PCORTF reauthorization, maternal mortality was identified as a strategic national research priority, reflecting the need to address the U.S.'s persistently high maternal death rates and large racial disparities. Four OS-PCORTF projects have been funded to continue ASPE's efforts to address gaps in maternal health research, both through their individual data infrastructure projects and their contributions to ASPE's Maternal Health Consortium.

Maternal Health Consortium. The Consortium is led by NIH/the National Institute of Child Health and Human Development (NICHD) and includes the three projects described below. The Consortium held regular meetings that included presentations from participating projects and agencies as well major EHR vendors that were invited to discuss their own approaches to data standards and linkages, providing context on activities in the larger health ecosystem.

The Consortium's focus is to assess data exchange standards, methodologies, projects, and other initiatives to inform development of new tools for longitudinal maternal and infant health information for research. At the conclusion of its Phase 1 work, Consortium members will produce a final implementation guide that will describe a standardized approach to link electronic data to study the medical conditions, interventions, and outcomes for pregnant, postpartum, or lactating women and their infants. The implementation guide will be publicly released and made available by HL7®.

Individual Project Contributions. In addition to their contributions to Consortium activities, the individual project teams are actively building data infrastructure components to support maternal health research. These projects, both as individual activities and through their work as a consortium, offer solutions that will leverage EHRs and other critical data sources and improve the scope and standardization of maternal health data resources available to support patient-centered outcomes research.

- **Severe Maternal Morbidity and Mortality EHR Data Infrastructure** (NIH/NICHD, 2021). This project will develop a set of standard data elements for EHRs and a FHIR® application programming interface (API) that researchers can use to access data on maternal morbidity/mortality from pregnancy through 1-year post-partum. An implementation guide will detail methods other researchers can use to create longitudinal datasets that support research on pregnancy risk factors, outcomes, and the effects of complications during pregnancy on longer-term post-pregnancy health outcomes.
- **MAT-LINK: MATernal and Infant Network to Understand Outcomes Associated with Medication for Opioid Use Disorder during Pregnancy** (CDC, 2019) has been expanded as **MAT-LINK2: Expansion of MATernal and Infant Network to Understand Outcomes Associated with Medication for Opioid Use Disorder during Pregnancy** (CDC, 2021). [MAT-LINK](#) established a surveillance network to address the lack of national-level data on maternal, infant, and child health outcomes following treatment for opioid use disorder (OUD) during pregnancy. MAT-LINK2 will expand the network from four clinical sites to seven, which will generate longitudinal data on over 4,000 mother-infant pairs; obtain follow-up data on children from infancy through age six; and increase the geographic and demographic representativeness of MAT-LINK. These data will improve understanding of effective, patient-centered care for pregnant women with OUD and for infants and children with prenatal opioid exposure.
- **Enhancing Surveillance of Maternal Health Clinical Practices and Outcomes with Federally Qualified Health Centers' (FQHCs) Electronic Health Records Visit Data** (CDC, 2021). This project focuses on creating a new nationally representative data source for research on maternal health treatment and outcomes at FQHCs. The project team will upgrade manual data collection methods for the National Ambulatory Medical Care Survey (NAMCS), the only source of nationally representative visit-level data on ambulatory health services. Automated electronic data transfers from provider EHRs as part of NAMCS will create a consolidated file of maternal health visits. The data will then be linked to mortality data from the NDI and administrative data from HUD to create a more complete picture of maternal health outcomes and SDOH.

Other projects, featured elsewhere in this report, contribute substantially to the availability of longitudinal, real-world data to improve patient centered outcomes research related to maternal health.

- **Developing a Multi-State Network of Linked Pregnancy Risk Assessment Monitoring System (PRAMS) and Clinical Outcomes Data for Patient-Centered Outcomes Research** (CDC, 2020) will link state-level [PRAMS data](#) with birth certificates and clinical outcomes data (e.g., hospital discharge, Medicaid claims, all-payer claims databases). These linkages will create a more comprehensive dataset to study interventions 1) prior to pregnancy; 2) during the perinatal period; and 3) through the post-partum period. It also will provide information on how social context and SDOH affect maternal health—data that are not often available in clinical datasets. This project may join the Maternal Health Consortium as the Consortium's focus expands to other data sources such as claims and data linkages.



