



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

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, Dhopeshwarkar 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 Gherther (ASPE)	Steven Schwartz (CDC)
Aly Goodman (CDC)	Heather Stone (FDA)
Violanda Grigorescu (CDC)	Marissa Sucosky (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
Background.....	1
Portfolio Contributions to HHS Priorities	1
II) PCORTF Reauthorization and Strategic Plan.....	2
III) Portfolio Contributions to Key HHS Priorities	4
Addressing COVID-19 through Data, Research, and Infrastructure Investments.....	4
Improving Maternal Health Research and Outcomes	7
Data and Data Infrastructure to Combat the Opioid Crisis	9
Advancing Equity and Inclusion through Improved Collection and Use of SDOH and Other Non-Health Care Data to Enhance Person-Centered Care.....	11
Summary of Contributions to National Health Priorities.....	12
IV) Completed Projects' Major Accomplishments.....	13
Capstone for the Outcome Measures Harmonization (OMH) Project (AHRQ)	14
Strengthening the Data Infrastructure for Outcomes Research on Mortality Associated with Opioid Poisonings (CDC)	14
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)	15
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)	16
Clinical and Community Data Initiative (CODI) (CDC).....	16
Conclusion	17
V) Agency for Healthcare Research and Quality (AHRQ)	18
Capstone for Outcomes Measures Harmonization Project.....	18
Enhancing Patient-Centered Outcomes Research: Creating a National Small-Area Social Determinants of Health Data Platform	22
VI) Centers for Disease Control and Prevention (CDC)	25
Augmenting the National Hospital Care Survey Data through Linkages with Administrative Records	26
Building Infrastructure and Evidence for COVID-19 Related Research Using Integrated Data..	28
Clinical and Community Data Initiative (CODI)	30

Clinical and Community Data Initiative (CODI): Integrated Data for Patient-Centered Outcomes Research Project 2.0	33
Data Linkage: Evaluating Preserving Privacy Methodology and Augmenting the National Hospital Care Survey with Medicaid Administrative Records	35
Developing a Multi-State Network of Linked Pregnancy Risk Assessment Monitoring System (PRAMS) and Clinical Outcomes Data for Patient-Centered Outcomes Research.....	37
Enhancing Identification of Opioid-Involved Health Outcomes Using Linked Hospital Care and Mortality Data	40
Enhancing Surveillance of Maternal Health Clinical Practices and Outcomes with Federally Qualified Health Centers' Electronic Health Records Visit Data.....	43
Identifying Co-Occurring Disorders among Opioid Users Using Linked Hospital Care and Mortality Data.....	44
Making Electronic Health Record Data More Available for Research and Public Health (MedMorph).....	47
MAT-LINK: MATernal and Infant Network to Understand Outcomes Associated with Medication for Opioid Use Disorder during Pregnancy	50
MAT-LINK2: Expansion of MATernal and Infant Network to Understand Outcomes Associated with Treatment for Opioid Use Disorder during Pregnancy	53
Strengthening the Data Infrastructure for Outcomes Research on Mortality Associated with Opioid Poisonings.....	55
VII) Food and Drug Administration (FDA)	59
Bridging the PCOR Infrastructure and Technology Innovation through Coordinated Registry Networks Community of Practice	59
CURE ID: Aggregating and Analyzing COVID-19 Treatments from Electronic Health Records and Registries Globally	63
SHIELD: Standardization of Lab Data to Enhance Patient-Centered Outcomes Research and Value-based Care.....	65
VIII) National Institutes of Health (NIH).....	69
Creating a Federal COVID-19 Longitudinal Patient Outcomes Research Database Linked to Health Systems and Clinical Data	69
AMNET: An Addiction Medicine Network to Address the United States Opioid Crisis.....	72
Severe Maternal Morbidity and Mortality Electronic Health Record Data Infrastructure	74
IX) Office of the Assistant Secretary for Planning and Evaluation (ASPE).....	77
Dataset on Intellectual and Developmental Disabilities: Linking Data to Enhance Person-Centered Outcomes Research.....	77
Child and Caregiver Outcomes Using Linked Data (CCOULD).....	79

Multistate EMS and Medicaid Dataset (MEMD): A Linked Dataset for Patient-Centered Outcomes Research.....	82
X) Office of the National Coordinator for Health Information Technology (ONC)	85
A Synthetic Health Data Generation Engine to Accelerate Patient-Centered Outcomes Research.....	85
Using Machine Learning Techniques to Enable Health Information Exchange to Support COVID-19-Focused Patient-Centered Outcomes Research	87
XI) Cross-Agency Funded Projects	90
Data Capacity for Patient-Centered Outcomes Research through Creation of an Electronic Care Plan for People with Multiple Chronic Conditions	90
Data Capacity for Patient-Centered Outcomes Research through Creation of an Electronic Care Plan for People with Multiple Chronic Conditions 2.0: Development of the Patient-facing Application	94
Enhancing Data Resources for Researching Patterns of Mortality in Patient-Centered Outcomes Research: Projects 1 & 2	96
Making Medicaid Data More Accessible Through Common Data Models and FHIR APIs	101
Training Data for Machine Learning to Enhance PCOR Data Infrastructure.....	103
Understanding Long-term Outcomes in COVID-19 Survivors with Multiple Chronic Conditions through e-Care Plan Development	107
Validating and Expanding Claims-based Algorithms of Frailty and Functional Disability for Value-based Care and Payment.....	109
Appendix A. OS-PCORTF Project Portfolio.....	113
Appendix B. Abbreviations	118
Appendix C. Glossary	120
References	122

List of Exhibits

Exhibit 1.	OS-PCORTF Framework for Building Data Capacity for Patient-Centered Outcomes Research	3
Exhibit 2.	NCHS Data Linkage Program	26
Exhibit 3.	Distributed Health Data Networks	30
Exhibit 4.	EHR Data Accessibility: Cancer Use Case	47
Exhibit 5.	Maternal, Infant, and Child Outcomes of Interest	50
Exhibit 6.	SHIELD Project Overview	66
Exhibit 7.	Who Uses Patient Data from the MCC eCare Plan App.....	91

I) Introduction

The Office of the Secretary's Patient-Centered Outcomes Research Trust Fund (OS-PCORTF) Annual Report highlights the accomplishments of 34 multi-agency projects funded through the OS-PCORTF to build data capacity and support the Department of Health and Human Services (HHS) to advance science and improve knowledge about the comparative effectiveness of health care interventions. For Fiscal Year 2021 (FY 2021), the report focuses on projects aligned with four national health priorities—Coronavirus Disease 2019 (COVID-19), maternal health, social determinants of health (SDOH) and health equity, and the opioid crisis. The report also highlights 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

Background

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.

Portfolio Contributions to HHS Priorities

The FY 2021 report centers on the accomplishments of the multiagency HHS projects that are building PCOR data capacity to address critical and urgent national priorities—COVID-19, maternal health, , the opioid crisis, and health equity and social determinants of health.

COVID-19. These OS-PCORTF projects focus on standardization, harmonization, interoperability, and integration of data about the effects of the COVID-19 pandemic on patient health outcomes and the effectiveness of interventions for patients affected by COVID-19.

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).

Maternal Health Data and Linkages. This past year, the OS-PCORTF portfolio added projects on enhancing maternal health data to improve the timeliness and comprehensiveness of maternal health metrics, such as enhancing Healthcare Cost and Utilization Project (HCUP) data to produce quarterly reports of hospital-discharge metrics; using electronic health records (EHRs) for surveillance of maternal morbidity/mortality; and collecting signs, symptoms, and patient-reported outcomes (PROs) from pregnant women at the point of care.

Opioids and Mental Health. The OS-PCORTF portfolio also supported projects to address the opioid crisis with particular emphasis on improving the quality and timeliness of outcomes data, increasing collection of PRO information, and building linkages to address co-morbid conditions that affect patient outcomes.

Equity and SDOH. A combination of new and existing portfolio projects focuses on enhancing equity by strengthening data infrastructure for underrepresented, underserved, and at-risk populations through the collection and use of SDOH data.

The sections that follow highlight areas of potential interest to researchers and policymakers, namely:

- **Section II. PCORTF Reauthorization and Strategic Plan:** This section discusses the 2019 reauthorization of the OS-PCORTF for ten more years and ASPE's subsequent development of a new strategic plan to guide funding.
- **Section III. Portfolio Contributions to HHS Priorities:** This section highlights a selected set of projects that are focused on HHS's national health priorities, COVID-19, maternal health, the opioid crisis, and health equity and SDOH.
- **Section IV. 2021 Major Accomplishments:** Five OS-PCORTF projects completed their work in 2021. This section summarizes these projects and their publicly available products.
- **Sections V-XI. Project Profiles:** Organized by agency, the project profiles summarize project objectives and rationale, partnerships, accomplishments, and publicly available products.

II) PCORTF Reauthorization and Strategic Plan

The OS-PCORTF was created to help build national data capacity and infrastructure to support patient-centered outcomes research that provides decision-makers with high-quality evidence on the effectiveness of treatments, services, and other interventions used in health care.

The Office of the Secretary (OS) of HHS delegated authority to ASPE to coordinate “relevant Federal health programs to build data capacity for comparative effectiveness research (CER), including the development and use of clinical registries and health outcomes research networks.” In keeping with this charge, ASPE’s Office of Health Policy, through the OS-PCORTF, has funded and supported a portfolio of over 100 individual projects aimed at building data infrastructure capacity for patient-centered outcomes research. Together, these projects have produced approximately 120 data and technical products and 135 communication and dissemination products.

On December 20, 2019, Public Law 116-94 reauthorized the OS-PCORTF through 2029.¹ The reauthorization extends the Trust Fund’s work to include two new research priorities—intellectual and developmental disabilities 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.

Following the reauthorization of the OS-PCORTF in 2019, ASPE embarked on the development of a strategic plan to continue strengthening PCOR data capacity over the next decade. This plan charts a course for working in partnership with federal agencies and the broader PCOR community to harness the increasing volume and types of data being produced within and outside the health care system, along with advances in analytic methods.

The previous decade of OS-PCORTF investments focused on five functionalities needed to build PCOR data infrastructure (see Exhibit 1). ASPE intends to build upon this infrastructure to ensure responsiveness to emergent health challenges and evolving HHS priorities over the next decade.

The Strategic Plan sets forth a long-term plan for strengthening data capacity while supporting flexibility and ensuring responsiveness, consistent with ASPE's commitment to addressing emergent challenges and evolving HHS priorities. The agency's new strategic plan, which is expected to be released in early 2022, is the culmination 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.
- **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 conducted 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; 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.
- **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

Exhibit 1. OS-PCORTF Framework for Building Data Capacity for Patient-Centered Outcomes Research



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, OS-PCORTF and CMS' Health Federally Funded Research and Development Center (FFRDC) conducted 32 interviews with 62 participants, including HHS agency leaders, leaders of OS-PCORTF projects, and agency data experts.

This multi-stakeholder consensus process resulted in the creation of four priority goals for the OS-PCORTF, which reflect ASPE's role in building PCOR data infrastructure: 1) data capacity for national health priorities; 2) longitudinal data standards and linkages; 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 like the COVID-19 pandemic. 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 attention on the improvement and modernization of the health data infrastructure needed for effective pandemic response. ASPE funded seven COVID-related projects that are described in this report section, in the project profiles, and in a recent [white](#)

[paper](#) by ASPE.² 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 to establish new cross-agency connections or enhance existing ones.

The COVID-19 Collaborative. The collaborative brings together the portfolio's seven COVID-related projects, each of which are creating the building blocks of data infrastructure that will benefit current and future COVID-19 research. The COVID-19 Collaborative functions as a learning network in which these projects share project updates, feedback, and knowledge among fellow researchers. They meet monthly as a group and have also formed communities of practice (CoP) around two areas of need, SDOH and data linkages, with plans to form additional CoPs in the coming year.

- 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 team will present preliminary findings at AcademyHealth's Annual Research Meeting in 2022, with the final report being published on ASPE's website and as a peer reviewed manuscript.
- The Data Linkage CoP is focused on supporting the project **Building Infrastructure and Evidence for COVID-19 Related Research** (CDC, funded in 2021). Using integrated data from the National Center for Health Statistics (NCHS) Data Linkage Program, this project will rely on linked data from NCHS surveys with 2014-2019 Medicaid data from the Transformed Medicaid Statistical Information System (TMSIS). The linkage methodologies and resulting linked data can be used to expand data infrastructure capacity that can further enable PCOR goals, including examining the effects of different treatment regimens among vulnerable populations. The CoP is using this work as a basis for developing a data linkage paper focused on how other researchers can apply 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.

The National Institutes of Health (NIH)'s National Center for Advancing Translational Sciences (NCATS) project is **Creating a Federal COVID-19 Longitudinal Patient Outcomes Research Database Linked to Health Systems and Clinical Data** (2021). The 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.

AHRQ and the NIH's National Institute of Diabetes and Digestive and Kidney Diseases (NIDDK) have launched a new project focused on **Understanding Long-term Outcomes in COVID-19 Survivors with Multiple Chronic Conditions (MCC) through e-Care Plan Development** (2021). Their prior collaboration created provider-facing and patient-facing apps to facilitate data collection and care coordination for adults with MCC. The new project work will add data elements on COVID-19 and long COVID, with the goal of identifying risk and protective factors, natural history, and near- and long-term outcomes including symptoms and impact on functional status in COVID-19 survivors. The project will

also develop an app for caregivers to share real world data that helps improve care quality and coordination for patients with MCCs. In April 2022, NIH released a funding announcement 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.

The Food and Drug Administration (FDA)'s project, **CURE ID: Aggregating and Analyzing COVID-19 Treatment from EHRs and Registries** (2021), aims to identify promising COVID-19 treatments by finding new uses for approved therapeutics ("repurposed" drugs). 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.

A new project led by the Office of the National Coordinator (ONC) focuses on COVID-19 research, **Using Machine Learning Techniques to Enable Health Information Exchange (HIE) Data Sharing to Support COVID-19-focused PCOR** (2021). The 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 implement United States Core Data for Interoperability (USCDI) standards and HL7® Bulk FHIR® (Fast Healthcare Interoperability Resources) application programming interfaces (APIs) to support easy, efficient data access. 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.

Through its project, **Building Infrastructure and Evidence for COVID-19 Related Research Using Integrated Data** (2021), the Center for Disease Control and Prevention (CDC)'s NCHS will address the need for publicly available datasets that protect individual privacy. The project team plans to link data sources that address central PCOR questions, such as how health disparities affect infections, care, and outcomes, particularly for the COVID-19 pandemic. Data from the National Health Interview Survey (NHIS) and the National Hospital Care Survey (NHCS) will be 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 team will create synthetic data that can be linked to these datasets to preserve the privacy of individual records. The project will also 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.

Through its project, **Building Infrastructure and Evidence for COVID-19 Related Research Using Integrated Data** (2021), the Center for Disease Control and Prevention (CDC)'s NCHS will address the need for publicly available datasets that protect individual privacy. The project team plans to utilize existing NCHS linked data sources that address central PCOR questions, such as how health disparities affect infections, care, and outcomes, and synthetic data generation methods. These methods will then be applied to future linked files that focus on the COVID-19 pandemic. The project will use 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 team will create synthetic linked data files that preserve the privacy of individual records to be used to answer key PCOR questions. The project will also 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.

ASPE's project, **Multistate Emergency Medical Services (EMS) and Medicaid Dataset (MEMD): A Linked Dataset for PCOR** (2021), seeks to fill data infrastructure gaps that prevent researchers from following patient outcomes after emergency services. 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. While the project team will focus on their research on behavioral health emergencies (e.g., drug overdoses, psychotic episodes, suicidal ideation, and panic attacks), the database will include EMS data from patients with all conditions, meaning it can support a broad range of research on emergency care, including emergency responses that involve COVID-19.

For the **Data Set on Intellectual and Developmental Disabilities (ID/DD): Linking Data to Enhance Person-Centered Outcomes Research** (2021) ASPE will partner with four to six states to link state-level data sources: the National Core Indicators In-Person Survey, Supports Intensity Scale, Medicaid claims, and other relevant state-level data sources. The linked dataset will 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.

Examples of OS-PCORTF Contributions and Impact

- ▶ **Real-World 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 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 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 OS-PCORTF's Maternal Health Consortium formed as a natural extension of ASPE's 2020 funding process, which ultimately funded three project teams that share a focus on collecting maternal health data from EHRs. In the Maternal Health Consortium, the project teams exchange information that benefits their individual projects while also collaborating on shared, multi-agency activities.

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 as major EHR vendors that were invited to discuss their own approaches to data standards and linkages, providing a broader health ecosystem context for Consortium activities. The Consortium's focus is on assessing 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.

The NIH/NICHD project on **Severe Maternal Morbidity and Mortality EHR Data Infrastructure** (2021) focuses on improving the timeliness and comprehensiveness of the data needed to examine the effect of medical conditions and/or interventions on pregnant, post-partum, or lactating women and their infants. This project will develop a set of standard data elements for EHRs and a FHIR® API that researchers can use to access data on maternal morbidity/mortality from pregnancy through 1-year post-partum. The project will also create an implementation guide that details linkage methods to create longitudinal datasets that support research on pregnancy risk factors, outcomes, and the effect of complications during pregnancy on longer-term post-pregnancy health outcomes.

CDC's 2019 project, **MAT-LINK: MATernal and Infant NetworK to Understand Outcomes Associated with Medication for Opioid Use Disorder during Pregnancy**, has been expanded as **MAT-LINK2: Expansion of MATernal and Infant NetworK to Understand Outcomes Associated with Medication for Opioid Use Disorder during Pregnancy**. MAT-LINK established a surveillance network of four clinical sites to address: 1) the lack of national-level data on maternal, infant, and child health outcomes associated with different treatments for opioid use disorder (OUD) during pregnancy, and 2) a dearth of data on mother and child outcomes post-partum. MAT-LINK2 will expand the network from four clinical sites to seven, which will generate longitudinal data on OUD-related outcomes for over 4,000 mother-infant pairs; extend the follow-up period with children from infancy through age six; and increase the geographic and demographic representativeness of MAT-LINK. These data will support research that informs patient-centered care for pregnant women with OUD and infants and children with prenatal opioid exposure.

CDC's 2021 **Enhancing Surveillance of Maternal Health Clinical Practices and Outcomes with Federally Qualified Health Centers' (FQHCs) Electronic Health Records Visit Data** project focuses on upgrading data collection methods for FQHCs participating in the National Ambulatory Medical Care Survey (NAMCS), the only source of nationally representative visit-level data on ambulatory health

services provided at FQHCs. NAMCS will transition from manual processes to automated electronic data transfers from provider EHRs and create a consolidated file of maternal health visits. The project will also explore the linkage of this EHR data to the NDI and administrative data from HUD. The resulting data source will be a nationally representative set of data on maternal health treatment and outcomes at FQHCs that can be used for patient-centered outcomes research and CER.

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

The 2020 CDC project, **Developing a Multi-State Network of Linked Pregnancy Risk Assessment Monitoring System (PRAMS) and Clinical Outcomes Data for Patient-Centered Outcomes**

Research, will link state-level PRAMS data with birth certificates and clinical outcomes data (e.g., hospital discharge, Medicaid claims), with the potential to connect its data to other maternal and child health surveillance systems. 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 focus will expand on other data sources such as claims and data linkages.

Examples of OS-PCORTF Contributions and Impact



- ▶ **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.

ASPE's and the Administration for Children and Families (ACF)'s 2019 project, **Child and Caregiver Outcomes Using Linked Data (CCOULD)**, is creating a single harmonized dataset of Medicaid and child

welfare records. This involves linking patient-level data, including Medicaid enrollment, patient diagnoses, services, and claims, with child welfare data. The longitudinal data will allow researchers to better understand treatment 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. The project involves a significant technical assistance component to help states overcome data siloes and policy barriers and to build relationships that support data sharing.

Through the CDC's project, **Strengthening the Data Infrastructure for Outcomes Research on Mortality Associated with Opioid Poisonings**, the agency standardized the opioid-related mortality from death certificates with a new electronic coding system. The system allows for electronic coding of death certificates, incorporating details related to drug use and overdose, as well as 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 the data management systems used by medical examiners, coroners, and the state. This effort has improved the availability and shareability of information on deaths related to opioids, which will improve public health surveillance and patient-centered outcomes research.

The NIH/NIDA's project, **An Addiction Medicine Network (AMNET) to Address the United States Opioid Crisis**, is establishing a platform for research on OUD and SUD. It connects office-based practices and their data to other clinicians and researchers, enabling clinical trials and CER. AMNET has harmonized the measures used and is developing processes to link AMNET to other databases and registries. The resulting network will provide real-time data on patient characteristics, care delivery, and recovery service utilization that can be used for PCOR and efforts to improve addiction treatment and patient outcomes.

A 2018 CDC project also seeks to improve data on opioids and mortality, **Enhancing Identification of Opioid-Involved Health Outcomes Using Linked Hospital Care and Mortality Data**. 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 visits, inpatient hospitalizations, and overdose deaths, and to develop strategies to reduce the morbidity and mortality associated with opioid misuse.

CDC's 2021 **MAT-LINK** project has been expanded to additional clinical sites and age groups as **MAT-LINK2: Expansion of MATernal and Infant NetworK to Understand Outcomes Associated with Medication for Opioid Use Disorder during Pregnancy** (described in the maternal health section above) also contributes substantially to the availability of longitudinal, real-world data on opioid use. The CDC is conducting another opioid project (2019) focused on **Identifying Co-Occurring Disorders among Opioid Users Using Linked Hospital Care and Mortality Data: Capstone to an Existing FY18 PCORTF Project** that is profiled in the second half of the report.

These projects aim to improve the data infrastructure that can be used 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 enable better research.



- ▶ **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-Health Care 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 four 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.

CDC's projects, **Clinical and Community Data Initiative (CODI)** (2018) and **the Community and Clinical Data Initiative (CODI 2.0,** 2020), focus on building tools to link data from the health sector with community data. The project initially focused on creating a multi-sector, longitudinal data set that could improve child obesity programs. CODI 2.0 will further develop this infrastructure to create linkages across the health and social sectors for both adult data and pediatric data. These linkages will collect data from larger, more diverse geographic areas and data networks, and expand the data available for obesity, diabetes, and heart disease research.

AHRQ is addressing the need for granular, standardized SDOH data through their project, **Enhancing Patient-Centered Outcomes Research: Creating a National Small-Area Social Determinants of Health Data Platform** (2019). This project has combined existing federal datasets with other publicly available data into a national standardized database. The database contains extensive SDOH information that can be used by researchers to incorporate community SDOH characteristics in health outcomes research. AHRQ is actively maintaining and improving the beta version of the database, which is ready to support PCOR studies.

ASPE is leading the development of a publicly available dataset through the 2021 project **Data Set on Intellectual and Developmental Disabilities (ID/DD): Linking Data to Enhance Person-Centered Outcomes Research**. The 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 resulting dataset will fill a gap in the data available to conduct ID/DD-focused patient-centered outcomes research and generates evidence to support person-centric health decision-making and equitable policymaking.

The National Center for Health Statistics (NCHS) Data Linkage Program's 2020 project **Data Linkage: Evaluating Preserving Privacy Methodology and Augmenting the National Hospital Care Survey with Medicaid Administrative Records** improved access to linked federal data assets by linking data from the NCHS National Hospital Care Survey (NHCS) with T-MSIS data. The linkage expands research capacity on a wide range of HHS priority issues, 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.

Examples of OS-PCORTF Contributions and Impact

- ▶ **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 these key national health priorities:

- COVID-19
- maternal health
- opioids
- health equity and SDOH

As evidenced by the range of project work, the data infrastructure being created and enhanced 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, state, and 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 of projects 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.”⁵

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 innovations in the following areas:

- Linking clinical data to mortality data, for both for Medicare and commercially insured populations
- 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 study opioid-related mortality
- Linking EHR data, weight management intervention data, and community-level census data to facilitate patient-centered outcomes research 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 the five projects that concluded in FY 2021. Detailed project descriptions are provided in individual project profiles presented later in this report (see text box above for project names and links to the ASPE website).

 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)

Capstone for the Outcome Measures Harmonization (OMH) Project (AHRQ)

Registry data can be useful to researchers studying patient outcomes across diverse clinical sites. However, implementing harmonized outcomes measures can be difficult due to: 1) burden on clinical sites; 2) challenges extracting outcomes data from clinical records; and 3) challenges with working with EHRs.⁶ 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 these barriers.

The team created processes to help standardize data that patients and clinical providers collect for 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 Substitutable Medical Applications and Reusable Technologies (SMART) on FHIR® app that integrates clinical and patient-reported data from multiple sources; the app provides a ‘snapshot’ view of a patient’s depression treatment and harmonized outcomes, making it easier for clinicians to view outcomes over time and adjust treatment as needed. 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, an 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.⁷ 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.

Strengthening the Data Infrastructure for Outcomes Research on Mortality Associated with Opioid Poisonings (CDC)

Researchers and public health officials often use death certificate data, such as cause of death information, for programmatic, policy, and outcomes research; however, the quality of this data is limited by lack of data standardization, data extraction practices, and interoperability issues. This project aims to modernize the capture of death certificate data, improving the quality, timeliness, and amount of supplemental information on drug overdose deaths involving opioids.

In collaboration with National Center for Health Statistics (NCHS), the CDC team redesigned two subsets of the National Vital Statistics System (NVSS): the Vital Statistics Rapid Release (VSSR) and Medical Mortality Data System (MMDS). Enhancements to the VSSR include the addition of geographic, demographic, and drug information and automated monthly reporting, which has improved public health surveillance and research. The MMDS was also redesigned to electronically code death certificates; and to incorporate supplemental drug information into the NDI and the NVSS’ restricted-use multiple cause of death mortality files (NVSS-M).

The team also developed a FHIR® API to ease the exchange of mortality data between state’s electronic death registration system (EDRS) and local medical examiner/coroner offices. The FHIR® application is currently in use by Georgia and Washington, DC. The team established an advisory committee to ensure that these system improvements aligned with end-users’ needs in studying drug overdose deaths involving opioids.

Products: Access to the supplemental data files in the NDI and NVSS-M is restricted to approved users; however, a NVSS Modernization Tool Kit is publicly available, as are enhanced NVSS data dashboards. The team created a Vital Records Death Reporting FHIR® API that allows states to securely exchange mortality information within their jurisdiction; and developed an implementation guide to support users. The CDC team produced three National Health Statistics Reports, one Vital Statistics Rapid Release Report, and a research letter using the enhanced data reporting systems.

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 including fact and cause of death for all deaths within the U.S. This project aimed to create 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). 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.

The datasets in this seminal project allow researchers to analyze cause-specific death rates following hospital care. The data linkage processes developed for this project will continue to be used to link the NHCS with multiple administrative data sources including Medicare data, HUD housing assistance data, and Medicaid Transformed Medicaid Statistical Information System (T-MSIS) files to support PCOR data infrastructure goals. Additionally, this project served as the foundation for two subsequent OS-PCORTF projects to further enhance the NDI for patient-centered outcomes research, described later in this Portfolio Report.

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 one 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)

Patient-centered outcomes research often benefits from information on death and cause of death; however, publicly available data, including claims and EHR data, often lack complete information on the occurrence, causes, or manner of death for individuals. Project 2 piloted the linkage process described in Project 1, linking 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 health plans and piloted a process for sharing data between multiple health plan databases, while safeguarding protected health information. The team used the linked data to estimate the incidence of mortality and sudden cardiac death among two cohorts of antiarrhythmic medication users with differing underlying risks. This analysis demonstrated 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.

Clinical and Community Data Initiative (CODI) (CDC)

Limited information exists concerning the effectiveness of pediatric obesity intervention programs, given information silos that separate clinical records from data on health behaviors, clinical and community interventions, health outcomes, and SDOH and community-level factors (e.g., Census-level information). The CDC team built tools to collect, combine, and query these data sources, including: standard data elements for all settings to use when collecting data related to SDOH, geography, and demographic information. The team developed Privacy Preserving Record Linkage (PPRL) techniques to protect individual level data when sharing among organizations. Finally, the team developed data cleaning tools for large datasets, which will be made publicly available to help other organizations automate data cleaning processes. Pilot studies were then conducted in a distributed health data network, Colorado Health Observation Regional Data Service (CHORDS), to test the linkages and data sharing processes across three health systems and two community-based organizations. The integration of multi-sector data related to childhood obesity will help address research gaps in assessing childhood weight management programs⁸ and advance health equity by improving the study of SDOH, including racial/ethnic, economic, geographic disparities in childhood obesity prevention and treatment. In turn, the evidence generated will benefit children and families dealing with obesity by enhancing health care professionals' capacity to tailor treatment to patient needs.

Products: The project team has created a project infographic and technical information sheet as an overview of the project. Technical products developed under CODI include data models, data linkage techniques, and the implementation guides to support their use. Several manuscripts describing CODI's governance process, common data model, and PPRL methodology are in development.



- ▶ **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.

Conclusion

The overarching goal of the OS-PCORTF is to build lasting data infrastructure that researchers can use to generate PCOR 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 known 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 myriad 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. The projects featured in this report have advanced knowledge through their individual work and have also developed methods, tools, and data sets intended to provide a steppingstone for others PCOR researchers.

In the remaining sections, the report profiles the 34 OS-PCORTF projects that were active in 2021. Each profile describes project objectives, accomplishments, and contributions to the PCOR data infrastructure goals set forth in HHS' Strategic Plan for OS-PCORTF.

V) Agency for Healthcare Research and Quality (AHRQ)

AHRQ is administering a total of six active projects including four cross-agency projects described later in Section XI.

- Capstone for Outcomes Measures Harmonization Project
- Data Capacity for Patient-Centered Outcomes Research through Creation of an Electronic Care Plan for People with Multiple Chronic Conditions*
- Data Capacity for Patient-Centered Outcomes Research through Creation of an Electronic Care Plan for People with Multiple Chronic Conditions 2.0: Development of the patient-facing application*
- Enhancing Patient-Centered Outcomes Research (PCOR): Creating a National Small-Area Social Determinants of Health Data Platform
- Understanding Long-term Outcomes in COVID-19 Survivors with Multiple Chronic Conditions (MMC) through e-Care Plan Development*
- Validating and Expanding Claims-based Algorithms of Frailty and Functional Disability for Value-based Care and Payment*

* Denotes a cross-agency-funded project that involves more than one federal agency; these projects are described in the Cross-Agency-Funded Projects section (Section XI).

Capstone for Outcomes Measures Harmonization Project

Period of Performance	Federal Point of Contact
6/1/18 – 5/1/21	Elise Berliner

There are over 8,000 patient registries listed in [ClinicalTrials.gov](#) that can be used by researchers, physicians, and policymakers to advance patient outcomes.⁹ Registries provide a bridge connecting research and clinical practice. They collect data from across health systems and house data to support patient-centered outcomes research, population health management, and quality improvement initiatives. However, the ability to use data across patient registries for patient-centered outcomes research has been limited by the lack of standardized definitions for outcome measures captured within these registries. A prior OS-PCORTF funded project led by AHRQ titled [Harmonization of Clinical Data Element Definitions for Outcome Measures in Registries](#) harmonized outcome measures in five clinical areas including depression. Harmonized measures use standard definitions and concepts to classify outcomes, enabling cross-registry data comparisons. During this prior project, stakeholders identified three major barriers to the implementation of harmonized outcome measures: 1) burden on clinical sites to collect data; 2) disruption to clinical care and challenges in extracting data from the clinical records; and 3) challenges related to working with electronic health records (EHRs). While stakeholders recognize the importance of harmonized outcome measures, they are reluctant to implement and adopt harmonized outcome measures.¹⁰

Project Purpose and Goals

Using depression treatment as a use case, this capstone project addressed these barriers to information exchange between clinical sites and patient registries. The purposes of this project were to pilot harmonized depression outcome measures to 1) collect EHR data from multiple clinical sites on patient outcomes associated with depression treatment, 2) transfer the collected data to electronic patient registries for research, and 3) develop infrastructure and tools for other clinical registries and clinical systems to use. Through these activities, the AHRQ team demonstrated the feasibility of incorporating standardized data collection into the workflow and communication channels of busy clinicians, health systems, and patient registries. The project achieved these goals through the following objectives:

- Developing tools for clinicians and patients to facilitate integration of the harmonized depression outcome measures into EHRs and registries so that these data will be available for clinical research, patient-centered outcomes research, quality improvement, and implementation research.
- Providing proof-of-concept for a standards-based approach for collecting and reporting patient outcomes information to clinicians within their workflow and simultaneously transmitting the data to registries to make it available for research.
- Developing tools, such as instructions and code, to make it easier for researchers and registry developers to replicate the study's methods and integrate registries and clinical systems.

By creating standardized measures and linking clinical data from EHRs to patient registries to advance patient-centered outcomes research, these objectives align with ASPE's **Goal 2: Longitudinal Data Standards and Linkages** outlined in HHS' Strategic Plan for OS-PCORTF. Additionally, this project's use of natural language processing (NLP) to extract unstructured text data from EHRs for use in calculating harmonized depression outcome measures aligns with **Goal 3: Technology Solutions to Advance Research**.

Accomplishments

The *Capstone for Outcomes Measures Harmonization Project* concluded in FY 2021. The project team's major accomplishments are summarized below. The team began by selecting six harmonized depression outcome measures developed under the prior OS-PCORTF project to study: response, remission, recurrence, suicide ideation and behavior, adverse effects of treatment, and death from suicide. The AHRQ team then developed standardized processes to collect depression measures using clinical data from: EHR, patient-reported outcome measures obtained from the Patient Health Questionnaire-9 (PHQ-9), and data extraction from clinical notes using NLP.

The project team successfully configured the registry platforms of the American Board of Family Medicine's PRIME Registry™ and the American Psychiatric Association's PsychPRO to calculate the six harmonized depression measures. The team then demonstrated the feasibility of using the registries to extract standardized depression data from EHRs to calculate the harmonized outcome measures. AHRQ determined that calculating harmonized outcome measures using registry data is feasible and can be used to support patient-centered outcomes research. As part of this feasibility study, they identified and published important lessons learned on the technical and operational barriers to standardized data collection, as well as how they ensured that measures were meaningful to clinicians.

After assessing the feasibility of calculating these measures in the PRIME Registry™ and PsychPRO, AHRQ assessed the availability and quality of data in these registries for conducting robust PCOR studies. They found that, despite some gaps in contextual variables for describing patients with depression, it is feasible to conduct patient-centered outcomes research using the harmonized outcome measure definitions and data from the two registries.

The AHRQ team then sought to determine if health systems could collect the patient-reported data needed to calculate the harmonized depression outcome measures, and then return these results back to clinicians via the EHR. To do so, AHRQ successfully developed and implemented the SMART on FHIR® application across 21 clinical sites within one integrated health system, Baystate Health. The app integrates clinical and patient-reported data from multiple sources to provide a “snapshot” view of a patient’s depression treatment and harmonized outcomes, thus making it easier for clinicians to view outcomes over time and adjust treatment as needed. The team also assessed the value and burden of calculating the harmonized outcome measures at the clinician level, health system level, and registry level. They found that the introduction of measures collection increased clinician, health system, and registry operator burden, but not overwhelmingly so.

Implications for Harmonized Outcome Measure Implementation

Lessons learned from this project can be used to inform future implementation of harmonized outcome measures, beyond those focused on depression. AHRQ researchers recommend that future projects on harmonized measures:

- Develop flexible, patient-centered plans for collecting patient-reported outcomes (PROs)
- Use a multi-pronged approach to capturing data
- Build on existing standards-based resources

Finally, the team produced a prioritized research agenda for patient-centered outcomes research in depression that harmonized depression outcome measures can address. The research agenda covers eight priority areas including treatment effectiveness; variation across care settings; screening, diagnosis, and prevention; treatment-resistant depression; impact of race, ethnicity, culture, and other factors on outcomes; depression and comorbidities; perinatal and postpartum depression; and suicidality. Additionally, the research agenda identifies patient-centered questions of interest covering four topics: screening, diagnosis, and prevention; social determinants of health; treatment effectiveness; and variation across care settings. To support future use of harmonized depression outcome measures, the team also developed a data use and governance toolkit in collaboration with a stakeholder panel, as well as standardized resources including FHIR® profiles and implementation guides for the harmonized outcome measures as well as libraries, app source code, and documentation. These resources can be used by other clinical sites interested in adopting harmonized outcome measures.



Data and Technical Products

- **SMART on FHIR® App:** Source code and related technical documentation for the Major Depression Outcomes app are publicly available. The purpose of this app is to provide a “snapshot” view of a patient’s depression symptoms and outcomes. The app uses scores from the PHQ-9 and clinical data to calculate and display standardized depression outcome measures.¹¹
- FHIR® Outcome Criteria [Framework](#) and [Implementation Guide](#) (version 1.0): Standardized implementation models, FHIR® profiles, and implementation guide for AHRQ’s Outcome Measures Framework (OMF) are publicly available. These resources can be used by other clinical sites to implement harmonized outcome measures in their systems.^{12,13}

- [**FHIR® Library for Depression Outcome Measures**](#): This GitHub library includes the harmonized depression outcome measures, including CQL and links to appropriate value sets. This library can be used by others to reproduce code for condition outcome definitions and criteria for use cases beyond those studied by the AHRQ team.¹⁴

Communication and Dissemination Products

- [**Standardized Library of Depression Outcome Measures White Paper**](#): This research paper describes the technical approach used to prepare the Standardized Library of Depression Outcome Measures.¹⁵
- [**Beyond Harmonization: Implementing Standardized Outcome Measures to Support Value-Based Care Research Article**](#): This article outlines AHRQ's Outcomes Measures Framework, used to produce standardized outcome measures.¹⁶
- [**Methodology Report: Outcome Measure Harmonization and Data Infrastructure for Patient-Centered Outcomes Research in Depression**](#): This methodology report summarizes processes for data extraction and exchange and includes lessons learned that can be applied in other clinical settings for reporting data to registries and clinicians.¹⁷
- [**A Prioritized Research Agenda for Using the Harmonized Outcome Measures in the Support Patient-Centered Outcomes Research in Depression Research White Paper**](#): This report summarizes research priorities and questions for future work on harmonized outcome measures in depression, as determined by a stakeholder panel. The report also lays out a roadmap for implementing the research agenda.¹⁸
- [**Data Use and Governance Toolkit Research White Paper**](#): This toolkit summarizes current best practices in sharing registry data and provides additional information to assist registries in sharing data including regulatory and privacy precautions, data governance structure and procedures, and procedures for reviewing and responding to data requests.¹⁹
- [**Final Report**](#): This final project report describes project objectives; findings and conclusions from major activities; and lessons learned. This report also provides insight on implications for future use of harmonized outcome measures and research needs in the field.²⁰
- [**Project Website**](#): This AHRQ webpage provides an overview of the capstone project and includes links to all reports produced under this work.²¹

Enhancing Patient-Centered Outcomes Research: Creating a National Small-Area Social Determinants of Health Data Platform

Period of Performance

3/1/19 – 9/1/22

Federal Point of Contact

Patricia Keenan

To deliver high-quality health care, it is important to understand the social determinants of health (SDOH) of patients and their communities. Considerable evidence exists about the relationship between the inequities in SDOH factors and poor health outcomes (e.g., mortality, acute and chronic disease, disability).^{22,23} Some studies have shown a relationship between SDOH factors and health care utilization, and studies are beginning to emerge on the effectiveness of the health care interventions that integrate patient and community SDOH information on patient and community health outcomes. As such, there is a growing demand for data that integrates information about SDOH, health service utilization, and systems of care. Research has demonstrated that for many SDOH factors, small-area data (i.e., data at the community or sub-county level) may be necessary to conduct meaningful analyses, and that for other SDOH factors, data at other geographic levels are more meaningful.

Researchers can spend substantial resources linking multiple datasets to create data files suitable for analyses because the current data are spread across many sources, topics, and levels of geography. There is no complete source of longitudinal information with uniformly formatted community-level data on SDOH readily available for health services research.

Project Purpose and Goals

The goal of the project is to develop a national standardized database on valid and reliable SDOH factors at the small-area and other geographic levels, building on existing databases developed by federal agencies (e.g., ASPE, Centers for Disease Control and Prevention, Health Resources and Services Administration, and National Institutes of Health), and other publicly available sources. This database can serve as a central place for researchers to access SDOH data elements that correspond to different SDOH domains. Data elements will span the SDOH landscape and include measures of income, employment, food, housing, environment, economics, education, safety, transportation, justice system, market structure, health status, health care access, and utilization. The SDOH data platform will be constructed longitudinally with 10 years of retrospective data available. Public use files and supporting documentation in a standardized, structured format will be developed and made publicly available.

By expanding data capacity, this project will enable research that examines interventions that can prevent disease, the effectiveness of interventions tailored to the whole person including the community in which they live, and the utility of system-focused interventions. These types of PCOR studies will ultimately inform the value transformation of the health care system, enabling policymakers, providers, and payers to make better decisions for patients and the health care system. The overall objectives of the project are to:

Examples of federal datasets being leveraged to create the SDOH database

- American Community Survey
- Area Health Resources Files
- Civil Rights Data Collection
- County Health Rankings
- Nursing Home Compare
- U.S. Cancer Statistics

- Identify a comprehensive set of datasets with existing or analyzable small-area level and other geographic level data on high priority SDOH data elements.
- Design and create a publicly available database of valid and reliable standardized sets of SDOH metrics at various geographic areas. Hosted on an AHRQ website, the SDOH database will contain data files that can be linked to other data sources through geographic identifiers. In addition, users will have web access to documentation about the SDOH factors, methodological reports, and interactive web queries.
- Coordinate and expand the data collection efforts on SDOH across HHS.
- Use the new data to conduct PCOR studies, demonstrating their research utility.
- Disseminate the SDOH database for use across the federal government, PCOR researchers, and health services researchers.
- Establish a sustainability and growth plan for the SDOH data for future development of the depth and breadth of SDOH information for use with health services research.

SDOH Beta Data Files

The beta version of the SDOH database is available for research use. The data files encompass 10 years of county-level files (2009-2018) and 8 years of zip code tabulation area-level files (2011-2018) including SDOH on social context, economic context, education, physical environment, and healthcare context.

The team has produced a codebook and supporting documentation for these data files.

A 2022 update will add Census tract-level data files, and additional datasets and variables.

By expanding the collection and analysis of socioeconomic, environmental, and other data, the aims of **Enhancing PCOR: Creating a National SDOH Data Platform** project align with HHS' Strategic Plan for **OS-PCORTF Goal 4: Person-Centeredness, Inclusion, and Equity**. With 10 years of longitudinal, retrospective data from multiple HHS agencies, it also aligns with **Goal 2: Longitudinal Data Standards and Linkages**.

Coordination with Other Federal Agencies

To build a comprehensive SDOH database, AHRQ will work with numerous departmental and federal agencies to access and link their datasets. Coordination with those agencies will be critical to continuing to support health care services research. AHRQ held several preliminary conversations with HHS agencies about including data from the SDOH database in their data resources to facilitate access and promote use of SDOH data. Additionally, since the SDOH beta files became available, AHRQ held multiple meetings to coordinate and gather feedback from other agencies on what additional SDOH measures can be included in the database.

Accomplishments

The AHRQ team has completed development of the beta version of the SDOH database and is working to complete a final version in Spring 2022.

- At present, the SDOH database contains 10 years of county-level data and 8 years of zip code tabulation area-level data on topics spanning the SDOH domains of social context, economic context, education, health care, and physical environment.

- Through approximately 25 additional conversations with internal and external stakeholders and end-users, the team has disseminated information about the database and identified several areas of expansion that will be reflected in the 2022 update.
- The Spring 2022 database will add census tract-level data files, and additional datasets and variables. We anticipate that the Spring 2022 version of the database will be a valuable resource for the patient centered outcomes research community, local health systems, and policymakers.
- The AHRQ team has established and maintained relationships across HHS in the planning and development of the database. This includes obtaining input on the project itself as well as ongoing participation in meetings supporting the HHS SDOH workgroup.
- The team has begun work on a 5-year strategic plan for the SDOH database. This plan is intended to ensure the sustainability of the SDOH platform and integration into HHS data enterprise structure informed by extensive conversations and data linkages/applications for PCOR and HHS programs.

Data and Technical Products

- **The SDOH Database (Beta Version)** is available for researchers and is accompanied by supporting documentation and codebook.²⁴ These SDOH beta data files are curated from existing federal datasets and other publicly available data sources. The purpose of the files is to make it easier to find a range of well documented, readily linkable SDOH variables across domains without having to access multiple source files, facilitating SDOH research and analysis. As of November 2021, there were about 10,000 file downloads from AHRQ's SDOH database web page. An update in the spring of 2022 will further expand the available data.
- **The Poverty and Access to Internet, by County** visualization was created to demonstrate how the SDOH database can be used to generate insights on SDOH. Given the growing importance of internet in accessing health care, this visualization shows county-level percentages of households with computers and smartphones and percentages of households with any type of broadband (including cellular data plans).
- **Tools for practice improvement** have been published to help health care organizations address SDOH. These include resources for forming community linkages, care coordination, delivering culturally and linguistically competent care, delivering health literate care, and screening and referral support for primary care providers.²⁵

Communication and Dissemination Products

- An **SDOH Environmental Scan** that includes information on SDOH databases categorized by SDOH domains was published. The environmental scan is organized in an Excel spreadsheet to maintain the filter functionality of each column in the scan.²⁶
- AHRQ has conducted one presentation at Academy Health, two presentations to the ASPE PCORTF seminar series, one briefing to bipartisan staff of House Energy and Commerce Committee, one briefing to the Senate Finance Committee majority staff, and two posters at Academy Health about the project.
- **A summary of an SDOH data visualization** was included in the *Journal of the American Medical Association* in July, in the "Health Agencies Update" section.²⁷

VI) Centers for Disease Control and Prevention (CDC)

CDC is administering 15 active projects including two cross-agency-funded projects described later in Section XI.

- Augmenting the National Hospital Care Survey Data through Linkages with Administrative Records: A Capstone Project
- Building Infrastructure and Evidence for COVID-19 Related Research, Using Integrated Data
- Community and Clinical Data Initiative
- Community and Clinical Data Initiative (CODI) 2.0: Integrated Data for Patient-Centered Outcomes Research Project
- Data Linkage: Evaluating Preserving Privacy Methodology and Augmenting the National Hospital Care Survey with Medicaid Administrative Records
- Developing a Multi-State Network of Linked Pregnancy Risk Assessment Monitoring System (PRAMS) and Clinical Outcomes Data for Patient-Centered Outcomes Research
- Enhancing Data Resources for Researching Patterns of Mortality in PCOR: Project 1 – Adding Cause-Specific Mortality to NCHS's National Hospital Care Survey by Linking to the National Death Index *
- Identifying Co-Occurring Disorders among Opioid Users Using Linked Hospital Care and Mortality Data
- Making Electronic Health Record Data More Available for Research and Public Health
- MAT-LINK: MATernal and Infant Network to Understand Outcomes Associated with Medication for Opioid Use Disorder during Pregnancy
- MAT-LINK2: Expansion of MATernal and Infant Network to Understand Outcomes Associated with Medication for Opioid Use Disorder during Pregnancy
- Strengthening the Data Infrastructure for Outcomes Research on Mortality Associated with Opioid Poisonings
- Validating and Expanding Claims-based Algorithms of Frailty and Functional Disability for Value-based Care and Payment*

* Denotes a cross-agency-funded project that involves more than one federal agency; these projects are described in the Cross-Agency-Funded Projects section (Section XI).

Augmenting the National Hospital Care Survey Data through Linkages with Administrative Records

Period of Performance

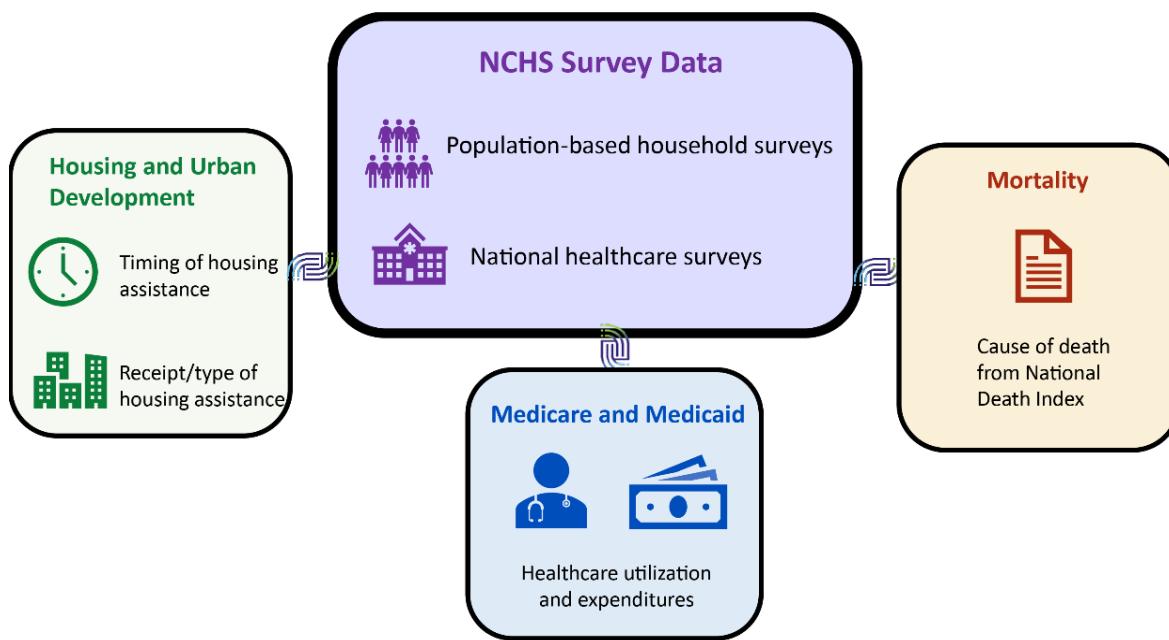
6/1/19 – 12/1/21

Federal Point of Contact

Lisa Mirel

The National Hospital Care Survey (NHCS), conducted by the National Center for Health Statistics (NCHS), is designed to provide accurate and reliable health care statistics, describing national patterns of health care delivery in hospital-based settings, including prevalence of conditions, health status of patients, and health services utilization. The NHCS collects patient-level identifiers, which enables linkage of patient episodes of care within hospital inpatient and emergency department settings to other administrative data sources, providing a more complete picture of patient care. Previously funded OS-PCORTF projects have linked the 2014 and 2016 NHCS to mortality data collected from the National Death Index (NDI), creating a new unique data resource to support the study of post-hospitalization mortality outcomes in more than 3.2 million patients. The Trust Fund also supported the linkage of the 2014 NHCS to the 2014-2015 Centers for Medicare & Medicaid Services (CMS) Master Beneficiary Summary File (MBSF), which links claims data from NHCS patients with Medicare coverage to Medicare program enrollment and summary cost and utilization data at the time of hospitalization and after.

Exhibit 2. NCHS Data Linkage Program²⁸



F Project Purpose and Goals

This project expands upon previously funded OS-PCORTF projects that increased the capacity of the NHCS to support a wide range of PCOR questions. This project links the 2016 NHCS with 2016-2017 CMS claims data (Medicare Fee-for-Service, Medicare Advantage encounters, and patient assessment data from long-term care facilities and home health providers) and federal housing assistance program data collected from the U.S. Department of Housing and Urban Development (HUD) (Exhibit 2). These linked data resources make it possible to integrate information on mortality, health care service utilization, prescription drug use, provider health assessments, and receipt of federal housing with a given patient's

hospital administrative claims and EHR data. The linkage between the NHCS to CMS and HUD data sources expands data capacity to support research studies focused on a wide range of patient health outcomes including initiatives targeting opioid use and mental health care services, efficacy of treatment protocols, medical interventions and prescription drugs, health outcomes associated with different types of post-acute care services utilization, and health disparities. The 2016 NHCS and 2016-2017 HUD linked data sources allows researchers to examine the role of federal social support programs in health outcomes and treatment efficacy for persons with stable housing, with the ability to focus on specific subpopulations, including persons with substance use disorders.

The project focuses on the following objectives:

- Conduct a patient-level record linkage of the 2016 NHCS hospital administrative claims and EHR data to 2016-2017 CMS Medicare claims encounter and health assessment data.
- Conduct a patient-level record linkage of the 2016 NHCS hospital administrative claims and EHR data to the 2016-2017 HUD administrative records on federal housing program participation.
- Refine linking algorithms and disseminate detailed statistical methodology reports to support high-quality future data linkage activities within and beyond the patient-centered outcomes research community.
- Create research files and user guidance documents to support researchers in using the new NHCS-linked data resources.
- Disseminate tools and analytic guidance to stimulate the broader use of this new data resources to expand the capacities of patient-centered outcomes researchers.

By linking clinical, intervention, and community data to create and standardize longitudinal individual-level records while preserving privacy, the aims of **Augmenting the NHCS Data through Linkages with Administrative Records** align with **Goal 2: Longitudinal Data Standards and Linkages** of HHS' Strategic Plan for OS-PCORTF. Additionally, this project's inclusion of individual-level social determinants of health data in linked datasets and emphasis on creating a more diverse sample will serve **Goal 4: Person-Centeredness, Inclusion, and Equity**.

Accomplishments

Since work on the project began in June of 2019, the project team has successfully accomplished project objectives, expanding data capacity to support research studies focused on a wide range of patient health outcomes related to opioid use and mental health care services and treatment protocols:

- The project team tested multiple methods for the sequential covering algorithm to link NHCS to CMS data, and subsequently completed the linkage of 2016 NHCS and 2016/2017 CMS data and published supporting documentation.
- The project team also linked 2016 NHCS data with 2016/2017 HUD research data files and published supporting documentation. The team completed an assessment of the quality of the linkage methodology using household survey data and demographic characteristics.



Data and Technical Products

- **Linked 2016 NHCS and 2016-2017 CMS Medicare Data:** The linked file was finalized in September 2020 and was made publicly available to researchers through the RDC. The NCHS Data Linkage website includes [a report](#) on the linkage methods and the analytic considerations,²⁹ and [codebooks](#) for the final 2016 NHCS data linked to 2016-2017 CMS Medicare.³⁰
- **Linked 2016 NHCS data to 2016-2017 HUD Data:** The linked data are now available to researchers through the NCHS Data Linkage website includes [a report](#) on the linkage methods analytic considerations,³¹ and the [codebooks](#) for the final 2016 NHCS data linked to 2016-2017 HUD files.³²
- Published algorithms from the 2018- 2019 NCHS Data Linkage projects: The algorithms from the [Opioid NLP](#),³³ [Medical Code-based \(translation from SAS\)](#),³⁴ and [SUD and MHI NLP](#)³⁵ projects are now available on GitHub.



Communication and Dissemination Products

- **Using Supervised Machine Learning to Identify Efficient Blocking Schemes for Record Linkage:** This journal article was published in the Statistical Journal of the International Association for Official Statistics in June 2021. The publication describes a case study using the Sequential Coverage Algorithm (SCA), a supervised machine learning algorithm. The SCA was used to develop the strategy for linking 2016 NHCS data with the CMS Enrollment Database.³⁶
- **Data Linkage Webinar:** Featuring members of the project team, this October 2020 webinar provides information on the NCHS Data Linkage Program including linked NCHS data sources created by the project.³⁷
- **NCHS Data Citation List:** The NHCS Data Linkage Team periodically updates this list of publications using the linked NHCS data, including the files produced by this project.³⁸

Building Infrastructure and Evidence for COVID-19 Related Research Using Integrated Data

Period of Performance

5/7/21- 5/7/24

Federal Point of Contact

Lisa Mirel

Conducted by the National Center for Health Statistics (NCHS), the National Hospital Care Survey (NHCS) provides robust data on health care delivery in hospital-based settings. Patient-level identifiers collected by the NCHS enable the linkage of data from hospital-based settings to other data sources. These data linkages are critical for patient-centered outcomes research that examines the complete picture of patient care. Previously funded OS-PCORTF projects have expanded research capacity by linking NHCS data to mortality data collected from the National Death Index (NDI), federal housing data from HUD, and Medicare data from CMS. To protect confidentiality, these linked data are only available in restricted-use files that must be accessed through the NCHS and Federal Statistical Research Data Centers (RDC). This creates barriers for researchers and reduces the utility of linked data.

Project Purpose and Goals

This project will expand the utility of current and future data linkages for researchers and reduce barriers to access by developing publicly available synthetic linked data products. These data products will protect participant privacy while integrating social determinants of health, health-related, and administrative data. Ultimately, the data products will not only support patient-centered outcomes research about COVID-19 symptoms, treatments, vaccinations, and long-term impacts, but also strengthen data infrastructure to examine and respond to future public health emergencies.

The objectives of this project are to:

- Engage stakeholders to identify the greatest data needs and prioritize data items for public-use release
- Establish a synthetic data creation methodology that can be applied to linked data resources
- Create publicly available synthetic linked data files and a data enclave validation process to confirm analytic results with restricted-use data
- Create a publicly available analytic dashboard using linked data on the NCHS website to broaden data accessibility and utility for a more wide-ranging community of users
- Disseminate tools and analytic guidance to encourage the broader use of this new data resources to expand PCOR capacity

By linking clinical, intervention, and community data to create and standardize longitudinal individual-level records while preserving privacy, **Building Infrastructure and Evidence for COVID-19 Related Research Using Integrated Data** aligns with **Goal 2: Longitudinal Data Standards and Linkages** of HHS' Strategic Plan for OS-PCORTF.

Accomplishments

Since work on the project began in May of 2021, the project team has started stakeholder engagement activities that will inform the development of the synthetic linked data products. The team awarded a contract to conduct expert interviews focused on identifying the variables to include in the synthetic linked data files.

Communication and Dissemination Products

- [American Statistical Association Links Lecture](#): A December 2020 presentation from project lead Lisa Mirel discussed this project in the context of advancing health data linkages.

Clinical and Community Data Initiative (CODI)

Period of Performance

5/30/18 – 9/30/21

Federal Point of Contact

Aly Goodman

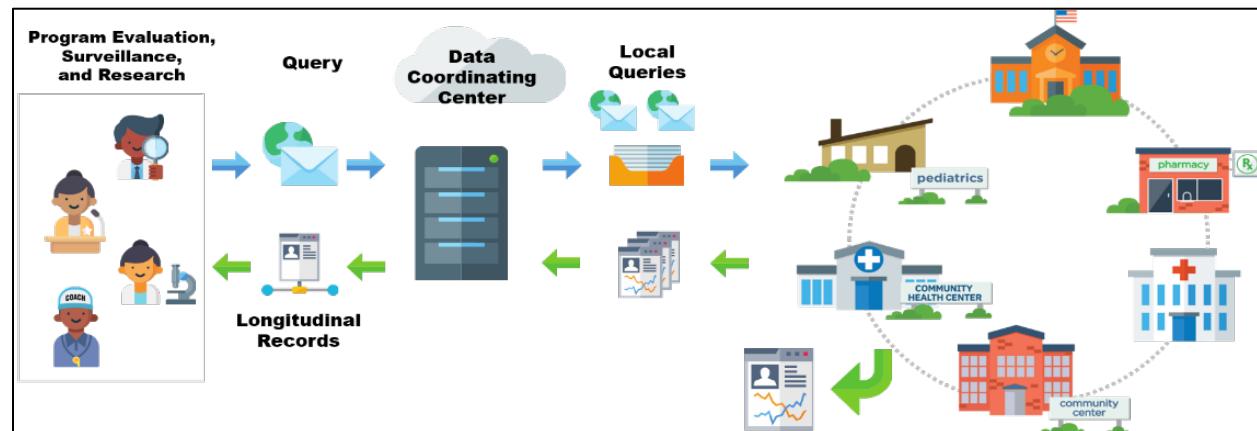
Marissa Sucosky

In 2012, the World Health Organization (WHO) identified childhood obesity as one of the most pressing public health challenges of the current century.³⁹ In the U.S., prevention and treatment of childhood obesity is a national priority. Evidence-based guidelines recommend screening all children for obesity and referring children with obesity to comprehensive, intensive, family-centered weight management programs (WMPs).⁴⁰ However, the research is unclear on which intervention components work best for different children. The fact that relevant outcomes data are siloed within three distinct sectors (health care, intervention programs, and communities) contributes to limited understanding. Distributed patient-centered research networks routinely gather data collected in health care settings and structure these data in a common way, known as common data models (CDMs), to link data from different sources. These networks, though capable of combining patient-level health intervention and community-level data, lack coding for children's data, so these types of linkages have been limited.

Project Purpose and Goals

The Clinical and Community Data Initiative (CODI), previously known as the Childhood Obesity Data Initiative, addresses gaps in technology and data accessibility by making multi-sector data accessible to local communities and researchers. With access to such data, programs and researchers can assess and compare the effectiveness of treatment and intervention programs. CODI developed processes to create individual-level records containing multiple years of data. CODI linked patient-level EHR data, intervention (both community and clinical) data, and community-level Census data (e.g., average household income in the patient's census block). The project piloted these tools in a National Patient-Centered Clinical Research Network (PCORnet®) distributed health data network (DHDN), the Colorado Health Observation Regional Data Service (CHORDS). See Exhibit 3 for a visual illustration of how a DHDN operates. By building data linkages and more advanced tools, CODI will help researchers and clinicians personalize interventions for children and their families' health, sociodemographic, and social factors.⁴¹ These data will strengthen comparative effectiveness research by helping researchers identify key characteristics of WMPs that are most effective, such as provider and service type, setting, timing, and duration.

Exhibit 3. Distributed Health Data Networks



In 2020, the OS-PCORTF awarded additional funding to the CDC team to expand the CODI project. Building upon the original CODI project, CODI 2.0 will further integrate social determinants of health (SDOH) data to address gaps in chronic disease patient-centered outcomes research for all ages. As part of this new phase, CODI's Colorado site expanded its focus from childhood obesity to other chronic diseases that impact individuals across the lifespan, including heart disease and diabetes. Please see the [Portfolio Report entry on CODI 2.0](#) for additional information about this new project.

The objectives of the CODI project are to:

- Establish a partnership of 15 subject matter experts, including health care providers, researchers, health systems, public health departments and federal agencies to guide the project and to ensure safety and security of all data.
- Leverage the widely adopted PCORnet® CDM to expand the ability to collect, combine, and query existing patient-level EHR and WMP data, as well as community-level census information (e.g., average household income in a patient or program's census geographic unit).
- Expand the linkage and de-duplication tools for integrating childhood obesity data, publicly available through CDC's cloud-based Surveillance Data Platform.
- Pilot and evaluate the expanded CODI technical services among multi-sector partners, testing the systems' ability to capture, link, and query clinical, intervention and community data to produce datasets for PCOR researchers to analyze.

By linking clinical, intervention, and community data to create and standardize longitudinal, individual-level records while preserving privacy, the aims of the CODI project are in alignment with HHS' Strategic Plan for OS-PCORTF **Goal 2: Longitudinal Data Standards and Linkages**. Additionally, CODI's analysis of the impact of socioeconomic and environmental factors on health outcomes and intervention effectiveness aligns with **Goal 4: Person-Centeredness, Inclusion, and Equity**.

Accomplishments

The CDC team not only successfully implemented the CODI pilot project to link clinical and community data across multiple partner organizations, but in doing so they also developed multiple resources that other clinical sites can use to assess effectiveness of childhood obesity treatments. CODI successfully partnered with three health systems and two community-based organizations in Denver, Colorado to conduct a pilot implementation program, which was completed in 2021 but continues its work as part of CODI 2.0. In this pilot, investigators linked child health-related data from Denver Health, Children's Hospital Colorado, Kaiser Permanente Colorado, Girls on the Run of the Rockies, and Hunger Free Colorado. To assess the technology infrastructure and needs at partner sites, the CDC team developed a clinical and community linkage assessment. The team has since used these activities to design external-facing resources for multisector collaboration, information sharing, and governance.

The project developed and piloted enhanced tools and services to 1) create individual-level, longitudinal records of linked multi-sector data relevant for childhood obesity patient-centered outcomes research and 2) promote sharing individual and community data across different sectors and organizations while preserving privacy. Specifically, the CDC team updated CDM tables to include new elements related to SDOH, geography, and clinical and community programs. These updates created standardized approaches to collecting data for studying childhood obesity. They also developed Privacy Preserving Record Linkage (PPRL) techniques that protect individual level data in a secure, private format before

sharing across organizations, through a process called data hashing. Finally, the team developed data cleaning tools for large EHR datasets, which will be made publicly available to help other organizations automate data cleaning processes. Health care and community organizations can use these tools to automate processes for standardizing and sharing data, thus reducing the burden for organizations looking to enhance their data infrastructure.

In addition to producing data linkage solutions, CODI functions as a DHDN. As such, CODI translates a question from researchers, clinicians, or community partners into a query that then requests and returns information from multiple databases as a single dataset to the requestor. To support external pediatric obesity research, the CDC team developed distributed queries that researchers and partners can use to ask the data pre-determined questions. The queries allow researchers to examine 1) the local burden of childhood obesity, and 2) the impact of a service or an intervention on individuals' health. The team will make guidance for running these distributed queries publicly available upon project conclusion.

Partners at the Colorado sites have implemented the CODI tools locally and are working with the CDC technical lead to identify and resolve issues. Once issues have been resolved, the longitudinal dataset will be publicly available; currently, external researchers can send queries to the Denver sites to request data. The CDC team is currently wrapping up their evaluation of the effectiveness of the CODI tools and products. As part of the evaluation activities, the CDC team will produce a final report detailing findings and a summary of lessons learned as well as a sustainability plan for CODI.



Data and Technical Products

- **[CODI Data Models Implementation Guide](#)**: This resource is available on GitHub and contains implementation guidance on the CODI Research Data Model (RDM) and CODI Record Linkage Data Model (RLDM), including descriptions of each and specific guidance regarding individual data elements of the CODI Data Models. This guide can be used by other organizations looking to implement data models developed under CODI.⁴²
- **[CODI Privacy Preserving Record Linkage \(PPRL\) Implementation Guide](#)**: This document, also available on GitHub, provides guidance on implementing PPRL for participating organizations, and includes a description of the PPRL process, specific guidance for each PPRL role, and performance evaluation guidance. This guide can be used by other organizations looking to replicate CODI's PPRL technique.⁴³

Products Under Development

- Website link to packaged query code and implementation guide for childhood obesity PCOR data
- Report on the processes and lessons learned from the CODI pilot implementation
- Sustainability plan
- Manuscripts covering five different topics, 1) foundation of CODI, 2) CODI's governance process, 3) the CODI CDM, 4) the EHR data cleaning tool, "growthcleanr", and 5) the PPRL tool



Communication and Dissemination Products

- **[CODI Fact Sheet](#)**: This information sheet from CDC provides additional background information on the CODI projects. It outlines the project vision, the CODI innovation, and the CODI pilot tests.

Clinical and Community Data Initiative (CODI): Integrated Data for Patient-Centered Outcomes Research Project 2.0

Period of Performance

8/1/20 – 3/21/24

Federal Points of Contact

Aly Goodman

Ray King

Addressing chronic diseases, such as obesity, is a top health concern, as chronic diseases are the leading causes of death and disability in the United States.^{44,45} Chronic diseases are also major contributors to healthcare utilization and spending.⁴⁶ Factors other than health behaviors and health care, such as environmental, economic, and social factors, are important drivers of chronic disease. Therefore, to develop effective programs and interventions that meet individuals' and families' needs, clinicians and researchers need to understand relationships between individual, community, and program factors, and characteristics of their patient population. To do so, researchers and clinicians need access to multi-sector data that captures information on health behaviors, interventions, social determinants of health (SDOH), and other community-level factors. However, due to existing challenges in standardizing, linking, and sharing data while still preserving patient privacy across settings, access to such data is limited.

Project Purpose and Goals

As a continuation of the first OS-PCORTF-funded CODI project, CODI 2.0 builds on the work of linking community and clinical data for evaluating intervention effectiveness. While maintaining a focus on childhood obesity, this project will augment the existing CODI infrastructure with a new site in North Carolina. The new site, STAR Clinical Research Network (STAR) at Duke University adds geographic and data network diversity to the samples used to develop CODI tools. CODI 2.0 also expands its sample population to include adults and additional chronic diseases, with data from both the North Carolina site and the original CODI Colorado site.