- ▶ **PCOR in Practice.** The draft value set and [preliminary FHIR® implementation guide](#) developed by the Severe Maternal Morbidity and Mortality EHR Data Infrastructure team focuses on two use cases: outcomes following pregnancy-induced hypertension and post-partum mortality within one year. These products were tested during the HL7® January 2022 Connectathon.
- ▶ **Sharing Resources, Increasing Knowledge.** A [Public Health Grand Rounds](#) on MAT-LINK's capabilities, which focused on reducing polysubstance use in pregnancy, was viewed by 1,174 people in 4 foreign countries, 49 states, and the District of Columbia. The scope of MAT-LINK's uses for maternal and infant health surveillance was discussed in an article published in the [Journal of Women's Health](#).
- ▶ **Data Linkage: Surveillance, Birth Certificate, and Outcomes Data to Improve Maternal Health.** [PRAMS](#) is the only surveillance system that provides population-based data about mothers and infants before and during pregnancy and the first few months following birth. PRAMS represents approximately 81 percent of all live births in the U.S. with information from 46 states, New York City, Puerto Rico, the North Mariana Islands, and the District of Columbia. PRAMS data are used by researchers to investigate emerging issues in reproductive health (e.g., Zika virus and other infectious disease risks) and to plan and review maternal health programs and policies.

Data and Data Infrastructure to Combat the Opioid Crisis

Combating the opioid crisis is a strategic priority for HHS. Recognizing the need for high quality data to inform research and the public health response, ASPE has supported a range of opioid projects that: 1) improve the quality and timeliness of data on opioid-related outcomes; 2) link data sources to create a comprehensive data source; and 3) address co-morbid conditions that affect patient outcomes. Below, we highlight projects that demonstrate the scope of settings and populations supported by this work.

- **Child and Caregiver Outcomes Using Linked Data (CCOULD)** (ASPE/Administration for Children and Families [ACF], 2019). The project is creating a single harmonized dataset of Medicaid and child welfare records by linking patient-level data, including Medicaid enrollment, patient diagnoses, services, and claims, with child welfare data. The linked longitudinal data will allow researchers to better understand the needs of parents with children in the child welfare system who also need treatment for substance use disorders (SUDs) like OUD and co-occurring mental health problems whose children are in the child welfare system. The project also involves a significant technical assistance component to help states overcome data siloes and policy barriers and to build relationships that support data sharing.
- **An Addiction Medicine Network (AMNET) to Address the United States Opioid Crisis** (NIH/National Institute of Drug Abuse [NIDA], 2019). AMNET is establishing a platform for research on OUD and SUD by connecting office-based practices and their data to other clinicians and researchers, enabling clinical trials and comparative effectiveness research (CER). AMNET has harmonized the measures used and is now developing processes to link AMNET to other databases and registries. The expanded network will provide real-time data on patient characteristics, care delivery, and recovery service utilization that can be used to improve addiction treatment and patient outcomes.

- **Strengthening the Data Infrastructure for Outcomes Research on Mortality Associated with Opioid Poisonings** (CDC, 2018). This project created a new electronic system for coding death certificates, incorporating details on drug use, overdose, and additional demographic and geographic data in the Vital Statistics Rapid Release Program. The project also created a FHIR® API that allows death information to be exchanged between medical examiners, coroners, and the state. Greater availability and shareability of information on opioid-related deaths will improve public health surveillance and patient-centered outcomes research.
- **Enhancing Identification of Opioid-Involved Health Outcomes Using Linked Hospital Care and Mortality Data** (CDC, 2018). The project team is developing methods to improve the identification of patients with opioid-involved hospital encounters and overdoses from three data sources: 1) the NHCS; 2) the NDI; and 3) the Drug-Involved Mortality (DIM) file (formerly known as the National Vital Statistics System-Mortality-Drug Overdose file). This linked data will allow researchers to identify the specific opioids involved in drug-related emergency department (ED) visits, inpatient hospitalizations, and overdose deaths, and to develop strategies to reduce the morbidity and mortality associated with opioid misuse.
- Other projects, featured elsewhere in this report, contribute substantially to the availability of longitudinal, real-world data on opioid use, including:
 - **MAT-LINK2: Expansion of MATernal and Infant Network to Understand Outcomes Associated with Medication for Opioid Use Disorder during Pregnancy** (CDC, 2021) (described in the maternal health section above), and
 - **Identifying Co-Occurring Disorders among Opioid Users Using Linked Hospital Care and Mortality Data (2019): Capstone to an Existing FY 2018 OS-PCORTF** (CDC, 2018) (described in the second half of the full report).