Working with community-based organizations, the CDC team will create standards to capture and link SDOH data obtained from community partners (e.g., food insecurity, housing needs) to clinical data. To further enhance CODI's ability to study SDOH, CODI 2.0 will link community and clinical data at the household-level. Within-household linkages will be used to understand how interventions, including clinical and community programs, social assets, and government benefits, impact the whole household. The CODI 2.0 sample will ensure that CODI tools and resources are generalizable to the

PCORnet® distributed data network, better enabling application of CODI to other networks. Additionally, CODI 2.0 will use FHIR® Bulk Data technology to help organizations populate the CODI data model when sharing data, to increase the ease of data sharing and the volume of patient-level dataset available for comparative effectiveness research.

The project objectives are to:

- Establish a North Carolina Collaborative Working Group comprised of subject matter experts to advise and assist on matters such as project direction, development, and execution, and assist in the development of national CODI governance policies and procedures.

CODI 2.0 Partners

STAR is a National Patient-Centered Clinical Research Network (PCORnet®) Data Coordinating Center (DCC). It adds a patient population with diverse sociodemographic characteristics and brings robust informatics infrastructure, including the first-ever statewide, electronic, closed-loop, clinical to community resource coordinated care network called NCCARE360.

Other CODI 2.0 partners include local and state health departments and community-based organizations in the Triangle region.

- Scale-up existing patient-centered outcomes research capacity through inclusion of individual-level SDOH in linked datasets; craft implementation guidelines to promote adoption and use of project resources for alternative use by health researchers, scientists, and community partners; and enhance CODI infrastructure to reduce end-user implementation requirements and cost of populating common data models.
- Integrate project infrastructure into a PCORnet® participating network and develop a linked longitudinal dataset including childhood obesity-related risk factors, comorbid conditions, interventions, and outcomes accessible by PCOR researchers.

By linking clinical, intervention, and community data to create and standardize longitudinal individual-level records while preserving privacy, the aims of CODI 2.0 align with HHS' Strategic Plan for OS-PCORTF

Goal 2: Longitudinal Data Standards and Linkages. Additionally, this project's inclusion of individual-level SDOH data in linked datasets and emphasis on creating a more diverse sample will serve **Goal 4: Person-Centeredness, Inclusion, and Equity.**

Accomplishments

CODI 2.0 has made progress on several project milestones thus far:

- The CDC team identified community partners to work with the North Carolina site and is continuing to bring on partners who can help expand project data on community interventions and SDOH.
- The CDC team finalized priority research questions for CODI 2.0 and developed a report describing methods used to develop the priority research questions.
- The team completed a programmatic environmental scan and technical and gaps analysis to determine which SDOH data could be included based on current data infrastructure capabilities. The CDC team also completed business process analyses for participating partners to assess partners' data infrastructure technology and needs.
- The CDC team developed a report describing recommendations for Privacy Preserving Record Linkage (PPRL) on FHIR® and developed a FHIR® implementation guide to allow others to use the methodology (see Data and Technical Products).
- The CDC team continues to update tools and resources generated by the first CODI project, based on activities and findings from CODI 2.0. The team has made updates to common data model data elements and models, privacy preserving record linkage techniques, data governance approaches, and resources. Scaling and refining the CODI data infrastructure will allow for broader external application of CODI tools.

Data and Technical Products

- An implementation guide on the project's PPRL on FHIR® will be developed and made publicly available.
- A draft of the CODI Data Model Implementation Guide has been developed and will be made publicly available upon completion.



Communication and Dissemination Products

- **CODI Fact Sheet:** This information sheet outlines the project vision, the CODI innovation, and the CODI pilot tests.

Data Linkage: Evaluating Preserving Privacy Methodology and Augmenting the National Hospital Care Survey with Medicaid Administrative Records

Period of Performance

5/12/20 – 12/30/22

Federal Points of Contact

Lisa Mirel

The National Hospital Care Survey (NHCS) provides precise and reliable health care statistics that can be used to reveal national patterns of care delivery in hospital settings. The NHCS collection includes patient-level identifiers that allow data linkages to be made between the patient episodes of care within hospital and emergency department settings represented by the NHCS and other administrative data sources, which together provide a fuller picture of patient care.

This project builds upon previously funded OS-PCORTF efforts to significantly expand the scope of data available for patient-centered outcomes research by linking the 2016 NHCS with 2016-2017 CMS Medicare Fee-for-Service claims, Medicare Advantage encounters, and patient assessment data from long-term care facilities and home health providers with federal housing assistance program data collected from HUD. These linkages created an integrated dataset that allows researchers to make connections between an individual's health, health care utilization, and outcomes. The current project focuses on patient health outcomes across the continuum of care through linkage of the NHCS to additional data sources that diversify the sample and utility for specific research areas: interventions for opioid use, evaluation of medication protocols, use of social programs as a health determinant, and health disparities among understudied demographic groups.



Project Purpose and Goals

This project has two components. The first component will focus on assessing privacy preserving record linkage (PPRL) tools that could facilitate data linkages without the exchange of personally identifiable information. To ensure the accuracy of linked datasets and to allow researchers to better investigate patient outcomes, linkage algorithms rely on the exchange and matching of personally identifiable information (PII). However, privacy concerns persist regarding the exchange and analysis of identifiable information. To lessen dependence on PII, groups such as Datavant and the OS-PCORTF project, **Clinical and Community Data Initiative (CODI)**, have been working to develop privacy preserving linkage techniques that can be used to link data across disparate organizations. This project will evaluate the PPRL technique against other OS-PCORTF linkage methodologies using unencrypted PII to isolate discrepancies and their bearing on analysis of the linked data.

The second component of the project will create new linked datasets to support patient-centered outcomes research through the linkage of Medicaid claims data from CMS's T-MSIS and the 2014 and 2016 NHCS. This expanded data source will diversify and widen researchers' investigative range across PCOR topics. In particular, the project seeks to enhance the data available to evaluate interventions for opioid use, evaluation of medication protocols, use of social programs as a health determinant, and health disparities among understudied demographic groups.

Given these objectives, the project activities will focus on:

- Evaluating PPRL technique utilizing past OS-PCORTF-funded NHCS-NDI linkages as a gold standard for comparison.
- Disseminating information on the PPRL evaluation as a linkage technique to facilitate the creation of new datasets for patient-centered outcomes research.
- Conducting patient-level record linkages of 2014 and 2016 NHCS hospital administrative claims and EHR data to CMS' T-MSIS data from 2014-2017 (and more recent data if available).
- Developing research and user guidance materials to encourage use of new and existing NHCS linked datasets for patient-centered outcomes research.

By linking clinical and administrative data between the NCHS and T-MSIS to create and standardize a set of longitudinal individual-level records while preserving privacy, the project aligns with **Goal 2: Longitudinal Data Standards and Linkages** of HHS' Strategic Plan for OS-PCORTF.

Accomplishments

Since the initial award date in mid-2020, the project team has completed the PPRL assessment and made progress toward the patient-level record linkages.

To evaluate the PPRL technique, the team completed the following activities:

- Installed the Datavant PPRL software on the NCHS secure server.
- Developed an approach to use the Datavant software to link the NHCS to the NDI and coded the analytic tables for linkage statistics.
- Used Datavant PPRL software to conduct a comparison of linkage techniques. The team linked the NHCS to the NDI, and the result shows the PPRL linkage produced high levels of precision and recall compared to the gold standard linkage.

To develop the patient-level data linkages, the team:

- Executed a contract with NORC to assist with this work and amended the CMS DUA to gain access to 2018 and 2019 T-MSIS data files.
- The team completed an analysis of available T-MSIS files for PII data quality issues and identified strategies for remediation. They have completed the person level linkage process and are preparing to extract TMSIS claims data for all linked survey participants.

Data and Technical Products

- The project team has begun drafting a report describing the methodology used for the NHCS-T-MSIS linkage including detailed statistical information (e.g., linkage error rates and match rates) on the linkage methodology employed to create the linked data files, available on NCHS and accessible through ASPE's websites.
- The team has completed a first draft of the analytic guidance and has begun to draft sections pertaining to analytic guidance for analyzing the linked T-MSIS files.



Communication and Dissemination Products

- The project team completed the PPRL analysis paper, which has been accepted to the Statistical Journal of the International Association for Official Statistics and will be published later in 2022.
- The work was presented at the [Joint Statistical Meeting](#) in August of 2021. Additionally, the work was highlighted in the [Innovations in Data Integration at the Data Foundation Symposium](#), an ASPE [webinar](#) in October 2021 and the [NCHS Board of Scientific Counselors](#) in May 2021.

Developing a Multi-State Network of Linked Pregnancy Risk Assessment Monitoring System (PRAMS) and Clinical Outcomes Data for Patient-Centered Outcomes Research

Period of Performance

6/24/20 – 6/23/23

Federal Points of Contact

Shanna Cox

Developed in 1987, the Pregnancy Risk Assessment Monitoring System (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. It was designed to improve the health of mothers and infants by identifying women and infants at high risk for adverse health outcomes and measuring progress toward health goals. PRAMS is a joint project between the state, local, and territorial departments of health and the CDC Division of Reproductive Health, and is an ongoing, site-specific, population-based surveillance system.⁴⁷ Several special projects use PRAMS data including the PRAMS Stillbirth Project (SOARS), PRAMS for Dads, PRAMS and Zika, and the Healthy Start Evaluation Project.⁴⁸

PRAMS data are used by:

1. **Researchers** to investigate emerging issues in the field of maternal and child health.
2. **State and local governments** to plan and review programs and policies aimed at reducing maternal and infant mortality and morbidity.

The data are collected using the PRAMS questionnaire which is a survey that asks new mothers about their experiences and behaviors before, during and after pregnancy. The questionnaire has two parts which includes core questions that are asked in all sites and include data about the following: maternal attitudes about most recent pregnancy, health behaviors and experiences in the time period before pregnancy, prenatal care, social services such as WIC (the Special Supplemental Nutrition Program for Women, Infants, and Children), breastfeeding initiation and duration, substance use including cigarette and alcohol use, health insurance status at various time points, intimate partner violence, safe sleep practices, and contraceptive use. The remaining questions are chosen from a list of questions developed by the CDC or by the state resulting in a unique PRAMS questionnaire for each site. The questionnaire can be self-administered by mail or interviewer-administered by telephone.⁴⁹

Data from PRAMS are valuable because the surveillance system is widely used, and it provides information not available from other sources for the period before and during pregnancy and the first few months after birth. PRAMS represents approximately 81 percent of all live births in the United States and includes information from 46 states, New York City, Puerto Rico, the North Mariana Islands, and the District of Columbia.⁵⁰ Because of its widespread use and level of detail, sites utilize PRAMS to review maternal and infant health programs. PRAMS allows comparison of data and outcomes among

participating sites because the same data collection methods are used. Additionally, findings can be applied to a site's entire population of women who have recently delivered a live-born infant.

Linking PRAMS data to other datasets will provide a more comprehensive understanding of multiple determinants of maternal and infant health outcomes.⁵¹ Because PRAMS information is acquired by interviewing the mother, it provides the patient voice and perception of care. It collects data about the mother and infant's social context including exposure to domestic violence, participation in Medicaid and WIC programs, maternal attitudes toward pregnancy, and barriers to postpartum visits. It is also unique in collecting data before, during, and after pregnancy including social stressors that women experienced in the 12 months before pregnancy.

Project Purpose and Goals

This project provides the opportunity to create multi-level information and analysis to answer a comprehensive set of maternal and child health research questions. The project will create linked datasets of the PRAMS, birth certificate, and clinical outcomes data (e.g., hospital discharge, Medicaid claims). Other maternal and child health surveillance systems (e.g., home visiting services, early intervention services) may also be linked to conduct patient-centered outcomes research. Linkage of clinical outcomes data with PRAMS self-reported data allows for analysis of interventions that occur in the perinatal period that may not be available from clinical data sources (e.g., home visitation services) on clinical outcomes. These linked data can also be used to investigate how clinical care (e.g., treatment for depression) is related to patient outcomes (e.g., postpartum depressive symptoms). In addition, analyses can account for the social context (e.g., intimate partner violence) and the social determinants of health (e.g., housing insecurity) that is reported in PRAMS data.

This project will address the following objectives:

- Establish a coordinating center to support a learning collaborative of multiple states that will link PRAMS with clinical outcome data and document project activities for sustainability and future replication.
- Develop standardized methodology for creation of linked datasets and provide technical assistance to states to use standardized methodology to link data.
- Conduct priority analyses for patient-centered outcomes research to improve maternal and infant health using the linked data.
- Create a process for hosting and accessing linked data for external researchers.

By linking individual level data from PRAMS to clinical and social service data to support patient-centered outcomes research for maternal health and disparities research, this project serves **Goal 1: Data Capacity for National Health Priorities**, **Goal 2: Longitudinal Data Standards and Linkages**, and **Goal 4: Person-Centeredness, Inclusion, and Equity** of HHS' Strategic Plan for OS-PCORTF.



Coordination with Other Federal Agencies

The team is working with the Association of State and Territorial Health Officials (ASTHO) to coordinate two workgroups of federal, academic, and public health partners: a technical assistance workgroup and an advisory workgroup. The team has reached out to senior leadership of various agencies to participate in these workgroups including Health Resources and Services Administration (HRSA), National Institutes of Health (NIH), CMS, and NCHS colleagues at the CDC.



Accomplishments

The CDC team has made significant progress toward achieving projects objectives:

- The CDC has selected ASTHO as the coordinating center for the Linking PRAMS and Clinical Outcomes Data Multi-Jurisdiction Learning Community. ASTHO is a nonprofit organization that represents public health agencies across the United States, the U.S. Territories, and the District of Columbia to formulate and influence public health policy and state-based public health practice.⁵² The CDC project team with support from ASTHO has established two inter-agency workgroups, a technical assistance workgroup and advisory workgroup, to assist ASTHO and jurisdictional teams in developing evidence-informed focus areas and generate a comprehensive list of subject matter experts (SMEs). These workgroups will also support team technical assistance needs by engaging ASTHO affiliates, national organizations, and federal agencies.
- In August of 2021, four awardee states were announced as participants in the Learning Community aimed at linking PRAMS data to clinical or social service data. These states (Alaska, New Mexico, Texas, and Washington) are pursuing linkages of PRAMS to a wide range of data including Medicaid claims; home visitation and Pre-Kindergarten data; and hospital discharge data. These linkages will support exploration of research questions related to a variety of clinical outcomes for women and children including early intervention and injury-related health care visits; home visiting, postpartum depression, and pre-Kindergarten outcomes; and provider counseling and severe maternal morbidity.
- The four state awardees are participating in a series of virtual learning sessions on data linkage and will share knowledge, discuss barriers, and explore data linkage sustainability plans. The Learning Community kickoff meeting occurred in November 2021 and the first virtual learning session occurred in January 2022. A funding opportunity announcement to invite additional states to join the Learning Community was posted on ASTHO's website and remains open through March 2022.



Data and Technical Products

- The project intends to make the linked PRAMS datasets, codebooks, data dictionaries, and guidance for developing data access proposals available to external researchers.



Communication and Dissemination Products

- **Project Description Webpage:** The project team has created a webpage on the CDC website that disseminates information about PRAMS special projects.⁵³ The section briefly details the PRAMS Data Linkage activities, including the new learning community.

Enhancing Identification of Opioid-Involved Health Outcomes Using Linked Hospital Care and Mortality Data

Period of Performance

4/15/18 – 10/31/22

Federal Point of Contact

Carol DeFrances

National-level statistics on opioid-related hospitalizations are limited and often incomplete. Between 2005 and 2017, the rate of opioid-related emergency department visits nearly tripled from 89.1 to 249.1 per 100,000.⁵⁴ Opioid overdose deaths in the emergency department also increased 27.7 percent from 2015 to 2016.⁵⁵ Researchers need comprehensive data on

opioid-related emergency department visits, inpatient hospitalizations, and deaths to identify and test strategies to reduce the morbidity and mortality from misuse and overdose of opioids. Through this OS-PCORTF-funded project, the National Center for Health Statistics NCHS is helping address this need by developing enhanced methods that link available structured and unstructured data from three data sources: 1) the National Hospital Care Survey (NHCS; 2) the National Death Index (NDI); and 3) the Drug-Involved Mortality (DIM) file (formerly known as the National Vital Statistics System-Mortality-Drug Overdose file) to identify the specific opioids involved in drug-related emergency department visits, inpatient hospitalizations, and overdose deaths.

Project Purpose and Goals

The NCHS houses three data sources that, when combined, will offer broad, national-level data on hospital care and death related to opioid-involved drug overdose: 1) the NHCS collects inpatient, emergency department, and outpatient claims and EHR data from a nationally representative sample of approximately 600 hospitals; 2) the NDI includes all deaths occurring within the United States, along with cause of death; and 3) the DIM file includes information on specific drugs involved in overdose death. Each source has limitations, such as identification of specific opioids and inclusion of deaths occurring outside of a hospital setting. Prior projects in the OS-PCORTF portfolio, including **Enhancing Data Resources for Researching Patterns of Mortality in Patient-Centered Outcomes in Research: Projects 1-4, and Improving the Mortality Data Infrastructure for Patient-Centered Outcomes Research**, have addressed improved specificity within each source. Now, the combination of the three will allow researchers to follow patients with an opioid event from presentation at a hospital to death (if applicable) and retroactively analyze previous encounters for more information. The project will produce several data files that will be available to researchers through the NCHS Research Data Center.

This project expands data capacity for patient-centered outcomes research on opioid use by: 1) creating a new research data file with specific opioid names involved in emergency department visits, hospitalizations, and deaths; and 2) developing data collection and reporting tools to support research on hospital encounters involving opioids. The overall goal of the project is to improve surveillance and

Examples of research questions that can be examined with the linked data include:

- What are patient and hospital characteristics for patients who had an opioid-related emergency department visit or inpatient hospitalization?
- What are common patterns of hospital use in the months prior to a death due to an opioid-related drug overdose?
- For patients that had an opioid-related emergency department visit or inpatient hospitalization, how do services received differ between those who died from an overdose and those who did not die from an overdose?
- How do patients with a history of repeated opioid-related emergency department visits or inpatient hospitalizations that die from an opioid overdose compared to those who did not die from an opioid overdose?

expand researchers' access to data on opioid-involved health outcomes and risk factors associated with opioid overdose deaths. The project's four major objectives are to:

- Develop and apply text mining strategies, such as natural language processing, to written and coded data to identify the specific opioids involved in hospital encounters and drug induced mortality.
- Link several data sources to create an enhanced, more comprehensive dataset on care and outcomes.
- Build infrastructure to report additional clinical information back to hospitals via a hospital web portal.
- Disseminate new data files, methods, and other outputs to the research community by providing access to analytic files through the federal and NCHS Research Data Center Network.

By improving data collection and reporting tools for patient-centered outcomes research related to opioid use and by linking multiple national-level data sources, this project aligns with both **Goal 1: Data Capacity for National Health Priorities** and **Goal 2: Longitudinal Data Standards and Linkages** of HHS' Strategic Plan for OS-PCORTF.

Coordination with Other Federal Agencies

NCHS will collaborate with ASPE, FDA, SAMHSA, and National Institutes of Health's National Institute on Drug Abuse (NIH/NIDA) in the development of algorithms and dissemination.

Accomplishments

Since work on the project began in April of 2018, the project team has made notable progress on several tasks in support of project objectives.

- The project team convenes quarterly technical expert panel (TEP) meetings. The TEP is comprised of subject matter experts across multiple agencies including ASPE, FDA, SAMHSA, and NIH/NIDA, who meet regularly to provide input on project methodology.
- The project team linked 2014 NCHS and 2014/2015 NDI data together using both deterministic and probabilistic record linkage methods. The resulting dataset was linked to the 2014/2015 DIM file, which is now available in the federal and NCHS Research Data Center. Similarly, the team has produced a linked 2016 NHCS/NDI/DIM dataset. These linked data files include information on hospital care, mortality post-hospital discharge, and specific drugs mentioned in the literal text on the death certificate.
- The team has successfully built the NHCS Annual Hospital Report (AHR) secure web-based portal, which allows participating hospitals that submit 12-months of NHCS data to NCHS to produce customizable reports on hospital encounter diagnoses, services, length of inpatient stay, discharge status, and post-acute mortality. AHR was first used to report 2019 NHCS data from participating acute care hospitals. Future versions of the AHR will include summary results of post-hospital mortality from the linked NHCS and NDI data. Participating hospitals will gain access to a patient mortality report for their hospital that presents results from the linked NHCS and NDI data.

- In October 2021, a Call for Research Proposals was released that offered to provide three research groups up to \$12,000 dollars to access the data created from this project in the NCHS or Federal Research Data Center (RDC). Additionally, up to \$5,000 would be provided for travel expenses to the RDC. Three research projects were accepted, and their proposals are currently undergoing RDC review.

Data and Technical Products

- [**AHR Portal**](#): Hospitals that submitted 12-months of UB-04 administrative claims to the 2019 NHCS can view their data in the AHR portal.
- [**The Linkage of the 2014 National Hospital Care Survey to the 2014/2015 National Death Index: Methodology Overview and Analytic Considerations**](#): This report describes the linkage process between the 2014 NHCS and the 2014/2015 NDI data.⁵⁶ It includes a brief overview of the data sources, a description of the methods used for linkage, and analytic guidance for researchers.
- The linked 2014/2015 Dataset and [**Summary Report**](#): This dataset is available through the [**NHCS Research Data Center**](#), and links 2014 NHCS data, 2014/2015 NDI data, and 2014/2015 DIM data. The summary report provides a detailed description of the data sources, the processes used to link the datasets, the variables available in the linked datasets, and the analysis potential of the linked file.⁵⁷
- The linked 2016/2017 Dataset and [**Report**](#): This dataset is available through the Federal and [**NHCS Research Data Center**](#), and links 2016 NHCS data, 2016/2017 NDI data, and 2016/2017 DIM data. The report briefly describes the data sources for the linked file.⁵⁸ The project team has also drafted an analytical report on the numbers of opioids found in hospitals and death certificates from the 2016 NHCS/NDI/DIM file.

Communication and Dissemination Products

- [**Enhancing Identification of Opioid-involved Health Outcomes Using National Hospital Care Survey Data**](#): This methodology report describes the development of techniques for identifying opioids in hospitals and death certificates.⁵⁹
- [**Opioid-involved Emergency Department Visits in the National Hospital Care Survey and the National Hospital Ambulatory Medical Care Survey**](#): this National Health Statistics report utilizes the enhanced opioid algorithm developed by this project by analyzing opioid-involved ED visits in the 2014 NHCS and the 2013-2016 National Hospital Ambulatory Care Survey.⁶⁰
- [**Assessing National Hospital Care Survey and National Hospital Ambulatory Medical Care Survey Data: A Comparison of Opioid and Respiratory Disease Encounters**](#): Staff gave this virtual presentation at the 2020 Joint Statistical Meetings.⁶¹ The presentation referenced the linked data files created from this project.
- [**Enhancing Identification of Opioid-involved Health Outcomes Using National Hospital Care Survey Data**](#): This report describes the development of the enhanced opioid-identification algorithm utilizing the 2016 NHCS data. It includes a summary of the NLP methods used and developed to identify opioid-involved mentions in the unstructured clinical data and identification of clinical codes used to identify patients with opioid-involved hospital visits.⁶²

Enhancing Surveillance of Maternal Health Clinical Practices and Outcomes with Federally Qualified Health Centers' Electronic Health Records Visit Data

Period of Performance

3/01/21-8/30/23

Federal Point of Contact

Brian Ward

NCHS has collected clinical visit data from physician offices and federally qualified health centers (FQHCs) through National Ambulatory Medical Care Survey (NAMCS) for over 10 years. NAMCS is the only source of nationally representative visit-level data on ambulatory health services, including maternal health care provided at FQHCs.

In CY 2021, NAMCS will be transitioning its data collection methods from manual patient record abstraction to electronic data transmission of EHR data. These infrastructure upgrades will include the collection of the data elements needed to enhance the study of maternal health by enabling data linkages between NAMCS data, the National Death Index (NDI), and HUD administrative data. This linkage opportunity was not previously possible using manual abstraction data collection since no personally identifiable information (PII) was collected on patients; whereas EHR-based clinical data will facilitate patient-level matching

HRSA-supported FQHCs continue to be an important source of maternal health care for women in the United States.

In 2018, over 560,000 women who delivered during the year received pre-natal care at these safety-net providers.⁶³ The most recent estimates from NAMCS show that over 2.1 million routine prenatal examinations were provided annually at health centers, which was one of the leading reasons for visits at these locations.⁶⁴

Project Purpose and Goals

This project will help enhance NAMCS data collection procedures from CY 2021 for maternal health visits to FQHCs and expand these collection procedures in CY 2022. This will include increasing the sample of FQHCs and creating a linked dataset of maternal health data from EHRs combined with outcomes data from the NDI and HUD administrative data sources.

The project will address the following objectives:

- Build upon FQHC patient visit data (e.g., clinical EHR data) collected and processed by the NAMCS in CY 2021 by collecting and processing patient visit clinical EHR data for all visits among a larger sample of FQHCs in CY 2022.
- Produce a nationally representative data file comprised of FQHC EHR data for all maternal health visits in CY 2022 in the sampled FQHCs. The file will include variables on patient characteristics, clinical care provided, and health outcomes.
- Evaluate the linkage of CY 2021 (and CY 2022 if available) FQHC maternal health EHR data with the NDI and HUD administrative data through the National Center for Health Statistics (NCHS) Data Linkage Program.
- Examine of relationships among the characteristics, care provided, and outcomes of the maternal health visits, including maternal health visits related to COVID-19.

By building data capacity for patient-centered outcomes research through clinical, intervention, and community data linkages, the aims of the **Enhancing Surveillance of Maternal Health Clinical Practices and Outcomes with Federally Qualified Health Centers' Electronic Health Records Visit Data** project align with **Goal 1: Data Capacity for National Health Priorities** of HHS' Strategic Plan for OS-PCORTF. Additionally, this project's inclusion of individual-level patient centered data in linked data sets and emphasis on creating a more diverse sample will serve **Goal 4: Person-Centeredness, Inclusion, and Equity**.

Coordination with Other Federal Agencies

Beginning in CY 2021, NCHS and HRSA are partnering to modernize the FQHC component of NAMCS. They will end the use of manual, in-person data abstraction to collect data from FQHCs and establish an electronic process to of these data using EHRs.

Accomplishments

Since the project launched in March of 2021, the project team has begun working towards the objectives:

- The team identified the data elements expected from the EHR data collection that will allow FQHCs to collect more data than existing manual processes.
- The sampling procedures and specifications for the NAMCS Health Centers in CY 2022 have been developed, and the CY 2022 sample has been drawn and delivered to the recruitment team. These procedures and specifications will be used in the recruitment process to establish clear expectation for the participating FQHCs which will launch in early 2022.
- The project team is participating in ASPE's Maternal Health Consortium, including the meetings. The team presented "Discussion on claims mother/baby record separation in the National Hospital Care Survey (NHCS)" during the September 2021 call to discuss progress and data insights.

Communication and Dissemination Products

- A manuscript focusing on the FQHC sampling procedures is currently under development.

Identifying Co-Occurring Disorders among Opioid Users Using Linked Hospital Care and Mortality Data

Period of Performance

5/1/19 – 10/31/21

Federal Point of Contact

Carol DeFrances

According to the 2017 National Survey on Drug Use and Health (NSDUH), the number of adults with Substance Use Disorders (SUDs) who had any mental illness was about 8.5 million, and the number with severe mental illness was about 3.1 million.⁶⁵ Based on 2015 NSDUH data, approximately 1.5 million adults with severe mental illness had misused opioids in the past year, which is equivalent to a co-occurrence of opioid misuse and severe mental illness in an estimated 1 in 8 adults (13 percent).⁶⁶ It is important for the National Hospital Care Survey (NHCS) to monitor the role that co-occurring disorders plays in opioid-related morbidity and mortality outcomes.

A previously funded OS-PCORTF FY 2018 project at CDC/NCHS, **Enhancing Identification of Opioid-Involved Health Outcomes Using Linked Hospital Care and Mortality Data**, created an enhanced methodology to accurately identify a hospital patient's use of opioids in any form (i.e., used as directed, misused in a manner contrary to provider instructions, used intentionally to become intoxicated or for the purpose of self-harm, taken accidentally, etc.). Additionally, that project identified the specific legal or illicit opioid agent taken. This OS-PCORTF capstone project built upon the FY 2018 project methodology to link key data sources and flag evidence of co-occurring mental health disorders.

Project Purpose and Goals

The goal of this project was to improve public health surveillance and expand researchers' access to data on health outcomes of patients with opioid-involved hospital encounters and co-occurring substance use and mental health issues. Both the original project and the Capstone Project used algorithms that determine the occurrence of an event (the use of opioids, type of opioid agent taken, and presence of a substance use or mental health issue) by selecting combinations of coded items (diagnoses, procedures, lab results, etc.) and terms contained in free text (clinical notes, cause of death). Both projects resulted in the creation of linked files that combine three data sources to enable access to data that follows patients with an opioid event for one year following hospital discharge: the National Hospital Care Survey (NHCS); the National Death Index (NDI); and the National Vital Statistics System restricted mortality data, drug specific information (NVSS-M-DO) files, referred to as the Drugs Mentioned with Involvement (DMI) program. This work ultimately allowed for retrospective analysis of the extent to which specific opioid agents and the co-occurrence of mental health disorders were involved in hospital encounters preceding post-discharge deaths.

The capstone completed the following activities:

- Developed a new algorithm that uses the linked NHCS/NDI/DIM files to identify opioid-involved hospital encounters for patients with co-occurring disorders using medical code-based data and natural language processing (NLP).
- Conducted a study to validate the algorithm from this project and the FY 2018 **Enhancing Identification of Opioid-Involved Health Outcomes Using Linked Hospital Care and Mortality Data** project to identify the use of opioids and the existence of co-occurring disorders.
- Applied the co-occurring disorders algorithm to identify prevalence of opioid-involved emergency department visits and co-occurring disorders among opioid users in the 2016 NHCS data.
- Provided data on opioid use and co-occurring disorders made available through: 1) the NCHS Research Data Center, and 2) a previously developed interactive web portal for NHCS participating hospitals.

By using NLP and newly developed algorithms to facilitate identification and research of opioid use and co-occurring disorders in hospital encounters, this project aligned with both **Goal 1: Data Capacity for National Health Priorities** and **Goal 3: Technology Solutions to Advance Research** of HHS' Strategic Plan for OS-PCORTF.

Accomplishments

The project team has made notable progress on several tasks since work began in May of 2019:

- The co-occurring disorders algorithm was developed and applied to the 2016 NHCS data to identify patients with opioid-involved hospitalizations and emergency department visits with co-occurring substance use disorders or mental health issues. The co-occurring disorders algorithm utilizes coded medical data and NLP methods to identify mentions of substance use disorders or mental health issues in the unstructured clinical data.
- The project team developed a report describing an integrated algorithm, including criteria, medical code and search term lists, and fields to be searched.
- The team completed a validation study to assess the performance of opioid use algorithms and co-occurring disorder algorithms.
- The 2016 NHCS data linked to the 2016 - 2017 NDI and DIM with enhanced information on opioid-involved hospital visits with co-occurring substance use disorders and mental health issues is available through the Federal and NCHS Research Data Center.

Data and Technical Products

- **Identifying Co-Occurring Disorders among Opioid Users Using Linked Hospital Care and Mortality Data:** This report describes the completed 2016 NHCS/NDI/DIM enhanced dataset, which includes additional information on co-occurring disorders.⁶⁷ The dataset itself is available for use for researchers through the NCHS Research Data Center.

Communication and Dissemination Products

- The project team drafted a final project report, “Identifying Co-Occurring Disorders among Opioid Users Using Linked Hospital Care and Mortality Data: Capstone to an Existing FY18 PCORTF Project: Final Report.” The report summarizes project accomplishments, lessons learned, and future implications of this work. This report has not been made publicly available yet.
- NCHS staff have conducted multiple presentations to raise awareness of the project and its resources, including: an overview of the project utilizing the NHCS data at the International Conference on Health Policy Statistics; an overview on the development of the enhanced opioid and co-occurring disorder algorithms at the Rx Drug Abuse and Heroin Summit; an overview of the PCORTF projects utilizing NHCS data at ASPE webinars and meetings.
- In October 2021, NCHS staff gave a virtual presentation entitled “Harnessing Natural Language Processing and Machine Learning to Enhance Identification of Opioid-Involved Health Outcomes in the National Hospital Care Survey” at the 2021 Statistics Canada Annual Conference. The presentation was attended by statisticians from international and Federal statistical agencies, private companies, and academic institutions.

Making Electronic Health Record Data More Available for Research and Public Health (MedMorph)

Period of Performance

4/15/19 – 10/18/23

Federal Point of Contact

Maria Michaels

Real-time data exchange between health systems, researchers, and public health is inconsistent and insufficient. Many patient-centered outcomes researchers and public health programs share a common challenge: they rely on clinical data that are frequently inaccessible. As a result, they may be unable to answer critical questions that could lead to health care that is more patient-centered or to using patient-level data for public health action.

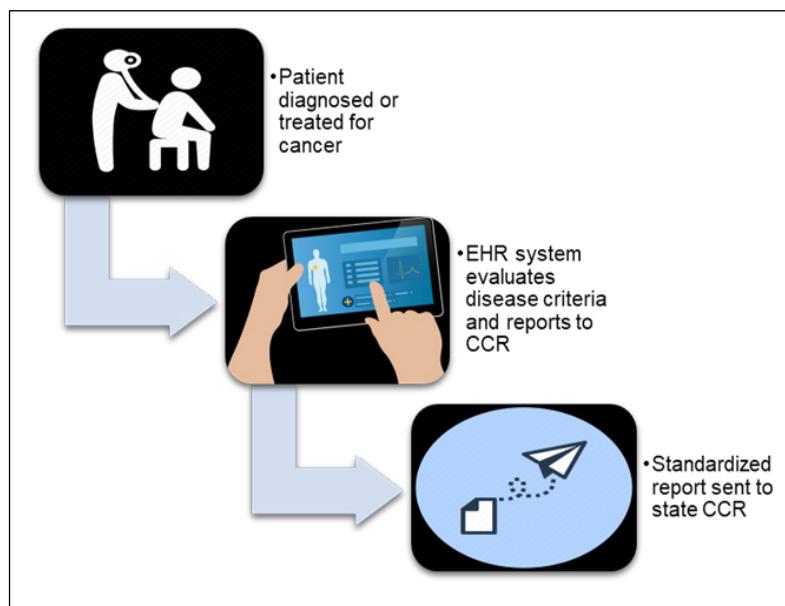
Similarly, lack of access to EHR data can preclude innovative partnerships between health care providers and public health that advance patient outcomes. Patient-centered outcomes researchers and public health professionals need better ways to access data from different EHR systems without creating additional burden on health care providers. In recent years, the maturation of standards such as FHIR® and the Office of the National Coordinator for Health Information Technology (ONC) EHR certification requirements—such as the United States Core Data for Interoperability (USCDI) and application programming interfaces (APIs)—have created an environment that is ripe for developing scalable and extendable solutions to overcome interoperability challenges.

Project Purpose and Goals

The purpose of this project, known as MedMorph, is to develop a reference architecture, app, and implementation guides to address some of the identified challenges of exchanging data between health care providers and endpoints in public health and research. The project aims to create a reliable and adaptable method for EHR data exchange to reduce burden, time, and effort for all. The project has selected three use cases to model and inform the development of the reference architecture, app design, and implementation guides:

- **Hepatitis C:** This use case builds a model for the collaborative use of EHR data for chronic hepatitis C surveillance, with the goal of improving patient and community health outcomes for hepatitis C. These enhancements will leverage a standardized reference architecture, automating access to more robust EHR datasets.
- **Cancer:** This use case enhances cancer surveillance by automating the transmission of cancer case data from EHRs to central cancer registries (CCRs) in current reporting protocols (Exhibit 4).

Exhibit 4. EHR Data Accessibility: Cancer Use Case



- **Health Care Surveys:** This use case automates reporting for health care utilization data from EHRs to CDC's National Center for Health Statistics (NCHS). NCHS will include this EHR data in the following national datasets: National Ambulatory Medical Care Survey (NAMCS) and National Hospital Care Survey (NHCS).⁶⁸

The MedMorph project team will fully test the app on the hepatitis C use case to ensure it can extract data from multiple clinical organizations using different EHR systems. The goal is to leverage a common reporting framework so that data is accessible to researchers and public health officials and can be applied to other health conditions in the future.

The project will focus on the following objectives:

- Define the use cases and research questions relevant to end-users.
- Use existing health data and exchange standards (e.g., FHIR®) to develop an app for real-time data exchange between EHRs and research and public health systems.
- Establish the infrastructure to create a collaborative network of clinical, research, and public health organizations and health IT vendors as a ready test bed for testing developed tools.
- Implement the app for at least one use case (hepatitis C) in both clinical research and public health surveillance contexts and evaluate it for improvements in the timeliness and completeness of data reported for research and public health.
- Develop a plan for broad use and long-term sustainability of the app, including publishing the app as open-source software.

With a focus on utilizing data exchange standards to develop an app that links EHRs and research and public health systems, MedMorph project objectives align with **Goal 2: Longitudinal Data Standards and Linkages** of HHS' Strategic Plan for OS-PCORTF.

Coordination with Other Federal Agencies

The project established regular collaborative meetings with FDA, NIH National Cancer Institute (NIH/NCI), NIH National Center for Advancing Translational Sciences (NIH/NCATS), and ONC to leverage other agencies' relevant expertise. For example, FDA is working with NIH/NCATS to adapt the PCORnet® Common Data Model to follow FHIR® standards, particularly for the hepatitis C use case. Additionally, monthly technical expert panel (TEP) meetings include presentations from other agency's OS-PCORTF-funded projects.

Accomplishments

This project has made progress by achieving the following objectives:

- The CDC team completed a landscape analysis to evaluate implementation of FHIR® and other standards among EHR vendors. This analysis informed development of an app capable of exchanging data in real time.
- The CDC team hosted a 100-member TEP kick-off meeting in late 2019 to gather information on the scientific, technical, and practical aspects of the app development. Over 100 TEP members have participated in either the Use Case Workgroup (three focus areas include hepatitis C,

Cancer, and Healthcare Surveys), Reference Architecture Workgroup (three focus areas include Clinical Workflows & Data Flows, Data Standards, Reference Architecture), or Evaluation Workgroup. These workgroups have met weekly since April 2020 to guide development of the implementation guides and related app as well as develop the evaluation plans for the pilot and beyond.

- With input from the TEP, the CDC team developed the HL7® FHIR® Reference Architecture implementation guide. This publicly available resource addresses data exchange needs with a common, streamlined approach (see Data and Technical Products). The RA implementation guide was submitted for HL7® balloting, an open feedback process prior to publication, in the January 2021 balloting cycle. Three content implementation guides have also been developed (Central Cancer Registry Reporting, Healthcare Surveys, and Research Data Exchange), which are being balloted in the January 2022 balloting cycle. The CDC team continues to test and refine the RA and content implementation guides through HL7® Connectathons and the hepatitis C use case pilot. The MedMorph team has participated in four HL7® Connectathons since 2020, including most recently the January 2022 event.
- The CDC team recruited two EHR vendors, MDL and InSync, to participate in the hepatitis C pilot in 2021. This pilot will allow the MedMorph team to test the MedMorph RA in real-world settings using chronic hepatitis C reporting and research as the primary use case, with optional tasks to pilot reporting and research for the cancer and health care survey use cases. The CDC team is performing ongoing pilot recruitment for additional clinical sites and/or EHR vendors.
- The CDC team completed an initial draft of a short-term evaluation plan for MedMorph. This evaluation plan will be executed during the hepatitis C use case pilot.



Data and Technical Products

- [MedMorph Software Code](#)**: This GitHub library includes code for products developed for this project. The code can be used by others to implement MedMorph in their own EHR system.⁶⁹
- [Reference Architecture HL7® FHIR® Implementation Guide](#)**: This resource is for external organizations interested in adopting the reference architecture. The implementation guide is a blueprint for use case scenarios and outlines how to use best practices to be more efficient in data exchange efforts.⁷⁰
- [Central Cancer Registry Reporting HL7® FHIR® Implementation Guide](#)**: This is the content implementation guide based on the MedMorph Reference Architecture that supports the cancer reporting use case. This resource is for external organizations interested in adopting the reference architecture.⁷¹
- [Healthcare Surveys HL7® FHIR® Implementation Guide](#)**: This is the content implementation guide based on the MedMorph Reference Architecture that supports the health care surveys use case. This resource is for external organizations interested in adopting the reference architecture.⁷²
- [Research Data Exchange HL7® FHIR® Implementation Guide](#)**: This is the content implementation guide based on the MedMorph Reference Architecture that supports research use cases. This resource is for external organizations interested in adopting the reference architecture.⁷³



Communication and Dissemination Products

- **MedMorph Webpage:** A public-facing website, detailing the project's goals and activities.⁷⁴
- In the brief communication “**Blueprint for aligned data exchange for research and public health**”, published in the December 2021 issue of the *Journal of the American Medical Informatics Association*, the MedMorph team summarized MedMorph’s activities to develop the RA implementation guide. This article also explores implications of improved data sharing by MedMorph for research and public health needs.⁷⁵

MAT-LINK: MATernal and Infant Network to Understand Outcomes Associated with Medication for Opioid Use Disorder during Pregnancy

Period of Performance

3/4/19 – 12/30/22

Federal Point of Contact

Shin Kim

From 1999–2014, the prevalence of opioid use disorder (OUD) among pregnant women in the United States quadrupled from 1.5 to 6.5 per 1,000 delivery hospitalizations.⁷⁶ Opioid use during pregnancy elevates the risk of an infant being born with neonatal abstinence syndrome (NAS). Recent evidence suggests that children born with NAS may experience developmental delays;^{77,78} however, the developmental trajectory of these children has not been systematically studied.

The American College of Obstetricians and Gynecologists recommends that pregnant women with OUD be offered an opioid agonist (methadone, buprenorphine) to treat OUD during pregnancy. In some clinical settings, pregnant women may also be offered medically supervised withdrawal (opioid detoxification).⁷⁹ There is limited information, however, comparing maternal, infant, and child health outcomes associated with these treatment regimens during pregnancy. Clinical treatment of OUD among this population relies on limited data, leaving clinicians and researchers with an incomplete picture of their efficacy.

Exhibit 5. Maternal, Infant, and Child Outcomes of Interest



Project Purpose and Goals

The MAT-LINK project will establish a surveillance network, consisting of four clinical sites across the United States, to collect data on maternal, infant, and child health outcomes associated with treatments for OUD during pregnancy. Results from MAT-LINK will be used to improve understanding of the spectrum of maternal, infant, and child health outcomes following treatment for OUD during pregnancy and the role of mediating and moderating factors on maternal and infant outcomes, including exposure to multiple substances, maternal comorbidities, and other psychosocial factors (Exhibit 5). Moreover, MAT-LINK partners with a network of active clinical settings, allowing for the collection of real-time behavioral data and outcomes for mother and child. These data can inform clinical practice in a way that artificially created clinical trials cannot given their controlled nature. MAT-LINK's unique collaboration with real clinical settings allows for the incorporation of data naturally drawn from health care visits, resulting in findings that are both evidence-based and practical.

MAT-LINK will be used to improve policies, clinical practice recommendations, and clinical decision-making. In addition, this project will develop and pilot a data platform to collect and link maternal, infant, and child data across clinical sites which can be modified to collect linked data on other exposures during pregnancy. Future MAT-LINK expansion efforts will also widen the reach of diverse populations studied and the breadth of data collected as the project plans to collaborate with several additional clinical sites and follow outcomes of older children (through six years of age), enabling researchers to answer questions that cannot be answered by data collected to date (see MAT-LINK2 profile, below).

The project will address the following objectives:

- Develop a data platform to collect linked maternal and infant data among women treated for OUD during pregnancy.
- Create a core set of variables for surveillance of OUD during pregnancy that can be analyzed for national, state, local, or healthcare system surveillance.
- Analyze and disseminate preliminary results to inform patient-centered care for pregnant women with OUD and for infants and children with prenatal opioid exposure.

By linking clinical, intervention, and community data to create and standardize longitudinal individual-level records while preserving privacy, the aims of MAT-LINK project align with **Goal 2: Longitudinal Data Standards and Linkages** of HHS' Strategic Plan for OS-PCORTF.

Accomplishments

Since MAT-LINK's inception in 2019, the CDC team has accomplished the following tasks:

- Established an organizational structure, including a CDC Steering Committee and Partners Group comprised of federal, clinical, and public health partners. After an inaugural meeting, the Steering Committee has convened regularly to provide guidance for the project.
- Published a Notice of Funding Opportunity to identify an implementation partner. After reviewing applications, CDC selected the Public Health Informatics Institute (PHII).

- PHII released the request for proposals (RFP) for the pilot clinical sites. Following the release of the RFP, PHII and CDC received 25 Letters of Intent from clinical sites located across the continental United States and awarded four of the 25 applicants to participate in the collection of maternal and infant data. The four awardees are Boston Medical Center Corporation, Kaiser Foundation Research Institute, University of Utah, and The Ohio State University.
- Identified and developed a list of core and standard variables that are being piloted at clinical sites through the dissemination of all data extraction tools. Through these data extraction tools, the team has been receiving and analyzing maternal health data.
- Developed an analytic plan with federal and other partners that included a prioritized list of research questions and a set of analyses. The team holds monthly calls with clinical site Principal Investigators (PIs) to discuss additional ideas for analyses. A more detailed document that includes table shells will be refined throughout the duration of the project.
- Drafted a data sharing agreement describing the process for external researchers to access MAT-LINK data (either directly at CDC or through the RDC).
- Initiated work on a user guide that will assist researchers who access data collected under the project. The guide will include: a detailed description on how data were collected, what variables are available for analysis, description of the sample, data dictionary, mother-infant linkage process methodology and any other relevant documentation that will facilitate the use of the collected data for research projects by external researchers.
- Identified different patterns of use of medications for OUD. The project can inform gaps in knowledge on which populations are more likely to use different medications for OUD using pilot data on demographic characteristics for different OUDs.
- Disseminated project work through several presentations, including a presentation at the Council of State and Territorial Epidemiologists Neonatal Abstinence Syndrome Workgroup and at the American Medical Informatics Association 2020 Annual Conference.
- Participated in monthly OS-PCORTF Maternal Health Consortium meetings, which help coordinate maternal health work among federal partners. The project team presented an overview of MAT-LINK in April, XML schemas and data collection tools in May, and methods for linking mother/child dyads in July.

A MAT-LINK report details procedures for accessing project data collected by external researchers, including:

- Accessing the data via a Research Data Center (RDC),
- Obtaining data at CDC locations in Atlanta and Washington, D.C., or
- Submitting a proposal for data access to CDC for analysis.

This report also includes a list of core variables for OUD surveillance during pregnancy that can be leveraged for national, state, local, or healthcare system surveillance.



Communications and Dissemination Products

- **[MAT-LINK Project Webpage](#)**: This project webpage provides background information about the project, including partners, clinical sites, goals, and example variables.⁸⁰
- **[CDC Public Health Grand Rounds Presentation](#)**: This presentation, “Reducing Polysubstance Use in Pregnancy,” from MAT-LINK PIs at the University of Utah discusses adverse maternal and child health outcomes caused by polysubstance use.⁸¹

- **Journal of Women's Health Publication:** Published in 2020, this article describes medications and herbal supplements used by pregnant women for OUD and how MAT-LINK can address gaps in knowledge about the management and treatment of OUD during pregnancy.⁸² A second manuscript detailing the methods used for the MAT-LINK project is under development.

MAT-LINK2: Expansion of MATernal and Infant Network to Understand Outcomes Associated with Treatment for Opioid Use Disorder during Pregnancy

Period of Performance	Federal Points of Contact
3/15/2021 – 2/28/23	Shin Kim

The opioid epidemic has had devastating consequences in the United States. Among pregnant women, the use, misuse, and abuse of opioids have increased dramatically in recent years.⁸³ Opioid use disorder (OUD) during pregnancy is associated with a risk for maternal overdose, maternal infection, and adverse infant outcomes (e.g., poor fetal growth, preterm birth).⁸⁴ In parallel with increases in opioid use during pregnancy, the incidence of neonatal abstinence syndrome (NAS)—when babies go through withdrawal from drugs that have been exposed to in the womb—has increased and is now being monitored in many regions across the U.S.

This project expands on a previously funded OS-PCORTF project, MAT-LINK, that established a maternal and infant network to examine practice patterns and outcomes associated with the treatment for OUD during pregnancy. MAT-LINK created a data platform and standard maternal and infant data elements to collect linked maternal and infant data among women treated for OUD during pregnancy in four clinical sites: Boston Medical Center, Ohio State University, University of Utah, and Kaiser Foundation Research Institute Northwest in Oregon and Washington. MAT-LINK has successfully defined key data elements for analysis of clinical practice and maternal and infant outcomes.

Project Purpose and Goals

Building upon the work of the first OS-PCORTF-funded MAT-LINK project, MAT-LINK 2 will strengthen the existing surveillance approach, increase the data available to answer key scientific and clinical questions, and better inform evidence-based, clinical best practices to enhance the lives of mothers, infants, and children affected by OUD.

To strengthen the robustness and representativeness of future findings, MAT-LINK 2 will add two to three additional clinical entities to increase the study population of women with varied racial, ethnic, and socioeconomic characteristics and expand the geographic reach. In addition, the period of collection of outcome data for infants will be expanded up to age six. This will allow for a more comprehensive assessment of cognitive and motor developmental delays that may not be apparent in younger children.

The purpose of this project is to address the following objectives:

- Increase diversity of the MAT-LINK network through expanded geographic representation and greater representation of underrepresented racial and ethnic populations.
- Assess cognitive and motor development through extending follow-up of children from two years to six years of age.

- Improve knowledge about maternal overdose, maternal infection, and infant outcomes, including NAS, and enhance clinical care of pregnant women with OUD and infants with prenatal opioid exposure.

By linking clinical, intervention, and community data to create and standardize longitudinal individual-level records while preserving privacy, the aims of MAT-LINK 2 align with the Strategic Plan for OS-PCORTF (2021-2029) **Goal 2, Longitudinal Data Standards and Linkages**. Additionally, this project's inclusion of individual-level SDOH data in linked data sets and emphasis on creating a more diverse sample will serve **Goal 4, Person-Centeredness, Inclusion, and Equity**.

Accomplishments

MAT-LINK 2 is working towards the expansion through the following tasks:

- Working with their implementation partner, the Public Health Informatics Institute, the project team sent letters of interest to clinical sites that had previously applied for MAT-LINK Phase I but were not funded. In response, the team received 13 applications from a range of private, public, and integrated care institutions. The project team selected three additional institutions to join MAT-LINK as Phase II clinical sites: University of New Mexico, University of Rochester, and University of South Florida.
- The project team began working with the new clinical sites and conducted a kickoff meeting in August 2021 with all seven sites. During the meeting, the project team presented the MAT-LINK data architecture and data analytic plan and Phase I sites shared lessons learned with Phase II sites.
- MAT-LINK Phase II sites are currently in the process of confirming patient eligibility and finalizing their case lists with basic information about prenatal OUD treatment, pregnancy outcomes, and follow-up status.
- To assess cognitive and motor development in all clinical sites, the team has finalized a list of variables for follow-up in children. A data dictionary is under development.
- The project team participated in the monthly OS-PCORTF Maternal Health Consortium, sharing information about MAT-LINK, sharing site specific methods for maternal/child EHR data linkages, and discussing cross-project synergies related to the use of EHRs for maternal health research.

Strengthening the Data Infrastructure for Outcomes Research on Mortality Associated with Opioid Poisonings

Period of Performance

5/1/18 – 5/14/20

Federal Points of Contact

Steven Schwartz

Kate Brett

In 2017, the White House declared the opioid epidemic a public health emergency.⁸⁵ Monitoring drug-related deaths can help researchers track trends and understand the impact of opioid use on Americans to ultimately inform policy, programs, and research. Researchers and public health officials often use cause of death information from death certificate data for such research; however, the quality of this data is limited by interoperability issues, lack of data standardization, and data extraction practices.

All death records in the U.S. are processed by the National Center for Health Statistics (NCHS). Developed in the 1980s, the current system uses algorithms to assign underlying cause of death and multiple causes of death from information in the medical and demographic portions of the death certificate.⁸⁶

From 2008 to 2010, nearly 25 percent of death certificates did not specify the drugs involved in the death.⁸⁷ However, the quality of data on death certificates improves when a physician completes the certificate using their EHR.⁸⁸ Medical examiners and coroners typically certify death in the case of drug overdose, so promoting data interoperability between the case management systems used by medical examiners and coroners with state electronic death registration systems (EDRS) will improve the quality of data for the deaths that physicians certify.

P Project Purpose and Goals

The purpose of this project is to strengthen the mortality data infrastructure for outcomes research on deaths associated with opioid poisoning. It achieved this goal by replacing the existing Medical Mortality Data System (MMDS) that codes and processes death certificate records with a new medical coding system, MedCoder. MedCoder increases the proportion of death certificates coded electronically and can identify new text, such as new synthetic opioid names, and automatically applies the correct codes through the application of machine learning and natural language processing (NLP) techniques. Using MedCoder, the project enhances the quality of death information data captured in the National Vital

CDC aims to increase understanding of hospital care patterns and risk factors associated with opioid misuse and overdose deaths by enhancing death certificate data collection methods and availability for research.

Statistics System multiple cause of death mortality files (NVSS-M) and the National Death Index (NDI) by collecting additional drug information. The CDC team created supplemental data files that incorporate details about drugs, beyond what is captured in medical classification codes (e.g., ICD-10), that caused or contributed to death. This project also strengthened the Vital Statistics Rapid Release (VSRR) program to produce and release more in-depth information on drug overdose data for public health surveillance and research. Enhancements allow the VSRR to capture a broader array of geographic and demographic data in vital statistics records for surveillance purposes.

The project objectives are to:

- Create a new system to electronically code and incorporate specific drug information captured in the literal text fields of death certificate records using machine learning and NLP techniques.

- Incorporate supplemental drug information from the literal text fields of death certificate records, especially information related to deaths involving opioids, as new variables in the NDI and the NVSS-M.
- Annually produce the NDI and the NVSS-M data files containing the supplemental information for deaths involving drugs such as opioids for use by approved researchers.
- Improve the specificity of drug information on death certificates supplied by states by developing and pilot-testing a FHIR® application programming interface (API) for the exchange of information from between medical examiner and coroner case management systems and EDRS.
- Improve the depth and timeliness of national reporting on drug deaths involving opioids by re-architecting the data system to produce and release more in-depth information about drug overdose data (e.g., specific drugs, demographic information) on a monthly basis for public health surveillance and research.
- Establish an advisory committee of the [NCHS Board of Scientific Counselors](#) to align changes in the mortality data system with end-users' (i.e., researchers') needs.

By leveraging machine learning and NLP techniques to strengthen quality of opioid overdose mortality data, these objectives serve **Goal 1: Data Capacity for National Health Priorities** and **Goal 3: Technology Solutions to Advance Research** of HHS' Strategic Plan for OS-PCORTF.

Coordination with Other Federal Agencies

CDC worked closely with NIH National Library of Medicine (NIH/NLM) and the Drug Enforcement Agency (DEA) on this project. NIH/NLM holds valuable data on pharmaceutically manufactured drugs. DEA agreed to make a reference list of illicitly manufactured drugs, which helped the NIH/NLM project team create a supplemental drug file, which includes all substances, both illicit and pharmaceutically manufactured, for accurate death certificate reporting.

Accomplishments

This project was completed in 2021 and has taken several important steps in enhancing mortality data infrastructure for opioid poisoning outcomes research. The project employed a multi-pronged approach to modernize the capture of death certificate data, leading to higher quality, timelier, and supplemented information on drug overdose deaths involving opioids. The CDC team developed and implemented the new coding system, MedCoder, to replace the existing MMDS. This new coding system leverages NLP and machine learning algorithms to capture information from literal text fields on death certificates and create new variables on opioid-involved deaths. These new variables have been added to NDI and NVSS-M 2021 datasets. Additionally, the team created datasets to train MedCoder to learn COVID-19 and maternal mortality auto-coding based on newly developed coding rules specific to these two areas.

NCHS compared the MedCoder system to the prior system to identify gaps and areas of improvement in MedCoder. The results of this evaluation have allowed the project team to fine tune the coding rules for processing and capturing information from death certificates to optimally identify drugs involved in deaths. The Division of Vital Statistics plans to switch to using MedCoder as the coding system of record in early 2022. A forthcoming report will describe the MedCoder system features and new supplemental drug data available to researchers, and the report will provide sample research questions that can be answered with MedCoder.

The VSRR program provides access to the timeliest vital statistics data for public health practitioners, researchers and policymakers using provisional data.⁸⁹ This project's enhancements to the VSSR included the addition of geographic, demographic, and drug information and automation of monthly reports. The CDC has released expanded reports since January 2021. The team also developed public web-based dashboards on monthly provisional drug overdose death counts and quarterly provisional estimates of national death rates. These VSSR efforts produce more in-depth drug overdose data that furthers public health surveillance and research efforts.

Finally, the project has made great strides to improve the depth and timeliness of national reporting of drug deaths involving opioids. The team successfully developed and piloted a FHIR® API for the exchange of mortality data between states' EDRS and medical examiner/coroner offices. This app improves the specificity of drug information on death certificates that states can supply for research. The FHIR® application is currently in use by Georgia and Washington, D.C., and a FHIR® implementation guide and Toolkit are available (see Data and Technical Products). The team established an advisory committee to ensure that the design of system improvements throughout the project aligned with end-user's needs in studying drug overdose deaths involving opioids.

To demonstrate the feasibility of these enhanced tools as well as their applicability to patient-centered outcomes research, the CDC team published three National Health Statistics Reports and one Vital Statistics Report (see Communication and Dissemination Products).

Data and Technical Products

- **The Vital Records Death Reporting (VRDR) FHIR® Implementation Guide:** This guide describes interoperability enhancements to allow for exchange of data between medical examiner/coroner offices and EDRS.⁹⁰
- **NVSS Modernization Tool Kit:** This resource was designed to help jurisdictions and technical partners utilize the modernized NVSS tools. Resources include training materials, technical tools, and guidance documents to prepare systems to be compatible with CDC tools and to utilize EDRS and FHIR® APIs.⁹¹
- **The Vital Statistics Rapid Release Quarterly Provisional Estimates Dashboard:** This public dashboard includes crude and age-adjusted death rate quarterly data in all states and can be sorted by cause of death, age, sex, and state.⁹²
- **Provisional Drug Overdose Death Counts Dashboard:** This public dashboard illustrates national and state-level monthly provisional counts for drug overdose deaths based on mortality data in the National Vital Statistics System.⁹³

Communication and Dissemination Products

- **National Health Statistics Reports: Opioid-involved Emergency Department Visits in the National Hospital Care Survey (NHCS) and the National Hospital Ambulatory Medical Care Survey:** This report compares emergency department (ED) NHCS data with national estimates of ED visits due to opioid use (i.e., "opioid-involved visits") to determine the potential of researching the impact and outcomes of opioid use on hospital EDs with NHCS data.⁹⁴

- **Vital Statistics Rapid Release Report: Timeliness of Death Certificate Data by Sex, Age, and Geography**: This report discusses the timeliness of death mortality data from the National Vital Statistics System.⁹⁵
- **National Health Statistics Report: Drug Overdose Deaths Involving Fentanyl, 2011–2016**: This report describes trends in drug overdose deaths involving fentanyl by demographic characteristics and geographic regions.⁹⁶
- **National Health Statistics Report: Drug-involved Infant Deaths in the United States, 2015–2017**: This report describes drug-involved infant deaths by type of drug and select maternal and infant characteristics, indicating where drugs were the underlying or a contributing cause of death.⁹⁷
- The research letter, titled “**Association of Medical Stimulants with Mortality in the U.S. From 2010 to 2017**” was published in *JAMA Internal Medicine* in February 2021. The authors used drug overdose data from the project to illustrate the association of medical stimulant use with mortality in the U.S. from 2010 to 2017.⁹⁸

VII) Food and Drug Administration (FDA)

FDA is administering five projects including one completed project and two cross-agency-funded projects described later in Section XI.

- Bridging the PCOR Infrastructure and Technology Innovation through Coordinated Registry Networks Community of Practice
- CURE ID: Aggregating and Analyzing COVID-19 Treatment from EHRs and Registries
- Enhancing Data Resources for Studying Patterns and Correlates of Mortality in Patient-Centered Outcomes Research: Project 2 - Pilot Linkage of NDI+ to Commercially and Publicly Insured Populations*
- Making Medicaid Data More Accessible Through Common Data Models and FHIR APIs*
- SHIELD—Standardization of Lab Data to Enhance Patient-Centered Outcomes Research and Value-based Care

* Denotes a cross-agency-funded project that involves more than one federal agency; these projects are described in the Cross-Agency-Funded Projects section (Section XI).

Bridging the PCOR Infrastructure and Technology Innovation through Coordinated Registry Networks Community of Practice

Period of Performance	Federal Point of Contact
4/16/19 - 9/30/22	Danica Marinac-Dabic

Comparative effectiveness research (CER) often relies on data captured at the point of care, re-entered into clinical research systems, and then consolidated and transformed for analysis and research purposes. The process is complex, labor-intensive, and expensive. It requires duplicate data entry and extensive data quality checks and restructuring to create research-ready datasets. As a result, study designs and study infrastructure for generating and analyzing real-world evidence are often limited.

FDA and its partners have invested significant resources in developing a strategic coordinated registry network (CRN), including a prior OS-PCORTF project, **Developing a Strategically Coordinated Registry Network to Support Research on Women's Health Technologies**. A collaboration between FDA, ONC, and NLM, the prior project aimed to align and connect existing registries of women's health technologies into a CRN, helping to address data-related barriers to conducting CER. As a continuation of this work, the FDA plans to strengthen existing CRNs as a real-world data source on medical devices in twelve clinical areas.

Unlike traditional clinical research networks, coordinated research networks overcome limitations of traditional registries by linking data from multiple sources, including Medicare and Medicaid claims, state-level datasets, and industry data to improve evaluation of medical device impact on quality of care.

Project Purpose and Goals

The FDA's current project will link data related to medical devices and improve standardized data collection to ease the information sharing process among networks as well as patients. The project team will harmonize minimum core datasets, incorporate unique identifiers for different types of medical

devices, create novel ways to include patient-generated data in the CRNs, and link CRNs to additional data sources. The resulting infrastructure will increase the availability of high-quality, relevant, reliable, timely, and actionable evidence to improve patient outcomes; it will also better support research on demographic (e.g., sex, age, race, and ethnicity) differences in patient use of medical devices and their health outcomes.

Building on the CRN Community of Practice (COP) that was developed in earlier phases of the FDA/MDEpiNet collaboration, this project will establish the CRN Learning Community (CLC). The CLC's goal is to create space for CRNs to share knowledge and ideas with each other in the areas of governance, informatics, methodologic approaches, data linkages, interoperability, and digital solutions. CRNs across twelve clinical areas make up the CLC. This collaborative initiative will create an opportunity to collect structured, standardized, analysis-ready patient data at the point of care. Strengthening the CLC and each individual CRN offers a more strategic approach to addressing the needs of the broader PCOR stakeholder community via harmonized and interoperable infrastructure. This strengthened infrastructure potentially supports more complex study designs.

The project aims to strengthen the CRNs as a national infrastructure for medical devices across the health care system through the following five objectives:

- Advance the CRNs' capacity for patient-centered outcomes research across twelve clinical areas by supporting their development in seven domains of maturity: patient engagement, unique device identification, data quality, efficiency, governance, sustainability, and fitness for use during the total product life cycle, and five levels of maturity: early learner, making progress, defined path to success, well-managed, and optimized.
- Pilot test and refine existing device specific FHIR® profiles (developed by the FY 2017 project **Developing a Strategically CRN for Women's Health Technologies**) in three to five CRNs to demonstrate the capture and exchange of CRN data using FHIR®.
- Pilot test and refine the instrument for capturing patient preferences in at least one clinical area (End Stage Kidney Disease) to evaluate scientifically valid data regarding patient uncertainty in accepting a variety of benefit/risk tradeoffs within a CRN.
- Advance CRN capacity to produce linked datasets, combine heterogeneous data, and develop machine learning techniques to validate and analyze the linked datasets.

CRNs are strengthening research capacity in twelve clinical areas:

- Women's health
- Vascular
- Cardiac
- Orthopedic
- Acute ischemic stroke
- Venous access
- Robotic surgery
- Temporomandibular joint
- Breast implants
- End stage kidney disease
- Abdominal hernia
- Prostate cancer

Example of a Mature CRN – VISION CRN

CRNs typically include data from national registry, claims data, EHRs, and patient-generated health data (PHGD). The VISION CRN has:

- Linked data from 2002-2019, with 15 years of follow-up, capturing over 400,000 people. Linkages cover 88% of all endovascular aneurysm repair (EVAR) patients and 93% of all abdominal aortic aneurysm (AAA) patients.
- Captured data on 880 clinical sites, 3,000 providers, and over 200 types of devices.
- Produced 30 publications, including 6 validation studies in high-impact journals.

- Develop a gender- and sex-specific outcome measure framework for devices and test it in the most mature CRNs.

These objectives meet several of the Goals in HHS' Strategic Plan for OS-PCORTF. By testing and refining FHIR® profiles to allow for the linking and exchange of standardized CRN data, this project addresses **Goal 2: Longitudinal Data Standards and Linkages**. The project's use of machine learning techniques to link CRN data aligns with **Goal 3: Technology Solutions to Advance Research**. Finally, this project's focus on capturing patient preferences and developing gender- and sex-specific outcome measure frameworks falls under **Goal 4: Person-Centeredness, Inclusion, and Equity**.

Accomplishments

Since FY 2019, this project has made progress to meet project objectives in the following areas:

- The FDA team developed a CRN assessment framework consisting of a novel tool to assess CRN data infrastructure across seven domains and five levels of maturity with input from a multi-stakeholder expert group of MDEpiNet collaborators from academia, clinical, industry, regulatory settings, and the patient community.
- Using results from CRN assessments, the FDA is currently developing an Implementation Roadmap. This document will create individualized implementation roadmaps for each CRN that helps them continue to develop their data infrastructure. The project team convened the CRN Architecture Taskforce (CAT) to facilitate harmonization and adoption of the Implementation Roadmap in 2021.
- The FDA team, in collaboration with multi-stakeholder community, continues to develop minimum core dataset of elements for each CRN. Three CRNs (End Stage Kidney Disease, Temporo-mandibular Joint, Venous Access CRN) currently have datasets under development. The final minimum core datasets will include a data dictionary and value sets.
- The FDA team is currently pilot testing and refining FHIR® profiles for devices, so that patient outcomes can be more easily associated with specific devices. These profiles will be included in CRN FHIR® implementation guides that will be made available to external researchers at the project's conclusion.
- Two CRNs have applied novel data linkage processes to link registry and claims data.
- The FDA team is currently working on developing open-access materials and implementation guides with four CRNs. Areas of focus include machine learning methodologies for predictions; augmentation approaches to address missing data in the CRN big data settings; transporting the results to specific CRN target populations; and matching CRN target populations with claims data. To demonstrate the feasibility of its machine learning algorithms, the project team completed an example study on cardiac ECMO and SUI analysis, and a liver cancer analysis is ongoing.
- The team is working on several approaches to enhance capacity to collect patient-generated health data and patient preferences: a mobile app that integrates with the CRN web platform; a proposal to pilot a patient-generated information instrument using HIVE™ software technology; and an instrument to capture patient preferences. The team is also currently preparing to analyze survey data collected to evaluate patient uncertainty in accepting a variety of benefit/risk tradeoffs.