These projects aim to improve the data infrastructure that can be used to support research, analyses, and policies to address the opioid crisis. These improvements encompass better data, networks, and registries; standardization of metrics and indicators of patient health and care quality; and robust linkages that enrich individual data sources and improve PCOR studies.

Examples of OS-PCORTF Contributions and Impact

- ▶ **Expanding Opioid Data to Drive Faster Action.** Expanding drug overdose data from the Vital Statistics Rapid Release Program has allowed rapid computation of provisional death rates by age, sex, and state, which enables researchers and policymakers to analyze – and act on – findings related to early mortality data (final mortality data are released approximately 11 months after the end of the data year).
- ▶ **Linking Opioid Hospital Data to Improve Outcomes.** Researchers outside of CDC have applied to the [National Center for Health Statistics \(NCHS\) Research Data Center](#) to use linked NHCS/NDI/DIM datasets for 2014/2015 and 2016/2017 to research the following topics:
 - Developing a method to better characterize transitions from nonfatal opioid hospital encounters to fatal overdose.
 - The associations between patient, surgical, hospital-related factors and opioid overdose hospitalizations following orthopedic procedures.
 - Identifying risk factors for opioid overdose deaths of concomitant drugs for patients with hospital and ED visits.

Advancing Equity and Inclusion through Improved Collection and Use of SDOH and Other Non-Healthcare Data to Enhance Person-Centered Care

Demographic, socioeconomic, and other data can provide a more comprehensive and holistic understanding of the range of factors that affect an individual's health, functioning, and quality-of-life outcomes than traditional health care data alone. The collection and use of these data are necessary to ensure person-centeredness, inclusion, and equity by improving the available evidence for supporting underrepresented, underserved, and at-risk populations, as described in a series of Executive Orders on health equity, issued in January 2021.³ Increased recognition of the effects of SDOH and differential access to care on health outcomes has attracted corresponding attention to the data and infrastructure needed to support patient-centered outcomes research on SDOH. These OS-PCORTF projects address the imperative for expanded access to demographic and social risk data, environmental data, and other non-health sector data to support health equity and research.

- **The Community and Clinical Data Initiative (CODI): Integrated Data for Patient-Centered Outcomes Research Project (2018) and the Community and Clinical Data Initiative (CODI 2.0)** (CDC, 2020) focus on increasing access to longitudinal clinical and community data with linkage and de-duplication tools. CODI 2.0 will further develop its infrastructure to create linkages across the health and social sectors for both adult data and pediatric data. These linkages will bring in data from larger and more diverse geographic areas and data networks, and it will allow the project to expand its clinical focus from obesity into diabetes and heart disease.
- **Enhancing Patient-Centered Outcomes Research: Creating a National Small-Area Social Determinants of Health Data Platform** (AHRQ, 2019). This project leverages publicly available data sources to develop a national standardized database with readily linkable SDOH variables at the small-area and other geographic levels. The beta version of the SDOH database is currently available for researchers to incorporate community SDOH characteristics in health outcomes research.
- **Data Set on Intellectual and Developmental Disabilities (ID/DD): Linking Data to Enhance Person-Centered Outcomes Research** (ASPE, 2021). The planned dataset will link key sources of ID/DD data, including Support Intensity Scale scores, Medicaid claims, National Core Indicators survey data, National Core Indicators COVID-19 supplement survey data, and potentially, other relevant state-level data for four to six states. The dataset will fill a gap in the data available to conduct ID/DD-focused patient-centered outcomes research and generate evidence to support person-centric health decision-making and equitable policymaking.
- **Data Linkage: Evaluating Preserving Privacy Methodology and Augmenting the National Hospital Care Survey with Medicaid Administrative Records** (CDC/NCHS, 2020). This project is currently linking data from the NCHS NHCS with Transformed Medicaid Statistical Information System (T-MSIS) data to expand the capacity to conduct studies of a wide range of HHS priority issues, particularly among the Medicaid covered population. This will improve access to linked federal data assets, enabling patient-centered research on health outcomes for a wide range of research priorities, including the outcomes of initiatives targeting opioid use and mental health services, efficacy of treatment protocols and drugs, disparities in efficacy for vulnerable subpopulations, and the role of social programs and SDOH in health outcomes.