- The team has produced more than 30 publications in peer-reviewed journals, and a special CRN Supplement featuring seven CRN papers was accepted for publication in the *British Medical Journal* (expected in March 2022).

Data and Technical Products

- A total of 16 [tools](#) have been developed, including data dictionaries, implementation guides, road maps, and lessons learned, and have been published on MDEpiNet website.⁹⁹
- CRN Assessment Tool: This tool assesses the maturity of CRN data infrastructure. It allowed the project team to develop individualized roadmaps to guide CRN development. A manuscript featuring the tool was accepted for publication in *British Medical Journal* (expected publication in March 2022).

Communication and Dissemination Products

- [MDEpiNET Project Website](#): This website summarizes the scope, activities, and tools developed under each of the twelve CRNs.
- MDEpiNet collaborators launched two task forces: 1) [Preparedness and Emergency Preparedness](#), and 2) [Blockchain and Artificial Intelligence](#). Activities for both task forces were conducted and concluded in 2020. The MDEpiNet webpages on each task force provide information on taskforce objectives, activities, and products.
- In 2021 and early 2022, the project team published 9 peer-reviewed journal manuscripts, demonstrating successful use of the network's data resources to support research.
 - Four manuscripts documenting safety and efficacy of medical devices and procedures were published, including [Outpatient Surgical Procedures for the Treatment of Benign Prostatic Enlargement](#) (*The Journal of Urology*, March 2021),¹⁰⁰ [Sling Mesh Implants for Stress Incontinence](#) (*The Journal of Urology*, January 2021),¹⁰¹ [TURP versus laser prostatectomy for the treatment of benign prostatic hyperplasia](#) (*World J Urol.*, December 2021),¹⁰² and [Femoropopliteal Artery Paclitaxel Devices](#) (JACC: *Cardiovascular Interventions*, December 2021).¹⁰³
 - Two validation studies were published, including [Validation of an indirect linkage algorithm to combine registry data with Medicare claims](#) (*Journal of Vascular Surgery*, February 2022)¹⁰⁴ and [Development and Usability Testing of a Mobile Application to Monitor Patient-Reported Outcomes after Stress Urinary Incontinence Surgery](#) (*Urology*, 2021).¹⁰⁵
 - Three studies assessing the impact of procedures on outcomes of interest, including sex-specific outcomes in [Sex Disparities in Long-Term Mortality after Paclitaxel Exposure in Patients with Peripheral Artery Disease: A Nationwide Claims-Based Cohort Study](#) (*J Clin Med*, July 2021)¹⁰⁶ and [Association of Sex with Risk of 2-year Revision for Patients Undergoing Total Hip Arthroplasty](#) (*JAMA Network Open*, June 2021)¹⁰⁷ and mortality in [Elective Repair of Intact Abdominal Aortic Aneurysms](#) (*Ann Surgery*, October 2021),¹⁰⁸ were published.

CURE ID: Aggregating and Analyzing COVID-19 Treatments from Electronic Health Records and Registries Globally

Period of Performance

5/1/21 – 10/31/23

Federal Point of Contact

Heather Stone

With the rapid emergence and exponential spread of COVID-19, the medical community has tried numerous treatments alone and in combination with others; only a small subset of these global clinical experiences have been published. However, there is an opportunity to learn from global COVID-19 treatment efforts more effectively by strengthening centralized data infrastructure platforms so that they can support data aggregation from a variety of global data sources. Building upon FDA's and the NIH's National Center for Advancing Translational Sciences' (NCATS) CURE ID platform, an existing repository focused on collecting information on drugs with potential off-label uses for infectious diseases, this project will expand the platform's infrastructure to make urgently needed data available to accelerate understanding of effective treatment options for COVID-19 and to avoid those that are ineffective or harmful.

This project will enhance **CURE ID's capabilities to support COVID-19 treatment research** by:

- Including specific fields to accommodate COVID-19 data elements.
- Developing and implementing an automated extraction tool.
- Creating a patient portal/form on the CURE ID platform to capture patient treatment experiences.
- Updating the CURE ID platform to include summary statistics and associated data visualizations.

Project Purpose and Goals

Initiated in 2012, CURE ID is an internet-based repository with approximately 7,000 registered users that allows clinicians to report and access case information related to the use of existing drugs for difficult-to-treat infectious diseases. The platform is a collaboration between the FDA and NIH/NCATS that enables the crowdsourcing of medical information from health care providers around the globe to facilitate the development of new treatments for neglected diseases.

The expansion of the CURE ID platform into COVID-19 will produce a rich repository of de-identified patient-level data based on case reports of individuals treated for COVID-19 infection with repurposed drugs. These data will be gathered via automated extraction and manual data collection from EHRs and clinical disease registries. Through the platform, health care providers will then be able to use data visualization tools to obtain real-time insight into the effectiveness of existing drugs, helping to distinguish those which are promising from those which may be unhelpful or even harmful. The information gathered through the platform can also help FDA and NIH work with sponsors of medical products, drug developers, and policymakers to develop the scientific data needed to support new treatment indications for COVID-19 and other unmet medical needs. Finally, more advanced analytics will serve the greater scientific community to guide new drug development and lead optimization based on clinical experience.

While the data infrastructure being built is specific for COVID-19, it will be sustainably designed so that it can be promptly deployed for future uses, including outbreaks of existing and emerging infectious diseases. This will provide ongoing, real-time access to the global clinical experience of repurposing drugs when there is an immediate need to identify potential existing treatments in the absence of novel drug development.

The project has the following specific objectives:

- Make large quantities of de-identified patient-level data on COVID-19 treatment from many different sources rapidly and openly available.
- Expand clinician engagement and creation of global treatment networks.
- Increase patient involvement in the platform.
- Identify promising drugs, drug combinations, or treatment regimens for COVID-19 and other diseases with inadequate therapy from the data in CURE ID.
- Ensure the sustainability and use of the CURE ID Platform's EHR/registry expansion beyond COVID-19.

By expanding the existing CURE ID platform to include additional COVID-19 patient data, this project aligns with **Goal 1: Data Capacity for National Health Priorities** of HHS' Strategic Plan for OS-PCORTF; this work also aligns with **Goal 3: Technology Solutions to Advance Research**, as it is utilizing advanced methods like automated extraction to gather and display de-identified patient data.



Coordination Across Federal Agencies and Additional Stakeholders

As part of this work, FDA and NIH/NCATS will continue to manage the CURE ID platform alongside the Critical Path Institute (C-Path). These three organizations represent the existing CURE Drug Repurposing Collaboratory (CDRC) consortium; C-Path will maintain its role as a coordinating center. The CDRC will launch the Partnership Coordinating Center (CDRC PCC), a collaboration between institutions that have expertise in the clinical management of infectious diseases and COVID-19, as well as experience in automated data extraction from EHRs and registries. Partners include the Johns Hopkins University School of Medicine, the Society of Critical Care Medicine/Mayo Clinic's VIRUS Registry, and Emory University School of Medicine. The FDA team will also work closely with the Infectious Diseases Data Observatory of Oxford University to ensure project efforts are aligned with other global EHR data standardization and harmonization efforts.



Accomplishments

By the conclusion of this project, the joint NIH/NCATs and FDA team will have improved the existing CURE ID platform by expanding the electronic data sources available to researchers on CURE ID to include EHRs and clinical disease registries.

This project will enhance CURE ID's capabilities to support COVID-19 treatment and/or future infectious disease research by:

- Developing partnerships with clinical consultants, technical consultants, and data providers to capture the most critical treatment and patient outcomes data from EHRs and registries.
- Customizing CURE ID data fields to accommodate COVID-19 data elements in the CURE ID platform.
- Building the infrastructure, technology, and methodology needed to extract and aggregate targeted clinical data from many global sources (EHRs, registries, and clinician-submitted and published cases); this will include developing and implementing an automated extraction tool.

- Creating a patient portal/form on the CURE ID platform to capture patient treatment experiences.
- Updating the CURE ID platform to include summary statistics and associated data visualizations.
- Capturing social determinants of health from the medical records.
- Expanding capacity to gather patient and parent-reported outcomes from other vulnerable populations, including pregnant women and neonates.

Anticipated products are described in more detail below.

Data and Technical Products

- A data dictionary that describes the newly incorporated data elements and algorithms related to COVID-19 that are added to the CURE ID platform.
- Data extraction tools for EHR and registry data sources along with the associated open-source code that will be publicly available.
- Interactive data visualizations on the newly added COVID-19 case reports will be available to users of CURE ID.
- Data briefs will be published describing the data that researchers can access through the CURE ID platform including descriptions of SDOH and patient-reported data.

Communication and Dissemination Products

- Publications and journal articles that describe the methodological learnings and development of the automated data extraction tools.
- Publications and journal articles describing analyses conducted on data from CURE ID.

SHIELD: Standardization of Lab Data to Enhance Patient-Centered Outcomes Research and Value-based Care

Period of Performance

4/1/19 - 10/21/21

Federal Point of Contact

Greg Pappas

The COVID-19 pandemic has underscored the need for reliable diagnostic laboratory data. Public health surveillance of diseases and conditions of national concern, such as SARS-CoV-2, relies on standardized reporting of laboratory data. However, mapping test results across electronic laboratory reporting systems is challenging due to the use of numerous testing platforms that do not use similar coding structures. Different institutions often code data from in vitro diagnostic (IVD) tests differently, which leads to ambiguity and prevents researchers from being able to readily access and use the data. These challenges have impeded public health efforts to accurately document the spread of SARS-CoV-2 and inform clinical and public health decisions concerning prevention and treatment.

Prior to the pandemic, FDA was already working to improve laboratory data infrastructure. Public workshops in 2015 and 2016 lead to the formation of the Systemic Harmonization and Interoperability Enhancement for Laboratory Data (SHIELD) collaborative. SHIELD is a public-private collaborative with

over 70 partners at FDA, CDC, NIH, ONC, CMS, U.S. Department of Veterans Affairs (VA), Pew Charitable Trusts, National Evaluation System for Health Technology (NEST), College of American Pathologists, as well as IVD manufacturers, EHR vendors, laboratories, standards developers, and academics.

F Project Purpose and Goals

This project expands the collaborative efforts of FDA and other stakeholders involved in the SHIELD initiative to standardize laboratory results across different EHR systems. This project will encourage device manufacturers to use specific Logical Observations Identifiers Names and Codes (LOINC) mappings for certain IVDs by developing manuals to consistently map the LOINC codes to the same type of IVD. This is important because without the specific guidance, laboratories and registries often assign different (and frequently incorrect) LOINC codes for the same type of IVD test. This project will also test standard digital formats that update the infrastructure of participating provider institutions and registries. These formats will include the seamless distribution of LOINC and SNOMED-CT

(Systematized Nomenclature of Medicine—Clinical Terms) coding. This project will improve confidence in the quality of data used in laboratory studies that cover multiple clinical sites. The resulting data from the SHIELD project can be used to study improvements to the timeliness, cost, and quality of patient care.

Previous Work on LOINC

- FDA has been actively engaged in efforts to develop tools necessary to decrease ambiguity in mappings between IVD tests and LOINC.
- In 2018, the FDA released [guidance](#) for using LOINC for IVD devices.
- This OS-PCORTF project will replicate processes used to develop the [Guide for Using LOINC Microbiology Terms](#) to create LOINC mapping manuals for six IVD domains.

Exhibit 6. SHIELD Project Overview

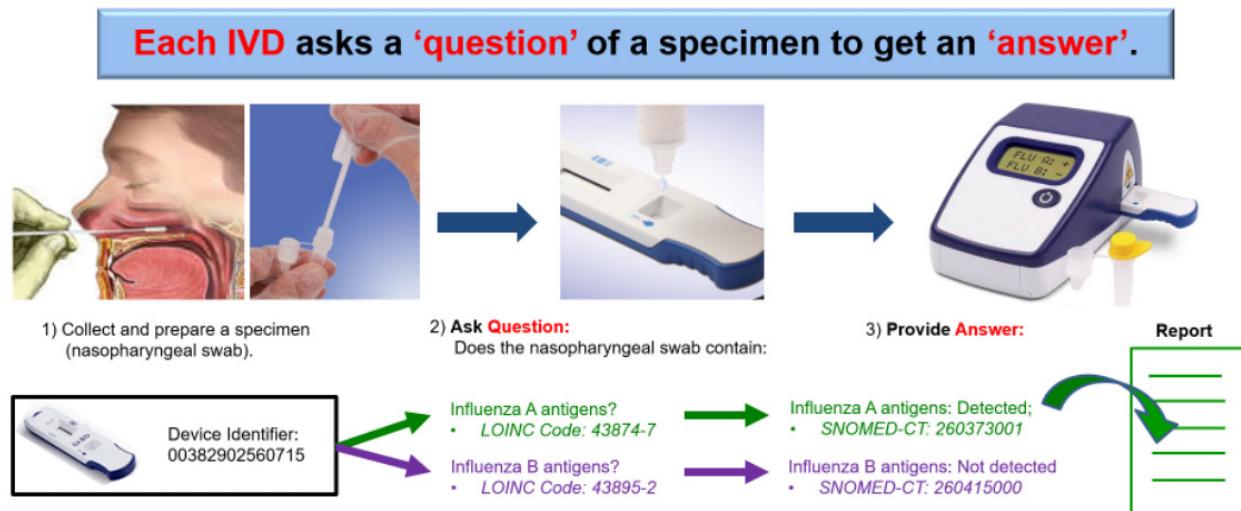


Image Source: <https://mdic.org/program/systemic-harmonization-and-interoperability-enhancement-for-lab-data-shield/#toggle-id-1>

In addition to improving quality, interoperability, and portability of laboratory data sharing, the COVID-19 pandemic required a rapid SHIELD focus on SARS-CoV-2 testing. The Coronavirus Aid Relief and Economic Security (CARES) Act requires “every laboratory that performs or analyzes a test that is intended to detect SARS-CoV-2 or to diagnose a possible case of COVID-19” to report the results from

each test to HHS.¹⁰⁹ On June 4, 2020, HHS announced new laboratory data reporting guidance for standardized COVID-19 test reporting, which was then updated on January 8, 2021.¹¹⁰ SHIELD developed and continues to maintain the coding specification for standardized reporting of SARS-CoV-2 test results.

The overall objectives of the project are to:

- Develop LOINC code mapping manuals for the following IVD domains: Chemistry, Drug/Toxicology, Allergy, Serology/Hematology, Cell Markers, and Molecular Pathology.
- Conduct pilot implementations and testing of an interoperability upgrade to laboratory information systems and registries by incorporating a SHIELD-approved, high-quality, industry-defined, and supported format to facilitate the publication and exchange of LOINC codes for vendor IVD test results.
- Assess the interoperability and value of the systems and tools tested in six pilot laboratory settings and EHR systems pre- and post-implementation.

This project aligns with several of the goals in HHS' Strategic Plan for OS-PCORTF. SHIELD's efforts to standardize laboratory data using harmonized measures fits within the purview of **Goal 2: Longitudinal Data Standards and Linkages**. Additionally, the focus on improving standards for COVID-19 laboratory test data aligns with **Goal 1: Data Capacity for National Health Priorities**.

Accomplishments

Since the project began in 2019, the project team has completed numerous activities to improve the quality, interoperability, and portability of laboratory data. In response to the COVID-19 pandemic, the project redirected its efforts to address COVID-19 laboratory test data.

- FDA completed drafts of all six LOINC IVD mapping manuals. Final versions and guides under development are published online. Draft manuals undergo public comment periods, testing, and revision (see Data and Technical Products for available manuals).
- The project team completed its demonstration of the SHIELD IVD LOINC mapping infrastructure. The project partnered with 5 clinical laboratories to pilot test SHIELD standards in active laboratory information systems and registries. Findings from the demonstration will be used to inform federal policy and regulations.
- The FDA team has met with the Clinical Laboratory Standards Institute (CLSI) to develop a document to describe to laboratories how to consistently implement the laboratory informatics standards developed through the project.
- The project team's work on laboratory data interoperability is being leveraged heavily in response to the COVID-19 pandemic. SHIELD's efforts resulted in the release of HHS guidance requiring the use of SHIELD-harmonized standards: LOINC In Vitro Diagnostic (LIVD) Test Code Mapping for SARS-CoV-2 Tests.¹¹¹
- The project team successfully completed strategic planning efforts. They have worked with the FDA and SHIELD collaborators to develop the SHIELD Strategic Plan, which will support the implementation of SHIELD standards nationally.



Data and Technical Products

- **LOINC Mapping Guides**: SHIELD partners released a **Quick Start Guide for Mapping Laboratory LOINC** to teach users how to select the correct LOINC term for lab tests. The Quick Start Guide can be used in tandem with domain-specific guides, which are in varying stages of development by the SHIELD team. The domain-specific guides contain mapping principles, examples, and exercises for each IVD domain. Guides in **Chemistry** and **Drug/Toxicology terms** are undergoing the final stages of review and a guide on **Allergy terms** is currently being tested.
- **LOINC In Vitro Diagnostic (LIVD) Test Code Mapping for SARS-CoV-2 Tests**: This LIVD mapping tool defines LOINC and SNOMED-CT codes for all FDA-approved SARS-CoV-2 diagnostic tests. Use of this tool will harmonize code selection between health care systems and reduce inaccuracies and time required to share laboratory test results with public health agencies.¹¹²



Communication and Dissemination Products

- SHIELD Websites: The project team maintains two SHIELD websites. The two websites on **Medical Device Innovation Consortium** and **MDEpiNET** includes SARS-CoV-2 and COVID-19 test coding resources for emergency use authorization authorized tests. Detailed information can be found on either project website.¹¹³
- **FDA Voices Blog Post**: This blog post, titled “FDA’s Ongoing Work to Support and Advance COVID-19 Diagnostic Test Accuracy and Availability”, describes FDA initiatives to address the COVID-19 pandemic, including SHIELD contributions. The FDA has been leveraging SHIELD’s initiative to harmonize COVID-19 test data.^{114; 115}

VIII) National Institutes of Health (NIH)

NIH is administering eight active projects including five cross-agency-funded projects described later in Section XI.

- Creating a Federal COVID-19 Longitudinal Patient Outcomes Research Database Linked to Health Systems and Clinical Data
- Data Capacity for Patient-Centered Outcomes Research through Creation of an Electronic Care Plan for People with Multiple Chronic Conditions*
- Data Capacity for Patient-Centered Outcomes Research through Creation of an Electronic Care Plan for People with Multiple Chronic Conditions 2.0: Development of the Patient-facing Application*
- Making Medicaid Data More Accessible Through Common Data Models and FHIR APIs*
- NIH/NIDA's AMNET: An Addiction Medicine Network to Address the United States Opioid Crisis
- Severe Maternal Morbidity and Mortality EHR Data Infrastructure
- Training Data for Machine Learning to Enhance Patient-Centered Outcomes Research Data Infrastructure*
- Understanding Long-term Outcomes in COVID-19 Survivors with Multiple Chronic Conditions through e-Care Plan Development*

* Denotes a cross-agency-funded project that involves more than one federal agency; these projects are described in the Cross-Agency-Funded Projects section (Section XI)

Creating a Federal COVID-19 Longitudinal Patient Outcomes Research Database Linked to Health Systems and Clinical Data

Period of Performance	Federal Point of Contact
8/13/21-8/12/24	Ken Gersing

The COVID-19 pandemic has precipitated an urgent need to have a near real-time centralized research dataset to conduct patient-centered outcomes research on COVID-19 and generate evidence on effective interventions, especially in vulnerable patient subgroups. Using linked clinical and claims data, researchers can address important questions on COVID therapeutics and vaccines, long-term complications of COVID, disparate impacts of COVID-19 on vulnerable populations, and the safety-net providers who serve them. To achieve this goal, this project will link CMS claims data, community characteristics, and social determinants of health (SDOH) data to the EHR within the NIH/NCATS National COVID Cohort Collaborative (N3C) Data Enclave, a racially and ethnically representative national sample of patients from a variety of contributors. These contributors include several NIH institutions as well as community data sources, i.e., community hospitals and federally qualified health centers. N3C is large, with over 5 million patients and 6 billion rows of data, but it is not comprehensive because the data contributors to the N3C Data Enclave are part of an “open” system, meaning patients may get their care from a variety of places and not necessarily all data are part of the N3C dataset. Additionally, the NIH/NCATS N3C Data Enclave has obscured tribal reported data held within the enclave.

Linkage of the N3C Data Enclave with CMS claims data enriches the dataset by granting access to claims filed by providers on behalf of Medicare patients, filling in certain knowledge gaps in the process. The linked data sets will facilitate more robust scientific inquiries by allowing researchers to construct and evaluate more comprehensive longitudinal COVID-19 care utilization patterns prior to and during the COVID-19 pandemic, including for patients who continue to suffer symptoms after the acute infections called Long Covid (also known as Post-Acute Sequelae of COVID). By linking the N3C Data Enclave to CMS claims data, this dataset will include additional linkages to SDOH on individuals, as well as aggregate community and provider characteristics. The additional linkages will also allow researchers to analyze relationships between sociodemographic information, provider/health system/community characteristics, health system-level interventions, and their relationships with patient health outcomes. Additionally, NIH/NCATS is

working with the NIH Tribal Health Research Office and consulting with the Tribal Advisory Committee to develop an appropriate process and structure to enable direct access to tribal data for research.

To ensure patient privacy, investigators will only be able to access the linked CMS-N3C national COVID longitudinal research dataset via the secure, cloud-based N3C Data Enclave. Enclave access is exclusively granted to investigators who are covered by an N3C Data Use Agreement, have an approved Data Use Request (DUR), and have had their research proposals reviewed by their Institutional Review Board. Investigators must complete an additional DUR process to access row-level data. All DURs will be carefully reviewed by NIH representatives serving on the N3C data access committee (DAC). The DAC will only grant access to investigators that have appropriate justifications for requesting Level 3 data as part of COVID-19 research.

Project Purpose and Goals

The overall goal of this project is to enhance COVID-19 data infrastructure for patient-centered outcomes research by producing a national research dataset on COVID-19 that links EHR data to CMS claims data, community characteristics and SDOH in the NIH/NCATS N3C's privacy-preserving, secure cloud-based data enclave. NIH/NCATs will use innovations to produce more comprehensive, timely, longitudinal, and supplemented information on COVID-19.

The project has the following goals:

- Assess and demonstrate the feasibility of linking clinical EHR data with Medicare claims data using the proposed N3C data linkage strategy and engage patient-centered outcomes researchers in using the linked research dataset;
- Provide researchers secure access to the NCATS N3C Data Enclave so they can access Medicare Claims Data linked to the N3C Clinical EHR Data;

When completed, this dataset could be used for studying PCOR issues such as:

- Prescription drugs (both COVID-19 and for underlying health conditions) whose utilization may affect COVID-19 outcomes
- Health outcomes after COVID-19 vaccinations, such as rates of infection, hospitalizations for COVID-19 or serious side-effects and death, especially in vulnerable patient subgroups who are less well-represented in clinical trials.
- Disparities in COVID-19 treatment and outcomes by race/ethnicity, individual/community characteristics on social determinants of health, or other patient demographic characteristics.

- Demonstrate the utility of the linked Medicare claims-N3C clinical data to conduct patient-centered outcomes research on COVID-19, including potential evaluation of economic outcomes; and
- Support the joint activities of the OS-PCORTF COVID Collaborative.

The project supports HHS's priorities to build a national longitudinal data infrastructure on COVID-19 and this enhanced dataset will serve as a critical resource for patient-centered outcomes researchers (both federal and non-federal), policymakers, and other end-users for surveillance and research on the national COVID-19 pandemic in the United States. It also supports **Goal 2: Longitudinal Data Standards and Linkages** of OS-PCORTF's Strategic Plan (2022-2029).

Accomplishments

Before the end of the Period of Performance, the project team will have produced and provided access to a national longitudinal research dataset on COVID-19 that links EHR data to CMS claims data, community characteristics, and SDOH in the NIH/NCATS N3C's data enclave. Additionally, the project team will produce multiple COVID-19 use cases that demonstrate the utility of the dataset. Anticipated products are described in more detail below.

Data and Technical Products

- The CMS Medicare-N3C Linked National Longitudinal COVID Research Dataset will be available to approved N3C users via the N3C Data Enclave and restricted access procedures; it will be updated monthly.
- CMS Medicare Data Preparation Methods Documentation will document how the CMS Medicare claims data was prepared prior to linkage to N3C, including a data dictionary and definitions for constructed variables (i.e., provider/health system and community characteristics).
- Final Data Methods Documentation will be available to linked dataset users and will document how the CMS Medicare claims data was linked to N3C, including results on data linkage accuracy.

Communication and Dissemination Products

- The team will publish a N3C web page that will feature key statistics on the CMS Medicare-N3C linked data including numbers, characteristics, representativeness of the sample compared to the CMS population, and sampling weights.
- Medicare Use Case Reports will be publicly available and elaborate on two Medicare use cases. The intent of disseminating them is to demonstrate the utility of the linked Medicare-N3C data to address PCOR questions on COVID-19. Economic outcomes may be incorporated into the two use case reports, or it may be explored in a separate use case. Use case manuscripts will be submitted for journal review.
- A White Paper on the CMS-N3C Linked Dataset and Use Cases will be developed with input from patient-centered outcomes researchers on the proposed linked dataset and the potential Medicare and Medicaid PCOR COVID-19 use cases. The final product will be disseminated to researchers via NIH and other HHS websites.
- Final Report will include Medicaid-N3C claims validation and EHR-claims concordance findings for dissemination to N3C users.

AMNET: An Addiction Medicine Network to Address the United States Opioid Crisis

Period of Performance	Federal Point of Contact
4/15/19 – 10/15/22	Carlos Blanco

Only about one-quarter of the over 2 million Americans with opioid use disorder (OUD) receive treatment.¹¹⁶ Individuals with OUD can receive OUD treatment in two settings, either office-based treatment or in opioid-treatment programs. Office-based treatment programs allow certified clinicians to offer care in an outpatient setting, such as a primary care office, alongside regular medical services. In contrast, opioid treatment programs are federally certified programs that only treat OUD and often involve daily attendance by patients.¹¹⁷ Office-based opioid treatment programs play a key role in the response to the opioid epidemic given the limited number of federally certified opioid treatment programs available to patients, especially in rural and other underserved communities. Of the three FDA-approved medications for treatment of OUD—buprenorphine, methadone, and extended-release naltrexone¹¹⁸—office-based practices can provide buprenorphine and naltrexone in the United States. Despite the increased access to care that office-based treatment programs provide, little is known about the outcomes for the patients treated with buprenorphine and naltrexone in these settings. These practices typically do not collect standardized data on patients' characteristics, treatments, and outcomes and have not been harnessed to conduct patient-centered outcomes research.

In addition to developing the AMNet practice-based research network and electronic patient registry, the team is identifying:

- Standardized clinician-rated and patient-reported assessment and quality measures, available via the PsychPRO portal
- Standardized and validated common data elements (CDEs) relevant to OUD and its treatment that will be collected and made available through PsychPro

Project Purpose and Goals

This NIH/NIDA project will establish a new practice-based research network and an electronic patient registry named the Addiction Medicine Network (AMNet) that connects to office-based practices. AMNet will operate as a platform for OUD treatment research, rapidly providing data to clinicians, researchers, and other stakeholders on OUD patient outcomes and routine clinic-based treatment delivery. AMNet will also collect data for clinical and health services research focused on addiction. AMNet will adapt the American Psychiatric Association's (APA) clinical data registry (PsychPRO) to collect OUD treatment data. AMNet's addiction medicine practitioners will collect standardized treatment and outcomes data, including clinical and patient-reported data from routine clinical practice (including in understudied populations). These improvements to addiction data infrastructure will increase data capacity to combat the opioid epidemic.

Participation in AMNet will improve clinical decision-making among addiction treatment providers and support research by AMNet participants and extramural scientists. With a focus on practice-based research, AMNet will develop the foundations of a research network for community-based clinical trials (e.g., comparative effectiveness of extended-release formulations of OUD medications). AMNet will provide real-time data on patient characteristics, care delivery, and recovery service utilization to aid performance improvement efforts and patient outcomes.

The specific objectives of this project are:

- Establish AMNet, an addiction medicine practice-based research network.
- Adapt PsychPRO to support data collected for AMNet.
- Perform feasibility and validity testing of AMNet measures and OUD CDEs.
- Expand addiction medicine research capacity and outreach.
- Develop business requirements for linking AMNet to other databases and registries. The first data linkage project will link AMNet data for providers participating in CMS' Merit-Based Incentive Program (MIPS).

With its goal of standardizing patient-provided information to improve data capacity for OUD in underserved and understudied populations, this project aligns with **Goal 4, Person-Centeredness, Inclusion, and Equity** of HHS' Strategic Plan for OS-PCORTF.

Accomplishments

The project team has made progress on the following activities:

- The NIDA funded research team, led by investigators at the Friends Research Institute, APA, and American Society of Addiction Medicine (ASAM), has developed and offered training materials for AMNet participants and clinicians interested in AMNet research. Materials include a newsletter for participating providers containing training materials, and AMNet updates and resources; the research team has also developed seven webinars on various educational topics related to addiction medicine.
- The project established an external Steering Committee comprised of experts in the field of addiction psychiatry and addiction medicine in addition to patient and Federal agency representatives. The project team has recruited 78 of the planned 120 addiction treatment practitioners.
- To advance standardized measurement and collection of assessment measures in a registry environment, AMNet investigators selected patient-reported outcome measures (PROMs) and quality measures for feasibility and validity testing. The APA information technology provider integrated the selected PROMs into PsychPRO, which will support data collection via AMNet.
- The project team is currently analyzing the results from beta testing of the clinical utility and feasibility of AMNet's assessment tools within five participating sites. A manuscript detailing the qualitative and quantitative findings will be published.

Communication and Dissemination Products

- A journal article, titled "[Addiction Medicine Practice-Based Research Network \(AMNet\): Building partnerships](#)" was published in *Psychiatric Services* in April 2021. This manuscript describes the project's collaborative efforts with the APA, American Society of Addiction Medicine, Friends Research Institute, and NIH/NIDA to create AMNet.¹¹⁹

- A journal article, titled “[**Addiction Medicine Practice-Based Research Network \(AMNet\): Assessment Tools and Quality Measures**](#)” was published in *Substance Abuse and Rehabilitation* in June 2021. This manuscript summarizes results of the environmental scan and efforts to identify, review, and select assessment tools and quality of care performance measures for OUD and SUD for inclusion in AMNet.¹²⁰

Severe Maternal Morbidity and Mortality Electronic Health Record Data Infrastructure

Period of Performance	Federal Point of Contact
3/1/2021-3/1/2024	Nahida Chakhtoura

Between 1987 and 2016, pregnancy-related mortality in the United States, as reported by the CDC, increased from 7.2 to 16.9 maternal deaths per 100,000 live births.¹²¹ Moreover, racial disparities in maternal mortality in the United States are staggering, with a greater than three-fold higher maternal mortality rate of 41 deaths per 100,000 live births for non-Hispanic black women compared with 13 deaths per 100,000 live births for non-Hispanic white women.^{122 123 124} Of the approximately 700 pregnancy-related deaths occurring in the U.S. yearly, more than half are attributable to preventable causes.¹²⁵

The U.S. lacks an efficient means to analyze how a woman’s longitudinal medical history and basic socioeconomic and demographic characteristics affect pregnancy outcomes for both the mother and the infant. HL7® FHIR® provides new opportunities for extracting and linking health data from EHRs; however, information requirements and definitions must be aligned. An HL7® FHIR® implementation guide (IG) can help create the infrastructure needed to link maternal and infant EHR data, researchers, policymakers, clinicians, and other stakeholders. IGs are “a set of rules about how FHIR® resources are used (or should be used) to solve a particular problem, with associated documentation to support and clarify the usage.”¹²⁶ An HL7® FHIR® IG can facilitate data linkages to support research that can identify and understand opportunities for intervention and prevention of adverse maternal and infant events. It can improve data standardization across EHRs, create a more complete profile of life course health for mother and infant, and include additional information, including some social determinants (age, race and ethnicity, nativity, marital status, and education) that impact maternal and infant health.

Project Purpose and Goals

The purpose of this project is to strengthen both maternal and infant health data needed to examine the effect of medical conditions and/or interventions on pregnant, postpartum or lactating women, and their infants. The project will address this need by creating an HL7® FHIR® IG, which supports key data exchanges and linkages. The guide will encompass pregnancy, pregnancy outcomes, and pregnancy-related conditions, co-morbidities, and procedures for 1-year postpartum.

The project, designed to first engage stakeholders from the research, public health and providers communities, and members of the OS-PCORTF Maternal Health Consortium, is aimed at reaching a consensus on harmonized information requirements and definitions to create a means of accessing these data through a FHIR® application programming interface (API) by:

- Developing HL7® FHIR® standards and guidance for pregnancy, including pregnancy outcomes, comorbid conditions, and pregnancy-related conditions and procedures, hereafter called standards. These standards will support research on maternal morbidity and mortality, pregnancy, pregnancy outcomes (e.g., miscarriage, stillbirth, live birth), and pregnancy-related conditions and procedures (e.g., pre-eclampsia, gestational diabetes).
- Conducting a proof-of-concept pilot study to assess the feasibility of the standards and Implementation Guidance in two existing NIH data systems that contain EHR data and utilize FHIR® standards.
- Conducting a feasibility pilot study in health departments in one to two states and the District of Columbia using the project standards to link EHR data with maternal mortality data in vital records.
- Producing a report that includes the final standards (standards, data models and vocabularies, and implementation guidance), making it available on the HL7® website with references on NIH, ONC, and CDC websites, and distributing it to key research and public health partners.

The HL7® FHIR® IG will provide researchers with the ability to identify the full range of risk factors, pre-existing conditions, and causes of severe maternal morbidity and pregnancy-related mortality, including longitudinal medical history and basic socioeconomic and demographic characteristics, that affect pregnancy outcomes for both the mother and the infant.