- ▶ **Sharing Tools, Improving Research Efficiency.** PCORnet® and the Robert Wood Johnson Foundation have already started using CODI's Toolbox, specifically the growthcleanr tool to clean EHR data.
- ▶ **Improving Data Accessibility for a More Comprehensive Understanding of Health.** The [AHRQ SDOH database](#) makes community-level SDOH data on five SDOH domains (social context, economic context, education, physical infrastructure, and health care context) more readily available for researchers and policymakers.⁴ As of November 2021, there were about 10,000 file downloads from AHRQ's SDOH database web page.
- ▶ **Disseminating Tools for Enhanced Data Analysis.** Multiple algorithms developed from the 2018-2019 NCHS Data Linkage projects are now publicly available, including:
 - [The Opioid NLP component](#), which can search clinical notes text for opioid involvement.
 - [The SUD and MHI NLP component](#), which flags mentions of SUD and mental health issues in clinical note text.
 - [The Medical Code-based translation from SAS](#), which flags ICD-10-CM codes related to opioid involvement, SUD, and mental health issues in structured hospital data.

Summary of Contributions to National Health Priorities

Through its administration of the OS-PCORTF, ASPE has made significant investments in data infrastructure that supports research on key national health priorities: COVID-19, maternal health, the opioid crisis, and health equity and SDOH. The data infrastructure being created and enhanced through a range of projects will improve patient-centered outcomes research through a combination of more complete, standardized, and higher quality data on priority outcomes and diverse populations, improved methods for combining and analyzing datasets, and collaboration among federal and state agencies and the social sector to link their data sources so that pressing research questions can be answered more efficiently.

IV. Completed Projects' Major Accomplishments

Across the portfolio, OS-PCORTF projects have made contributions toward building data capacity to conduct patient-centered outcomes research. Each year, the portfolio produces tools and resources that improve the data infrastructure that supports patient-centered outcomes research by enhancing capacity to collect, link, and analyze data on outcomes and effectiveness.

To illustrate the impact and scope of the OS-PCORTF project contributions, the 2021 Portfolio Report highlights five projects that concluded in FY 2021 whose activities offer practical solutions, tools, and resources for researchers. These projects offer advancements in data capacity in the following areas:

- Linking clinical data for both for Medicare and commercially insured populations to mortality data
- Standardizing clinical care data and creating a process for extracting and linking EHR data to electronic patient registries to facilitate patient-centered outcomes research

- Re-designing and creating enhanced data sources and data infrastructure to advance patient-centered outcomes research on opioid-related mortality
- Linking EHR data, weight management program intervention data, and community-level census data to facilitate PCOR on childhood obesity
- Linking health-related administrative and survey data to expand PCOR data infrastructure for opioid use and mental health services, efficacy of treatment protocols and drugs, and disparities and SDOH research.

Below, we highlight the aims, activities, and key tools and resources developed for three of the five projects that concluded in FY 2021. Two of the other projects were discussed above as part of HHS priorities. Detailed project descriptions for all five projects are provided in individual project profiles presented later in the full report.

Capstone for the Outcome Measures Harmonization Project (AHRQ)

Using depression as a use case, AHRQ's Capstone Project worked with 21 clinical sites within an integrated health system and two patient registries to address barriers to using harmonized outcome measures across and between clinical sites and patient registries. The team created processes to standardize patient and clinical data for calculating standardized depression measures in EHRs, processes for reporting these data to registries, and natural language processing (NLP) methods to extract data from clinical notes. The team also developed a SMART on FHIR® app that integrates clinical and patient-reported data from multiple sources, facilitating treatment over time. These tools also allow researchers and registry developers to integrate patient registry data with clinical systems more easily to support patient-centered treatment and quality improvement efforts.

Products: The AHRQ team produced standardized implementation models, FHIR® profiles, a FHIR® implementation guide, and a methodology report to allow clinical sites to implement the library of depression outcome measures in EHRs, as well as three white papers, one peer-reviewed publication, a methodology report, and a final report.⁵ Additionally, two forthcoming manuscripts will showcase 1) the feasibility of using the harmonized outcome measures for conducting depression-related patient-centered outcomes research based on pilot data, and 2) the development of the SMART on FHIR® [“Major Depression Outcomes”](#) app.



Projects That Completed Their Work in 2021

- ▶ [Capstone for the Outcome Measures Harmonization Project \(AHRQ\)](#)
- ▶ [Community and Clinical Data Initiative \(CODI\) \(CDC\)](#)
- ▶ [Enhancing Data Resources for Researching Patterns of Mortality in Patient Centered Outcomes Research: Project 1 - Adding Cause-Specific Mortality to the National Hospital Care Survey by Linking to the National Death Index \(CDC/NCHS\)](#)
- ▶ [Enhancing Data Resources for Researching Patterns of Mortality in Patient Centered Outcomes Research: Project 2 – Pilot Linkage of NDI+ to Commercially and Publicly Insured Populations \(FDA\)](#)
- ▶ [Strengthening the Data Infrastructure for Outcomes Research on Mortality Associated with Opioid Poisonings \(CDC\)](#)

Enhancing Data Resources for Researching Patterns of Mortality in Patient Centered Outcomes Research: Project 1 - Adding Cause-Specific Mortality to the National Hospital Care Survey by Linking to the National Death Index (CDC/NCHS)

The NDI is a centralized database of death record information for all deaths within the U.S. This project created three new data sources for studying mortality following ED visits and/or hospital stays. Collaborating with partners from CMS and FDA, the NCHS team linked EHR and hospital claims data for inpatient and ED encounters collected in the 2014 and 2016 NHCS with claims data to 2014/2015 NDI data and 2016/2017 data, respectively. The inpatient and ED claims and EHR data collected in the 2014 NHCS were then linked to 2014-2015 Medicare enrollment and summary utilization and cost data from the CMS Master Beneficiary Summary File (MBSF). Separately, the 2014 NHCS was linked to 2014/2015 NDI data. In doing so, the NCHS Data Linkage Program produced linkages of patient EHRs and national mortality data as well as created new methods to optimize data linkages when using very large national data files.

Products: The CDC team produced three datasets linking 2014 and 2016 NHCS and NDI data and 2014 NHCS to the CMS MBSF. Due to confidentiality requirements, researchers must apply for access through the NCHS Research Data Center Network ([RDC website](#)). Data linkage methodology reports are available to the public. The CDC team published six research reports to demonstrate the utility of the linked datasets: five National Health Statistics Reports and a National Vital Statistics Report. As of January 2022, researchers outside of CDC have applied to the CDC's Research Data Center to use linked data sets to research the following topics: developing a method to better characterize transitions from nonfatal opioid hospital encounters to fatal overdose; the associations between patient, surgical, hospital-related factors and opioid overdose hospitalizations following orthopedic procedures; and identifying risk factors for opioid overdose deaths of concomitant drugs for patients with hospital and ED visits.

Enhancing Data Resources for Researching Patterns of Mortality in Patient Centered Outcomes Research: Project 2 – Pilot Linkage of National Death Index+ to Commercially and Publicly Insured Populations (FDA)

Research studies would often benefit from information on death and cause of death; however, publicly available data, including claims and EHR data, often lack complete information. Project 2 linked NDI data that includes cause of death from death certificates—contained in the NDI+ dataset—to clinical datasets contributed by six different health plans. The project team developed processes for matching patient records across the plans and piloted a process for sharing data between multiple health plan databases while safeguarding protected health information. Using the linked data, FDA performed an analysis to

demonstrate the usefulness of a standardized NDI linkage process to assess drug safety and effectiveness, adverse event surveillance, and CER in distributed data networks.