By linking clinical, intervention, and community data to create and standardize longitudinal individual-level records while preserving privacy, the aims of the **Severe Maternal Morbidity and Mortality EHR Data Infrastructure** project align with **Goal 2: Longitudinal Data Standards and Linkages** of HHS' Strategic Plan for OS-PCORTF.

Accomplishments

Since the project launched in March of 2021, the project team accomplished the following tasks:

- Establishing an organizational structure, including a technical expert panel comprised of federal, clinical, and public health partners. After an inaugural meeting, the Steering Committee has convened regularly to provide guidance for the project.
- Participating in the OS-PCORTF Maternal Health Consortium of initially three funded projects, with potential to involve additional federal projects. The consortium focuses on developing an agreed upon method for the linkage of mother/child data, developing codes and value sets to define pregnancy in the implementation guide, and reviewing existing tools to capture more robust data to support maternal health research.

Data and Technical Products

- The project team has developed the [draft value set](#) and will be reviewed by the TEP. The team defined the scope of the HL7® implementation guide to include key data elements in a longitudinal record of a mother and child for two initial research use cases: (1) Pregnancy and subsequent death within a year of a pregnancy, regardless of cause of death or pregnancy outcome; and (2) Pregnancy-induced hypertension (PIH) focused on women with a diagnosis of pregnancy-induced hypertension.

- The project team has created and socialized the HL7® Project Scope Statement with interested stakeholders at the September 2021 HL7® Working Group Meeting and then shepherded it through the Consensus Review process, an initial step in taking an IG through ballot with HL7®.

Communication and Dissemination Products

- The team completed an environmental scan which included an inventory of maternal and child data requirements. The team mapped the requirements and will continue to refine them as the IG is being developed.

IX) Office of the Assistant Secretary for Planning and Evaluation (ASPE)

ASPE is administering four active projects including one cross-agency-funded project described later in Section XI.

- Dataset on Intellectual and Developmental Disabilities: Linking Data to Enhance Person-Centered Outcomes Research
- Child and Caregiver Outcomes Using Linked Data (CCOULD)
- Multistate Emergency Medical Services and Medicaid Dataset (MEMD): A Linked Dataset for PCOR
- Validating and Expanding Claims-based Algorithms of Frailty and Functional Disability for Value-based Care and Payment*

* Denotes a cross-agency-funded project that involves more than one federal agency; these projects are described in the Cross-Agency-Funded Projects section (Section XI).

Dataset on Intellectual and Developmental Disabilities: Linking Data to Enhance Person-Centered Outcomes Research

Period of Performance	Federal Point of Contact
2022-2026	Emma Plourde

An estimated 7.4 million people with Intellectual and Developmental Disabilities (ID/DD) live in the United States (U.S.).¹²⁷ In 2015, the U.S. spent over \$65.2 billion on ID/DD long-term services and supports through Medicaid.¹²⁸ While individuals with ID/DD receive services and supports in a variety of institutional and community-based settings, they are increasingly living in more community-integrated settings, sometimes with family members or independently.¹²⁹ Through various Medicaid authorities, it is estimated that states cover long-term services and supports for over one million people with ID/DD.¹³⁰

However, despite significant investments in supports and services for the ID/DD population, there is a lack of information on health experiences and person-centered outcomes in administrative and other data sources. As a result, there are significant barriers to studying health experiences and person-centered outcomes in administrative and other data sources for the ID/DD population. This lack of information limits the ability of the Federal government, states, and researchers to assess the quality and outcomes prior, during, and after services provided to people with ID/DD, especially during a public health crisis like the COVID-19 pandemic. A lack of evidence on home and community-based services' (HCBS) impact on person-centered outcomes also limits individuals' with ID/DD and their families' ability to strategically choose and invest in supports and services that promote their desired person-centered outcomes.

This project will build data infrastructure for the ID/DD population by linking the following state-level data sources: Supports Intensity Scale (SIS), Medicaid claims, and the National Core Indicators In-Person Survey (NCI IPS), and other potentially relevant state data sources. The linked data will enable researchers to analyze relationships between various sociodemographic information, need for HCBS, service utilization, Medicaid expenditures, and person-centered outcomes for the ID/DD population prior to and during the COVID-19 pandemic.

Project Purpose and Goals

The purpose of this project is to produce a de-identified publicly accessible dataset of linked data sources that include sociodemographic information, health services data, and person-centered outcomes for the ID/DD population prior to and during the COVID-19 pandemic. This dataset is intended to allow researchers to analyze relationships between these variables. Four to six states will participate in the pilot project and submit data to be included in the Dataset on Intellectual and Developmental Disabilities.

Project objectives include:

- Linking 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
- Producing the Dataset on Intellectual and Developmental Disabilities (DIDD), a publicly accessible de-identified dataset of the linked data
- Conducting exploratory analyses of the linked dataset to demonstrate utility of the dataset for person-centered outcomes research for the ID/DD population
- Disseminating data linkage methodologies and technical assistance guidance to help states replicate data linkages

By linking sociodemographic, environmental, and other non-health care datasets to support patient-centered research more fully, this work aligns with **Goal 4: Patient-Centeredness, Inclusion, and Equity** of HHS's Strategic Plan for OS-PCORTF.

Examples of patient-centered outcomes research for the ID/DD population that may be conducted using the linked dataset include:

- Determining appropriate and effective use of medication by evaluating relationships between need for behavioral supports, service utilization or expenditure rates for psychosocial interventions and psychotropic medications, and the individual's reported health status.
- Assessing predictors of gainful employment by exploring relationships between an individual's need for vocational rehabilitation services, utilization or expenditures for certain vocational rehabilitation supports, and the individual's report of having a paid job in the community with which they are satisfied.
- Evaluating community inclusion and participation by examining an individual's need for community living activities, utilization or expenditure rates for day habilitation services, and the individual's rated satisfaction with their level of participation in various community activities.

Accomplishments

This project will help to establish and promote access to relevant linked data sets specific to individuals with ID/DD. This larger goal will be bolstered by the potential deliverables detailed below.

Data and Technical Products

- A de-identified, public use or restricted use dataset of individuals with ID/DD, as well as a codebook and user guide.
- A dataset sustainability plan which will identify requirements, potential sources of additional funding and partnerships to support the dataset and its sustained use.

- A data linkage roadmap which will include approaches to partnering with states, identifying relevant data sources, establishing data use agreements, and navigating state and federal privacy laws and requirements related to the transfer, storage, linking and maintenance of the linked dataset.

Communication and Dissemination Products

- A policy brief summarizing the project goals and demonstrating how the linked data can address patient-centered outcomes research for individuals with ID/DD.
- Findings from the exploratory analyses will be disseminated through published reports and peer-reviewed journals. These materials will also be presented to the OS-PCORTF collaborative.
- A stakeholder webinar that demonstrates and describes dataset functionalities and utility.

Child and Caregiver Outcomes Using Linked Data (CCOULD)

Period of Performance	Federal Point of Contact
6/1/19 - 12/31/22	Robin Gherchner (ASPE) Emily Madden (ASPE) Valeria Butler (ACF)

Parental substance use negatively impacts a child's health and other outcomes later in life, including child maltreatment¹³¹ and medical and behavioral health issues.^{132, 133} Having a parent or caregiver who has a substance use disorder (SUD) is a leading determinant of child maltreatment and foster care placement.¹³⁴ Child welfare-involved parents with SUDs are an especially vulnerable population facing additional barriers (e.g., financial, treatment adherence, and inadequate recovery supports) to receiving SUD treatment. These barriers are important to understand in order to provide successful treatment and sustain recovery. A family's involvement in the child welfare system can be an opportunity for caregivers to connect to integrated, evidenced-based treatment and services to support their path to recovery.

Understanding the correlation between SUD diagnosis and treatment, additional supports provided by the child welfare system, and a return to successful daily functioning and family stability is central to assessing patient outcomes for parents receiving treatment. However, health outcomes for child-parent dyads are rarely tracked regularly by state child welfare agencies, and this impedes the ability to understand how a parent's SUD affects child-welfare involved youth.

Linking child welfare and Medicaid data can help ensure that agencies are leveraging existing data to best meet the assessment, treatment and service needs of parents and caregivers with SUDs and their children. Linking Medicaid and child welfare case records for children in the foster care system is required for new Comprehensive Child Welfare Information Systems (CCWIS)—the case management information systems for child welfare programs. CCWIS are currently under development in most states; however, states do not typically link parent Medicaid records to the CCWIS for research, and federal-level data is unavailable. Therefore, little is known about use of Medicaid for SUD or mental health treatment within the population of parents who have children in the child welfare system, or how that treatment may be associated with child welfare outcomes. There is also a gap in knowledge around how services provided by child welfare agencies align with those paid for by Medicaid. Few, if any, states have the capacity to monitor SUD treatment outcomes of parents with children in child welfare systems, as adequate data on parents are rarely collected.

Project Purpose and Goals

The Child and Caregiver Outcomes Using Linked Data (CCOULD) project aims to enhance PCOR data infrastructure and increase data availability for research on parents with a SUD or behavioral health issue who have children in the child welfare system. CCOULD provides technical assistance to states to develop state-specific datasets linking parental Medicaid claims and corresponding CCWIS data. The data will be combined into a multi-state, deidentified dataset for secondary data analysis. Federally linked parental claims data and CCWIS data has previously been unavailable, but the data harmonization created through CCOULD will allow national-level linkages. Potential end users of the dataset include researchers, state and federal agencies, national and state health policy groups, and child welfare advocates.

These data will be made available to researchers studying the relationship between involvement with child welfare services and SUD treatment outcomes among parents, focusing on the outcomes that are most important to parents (e.g., Will completing SUD treatment allow my child to return home?). By improving the ability to understand the relationship between child and family outcomes and parental SUDs, the resulting dataset will enable researchers to answer a multitude of questions relevant for SUD-focused patient-centered outcomes research and health equity research.

The CCOULD dataset will enable research on SUD and SDOH related to child welfare services. For example:

1. Receipt of OUD treatment and child welfare outcomes
 - How does participation in OUD treatment impact child outcomes (e.g., return to home, re-abuse rates)?
 - Which child welfare services mediate the effectiveness of SUD treatment?
2. Equitable treatment access
 - Who receives which services?
 - Where are there gaps in access to services?
 - How does treatment trajectory differ among sub-populations?
 - How does the intensity of treatment or services differ among sub-populations?

The overall objectives of the project are to:

- Develop datasets that link state Medicaid records of parents with their child's record from child welfare systems. These datasets will contain linked patient-level data including Medicaid enrollment, patient diagnoses, services, and claims data, along with child welfare outcomes.
- Prepare harmonized multi-state, de-identified research use datasets and archive the data where it may be accessed by external researchers.
- Document and publish findings from this project as case studies and overall lessons learned, as a public-facing document for other agencies interested in data sharing for research purposes and overcoming related challenges.
- Support these sites in creating sustainability in this linkage between Medicaid and child welfare data systems, including integrating Medicaid eligibility, enrollment, and claims data with the new CCWIS.
- Design, conduct, and encourage analyses on the linked datasets.

By linking child welfare records and parental Medicaid data, this project aligns with **Goal 2: Longitudinal Data Standards and Linkages** in HHS' Strategic Plan for OS-PCORTF. Through efforts to utilize data on child welfare involvement and services to assess outcomes for parents with SUDs, this project addresses **Goal 4: Person-Centeredness, Inclusion, and Equity** as well.



Coordination with Other Federal Agencies

This is a collaborative project between ASPE's Office of Human Services and Policy and the Administration for Children and Families (ACF) Office of Planning Research and Evaluation (OPRE). The project team is comprised of ASPE and OPRE personnel and has worked closely with OPRE to leverage expertise on child welfare data infrastructure, regulations, and programs. The project engages several HHS agencies as part of their Technical Expert Panel, including CMS, the Substance Abuse and Mental Health Services Administration (SAMHSA), and the Health Resources and Services Administration (HRSA).



Accomplishments

Since work on the project began in June 2019, the project team has made progress on several tasks.

- The project team selected two states (Kentucky and Florida) to participate in the project. They provide routine, in-depth technical assistance, and support to states to navigate the legal, policy, and privacy barriers to linking this data. The project team will use lessons learned from state and research partners to develop a roadmap, which will outline each state's experience as use cases that other states can leverage to implement data sharing and linking processes. The roadmap will contain technical information on linking Transformed Medicaid Statistical Information System (T-MSIS) and CCWIS data, considerations for implementing data linkages, data sharing and governance guidance, applicable research opportunities, and funding opportunities to support sustainability.
- As part of their technical assistance to states, the ASPE/ACF team provided specific guidance on 42 CFR Part 2 and HIPAA compliance to address challenges concerning privacy regulations and HIPAA compliance. States' challenges and strategies employed will be included in the CCOULD roadmap.
- The ASPE/ACF team built a common data model (CDM) to standardize data from multiple states to create a single research use file. This CDM was created in collaboration with states and experts in child welfare and Medicaid data.
- Linking methodology plans for each state have been finalized. Kentucky is nearing completion in data linking activities and mapping to the CDM. Research partners in Florida have obtained Medicaid data and are currently in the process of gaining access to child welfare records. Ultimately, both states will provide a linked dataset, accompanied by data documentation and a codebook.
- The ASPE/ACF team also finalized a data quality assessment plan and began testing data quality to ensure accuracy of the linking process in 2021.
- The ASPE/ACF team has secured a data archiving agreement with the National Data Archive on Child Abuse and Neglect (NDACAN). ASPE/ACF produced a data archiving plan that includes steps to provide access to researchers while maintaining strong statistical disclosure controls and other relevant security requirements related to archiving data.
- ASPE/ACF hosted a data linking summit for both states and research partners involved with project. States shared progress, asked questions, and discussed important issues related to the project. Multiple experts presented and were able to provide targeted technical assistance to project partners.

- To demonstrate the research applications of the linked dataset, the ASPE/ACF team has developed several analytic questions. The ASPE/ACF team is currently developing analysis plans for each of these questions and are documenting how certain variables can or cannot be used for different analyses.
- ASPE/ACF has worked with states throughout the project to mitigate barriers to sustainability and has fostered relationships between state agencies to ensure ongoing operation. In 2021, the ASPE/ACF team was awarded funding through the Office of the Assistant Secretary for Financial Resources (ASFR) Wedge Fund. This additional funding will help sustain the data linkage systems built under CCOULD, allowing the team to expand CCOULD in as many as 5 additional states.



Communication and Dissemination Products

- [**Child and Caregiver Outcomes Using Linked Data Website**](#): This webpage provides a summary of the project.¹³⁵
- [**Child and Caregiver Outcomes Using Linked Data Informational Sheet**](#): This project overview summarizes the project and will be used in outreach efforts to expand knowledge about the project.¹³⁶

Multistate EMS and Medicaid Dataset (MEMD): A Linked Dataset for Patient-Centered Outcomes Research

Period of Performance

3/29/21-9/29/23

Federal Point of Contact

Daniel Schwartz

There is a dearth of data on treatments and outcomes for patients who receive emergency medical services (EMS) that prevents analyses on the quality and effectiveness of the care they receive. Health insurance claims data generally exclude a portion of EMS encounters because some EMS agencies lack the infrastructure to bill insurance for their services. Furthermore, many state Medicaid programs do not reimburse EMS providers unless they transport a patient to a high-acuity health care facility (e.g., an emergency department), when a substantial number of these patients benefit from care at alternative settings (e.g., urgent care center, primary care office, or behavioral health clinic).

In the absence of sufficient data, researchers have limited ability to assess health care services delivered by EMS providers and the policies that affect service provision. The need for supportive data infrastructure is especially relevant amidst the COVID-19 pandemic. EMS providers have treated patients in place or transported them to alternative destinations at higher rates than before the pandemic, and some states have temporarily waived transportation-related coverage restrictions in Medicaid. It is crucial to analyze the outcomes of these patients for states considering whether to make these waivers permanent and for EMS agencies who may modify clinical protocols.

Project Purpose and Goals

This project will help fill the evidence gaps by linking EMS electronic patient care reports (ePCRs) and Medicaid claims data from up to five states in the calendar years 2018 to 2020. This linkage will create a Multistate EMS and Medicaid Dataset (MEMD) that enables longitudinal research on the effectiveness of EMS clinical interventions and the outcomes of Medicaid beneficiaries who are treated in place or transported to alternative settings. This project will demonstrate the conduct of such research, using the use case of behavioral health (i.e., mental health and substance use) emergencies; creating technical assistance materials that facilitate future linkages; and developing sustainability plans.

The purpose of this project is to enhance data infrastructure for research on the health outcomes of individuals who receive EMS by:

- Creating new data capabilities for federal and external researchers to longitudinally analyze the relationship between EMS provision, including treatment and transportation, and health outcomes among Medicaid beneficiaries.
- Conducting research using MEMD to provide evidence to inform practices and policies that promote high-quality care and prioritize patient needs, preferences, and outcomes in EMS delivery.
- Developing information resources to facilitate the continuation and expansion of efforts to link EMS and Medicaid data.

By creating new capabilities to conduct patient-centered outcomes research on Medicaid beneficiaries who engage with EMS, the aims of MEMD align with HHS' Strategic Plan for OS-PCORTF **Goal 1: Data Capacity for National Health Priorities**.

Coordination with Other Federal Agencies

The production of MEMD and associated deliverables will leverage expertise from ASPE, the Department of Transportation's (DOT) National Highway Traffic Safety Administration (NHTSA), and the Centers for Medicare and Medicaid Services (CMS).

Accomplishments

Since the launch of the project in 2021, ASPE has awarded a contract to National EMS Information System Technical Assistance Center (NEMSIS TAC) at the University of Utah. The NEMSIS TAC has begun identifying relevant EMS and Medicaid data elements and analyzing their quality and availability across states, which will ultimately inform the recommended candidate states.

Example research questions that can be answered using the Multistate EMS and Medicaid Dataset include:

- How do the number and type of related health care events (experienced and avoided) during the COVID-19 pandemic compare to the pre-pandemic period for patients who are administered naloxone by EMS?
- Which alternative destinations are associated with better outcomes for individuals who experience a behavioral health emergency?
- Has the increased proportion of treatment in place for behavioral health emergencies during the COVID-19 pandemic resulted in changes to rates of follow-up care?

Future deliverables for this project will include:

Data and Technical Products

- Developing a publicly available dataset containing linked Medicaid and EMS records from up to five states, the Multistate EMS and Medicaid Dataset (MEMD).
- Creating technical assistance materials to enable states to link their own EMS and Medicaid data

Communication and Dissemination Products

- Conducting and disseminating research using MEMD to inform clinical protocols, insurance coverage policies, and scope of practice laws for EMS.

X) Office of the National Coordinator for Health Information Technology (ONC)

ONC is administering three active projects including one cross-agency-funded projects described later in Section XI.

- A Synthetic Health Data Generation Engine to Accelerate Patient-Centered Outcomes Research
- Training Data for Machine Learning to Enhance Patient-Centered Outcomes Research Data Infrastructure*
- Using Machine Learning Techniques to Enable Health Information Exchange Data Sharing to Support COVID-19-focused PCOR

* Denotes a cross-agency-funded project that involves more than one federal agency; these projects are described in the Cross-Agency-Funded Projects section (Section XI).

A Synthetic Health Data Generation Engine to Accelerate Patient-Centered Outcomes Research

Period of Performance	Federal Point of Contact
4/1/19 - 4/30/22	Stephanie Garcia

Access to high-quality health care data is often difficult because of cost, patient privacy, or other legal and intellectual property restrictions. Even when using anonymized patient data, the risk of re-identification is impossible to eliminate completely, especially for patients with rare conditions. Further, due to a variety of interoperability issues, it is often difficult to bring data together from different sources to test theories, data models, algorithms, or prototype innovations. To protect patient privacy and overcome access and interoperability issues, synthetic data can serve as an alternative data source to initiate, refine, or test innovative research approaches more quickly.

Synthea™ is an open-source, free, and publicly available software program that creates high-quality, clinically realistic, synthetic patient health records in large volumes and in formats compatible with a variety of technologies, such as FHIR® and Consolidated-Clinical Document Architecture (C-CDA). A synthetic data engine is an important piece of the broader patient-centered outcomes research data infrastructure. Synthetic data engines offer a low-risk, readily available synthetic data source that complements researchers' use of real clinical data and enhances their ability to generate findings from rigorous analyses that can inform health care and treatment decisions.

Project Purpose and Goals

This project utilizes the Synthea™ software program to increase the amount and type of realistic, synthetic data available to researchers. These efforts present a novel and innovative data solution to the need for research-quality data. The ONC-led project team selected five clinical use cases with input from a multidisciplinary panel of experts, focusing on conditions involving complex care and where data can be limited: cerebral palsy, opioids for chronic pain and treatment of opioid use disorder, sepsis, spina bifida, and acute myeloid leukemia. The team will create clinical disease modules from a combination of clinical guidelines and care protocols, publicly available incidence and prevalence statistics, medical coding dictionaries, and clinical expert feedback. Synthea™ can then use these modules to generate individual

synthetic patient health records for use in patient-centered outcomes research. Potential applications of the synthetic data include simulating care interventions and analyzing longitudinal patient progress within each condition. The ONC team also aims to engage the broader community of researchers and developers to validate and test the utility of the synthetic health records for research.

The project is pursuing the following objectives:

- Enhance Synthea™ by developing or updating five data generation modules for the selected use cases to increase the number and variety of synthetic patient health records.
- Administer a prize competition (“challenge”) to encourage researchers and developers to validate the realism of the synthetic health records.
- Support awareness and use of Synthea™ including its updated modules, module builder, and the generated synthetic data through various dissemination mechanisms.

Through innovations that increase the availability and demonstrate the use of synthetic data, these objectives align with HHS’ Strategic Plan for OS-PCORTF **Goal 3: Technology Solutions to Advance Research.**

Accomplishments

Since the project launch in Spring 2019, the ONC team has progressed in several activities in support of project objectives:

- The ONC team was supported by a Technical Expert Panel (TEP) throughout the course of this project; from September 2019 to November 2021. The TEP was comprised of a group of diverse stakeholders representing viewpoints external to the federal government to provide their respective insights and subject matter expertise.
- After selecting the five use cases with the assistance of the TEP and other stakeholders, the ONC team developed the Synthea™ clinical modules for cerebral palsy, prescribing opioids for chronic pain and treatment of opioid use disorder, sepsis, spina bifida, and acute myeloid leukemia. The team has published four of the five modules and their associated resources online (see Data and Technical Products).
- ONC hosted the [Synthetic Data Validation Challenge](#) to engage the broader research community in the project work. Through the data challenge, ONC invited innovators, providers, researchers, and technology developers to develop solutions that enhance Synthea™ or demonstrate novel uses of Synthea™-generated synthetic health data. Phase I of the challenge required applicants to submit proposals for innovative models. Selected proposals moved on to Phase II, where they developed prototypes proposed in Phase I. Six winners were selected and a total of \$100,000 was awarded; winning solutions are posted to www.healthIT.gov and www.Challenge.gov.

Simulations are used in health care research when desired empirical data are not available. To evaluate the utility of Synthea™ for use in simulation studies, the ONC team conducted a demonstration study using data from the Acute Myeloid Leukemia Synthea™ module. A publication detailing project findings is nearing completion and will be submitted to a peer-reviewed scientific journal.



Data and Technical Products

- Synthea™ Disease-specific Modules: The projects modules for [Cerebral Palsy](#)¹³⁷, [Prescribing Opioids for Chronic Pain and Treatment of Opioid Use Disorder](#),¹³⁸ [Sepsis](#),¹³⁹ and [Spina Bifida](#)¹⁴⁰ are available on the Synthea™ Module Builder. Accompanying Companion Guides are available for all modules: [Cerebral Palsy](#),¹⁴¹ [Prescribing Opioids for Chronic Pain and Treatment of Opioid Use Disorder](#),¹⁴² [Sepsis](#),¹⁴³ and [Spina Bifida](#).¹⁴⁴ These Companion Guides provide summaries of each module's scope and intent, as well as references and data sources used to build the modules.



Communication and Dissemination Products

- Project Website:** A public-facing website was created on ONC's HealthIT.gov. This webpage summarizes ONC project activities and goals and includes links to resources, including project products, frequently asked questions, and Synthea™ technical guidance and tips.¹⁴⁵
- Synthetic Data Validation Challenge Informational Webinars: In 2021, ONC conducted two public informational webinars for the Synthetic Data Validation Challenge to introduce participants to Synthea™. Slides from the Phase I webinar is publicly available [here](#). A recording of the webinar that ONC conducted to announce Phase II winners is viewable [here](#).

Using Machine Learning Techniques to Enable Health Information Exchange to Support COVID-19-Focused Patient-Centered Outcomes Research

Period of Performance

5/21/21-5/21/24

Federal Point of Contact

Adam Wong

The COVID-19 pandemic has exposed many of the weaknesses in the health care data infrastructure for PCOR researchers, especially in data quality and fitness for research. Challenges include incomplete data, issues with data timeliness, missing demographic information such as race and ethnicity, and difficulties in aggregating data from different sources. Health information exchanges (HIEs), which receive EHR data from more than 60% of US hospitals, are a source of large-scale patient-level electronic clinical data that is currently underutilized.

HIEs receive data from health care providers across a specific local or state geography that vary in size. They are designed to allow clinicians and other health care providers to securely access and share patient health information to coordinate care, improve health care quality, and reduce costs. Many HIEs have the capability to retrieve, link, analyze, and aggregate data from clinical settings and laboratories. Additionally, some HIEs currently have access to vaccination data or are expanding their ability to collect vaccination data.¹⁴⁷

Using Machine Learning Techniques to Enable Health Information Exchange to Support COVID-19-Focused Patient-Centered Outcomes Research creates a foundation to use electronic health data from health information exchanges (HIEs) for research by addressing technical and privacy barriers with data standardization, application programming interfaces (APIs), and privacy-preserving machine learning (ML) techniques.¹⁴⁶

Despite the availability of robust patient-level electronic health data in state and regional HIEs, these datasets are rarely used for research purposes because of technical and privacy related barriers. Furthermore, the lack of data standardization across HIEs can result in data that is less useful for secondary research and application of advanced technologies such as artificial intelligence and machine learning.

Project Purpose and Goals

This project strengthens and modernizes the nation's data infrastructure by creating a foundation to use electronic health data from health information exchanges (HIEs) for research. It will result in three HIEs with expanded data capacity and interoperability. Each participating HIE will adopt two national data standards - the United States Core Data for Interoperability (USCDI) and the HL7® Bulk FHIR® API - to efficiently access large amounts of electronic health data. A privacy preserving machine learning technique called split learning will be used to test the suitability of HIE data for research, such as studying COVID-19. Valuable insights and lessons learned from this project will be disseminated as tailored resources for researchers and stakeholders that can apply these methods at other HIEs.

The project will achieve these goals by completing the following activities:

- Upgrading HIE infrastructure to support patient-centered outcomes research by
 - Implementing USCDI, a nationally recognized data standard for interoperable health information exchange.
 - Implementing Bulk FHIR® APIs to facilitate efficient access data from health systems and providers.
- Testing the use of split learning, a type of machine learning, to facilitate privacy-preserving data sharing to conduct patient-centered outcomes research and COVID-19 related analysis using HIE data.
- Disseminating resources and lessons learned to support the adoption of data standards, technology, and methods used in this project among HIEs and to encourage PCOR researchers to understand and explore HIE data as a source for research.

By building data capacity for patient-centered outcomes research that informs the greatest evidentiary needs of federal health programs and the people served by these programs, **Using Machine Learning Techniques to Enable Health Information Exchange to Support COVID-19 PCOR** aligns with **Goal 1: Data Capacity for National Health Priorities** of HHS' Strategic Plan for OS-PCORTF.

Accomplishments

Since the project kicked off in May 2021, the project has completed an internal preliminary research report in summer of 2021. This report will serve as an operational guide as the team executes the technical work of the project, guiding the process of selecting HIEs to participate in implementing USCDI and Bulk FHIR® APIs to facilitate efficient access data from health systems and providers and testing the use of split learning).

Data and Technical Products

ONC plans to make the following technical resources publicly available:

- Open-source code for split learning model that will be posted to ONC's GitHub.
- Implementation guide on implementing FHIR®/ USCDI data standards to upgrade HIE infrastructure and applying split learning techniques to HIE data.

Communication and Dissemination Products

- To disseminate the results, the project team will be developing a journal manuscript, webinars, blog posts, and a publicly available final project report.

XI) Cross-Agency Funded Projects

There are 7 active cross-agency funded projects.

- Data Capacity for Patient-Centered Outcomes Research through Creation of an Electronic Care Plan for People with Multiple Chronic Conditions
- Data Capacity for Patient-Centered Outcomes Research through Creation of an Electronic Care Plan for People with Multiple Chronic Conditions 2.0: Development of the Patient-facing Application
- Enhancing Data Resources for Researching Patterns of Mortality in PCOR: Projects 1 and 2
- Making Medicaid Data More Accessible Through Common Data Models and FHIR APIs
- Training Data for Machine Learning to Enhance Patient-Centered Outcomes Research Data Infrastructure
- Understanding Long-term Outcomes in COVID-19 Survivors with Multiple Chronic Conditions through e-Care Plan Development
- Validating and Expanding Claims-based Algorithms of Frailty and Functional Disability for Value-based Care and Payment

Data Capacity for Patient-Centered Outcomes Research through Creation of an Electronic Care Plan for People with Multiple Chronic Conditions

AHRQ Period of Performance

4/15/19 – 5/15/23

AHRQ Federal Points of Contact

Arlene Bierman
Janey Hsiao
Steve Bernstein

NIH/NIDDK Period of Performance

4/8/19 – 9/30/23

NIH/NIDDK Federal Points of Contact

Jenna Norton
Kevin Abbott

Interest in pragmatic research is growing as recognition increases that traditional randomized trials may not apply to real-world situations. Pragmatic trials represent a cost-effective and efficient research approach in which point-of-care (e.g., EHR) data informs understanding of real-world effectiveness of health interventions. However, lack of interoperability and exchange of data across EHRs creates barriers to pragmatic research; essential data on patient-centered outcomes are frequently missing, inconsistent, or difficult to compile across settings and conditions. Data aggregation is particularly important and challenging for people with multiple chronic conditions (MCCs). These individuals have complex health needs handled by diverse providers across multiple settings of care. Research is needed to optimize care for these complex patients, yet comprehensive data related to MCCs, care, and outcomes in real-world settings are largely unavailable.

In addition to data aggregation, care plans for patients with MCCs are an essential part of care coordination interventions for improving outcomes and disease management. Care plans are largely paper-based, lack standardization across settings, and tend to focus on a single disease or care setting. Developing and using electronic care plans based on structured data can enable electronic systems to aggregate and share data dynamically and automatically. In addition to providing information that allows

patients with MCCs and their providers to identify and achieve health goals, aggregated data from care plans can improve the quality of point-of-care data used in pragmatic research.¹⁴⁸

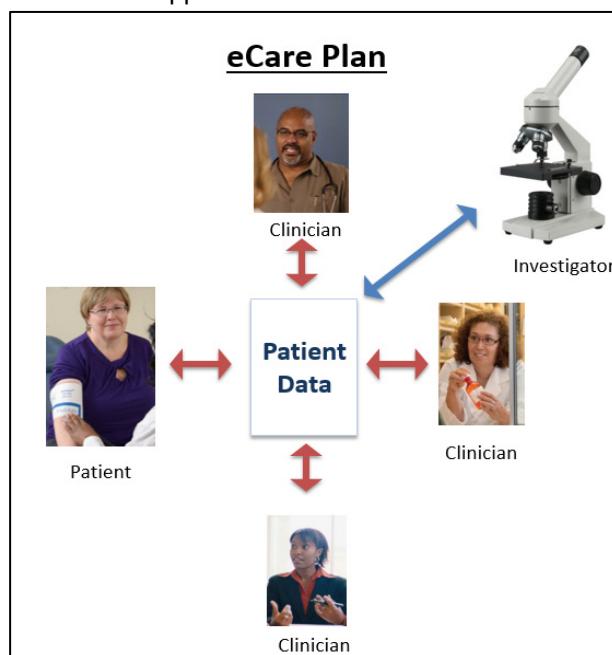
F Project Purpose and Goals

This project will build data capacity to conduct pragmatic patient-centered outcomes research by developing an interoperable electronic care (eCare) plan. The eCare plan will aggregate and share critical patient-centered data across home-, community-, clinic-, and research-based settings by extracting EHR data and exchanging that data across settings (Exhibit 7). The eCare plan will be an overarching, longitudinal blueprint of the prioritized health concerns, goals, interventions, and health status of an individual patient across care settings and care team members. The eCare plan app developed for this project will be designed for clinicians to use with patients who have chronic kidney disease, cardiovascular disease, diabetes, chronic pain, and/or OUD.

The overall objectives of the project are to:

- Expand an existing data element and standards set focused on chronic kidney disease to the following comorbid health conditions: cardiovascular disease, chronic pain, OUD, and diabetes.
- Develop an open-source, clinician-facing SMART on FHIR® eCare plan app for people with MCCs, as well as an accompanying HL7® FHIR® implementation guide, and pilot the app and implementation guide in patient populations with chronic kidney disease.
- Establish an eCare plan repository and app development collaborative Confluence website and a listserv to allow sharing of information and ideas about the project's development, testing, piloting, and implementation efforts, and provide an open-source repository on GitHub and available from AHRQ's CDS Connect repository website to store, enable search, access, sharing, and exchange of eCare plans.
- Disseminate all project products through free, open-source channels (e.g., federal government websites, open-source software exchanges such as GitHub).

Exhibit 7. Who Uses Patient Data from the MCC eCare Plan App



By increasing the exchange of eCare plans through the refinement of standardized data elements and the development of a SMART on FHIR® application, these objectives address HHS' Strategic Plan for OS-PCORTF **Goal 2: Longitudinal Data Standards and Linkages**.



Coordination with Other Federal Agencies

The project has convened a federal Stakeholder panel including diverse federal agencies and organizations with interest in the eCare plan to acquire additional input and guidance on the project. Members of this panel include Agency for Community Living (ACL), ACF, CDC, CMS, FDA, HRSA, Indian Health Service (IHS), numerous NIH Institutes and Centers, ONC, Patient-Centered Outcomes Research Institute (PCORI®), and the Veteran's Health Administration (VHA).



NIH/National Institute of Diabetes and Digestive and Kidney Diseases (NIDDK)

Accomplishments

Since the project began in April 2019, the NIH/NIDDK project team has made notable progress on several tasks.

- Between October 2019 and September 2021, the technical expert panel (TEP) has met monthly to identify and prioritize data elements for cardiovascular disease, chronic pain with OUD, and diabetes, and to develop value sets for use case conditions.
- The NIH/NIDDK team identified 272 chronic disease kidney data elements for incorporation in the eCare Plan app and implementation guide, using the Chronic Kidney Disease Data Element and Standards Set (CKD DESS). Using these data elements, the team developed 164 value sets, 175 clinical information models (CIMs) and appropriate FHIR® profiles focused on chronic kidney disease.
- The project team tested both the chronic kidney disease-specific, clinician-facing eCare Plan app and implementation guide during the September 2020 and January 2021 HL7® Connectathons and made revisions based on feedback. The first draft of the clinician-facing app and implementation guide are now being piloted for real-world implementation and testing at six Oregon Health & Science University (OHSU) clinical sites. The AHRQ team is leading the pilot testing efforts of the app and implementation guide, which is currently underway. See “AHRQ Accomplishments” for additional details.
- The NIH/NIDDK team identified more than 800 additional data elements and developed associated value sets for cardiovascular disease, chronic pain, pain associated OUD, and type 2 diabetes to expand on the CKD DESS. The data elements and value sets include approximately 70 data elements related to social determinants of health. These value sets are being aligned with FHIR® profiles for incorporation into revised versions of the eCare Plan application and implementation guide, which will be completed along with revisions based on lessons learned from pilot testing. The NIH/NIDDK team awarded a contract to a new subcontractor to expand and update the eCare Plan application and implementation guide based on pilot test findings.

MCC eCare Plan Application

- The project team has created a standardized eCare Plan that clinicians can use to collect standardized data from patients with MCCs. The code for the [MCC eCare Plan Provider-Facing App](#) is accessible online through GitHub.
- The accompanying [MCC eCare Plan Draft Implementation Guide](#) defines FHIR® profiles, structures, extensions, transactions and value sets needed to represent, query for, and exchange care plan information using clinical information models (CIMs).
- The project team created an initial version that focuses on chronic kidney disease and subsequently has drafted an expanded version that includes cardiovascular disease, pain with OUD, and diabetes.

Data and Technical Products

- **MCC eCare Plan Draft Implementation Guide**: The first draft of the HL7® FHIR® implementation guide for the clinician-facing eCare Plan software app defines how to represent coded EHR content that supports care planning activities for patients with multiple chronic conditions.¹⁴⁹
- **MCC eCare Plan Provider-Facing App Software Code**: The code for the first draft of the clinician-facing eCare Plan app is accessible online through GitHub.¹⁵⁰
- **Comprehensive MCC DESS Draft**: This spreadsheet documents data elements identified by the NIH/NIDDK team for chronic kidney disease, pain, type 2 diabetes, and cardiovascular disease. The project team routinely updates this document based on feedback.
- **MCC Value Sets**: Value sets included in the MCC DESS developed by the NIH/NIDDK team are housed on NLM Value Set Authority Center. The value sets cover chronic kidney disease, cardiovascular disease, diabetes, chronic pain, and/or OUD and can be used to develop FHIR® profiles.

AHRQ Accomplishments

The AHRQ project team continues to make significant progress to improve coordinated care for chronically ill patients.

- AHRQ awarded a contract to RTI to pilot and evaluate the eCare Plan app and implementation guide.
- The RTI/Oregon Health and Sciences University (OHSU) team completed three rounds of stakeholder work group meetings with patients/caregivers, clinicians, IT staff/vendors, and providers/leaders for the clinician-facing eCare Plan app and delivered a summary document with findings on workflow and usability. The findings were shared with the eCare Plan app developer.
- In preparation for pilot implementation, the RTI team worked with OHSU sites and the NIH/NIDDK team to configure and refine the eCare Plan app and implementation guide for use in real-world EHR environments. The RTI/OHSU team is currently working on documenting the team's approach to aligning the eCare Plan app and Implementation Guide with sites' EHR environments in a data element gap analysis report.
- The RTI/OHSU team is conducting two rounds of usability testing to assess the clinician-facing app functionalities and workflows between May 2021 and February 2022 with 9 providers. The clinician-facing app was moved into production (i.e., implementation and testing in a real-world setting) at six OHSU sites. The team is currently collecting the pilot data, which will inform a report that describes app implementation, opportunities for improvement, and recommendations for future activities.
- The RTI/OHSU team has updated training materials for the clinician-facing eCare Plan app to accommodate findings from usability testing.
- The AHRQ team maintains a public-facing Confluence project webpage, as part of its responsibilities to disseminate project materials. The Confluence page is regularly updated and maintained in partnership with the NIH/NIDDK team.



Data and Technical Products

- **Documentation of App and Implementation Guide Testing:** The AHRQ Confluence website houses materials, resources, and documentation for the MCC eCare Plan testing activities at HL7® Connectathons.¹⁵¹



Communication and Dissemination Products

- **eCare Plan Confluence Website:** This collaboration website was developed to allow for management of tasks, sharing of documents, and group discussions. The website also provides additional project information, resources and materials, timelines, and opportunities to get involved. The website is maintained by the AHRQ team.¹⁵²

Data Capacity for Patient-Centered Outcomes Research through Creation of an Electronic Care Plan for People with Multiple Chronic Conditions 2.0: Development of the Patient-facing Application

AHRQ Period of Performance

5/15/20 – 5/14/23

NIH/NIDDK Period of Performance

4/23/20 – 4/22/23

AHRQ Federal Points of Contact

Arlene Bierman
Janey Hsiao

NIH/NIDDK Federal Points of Contact

Jenna Norton
Kevin Abbott

The joint AHRQ and NIH/NIDDK project will build on AHRQ's and NIH/NIDDK's ongoing OS-PCORTF project, [Data Capacity for Patient-Centered Outcomes Research through Creation of an Electronic Care Plan for People with Multiple Chronic Conditions](#), which was awarded in 2019. The first eCare plan project develops an interoperable eCare plan to facilitate aggregation and sharing of critical patient-centered data across home-, community-, clinic- and research-based settings. The eCare 1.0 project is developing a *clinician-facing* eCare plan software app and implementation guide for use in research and the clinical care of people living with MCCs. The eCare plan 2.0 initiative will develop and pilot test a *patient-facing* eCare plan app for use with patients who have chronic kidney disease, cardiovascular disease, diabetes, and chronic pain with or without Opioid Use Disorder (OUD).

Importantly, the eCare plan is the first patient-centered eCare plan that crosses multiple sectors, enabling it to serve as an overarching, longitudinal blueprint of the prioritized health concerns, goals, interventions, and health status of individual patients. This ensures that the patient voice is heard across all care settings for patients with MCC.

The eCare plan 2.0 project will build upon the clinician-facing eCare plan to create an open-source patient-facing eCare plan app.

- The eCare plan will be available for use in MCC populations (including chronic kidney disease, cardiovascular disease, diabetes mellitus, and chronic pain, and pain-associated with OUD).
- It will integrate with the clinician-facing eCare plan app.
- Both apps and code will be made available through federal websites and other open-source channels (e.g., GitHub).

Project Purpose and Goals

The NIH/NIDDK portion of this project aims to:

- Develop a mobile patient-facing SMART on FHIR® eCare plan app.
- Expand the implementation guide to incorporate patient considerations.
- Expand the eCare implementation guide and patient-facing app to address pilot findings.
- Ballot the implementation guide through HL7® as standard for trial use (STU).

The AHRQ portion of this project focuses on:

- Implementing and evaluating the patient-facing app in clinical settings.
- Supporting workforce development in eCare planning.

This project's objectives to advance interoperability of electronic care management plans using standardized data elements and a SMART on FHIR® app addresses **Goal 2: Longitudinal Data Standards and Linkages** of HHS' Strategic Plan for OS-PCORTF. The project's focus on developing a patient-facing app to ensure the patient voice is included in care management aligns with **Goal 4: Person-Centeredness, Inclusion, and Equity** of HHS' Strategic Plan for OS-PCORTF.

NIH/NIDDK Accomplishments

- In collaboration with their Technical Expert Panels (TEPs), the NIH/NIDDK team created a panel of questions and answers to code, "[Challenges for treatment plan maintenance panel](#)." This panel is used to document the challenges faced by an individual patient that interfere with his or her ability to maintain an agreed upon treatment plan or health behavior.
- The NIH/NIDDK has completed development of the initial draft of the open-source patient-facing eCare plan app and HL7® FHIR® Implementation Guide for piloting in populations with chronic kidney disease. The draft materials have undergone multiple rounds of testing via HL7® FHIR® Connectathons and continue to be refined.
- NIH/NIDDK executed a contract for expanding the patient-facing application and Implementation Guide to include additional MCC use case (i.e., cardiovascular disease, diabetes, and chronic pain with or without OUD) data elements and value sets. Work to expand the patient-facing eCare Plan app and Implementation Guide began in late 2021. As part of this contract, the contractor will revise the patient-facing app materials based on findings from pilot testing with the chronic kidney disease use case conducted by AHRQ.

Data and Technical Products

- [**MCC eCare Plan Patient-Facing Application**](#): A draft version of open-source patient-facing eCare Plan app for use in chronic kidney disease populations that integrates with the clinician-facing app.¹⁵³

- **MCC eCare Plan Implementation Guide**: A draft version of eCare plan HL7® FHIR® implementation guide that incorporates both the clinician- and patient-facing eCare plan app to inform pilot site implementation.¹⁵⁴

AHRQ Accomplishments

- The RTI/OHSU team completed three rounds of stakeholder work group meetings with patients/caregivers, clinicians, IT staff/vendors, and providers/leaders for the patient-facing eCare Plan app and delivered a summary document with findings on workflow and usability. The findings were shared with the eCare Plan app developer.
- In preparation for pilot implementation, the RTI team worked with OHSU sites and the NIH/NIDDK team to configure and refine the eCare plan app and Implementation Guide for implementation in real-world EHR environments. The RTI team is currently working on documenting the team's approach to aligning the eCare plan app and Implementation Guide with sites' EHR environments in a data element gap analysis report.
- Testing of the patient-facing eCare Plan application began in October 2021. The testing focuses on usability and user acceptance with patients. The RTI team finalized a usability testing plan and protocol for chaperoned patient testing.
- The RTI team has updated training materials for the patient-facing eCare Plan app to accommodate relevant findings from usability testing of the clinician-facing eCare Plan app.

Enhancing Data Resources for Researching Patterns of Mortality in Patient-Centered Outcomes Research: Projects 1 & 2

CDC Project 1 Period of Performance
2/1/17-1/31/21

CDC Project 1 Federal Point of Contact
Carol DeFrances

FDA Project 2 Period of Performance
2/1/17-7/31/20

FDA Project 2 Federal Points of Contact
Greg Pappas
Robert Ball

An important objective of the OS-PCORTF is to build data capacity for patient-centered outcomes research in order to collect, link, and analyze data on outcomes and effectiveness from multiple sources. Efforts to better harmonize, connect, and enrich the federal mortality data through the two projects described below will accelerate its availability and utility for patient-centered outcomes research.

The projects described below are part of a group of four independently led projects with supportive components designed to enhance data resources for researching patterns of mortality. Two of the projects, one led by CMS (Project 3) and the other by CDC (Project 4), concluded in FY 2018 and are described in the [2018 OS-PCORTF Portfolio Report](#). The remaining projects, one led by the **CDC (Project 1)** and the other led by the FDA (Project 2), concluded in 2021. Together, these projects have built capabilities for systematic and more complete death information that is linkable, shareable across health systems, and more useful to PCOR researchers.

To accomplish these goals, the teams focused on developing linkages between the National Death Index (NDI) and other sources, including the National Hospital Care Survey (NHCS), the CMS Master Beneficiary Summary File (MBSF), and claims data (Project 1) and health plan data from commercially and publicly insured populations (Project 2).

The National Death Index (NDI) is the only central data source containing death information on both fact and cause of death (if using the NDI+ service) for all deaths occurring within the United States. It is used by researchers across medical and health studies for mortality ascertainment to assess the causes and risk factors of diseases and conditions and to evaluate the effectiveness of a wide range of interventions and drug therapies. The NHCS provides statistics on health and health care utilization based on hospital inpatient and emergency department visits, as well as personally identifiable information necessary for linkage. The MBSF contains data on all Medicare beneficiaries enrolled in or entitled to Medicare and includes data on costs and utilization. Finally, health plan data gathered from multiple sites providers further insight into patterns of care and patient outcomes among commercially and publicly insured populations.

- Project 1 (CDC) – Adding Cause-Specific Mortality to National Center for Health Statistics' National Hospital Care Survey (NHCS) by Linking to the NDI and CMS MBSF
- Project 2 (FDA) – Pilot Linkage of NDI+ to Commercially and Publicly Insured Populations
- Project 3 (CMS) – Enhancing Data Resources for Researching Patterns of Mortality in Patient-Centered Outcomes Research
- Project 4 (CDC) – NDI Workshop and Strategy Paper

Project Purpose and Goals

By linking multiple data sources into a single dataset in Project 1, researchers now have access to information on hospital and emergency department utilization, individual patient characteristics, and mortality outcomes (including fact and cause of death). Project 2 has extended this reach by developing linkage processes that enable multi-site outcomes research among health plans.

Project 1 (CDC) – Adding Cause-Specific Mortality to National Center for Health Statistics' National Hospital Care Survey (NHCS) by Linking to the NDI and CMS MBSF. This project leveraged data from the NHCS, NDI, and CMS MBSF to create new PCOR datasets that advances studies on mortality and post-acute care utilization following hospital care. This was accomplished by linking:

- the 2014 NHCS inpatient and emergency department claims data to the 2014 and 2015 NDI (to ensure at least 12 months of post-discharge follow-up for each hospital event).
- the 2014 NHCS inpatient and emergency department claims data to the 2014 and 2015 CMS MBSF.
- the 2016 NHCS inpatient and emergency department claims and EHR data to the 2016 and 2017 NDI.

Together, these linkages expand the capability of PCOR researchers to examine mortality; for example, comparing inpatient to emergency department discharge outcomes; 30-, 60-, 90-day post-acute hospital mortality for specific causes of death; and post-acute care utilization

Project 1 has made several novel contributions to PCOR data infrastructure:

- Provides the first-ever data linkage of EHR data from a national provider survey to the NDI
- Enables evaluations of EHR and claims data on their quality and complementarity
- Creates new approaches to optimize patient-level linkage by using the personally identifiable information available in claims and EHR data.

following emergency department visits and/or hospital inpatient stays for specific conditions and/or health care treatments and procedures.

By linking multiple data sources that represent patients across care settings, this project aligns with **Goal 2: Longitudinal Data Standards and Linkages** of the HHS' Strategic Plan for OS-PCORTF.

Project 2 (FDA) – Pilot Linkage of NDI+ to Commercially and Publicly Insured Populations.

Information on death and cause of death is often needed in patient-centered outcomes research, but administrative claims and EHR databases generally do not have complete information on fact, cause, and manner of death. This FDA project developed a standard, repeatable, and efficient process for linking a distributed data network of commercial and public health plans with the NDI+ (which includes cause of death data) with the goal of addressing the logistical challenges of data linkage across multiple health plans through its primary objectives:

- Develop standard, repeatable, and efficient technical solutions for linking the NDI's death and cause of death data to large, commercially, and publicly insured populations.
- Demonstrate the feasibility of linkage by using linked data to assess associations between select medications and death or cause of death as an outcome.
- The capability to link distributed data networks like Sentinel or the National Patient-Centered Clinical Research Network (PCORnet®) to the NDI+ enables many types of patient-centered outcomes research, including adverse event surveillance, predictive risk modeling, and comparative effectiveness research.

Project 2 answers the following key questions to support access to the NDI and enable more efficient and robust patient-centered outcomes research by linking mortality and administrative data.

- How to obtain administrative permissions across multiple sites to access NDI data.
- How to efficiently retrieve and link data from the NDI in a distributed manner without sharing PHI.
- How to analyze data from the NDI and provide aggregated results.

By linking multiple data sources that represent patients across care settings, this project aligns with **Goal 2: Longitudinal Data Standards and Linkages** of HHS' Strategic Plan for OS-PCORTF.

Project 1 (CDC) Accomplishments

This project concluded in 2021. The CDC team has completed the three data linkages enumerated in the project objectives: 1) the 2014 NHCS inpatient and emergency department claims to the 2014/2015 NDI; 2) the linkage of 2014 NHCS inpatient and emergency department claims to the 2014/2015 CMS MBSF; and 3) the linkage of the 2016 NHCS inpatient, emergency department claims, and EHR records to the 2016/2017 NDI. These linked data files and reports are currently available for research use. They have also produced numerous reports demonstrating the research applications of the linked datasets to clinical areas that include Alzheimer's disease, maternity mortality, respiratory illness, strokes, and opioids. Linkage of 2020 hospital data will also make this dataset an incredibly valuable resource in assessing COVID and mortality. Beyond the immediate use of this data, this project will continue to spawn follow-up projects and lay the groundwork for future patient-centered outcomes research.



Data and Technical Products

- **Final Report:** The project's final report summarizes the methodology used to link the NHCS to the NDI to obtain cause-specific mortality and is available on the OS-PCORTF website.¹⁵⁵
- The [**linkage of the 2014 NHCS inpatient and emergency department claims to the 2014/2015 NDI**](#) has been completed and the linked data files are currently available for research use.¹⁵⁶ Related materials include a [**codebook**](#) for the 2014 NHCS claims data linked to the 2014/2015 NDI file.¹⁵⁷ NCHS also published this [**report**](#) describing the methods used for linkage and the analytic considerations.¹⁵⁸
- The [**linkage of the 2014 NHCS inpatient and emergency department claims to the 2014/2015 CMS MBSF**](#) data files has been completed and the linked data files are currently available for research use.¹⁵⁹ NCHS published this [**report**](#) describing the methods used for linkage and analytic considerations for the 2014 NHCS linkage to the 2014/2015 CMS MBSF.¹⁶⁰
- The [**linkage of the 2016 NCHS inpatient and emergency department claims and EHR records to the 2016/2017 NDI**](#) has been completed and the linked data files are currently available for research use.¹⁶¹ NCHS published this [**report**](#) describing the methods used for linkage and the analytic considerations for linkage of the 2016 NHCS to the 2016/2017 NDI.¹⁶²



Communication and Dissemination Products

- [**National Hospital Care Survey Demonstration Projects: Characteristics of Inpatient and Emergency Department Encounters Among Patients with Any Listed Diagnosis of Alzheimer Disease**](#): This report demonstrating the use of NHCS data pertaining to inpatient discharges and emergency room encounters among patients with Alzheimer's disease was published in the National Health Statistics Reports.¹⁶³
- [**Respiratory Illness Emergency Department Visits in the National Hospital Care Survey and the National Hospital Ambulatory Medical Care Survey**](#): this [**report**](#) on respiratory illness emergency department visits in the NHCS and the National Hospital Ambulatory Medical Care Survey (NHAMCS) was published in the National Health Statistics Reports.¹⁶⁴
- [**Maternal Mortality in the United States: Changes in Coding, Publication, and Data Release, 2018**](#): this report describes changes in how the NCHS will code, publish, and release maternal mortality data and presents official 2018 maternal mortality estimates using a new coding method. It was published in the National Vital Statistics Reports.¹⁶⁵
- [**National Hospital Care Survey Demonstration Projects: Stroke Inpatient Hospitalizations**](#): this report on the NHCS demonstration projects to examine inpatient hospitalizations for stroke was published in the National Health Statistics Reports.¹⁶⁶
- [**Opioid-involved Emergency Department Visits in the National Hospital Care Survey and the National Hospital Ambulatory Medical Care Survey**](#): this report on opioid-involved emergency department visits in the NHCS and the NHAMCS was published in the National Health Statistics Reports.¹⁶⁷
- [**National Hospital Care Survey Demonstration Projects: Opioid-involved Emergency Department Visits, Hospitalizations, and Deaths**](#): This report on the NHCS demonstration projects related to opioid-involved emergency department visits, hospitalizations, and deaths was published in the National Health Statistics Reports.¹⁶⁸

Project 2 (FDA) Accomplishments

The FDA, in partnership with their Coordinating Center, the Harvard Pilgrim Healthcare Institute (HPHCI), achieved its goal of developing efficient processes for multi-site research centers to link their clinical data to the NDI+ to obtain death and cause of death data.ⁱⁱ The FDA team developed and demonstrated the feasibility of linkage processes in a cardiac use case designed to assess associations between select medications and mortality outcome.

Using the NDI+ dataset, the FDA team developed and tested distributed linkage and matching processes for data exchange between the NDI+ and six public and private health plans, a diverse group that included national insurers, regional health plans, and integrated delivery systems. Additionally, the team was able to develop a standardized and reusable process for efficiently obtaining and analyzing death and cause of death information from NDI+ across multiple health plan databases. Critically, they did so without sharing protected health information between health plans or with the Coordinating Center.

In addition to creating technical processes to enable multi-site data linkages to the NDI, the project team established successful administrative processes that are now publicly available to other researchers. For example, the project team developed and tested administrative workflows for utilizing a central Institutional Review Board (IRB), and applying and obtaining approval for NDI access, both of which are intended to help others minimize administrative burden.

In the final phase of the project, researchers used the linked health plan and NDI data to estimate the incidences of mortality and sudden cardiac death among two cohorts of antiarrhythmic medication users with differing underlying risks. This analysis was conducted to both validate the linkage process and demonstrate the use of the linked data to enhance the FDA's capability to assess post-market drug safety and effectiveness. Additional detail on the research use case and associated processes are captured in the project's final report (see below).

Data and Technical Products

The project team completed their final report, "Distributed Processes for Attaining Death and Cause of Death Information from the National Death Index (NDI) in Multi-Center Studies." It is available upon request by contacting OSPCORTF@hhs.gov. The report is a methods guide to the processes that the project team developed, including:

- Administrative and technical workflows
- Processes for requesting NDI data access
- Processes for NDI data use (e.g., records matching, data import, deduplication)
- A process for using a Central IRB to obtain multi-site research approval

ⁱⁱ The NDI provides death information including death date and death certificate number (referred to as NDI data), as well as cause of death from death certificates (referred to as NDI Plus or NDI+ data) upon request.



Communication and Dissemination Products

- [**Developing a Standardized and Reusable Method to Link Distributed Health Plan Databases to the National Death Index: Methods Development Study Protocol:**](#) This methods study protocol was recently published in the *Journal of Medical Internet Research* and describes the workflows for linking multisite health plan data and the NDI+, and use of a centralized IRB for multi-site research.¹⁶⁹

Making Medicaid Data More Accessible Through Common Data Models and FHIR APIs

FDA Period of Performance

4/1/21 – 12/31/24

FDA Federal Point of Contact

Sarah Dutcher

NIH/NLM Period of Performance

3/8/21 – 3/30/24

NIH/NLM Federal Point of Contact

James Mork

The movement toward more open-sourced standards throughout the health care infrastructure provides more opportunities to leverage standardized data formats, such as Common Data Models (CDMs). CDMs create a standard structure that helps enable data sharing across distributed research network participants by standardizing both the data structure and analytics layer. They are increasingly used by public health researchers, regulatory agencies, and others due to growing demand for rapid evidence generation using multiple databases.

In parallel, there has been increasing emphasis on the use of the Fast Healthcare Interoperability Resources (FHIR®) Application Programming Interfaces (API) designed to facilitate health data exchange and data use. The ability to link EHR data from large hospital systems, academic medical centers, or Federally Qualified Healthcare Centers to Medicaid through FHIR® APIs would help leverage EHR data for research and advance a data infrastructure to support the generation of patient-centered outcomes research.¹⁷⁰

Recognizing the role of CDMs and FHIR® APIs in enhancing data capacity for patient-centered outcomes research, the FDA and the NIH/NLM are working in tandem to 1) enhance Medicaid and CHIP data access by transforming T-MSIS data into two CDMs and 2) assess the feasibility of linking the Transformed Medicaid Statistical Information System (T-MSIS) data to electronic health record (EHR) data using FHIR® APIs.



Project Purpose and Goals

The purpose of this project is to increase accessibility to Medicaid and CHIP data by creating an open-source code to format T-MSIS data into two CDMs: Observational Medical Outcomes Partnership (OMOP) used by Observational Health Data Sciences and Informatics (OHDSI) and Sentinel used by the FDA. Both the FDA and NIH/NLM are creating the open-source documents and characterizing data quality based on a harmonized Data Quality Assessment Framework for electronic health care data, a framework developed in a previously funded OS-PCORTF project, **Standardization and Querying of Data Quality Metrics and Characteristics for Electronic Health Data**.¹⁷¹ These [data quality metrics](#) will be incorporated in characterizing the impact of transformation into two CDMs and assessing the completeness of electronic health data and fitness for use.

To demonstrate the value of CDM transformation, the open-source tools will be used to conduct studies of public health importance. The FDA is developing a mother-infant linkage to conduct these demonstration studies on questions related to maternal and fetal health. To inform future efforts to enrich Medicaid data with more detailed clinical information that can be used in descriptive and causal inference studies, this project will assess the feasibility of using FHIR® APIs to link EHR and T-MSIS claims data. Results will be disseminated and end-users will be trained on the open-source tools through a series of webinars hosted by this project.

The overall objectives of the project are to:

- Create an open-source code to format T-MSIS data files into the OMOP CDM by NIH/NLM and Sentinel CDM by the FDA.
- Develop data quality metrics to characterize the transformed Medicaid data and assess the completeness of these data and their fitness for use.
- Develop a mother-infant linkage within the T-MSIS data formatted into the Sentinel CDM and conduct demonstration studies on maternal health questions.
- Evaluate the feasibility of linking T-MSIS data with EHRs using FHIR® APIs.
- Train Medicaid researchers on the new data transformation tools and disseminate major findings through a webinar series.

The project will improve data access, accelerate analyses, and support multi-database studies thereby enhancing PCOR research efforts. By creating CDMs that facilitate clinical, intervention, and community data to data extraction and linkage, the aims of the **Making Medicaid Data More Accessible Through CDMs and FHIR APIs** project align with **Goal 1: Data Capacity for National Health Priorities** of HHS' Strategic Plan for OS-PCORTF.

FDA Accomplishments

Since the project started in mid-2021, the FDA team has started working towards the project goals by completing the following tasks:

- The FDA collaborated with NIH/NLM and ASPE to establish the technical expert panel (TEP), which consists of 5 individuals and guides both agencies.
- The team acquired T-MSIS data files through FDA's existing Data Use Agreement with CMS.
- The FDA Sentinel team collaborated with Duke Clinical Research Institute (DCRI) to conduct a review of the Medicaid Data Quality (DQ) Atlas and data source files. The review obtained information regarding the T-MSIS data characteristics to inform Sentinel CDM transformation code development.
- DCRI started the initial development of Sentinel CDM transformation specifications, including an outline of inclusion/exclusion criteria for the initial CHIP data based on review of the DQ Atlas and analytic data source files for the initial transformed T-MSIS data in the Sentinel CDM.

NIH/NLM Accomplishments

- The project team validated that Databricks—an ETL (“extract, transform, load”) platform that helps users extract from different sources, transform data into a usable format, and load into an accessible end user system¹⁷² —would be able to support the transformation of T-MSIS into OMOP CDM.
- The project team developed a mapping guide from T-MSIS data files to OMOP for a preliminary set of values.

To test the preliminary values, the project team developed a one state, one month Report and a one state, all years Report using a preliminary set of values.

Data and Technical Products

- **Preliminary ETL Code for Data Transformation:** A preliminary version of the ETL developed by NLM for transformation to the OMOP CDM has been uploaded to GitHub.

Training Data for Machine Learning to Enhance PCOR Data Infrastructure

NIH/NLM Period of Performance

8/1/19 – 2/28/22

NIH/NLM Federal Point of Contact

Stefan Jaeger

ONC Period of Performance

8/1/19 – 2/28/22

ONC Federal Point of Contact

Matthew Rahn

Artificial Intelligence (AI) and associated innovative technologies like machine learning have the power to consume large amounts of data to predict clinical outcomes. Machine learning techniques can be applied to accelerate the research process and the translation of findings to clinical settings. Machine learning techniques can also allow researchers to carry out complex studies that traditional analytical methods are unequipped to handle, combining previously unconnected data resources, extracting relevant data points, and elucidating complex issues. However, machine learning algorithms must be trained on large amounts of high-quality data to effectively “learn” how to identify patterns with little human intervention. High-quality training datasets must be well-labeled and structured, they must use common data models and common data elements, and they must be trained on representative data sets to avoid bias. The combination of machine algorithms and robust training datasets have the potential to hasten the discovery of novel disease-outcome correlations and associations and inform the design of prevention and treatment studies.

Project Purpose and Goals

This project will enhance capacity of using machine learning to improve patient-centered outcomes research by developing and disseminating methodological and training resources. This project will curate high-quality training datasets for two use cases: 1) drug resistance in patients infected with tuberculosis (TB), implemented by the NIH/NLM and the NIH National Institute of Allergy and Infectious Diseases (NIAID); and 2) kidney disease, led by ONC and the NIH/National Institute of Diabetes and Digestive and Kidney Disease (NIDDK). The datasets developed for each use case will be used to build and train AI

prediction models. The use cases were selected because despite the growing incidence of kidney disease and drug-resistance TB, there are limited evidence-based treatments for these conditions. Data and evidence generated from these datasets will support multiple federal initiatives in precision medicine for kidney and TB research programs. For example, these data can support individualized treatment based on a patient's health conditions, genetic/phenotypic profiles, and preferences. The project will also provide a blueprint for the use of AI in discovery and safety surveillance, including the potential uses and limitations of the technology.

This project will address the following objectives:

- Develop high-quality training datasets and generate lessons learned from implementing industry best practices; and compile insights on the data quantity and quality requirements for machine learning to be applied in patient-centered outcomes research.
- Develop machine learning algorithms that will be used to test the training datasets and validate approaches for evaluating their performance using conventional metrics.
- Develop an implementation guide detailing methods with parameters for each use case as well as sufficient detail to facilitate application to a wider array of use cases.
- Disseminate tools, training data, and lessons learned to stimulate the application of these methods to a wider array of use cases by researchers.

These activities will develop novel tools for applying machine learning to PCOR, aligning with **Goal 3: Technology Solutions to Advance Research** of HHS' Strategic Plan for OS-PCORTF.

NIH/NLM and NIH/NIAID Accomplishments

The NIH/NLM portion of the project, focused on TB, launched in August of 2019, and has made several achievements since award.

- To develop training data for the TB drug resistance use case, the NIH team investigated and identified 25 statistically significant features (two clinical and 23 radiological features) associated with drug-resistant TB cases. They also trained a decision tree model utilizing a sample of 2,622 drug-sensitive and drug-resistant patients to determine the most important features, among seven demographic features and 25 radiological features, for detecting drug-resistant TB.
- To obtain TB training data, NIH acquired access to approximately 6,000 patient records in NIH/NIAID's TB Portals. The downloadable data includes clinical data, radiographs, patient information, and links to genomic information.
- The project team successfully trained the machine learning models to distinguish between drug-resistant and drug-sensitive TB in chest X-rays using four classification problems. They are currently working on a fifth, and final, classification problem.
- Considering the COVID-19 crisis, the team identified several possibilities for how machine learning could help detect the disease in X-rays and CTs. They found that techniques similar to methods applied in this project could be applied to detect COVID-19 in radiographs and could increase the specificity of COVID-19 diagnosis. These efforts resulted in a manuscript on clinical and radiological manifestations of coronavirus pneumonia that was published in the 28th volume of *Journal of X-Ray Science and Technology* (see "Communication and Dissemination Products" for a link to the article).

- The team is currently working on the organization and documentation of machine learning software and the accompanying implementation guide, all of which will be made available at the conclusion of this project.

Data and Technical Products

- Training Data: The training data used in this project is available [online](#) and for [download](#) through NIH/NIAID's TB Portals. These training data can be used by others to develop models using machine learning.¹⁷³

Communication and Dissemination Products

- Journal publication titled "[**Clinical and Radiological Features of Novel Coronavirus Pneumonia**](#)" published in the *Journal of X-Ray Science and Technology*, June 2020. This paper describes the unique features of COVID-19, testing methods to strengthen clinical diagnoses, and the focus of future research efforts as the pandemic develops.¹⁷⁴
- Journal publication titled "[**Developing and Verifying Automatic Detection of Active Pulmonary Tuberculosis from Multi-slice Spiral CT Images Based on Deep Learning**](#)" published in the *Journal of X-Ray Science and Technology*, September 2020. This paper describes the successful development of an AI tool for automatic detection of active TB in chest CT imaging and the implications for use in clinical settings.¹⁷⁵
- Journal publication titled "[**Differentiating between Drug-Sensitive and Drug-Resistant Tuberculosis with Machine Learning for Clinical and Radiological Features**](#)" published in *Quantitative Imaging in Medicine and Surgery*, January 2022. This paper describes the identification and usage of readily available clinical data and data derived from chest X-rays in drug-resistant TB prediction and the application of machine learning (support vector machine) techniques to selection of clinical and radiological features to differentiate drug-resistant TB and drug-sensitive TB.¹⁷⁶
- Conference publication titled "[**Identifying Drug-Resistant Tuberculosis in Chest Radiographs: Evaluation of CNN Architectures and Training Strategies**](#)," presented at the Annual International Conference of the IEEE Engineering in Medicine and Biology Society (2021). This paper describes the application of deep learning (AI) to discriminate between drug-resistant and drug-sensitive TB on NIAID TB Portals data, achieving an Area Under the ROC Curve (AUC) of up to 85%.¹⁷⁷
- Workshop publication titled "Automated Drug-Resistant TB Screening: Importance of Demographic Features and Radiological Findings in Chest X-Rays", presented at the Applied Imagery Pattern Recognition (AIPR) workshop (2021). This paper describes the identification of demographic and radiological features that enable differentiation between drug-resistant and drug-sensitive TB cases using machine learning (random forest). (*anticipated*)
- Journal publication titled "[**Generalization Challenges in Drug-Resistant Tuberculosis Detection from Chest X-rays**](#)," published in *Diagnostics*, January 2022. This paper describes an evaluation of the generalizability of models across countries with respect to the task of differentiating between drug-resistant and drug-sensitive TB from chest X-rays. This includes a discussion of potential bias in the training data caused by image acquisition differences or other non-pathological features.¹⁷⁸

ONC and NIH/NIDDK Accomplishments

As of September 2021, ONC and NIH/NIDDK have completed their project activities on kidney disease. Major accomplishments include:

- The project team built high-quality training datasets and validated machine learning algorithms using data from the United States Renal Data System (USRDS) for their kidney disease use case: predicting mortality in the first 90 days following initiation of chronic dialysis in end-stage kidney disease patients. The training dataset includes 188 elements, including demographics, prior care, clinical variables, comorbidities, and patient education and contains over 1 million individual patients. The ONC team prepared two versions of the dataset using imputed and non-imputed values. The training datasets are now hosted by [USRDS](#), available by request.
- ONC hosted two webinars describing project accomplishments, one of which was public (see Communication and Dissemination Products). The second, internal webinar was presented for the Interagency Assembly quarterly meetings and featured interim project updates.