Products: This project produced an open-access methods protocol, processes for data linkages, and a distributed method for obtaining death and cause of death information to support multi-center research. FDA has shared the protocol with many of their stakeholders (e.g., FDA Sentinel, health plans) for studies that require NDI linkage to help ensure a successful NDI application and data linkage process. Please contact OSPCORTF@hhs.gov to obtain the final project report, which describes project objectives, methods, and accomplishments.



- ▶ **Improving the Efficiency of Research.** [The Methods Development Study Protocol](#), which was published in the *Journal of Medical Internet Research*, provides reusable, generalizable methods for linking multiple health plans' databases with NDI+ data. These standardized methods may assist researchers in assessing mortality-related safety questions in real-world settings. As of April 11, 2022, the published protocol has had 3,913 reads according to ResearchGate.
- ▶ **Addressing Important Research Questions.** As of January 2022, examples of topics that researchers have applied to use the linked NHCS and NDI data sets include: developing a method to better characterize transitions from nonfatal opioid hospital encounters to fatal overdose; studying the associations between patient, surgical, hospital-related factors and opioid overdose hospitalizations following orthopedic procedures; and identifying risk factors for opioid overdose deaths of concomitant drugs for patients with hospital and ED visits.

V. Conclusion

The overarching goal of the OS-PCORTF is to build lasting data infrastructure that researchers can use to conduct the PCOR studies needed to generate evidence and improve patient care. The projects highlighted in this report demonstrate the progress being made toward this goal and the scope of the projects' efforts to address areas of high research significance.

The OS-PCORTF addresses both longstanding and emergent needs in health research. In 2021, HHS priorities centered the portfolio's new projects on COVID-19, maternal health, the opioid epidemic, and SDOH and health equity. Projects that completed or expanded their work in 2021 contributed to addressing known data and evidential gaps in research and care (e.g., depression, opioid use disorder, weight management, diabetes).

The OS-PCORTF projects have improved the availability of data and linked a myriad of data sources to expand their research applications. These include federal data sources that contain important information on public health trends and health care utilization, as well as clinical data sources such as EHRs, hospitals, and registries. Evidence generated from these data sources will improve collective understanding of patient health and outcomes, and the cost and quality of health care being delivered. The projects are also prioritizing improvements in the quality and volume of data gathered directly from patients and communities to identify SDOH risk factors, increase representation of diverse populations, and continue to emphasize patient-centeredness in research and clinical care.

Most critically, the OS-PCORTF projects have developed research data infrastructure that is reusable, publicly available, and adaptable to meet researchers' needs. Building and strengthening the availability and suitability of data, as well as the analytical resources needed to address important questions, improves the quality and efficiency of PCOR studies, which, in turn, generates better evidence to improve health outcomes.

VI. References

¹ Public Law 116-94—Further Consolidated Appropriations Act, 2020. December 20, 2019. <https://www.congress.gov/116/plaws/publ94/PLAW-116publ94.pdf>

² Grigorescu V, Smith SR, Trevino M, Campbell NR. Building Data Capacity for Patient-Centered Outcomes Research (PCOR) for COVID-19: Highlights of Seven OS-PCORTF Funded Multiagency Data Initiatives. June 2021. <https://aspe.hhs.gov/sites/default/files/2021-07/covid-pcor-report.pdf>

³ The White House. Presidential Actions, Executive Order Executive Order On Advancing Racial Equity and Support for Underserved Communities Through the Federal Government. January 20, 2021. <https://www.whitehouse.gov/briefing-room/presidential-actions/2021/01/20/executive-order-advancing-racial-equity-and-support-for-underserved-communities-through-the-federal-government/>

⁴ Agency for Healthcare Research and Quality. December 2020. Social Determinants of Health Database (Beta Version). Retrieved from: <https://www.ahrq.gov/sdoh/data-analytics/sdoh-data.html>

⁵ Gliklich RE, Leavy MB, Cosgrove L, Simon GE, Gaynes BN, Peterson LE, ... & Trivedi MH. (2020). Harmonized outcome measures for use in depression patient registries and clinical practice. *Annals of internal medicine*, 172(12), 803-809. <https://effectivehealthcare.ahrq.gov/products/library-depression/white-paper>