The project team submitted a manuscript using the kidney use case training data to the scientific journal, American Society of Nephrology.

Data and Technical Products

- **[Training Dataset Code](#)**: The code used to develop the training datasets is available on GitHub for use by others.¹⁷⁹
- **[Machine Learning Algorithm Code](#)**: The code used to develop the machine learning algorithm is available on GitHub for use by others.¹⁸⁰
- “Training Data for Machine Learning to Enhance PCOR” Implementation Guide: This three-part guide provides PCOR and other researchers methodology for, and lessons learned from, building high quality training datasets and ML models. The guide includes an [overview](#) of the project and use cases selected, [implementation guidance](#), and a downloadable [data dictionary](#) for the training datasets.¹⁸¹

Communication and Dissemination Products

- **[Project Website](#)**: This project webpage provides additional information about the project’s background and goals, as well as links to several resources described below.¹⁸²
- **[Project Infographic](#)**: This infographic explains project goals, objectives, and activities.¹⁸³
- **[Training Data for Machine Learning to Enhance Patient-Centered Outcomes Research \(PCOR\) Data Infrastructure Webinar](#)**: This webinar was presented by the ONC team as the project concluded to provide an overview of the project and activities and provide recommendations for future applications of AI and machine learning in patient-centered outcomes research and health services research.¹⁸⁴
- **[The Application of Machine Learning to Address Kidney Disease Blog Post](#)**: This ONC blog post describes how machine learning techniques were applied to address kidney disease through the ONC project.¹⁸⁵

- **Training Data for Machine Learning to Enhance Patient-Centered Outcomes Research (PCOR) Data Infrastructure Final Report:** This report provides a detailed description of project activities, methodology, findings, and recommendations.¹⁸⁶

Understanding Long-term Outcomes in COVID-19 Survivors with Multiple Chronic Conditions through e-Care Plan Development

AHRQ Period of Performance

5/1/21– 5/1/24

AHRQ Federal Point of Contact

Arlene Bierman
Janey Hsiao

NIH/NIDDK Period of Performance

4/1/21– 3/31/24

NIH/NIDDK Federal Point of Contact

Jenna Norton
Kevin Abbott

A growing number of people are manifesting persistent, debilitating symptoms after COVID-19 infection. Post-Acute Sequelae of SARS-CoV-2 infection (PASC) (or “long COVID”) can occur across multiple organ systems and is potentially heightened by underlying chronic conditions. PASC may coexist with behavioral and psychosocial consequences of the pandemic (e.g., depression, financial strain, substance abuse), which compounds negative impacts on health and quality of life.

Leveraging clinical data (e.g., from the EHR), patient generated health data (PGHD), and health information technology systems can enable more robust pragmatic patient-centered outcomes research that directly applies to “real-world” situations like care for individuals with PASC. However, aggregating data across clinical and research settings is often difficult due to lack of interoperability and standardization across the multiple sites and settings where care is delivered. In addition, these data often lack essential, person-centered data such as social determinants of health and person-reported outcomes. This creates barriers to conducting patient-centered outcomes research that considers the needs of people with multiple chronic conditions (MCC) who receive care across multiple providers and settings, an overwhelming majority of people.

Project Purpose and Goals

Building upon AHRQ’s and NIDDK’s ongoing OS-PCORTF Data Capacity for Patient-Centered Outcomes Research through Creation of an Electronic Care (e-Care) Plan for People with Multiple Chronic Conditions 1.0 and 2.0 projects, this project will expand clinician- and patient-facing eCare plan applications to support generation and use of person-centered data for individuals receiving care for COVID-19 and PASC. Availability of comprehensive patient-centered data will enable researchers to probe a range of questions related to the long-term health, social, and quality of life outcomes of COVID-19; PASC outcomes, as well as opportunities to address them; duration of immunity following vaccination; clinical use patterns and long-term performance of products marketed through Emergency Use Authorization, racial and socioeconomic disparities in COVID-19; strategies to improve patient and caregiver engagement and care coordination; and other topics.

The objectives for NIH/NIDDK are to:

- Identify data elements & standards for COVID-19 risks, natural history, and sequelae.
- Expand the HL7® [**FHIR® MCC e-care Plan Draft implementation guide \(IG\)**](#) to address COVID-19 data elements and caregiver app considerations which will help health care systems implement the data standards required to facilitate this data exchange.
- Expand patient and clinician software applications to include COVID-19/PASC data elements.
- Develop an application for use by unpaid caregivers.
- Revise e-care plan applications and IG to address pilot findings.

The objectives for AHRQ are to:

- Test and evaluate the new COVID-19 patient, provider, and caregiver apps as well as the IG in real world settings.
- Disseminate project resources (i.e., data standards/value sets, applications, implementation guide, findings from pilot testing) for use by stakeholders.

By creating e-care plans that facilitate clinical, intervention, and community data to data linkages, the aims of the **Understanding Long-term Outcomes in COVID-19 Survivors with MCC through e-Care Plan Development** project align with the HHS' Strategic Plan for OS-PCORTF **Goal 1: Data Capacity for National Health Priorities**. Additionally, this project's inclusion of individual-level SDOH data in linked datasets and emphasis on creating a more diverse sample will serve **Goal 4: Person-Centeredness, Inclusion, and Equity**.

NIH/NIDDK Accomplishments

Since the project's inception in 2021, the NIH/NIDDK team has made progress towards meeting project objectives by completing the following tasks:

- The project team awarded a contract to expand the HL7® FHIR® MCC e-Care Plan Draft IG to address COVID-19/PASC data elements and caregiver app considerations, expand patient and clinician software applications to include COVID-19/PASC data elements, and develop a software application for use by unpaid caregivers.
- The project formed a technical expert panel (TEP) that is meeting regularly throughout the project. The TEP will provide expert insights and guide both the NIH/NIDDK and AHRQ through their tasks.
- The project team began to review existing COVID-19/PASC relevant value sets, including critical data elements required for research and care of individuals with COVID-19/PASC, history of COVID-19, chronic kidney disease, chronic pain, cardiovascular disease and/or type 2 diabetes. The TEP also began vetting identified standards.
- The project team completed an evidence review of caregiver needs for an app to support their role in care planning.

- The project team awarded contracts to develop an open-source, EHR-agnostic clinician- and caregiver-facing SMART on FHIR® e-Care plan application for use in COVID-19 survivors/people with PASC and/or MCC populations.

AHRQ Accomplishments

The main body of AHRQ's work on this project will be completed once the NIH/NIDDK tasks have been completed; however, AHRQ has moved forward with several tasks in preparation for this work:

- The project team is working to establish a contract to test and evaluate the new COVID-19 patient, provider, and caregiver apps as well as the IG in real world settings; revise e-care plan applications and IG to address pilot findings; and disseminate project resources.
- The AHRQ team, in collaboration with NIH/NIDDK, has also conducted a landscape review of caregiver applications designed to share health data.

Communication and Dissemination Products

- [Project Confluence Site](#)**: Used for the three e-Care plan projects, this project Confluence site provides information on the project team, timelines, and key deliverables. Materials from the project such as the expanded HL7® FHIR® MCC e-Care Plan Draft IG will be made available on this site.

Validating and Expanding Claims-based Algorithms of Frailty and Functional Disability for Value-based Care and Payment

AHRQ Period of Performance

5/28/19 – 3/15/23

AHRQ Federal Point of Contact

Arlene Bierman

ASPE Period of Performance

5/28/19 – 4/28/23

ASPE Federal Point of Contact

LokWong Samson

CDC Period of Performance

6/24/19 – 6/23/23

CDC Federal Point of Contact

Violanda Grigorescu

Patient function, both physical and cognitive, are important outcomes assessed by patient-centered outcomes researchers. Older adults who are frail and persons with functional disabilities (e.g., vision impairment, deaf or hard of hearing, difficulties with mobility) are at increased risk for poor health outcomes.¹⁸⁷ This project focuses on the functional risk factors that contribute to physical and cognitive decline and impairment. These functional risk factors play an important role in risk-adjustment of research studies as well as for evaluating performance and determining payments in value-based care programs. ASPE has conducted preliminary research on four categories of functional risk factors: frailty, predictors of functional dependence, mental and behavioral health disorders, and existing indicators in the Chronic Conditions Warehouse (CCW) of other potentially disabling conditions. ASPE found that these indicators can improve the predictive power of risk-adjustment models of Medicare outcomes measures used in value-based care programs, and they explain poorer outcomes in Medicaid-Medicare dually enrolled

beneficiaries. Validating frailty risk indicators would ensure broader use and acceptability in payment programs and research studies.

Claims and EHR-based algorithms are two approaches that can be employed to confirm the diagnosis of condition on record, perform risk-adjustment in research studies or for clinical quality measures, and to potentially identify at-risk patients for specified conditions in clinical intervention or quality improvement programs. While claims-based algorithms have been traditionally used, EHR data are increasingly used for research studies and quality reporting. EHR data can be used to further validate claims-based algorithms for frailty risk indicators by comparing EHR- and claims-derived indicators.

Project Purpose and Goals

The project is led by ASPE, in collaboration with AHRQ and CDC. The purpose of the overall project is to build HHS data capacity for conducting robust patient outcomes research by refining and validating claims-based algorithms of frailty and functional disabilities using EHR data across patient populations and payers.

This project will modify claims-based algorithms to develop EHR-based algorithms for use in different payer populations, including public and commercial payers. The project will then use the claims-based algorithms that rely on diagnoses and further incorporate data fields available in EHR systems to create modified versions of the algorithms that can be used to mine EHR data. The EHR-based algorithms will include data elements such as the results of patient disability screening tools, data fields on functional status and disability where available, and potentially physician notes.

ASPE is partnering with CMS to update and validate the current list of potentially disabling conditions in the CCW and potentially add these indicators in both Medicare and the Transformed Medicaid Statistical Information System (T-MSIS) data. ASPE is partnering with AHRQ and the CDC to develop and test EHR-based algorithms that identify patients with frailty and functional disabilities across patient populations (i.e., Medicare, Medicaid, and dual-eligible beneficiaries). The algorithm development and testing process will begin with a close look at descriptive statistics, including the demographics available, and types of EHR data fields available or used to capture patient function or functional disabilities. This process will also involve a direct comparison of the claims-based algorithms to the EHR-based algorithms and in detecting disability-related fields, and iterative refinement.

Specifically, the objectives of this project are to develop:

- Build a set of validated and refined claims-based algorithms using Medicare and/or Medicaid claims data that identify patients' functional risk and frailty, to be made available to the public through the CCW.
- Draft a set of draft EHR-modified algorithms that have been tested in a health system.
- Develop a final implementation guide to support users in implementing the EHR algorithms for research or quality-related risk adjustment, including researchers, health systems, and payers. The implementation guide may identify areas where users may need to modify the algorithms to adapt their EHR system. It may also include summaries of the EHR Learning Network meetings and background meeting materials for reference.

By building EHR algorithms to identify frailty, this project's activities align with **Goal 3: Technology Solutions to Advance Research** described in HHS' Strategic Plan for OS-PCORTF.



AHRQ Accomplishments

Since the project began, AHRQ has made progress towards testing and refining EHR-modified algorithms in a health system, as well as comparing and validating claims-based algorithms.

- AHRQ awarded a contract to Johns Hopkins University to test and refine the draft EHR-modified algorithms, compare with claims-based algorithms, and validate the claims-based algorithms of frailty and functional disabilities against EHR data.
- In addition to using data from OptumLabs Data Warehouse (OLDW) and Kaiser Permanente Mid-Atlantic States (KPMAS), the research team expanded data sources in 2021 to include Johns Hopkins Medicine (JHMI) claims data (a.k.a. Johns Hopkins Healthcare (JHHC) data). They have merged JHHC data with JHMI's EHR data.
- The research team completed several analytical subtasks. These accomplishments include analyzing the EHR data from all clinical sites; linking and comparing claims and EHR data to identify data quality issues; analyzing the frequency of frailty predictors and codes from JHHC and OLDW; selecting the patient denominator of interest across all data sources; and measuring Charlson¹⁸⁸ and Elixhauser¹⁸⁹ comorbidity indexes using the underlying data.



ASPE Accomplishments

Since 2019, the ASPE team has made progress in its role in developing validated and refined claims-based algorithms using Medicare and/or Medicaid claims and EHR-modified algorithms.

- ASPE selected the preferred EHR algorithm specification, and the project team compared the algorithm's performance to two models (Kim Frailty Index¹⁹⁰ and Faurot Frailty Index¹⁹¹) in the general Medicare population for two claims-based outcomes: hospitalizations and nursing facility stay. With input from the Project Advisory Taskforce, the team has chosen the Kim Frailty Index algorithm as the claims-based model to use in AHRQ's validation work.
- The team completed the validation, sensitivity, and outcome analyses of the frailty algorithm for two outcomes (activity limitations and memory recall) using post-acute care assessment data and Medicare claims. Using the results of the validation study, the ASPE team submitted a final recommendation to CMS to use the Kim Frailty Index algorithm with additional variables for age and sex included. The project team included detailed instructions for CMS on applying the algorithm in the CCW to ensure that CMS could use the proposed algorithm to make frailty data on Medicare beneficiaries available to researchers. The project team has begun drafting supplemental guidance and technical

EHR Learning Network

The EHR Learning Network is for providers, health services researchers, health system administrators, and EHR vendors to engage with ASPE on how EHR data can be used to identify patients with frailty and functional disabilities. The goals of the EHR Learning Network are to:

- Learn from administrators and providers about how EHR data are used to identify patients with or at risk of frailty or functional impairment.
- Provide EHR Learning Network participants the opportunity to support the development of claims- and EHR-based algorithms.
- Share additional ways that providers and health systems can use EHR data to target interventions and improve patient care.

documentation on using the Kim frailty index algorithm in the CCW to calculate frailty scores, which will be made available to CCW users.

- The ASPE team launched the EHR Learning Network. The EHR Learning Network engaged stakeholders in the development of the EHR-based frailty and functional disabilities algorithms through interactive webinars, optional small group interviews, and invitations to review and give feedback on draft algorithms and implementation guidance materials. The team hosted a public EHR Learning Network webinar in May 2021 where they summarized preliminary findings of the validation of the claims-algorithms and described the AHRQ-funded portion of this project to test the claims-algorithm against EHR data and test a EHR version of the algorithm.
- Based on feedback received from stakeholders in the EHR Learning Network, ASPE's contractor RAND has drafted the EHR implementation guide for using EHR frailty/impairment data for population management. This resource will support patient-centered care by providing guidance on using EHR data to identify patients with frailty or functional impairment.

Communication and Dissemination Products

- **[The EHR Learning Network on Algorithms to Identify Frailty and Functional Disability Webpage](#)**: This webpage provides additional information on The EHR Learning Network's goals, objectives, and planned activities.¹⁹²

CDC Accomplishments

The CDC team has made progress towards its objectives to examine the feasibility of using standardized ambulatory EHR data to expand claims-based algorithms of frailty and functional disabilities.

- CDC acquired access to ambulatory EHR data and provided descriptive statistics of the overall sample as well as demographic cohorts, co-morbidities, and frailty and functional disabilities.
- CDC has queried the ambulatory EHR data using the algorithm and codes provided by ASPE to identify patients with frailty or disability. The team drafted an internal report presenting findings from the descriptive analysis of these patients in ambulatory EHR data, which included a list of variables used to capture patient function and functional disability. The team also standardized EHR data fields from the screening tool.

Appendix A. OS-PCORTF Project Portfolio

Table A1. Active OS-PCORTF Projects

Funded Agency	Project Title
Agency for Healthcare Research and Quality	Data Capacity for Patient-Centered Outcomes Research through Creation of an Electronic Care Plan for People with Multiple Chronic Conditions*
	Data Capacity for Patient-Centered Outcomes Research through Creation of an Electronic Care Plan for People with Multiple Chronic Conditions 2.0: Development of the patient-facing application*
	Enhancing Patient-Centered Outcomes Research: Creating a National Small-Area Social Determinants of Health Data Platform
	Understanding Long-term Outcomes in COVID-19 Survivors with Multiple Chronic Conditions through e-Care Plan Development*
	Validating and Expanding Claims-based Algorithms of Frailty and Functional Disability for Value-based Care and Payment*
Centers for Disease Control and Prevention	Augmenting the National Hospital Care Survey with Medicaid Administrative Records
	Building Infrastructure and Evidence for COVID-19 Related Research, Using Integrated Data
	Clinical and Community Data Initiative (CODI) 2.0: Integrated Data for Patient-Centered Outcomes Research Project
	Data Linkage: Evaluating Preserving Privacy Methodology and Augmenting the National Hospital Care Survey with Medicaid Administrative Records
	Developing a Multi-State Network of Linked Pregnancy Risk Assessment Monitoring System (PRAMS) and Clinical Outcomes Data for Patient-Centered Outcomes Research
	Enhancing Identification of Opioid-Involved Health Outcomes Using Linked Hospital Care and Mortality Data
	Enhancing Surveillance of Maternal Health Clinical Practices and Outcomes with Federally Qualified Health Centers' Electronic Health Records Visit Data
	Making Electronic Health Record Data More Available for Research and Public Health
	MAT-LINK: MATernal and Infant NetworK to Understand Outcomes Associated with Medication for Opioid Use Disorder during Pregnancy
	MAT-LINK2: Expansion of MATernal and Infant NetworK to Understand Outcomes Associated with Medication for Opioid Use Disorder during Pregnancy
	Validating and Expanding Claims-based Algorithms of Frailty and Functional Disability for Value-based Care and Payment*

Funded Agency	Project Title
Food and Drug Administration	CURE ID: Aggregating and Analyzing COVID-19 Treatment from EHRs and Registries
	Bridging the PCOR Infrastructure and Technology Innovation through Coordinated Registry Networks Community of Practice
	Making Medicaid Data More Accessible Through Common Data Models and FHIR APIs*
National Institutes of Health	Creating a Federal COVID-19 Longitudinal Patient Outcomes Research Database Linked to Health Systems and Clinical Data
	Data Capacity for Patient-Centered Outcomes Research through Creation of an Electronic Care Plan for People with Multiple Chronic Conditions*†
	Data Capacity for Patient-Centered Outcomes Research through Creation of an Electronic Care Plan for People with Multiple Chronic Conditions 2.0: Development of the Patient-facing Application*†
	Making Medicaid Data More Accessible Through Common Data Models and FHIR APIs*
	NIH/NIDA's AMNET: An Addiction Medicine Network to Address the United States Opioid Crisis†
	Training Data for Machine Learning to Enhance PCOR Data Infrastructure*†
	Understanding Long-term Outcomes in COVID-19 Survivors with Multiple Chronic Conditions through e-Care Plan Development*
Office of the Assistant Secretary for Planning and Evaluation	Severe Maternal Morbidity and Mortality EHR Data Infrastructure
	Dataset on Intellectual and Developmental Disabilities: Linking Data to Enhance Person-Centered Outcomes Research
	Child and Caregiver Outcomes Using Linked Data (CCOULD)
	Multistate Emergency Medical Services and Medicaid Dataset (MEMD): A Linked Dataset for PCOR
Office of the National Coordinator for Health Information Technology	Validating and Expanding Claims-based Algorithms of Frailty and Functional Disability for Value-based Care and Payment*
	A Synthetic Health Data Generation Engine to Accelerate Patient-Centered Outcomes Research
	Training Data for Machine Learning to Enhance Patient-Centered Outcomes Research Data Infrastructure*
	Using Machine Learning Techniques to Enable Health Information Exchange Data Sharing to Support COVID-19-focused PCOR

* OS-PCORTF project funding awarded to multiple agencies.

† OS-PCORTF project funding awarded within NIH.

Table A2. Completed OS-PCORTF Projects

Funded Agency	Project Title
Agency for Healthcare Research and Quality	Assessing and Predicting Medical Needs in a Disaster*
	Advancing the Collection and Use of Patient-Reported Outcomes (PROs) through Health Information Technology (IT)*
	Capstone for Outcomes Measures Harmonization (OMH) Project
	Harmonization of Clinical Data Element Definitions for Outcome Measures in Registries
Assistant Secretary for Planning and Evaluation	Assessing and Predicting Medical Needs in a Disaster*
	Beta Testing of the Multi-Payer Claims Data^
	Comparative Effectiveness Research Inventory
Centers for Disease Control and Prevention	Augmenting the National Hospital Care Survey (NHCS) Data through Linkages with Administrative Records
	Clinical and Community Data Initiative (CODI): Integrated Data for Patient-Centered Outcomes Research Project
	Enhancing Data Resources for Researching Patterns of Mortality in Patient-Centered Outcomes Research: Project 1 - Adding Cause-Specific Mortality to NCHS's National Hospital Care Survey by Linking to the National Death Index*
	Enhancing Data Resources for Studying Patterns and Correlates of Mortality in Patient-Centered Outcomes Research: Project 4 – NDI Workshop and Strategy Paper*
	Expanding Data Collection Infrastructure of the National Program of Cancer Registries for Comparative Effectiveness Research
	Identifying Co-Occurring Disorders among Opioid Users Using Linked Hospital Care and Mortality Data
	Improving the Mortality Data Infrastructure for Patient-Centered Outcomes
	Improving Beneficiary Access to their Health Information through an Enhanced Blue Button Service
	Privacy and Security Blueprint, Legal Analysis and Ethics Framework for Data Use, & Use of Technology for Privacy*
	Strengthening the Data Infrastructure for Outcomes Research on Mortality Associated with Opioid Poisonings
Centers for Medicare & Medicaid Services	Enhancing Data Resources for Researching Patterns of Mortality in Patient-Centered Outcomes Research: Project 3*
	Maintenance and Support of the Chronic Conditions Warehouse for Comparative Effectiveness Research

Funded Agency	Project Title
Food and Drug Administration	Collection of Patient-Provided Information through a Mobile Device Application for Use in Comparative Effectiveness and Drug Safety Research
	Cross-Network Directory Service
	Development of a Natural Language Processing Web Service for Public Health Use
	Developing a Strategically Coordinated Registry Network (CRN) for Women's Health Technologies*
	Enhancing Data Resources for Researching Patterns of Mortality in Patient-Centered Outcomes Research: Project 2
	Harmonization of Various Common Data Models and Open Standards for Evidence Generation*
	Source Data Capture from EHRs: Using Standardized Clinical Research Data Utilizing Data from Various Data Partners in a Distributed Manner
	Standardization and Querying of Data Quality Metrics and Characteristics for Electronic Health Data*
	SHIELD - Standardization of Lab Data to Enhance Patient-Centered Outcomes Research and Value-based Care
Health Resources and Services Administration	Strengthening and Expanding the Community Health Applied Research Network (CHARN) Registry to Conduct Patient-Centered Outcomes Research
National Institutes of Health	Creation of LOINC Equivalence Classes
	Development of Data Infrastructure for Use of EHRs in Comparative Effectiveness Research*†
	Developing a Strategically Coordinated Registry Network to Support Research on Women's Health Technologies*
	Emergency Medicine Opioid Data Infrastructure: Key Venue to Address Opioid Morbidity and Mortality†
	Harmonization of Various Common Data Models and Open Standards for Evidence Generation*†
	Technologies for Donating Medicare Beneficiary Claims Data to Research Studies*
	Use of the ADAPTABLE Trial to Strengthen Methods to Collect and Integrate Patient-Reported Information with Other Data Sets and Assess Its Validity
Office of the National Coordinator for Health Information Technology	Advancing the Collection and Use of Patient-Reported Outcomes (PROs) through Health Information Technology (IT)*
	Conceptualizing a Data Infrastructure for the Capture and Use of Patient-Generated Health Data
	Creating the Foundational Blocks for the Learning Healthcare System: Data Access Framework
	Creating the Foundational Blocks for the Learning Healthcare System: Structured Data Capture
	Development of Data Infrastructure for Use of EHRs in Comparative Effectiveness Research*
	Developing a Strategically Coordinated Registry Network (CRN) for Women's Health Technologies*

Funded Agency	Project Title
	PCOR: Privacy and Security Blueprint, Legal Analysis and Ethics Framework for Data Use, & Use of Technology for Privacy*
	Privacy and Security Blueprint, Legal Analysis and Ethics Framework for Data Use, & Use of Technology for Privacy*
	Security and Privacy Standards for Patient Matching, Linking and Aggregation
	Strategic Opportunities for Building Data Infrastructure for Patient-Centered Outcomes Research

[^] The Multi-Payer Claims Data (MPCD) project was a \$16 million CMS project with a contract me of 09/14/2010 to 09/15/2013. On 09/24/2012, the contract was modified with ASPE-provided OS-PCORTF funding to conduct a Beta Test. ASPE was responsible for leadership oversight of the Beta Testing of MPCD.

* OS-PCORTF project funding awarded to multiple agencies.

+ OS-PCORTF project funding awarded within NIH

Appendix B. Abbreviations

Acronym	Description
ACF	Administration for Children and Families
AHRQ	Agency for Healthcare Research and Quality
APA	American Psychiatric Association
API	Application Programming Interface
ASPE	Office of the Assistant Secretary for Planning and Evaluation
ASPR	Office of the Assistant Secretary for Preparedness and Response
CCWIS	Comprehensive Child Welfare Information Systems
CDC	Centers for Disease Control and Prevention
CDE	Clinical Data Element
CDM	Common Data Model
CER	Comparative Effectiveness Research
CMS	Centers for Medicare & Medicaid Services
CODI	Childhood Obesity Data Initiative
CRN	Coordinated Registry Network
DIM	Drug-Involved Mortality (formerly known as the National Vital Statistics System-Mortality-Drug Overdose file)
EDRS	Electronic Death Registration Systems
HER	Electronic Health Record
FDA	U.S. Food and Drug Administration
FHIR®	Fast Healthcare Interoperability Resources
FQHC	Federally Qualified Health Center
HHS	U.S. Department of Health and Human Services
HL7®	Health Level Seven International
HRSA	Health Resources and Services Administration
ICD-10	International Classification of Diseases 10 th Edition
IRB	Institutional Review Board
IT	Information Technology
IVD	In Vitro Diagnostic
LOINC	Logical Observation Identifiers Names and Codes
MBSF	Master Beneficiary Summary File
MMDS	Medical Mortality Data System
NAMCS	National Ambulatory Medical Care Survey
NCHS	National Center for Health Statistics
NCATS	National Center for Advancing Translational Science
NCI	National Cancer Institute
NDI	National Death Index
NHCS	National Hospital Care Survey
NIDA	National Institute on Drug Abuse
NHTSA	National Highway Traffic Safety Administration
NIAID	National Institute of Allergy and Infectious Diseases

Acronym	Description
NIDDK	National Institute of Diabetes and Digestive and Kidney Diseases
NIMH	National Institute of Mental Health
NIH	National Institutes of Health
NLM	National Library of Medicine
NLP	Natural Language Processing
NVSS	National Vital Statistics System
OMOP	Observational Medical Outcomes Partnership
ONC	Office of the National Coordinator for Health Information Technology
OS-PCORTF	Office of the Secretary Patient-Centered Outcomes Research Trust Fund
OUD	Opioid Use Disorders
PCOR	Patient-Centered Outcomes Research
PCORI®	Patient-Centered Outcomes Research Institute
PCORnet®	PCORI's National Patient-Centered Clinical Research Network
PPRL	Privacy Preserving Record Linkage
PRO	Patient-Reported Outcome
RDC	Research Data Center
SAMHSA	Substance Abuse and Mental Health Services Administration
SDOH	Social Determinants of Health
SMART on FHIR®	Substitutable Medical Apps, Reusable Technology on Fast Healthcare Interoperability Resources
SMM	Severe Maternal Mortality
SNOMED	Systematized Nomenclature of Medicine
SUD	Substance Use Disorder
TEP	Technical Expert Panel
T-MSIS	Transformed Medicaid Statistical Information System
USCDI	United States Core Data for Interoperability
USPSTF	U.S. Preventative Services Task Force
USRDS	United States Renal Data System
VSRR	Vital Statistics Rapid Release
WHT-CRN	Women's Health Technologies Coordinated Registry Network
WMP	Weight Management Program

Appendix C. Glossary

Key Terms	Description
Common Data Elements (CDE)	Data elements shared between multiple datasets. ¹⁹³
Common Data Models (CDM)	An aggregated or centralized data model copies data from original sources and brings and standardizes these data in a centralized place. The copied data can then be queried and analyzed.
Data Governance	The process by which stewardship responsibilities are conceptualized and carried out, that is, the policies and approaches that enable stewardship.
Distributed Research Network (DRN)	A DRN is an approach in which data holders maintain control over their protected data and its uses. A DRN features a central portal that performs network functions, such as operations (e.g., workflow, policy rules, auditing, query formation, distribution) and security (e.g., authentication, authorization) and distributed data marts that remain under the control of the data holders.
Electronic Health Record (EHR)	An electronic record of health-related information for a patient that contains information captured in clinical visits, lab and imaging studies, and other information important to the patient's medical past.
Fast Healthcare Interoperability Resources (FHIR®)	A standard for translating health information data into a structure that can be accepted by a wide range of apps.
GitHub	A web-based service for developers to build software.
International Statistical Classification of Diseases and Related Health Problems, 10th revision (ICD-10)	ICD-10 is the diagnostic classification standard for all clinical and research purposes.
Interoperability	The ability of health information technology (health IT) systems from different vendors to communicate and share information.
Institutional Review Board (IRB)	A group that follows federal regulations, state laws, and institutional policy to review, monitor, and approve research in order to protect the ethical rights and privacy of the subjects involved.
Logical Observation Identifiers Names and Codes (LOINC)	A universal coding system for laboratory tests and other clinical observations. It is a national and international standard with widespread adoption and recognition of its utility.
Natural Language Processing (NLP)	A computational model that analyzes texts using several linguistics approaches, such as syntax, semantics, and pragmatics, for the purpose of achieving human-like language results.
Patient-Centered Outcomes Research (PCOR)	Patient-Centered Outcomes Research helps people make informed health care decisions and allows their voices to be heard in assessing the value of health care options. It answers four patient-focused questions: “Given my personal characteristics, conditions and preferences, what should I expect will happen to me?” “What are my options and what are the benefits and harms of those options?” “What can I do to improve the outcomes that are most important to me?” “How can the health care system improve my chances of achieving the outcomes I prefer?”

Key Terms	Description
PCORI's National Patient-Centered Clinical Research Network (PCORNet®)	A “network of networks” that brings together Clinical Data Research Networks and Patient-Powered Research Networks to support patient-centered outcomes research. ¹⁹⁴
Sentinel	A distributed research network, using existing electronic health care data from multiple sources to support monitoring FDA regulated medical products and devices.
SMART on FHIR®	The SMART on FHIR® application programming interface (API) is a standards-based API that builds on FHIR® profiles and resource definitions. ¹⁹⁵
Systematized Nomenclature of Medicine (SNOMED)	A standard for the electronic exchange of clinical health information that has been designated for use by U.S. Federal Government systems. ¹⁹⁶

References

- ¹ Public Law 116-94—Further Consolidated Appropriations Act, 2020. DEC. 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 Public Health Services Act Section 937(7). (2019). Dissemination and Building Capacity for Research.
<https://www.congress.gov/111/plaws/publ148/PLAW-111publ148.pdf>
- ⁶ Assistant Secretary for Planning and Evaluation. (2021). Harmonization of Clinical Data Element Definitions for Outcome Measures in Registries. Retrieved from: <https://aspe.hhs.gov/harmonization-clinical-data-element-definitions-outcome-measures-registries>
- ⁷ 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. Retrieved from: <https://effectivehealthcare.ahrq.gov/products/library-depression/white-paper>
- ⁸ U.S. Preventive Services Task Force. (2017). Screening for Obesity in Children and Adolescents: U.S. Preventive Services Task Force Recommendation Statement. JAMA, 317(23), 2417–2426. doi:10.1001/jama.2017.6803
- ⁹ U.S. National Library of Medicine. (n.d.). Search of: Observational, Patient Registry studies - list results. Retrieved from:
https://clinicaltrials.gov/ct2/results?cond=&term=&type=PReg&rslt=&age_v=&gndr=&intr=&titles=&outc=&spons=&lead=&id=&cntry=&state=&city=&dist=&locln=&rsub=&strd_s=&strd_e=&prcd_s=&prcd_e=&sfpd_s=&sfpd_e=&rfpd_s=&rfpd_e=&lupd_s=&lupd_e=&sort
- ¹⁰ Assistant Secretary for Planning and Evaluation. Harmonization of Clinical Data Element Definitions for Outcome Measures in Registries. Retrieved from: <https://aspe.hhs.gov/harmonization-clinical-data-element-definitions-outcome-measures-registries>
- ¹¹ Major Depression Outcomes App. Retrieved from: <https://apps.smarthealthit.org/app/major-depression-outcomes-app>
- ¹² Agency for Healthcare Research. (2021). Outcome Criteria Framework – 1.0 CI Build. Retrieved December 13, 2021, from: <http://build.fhir.org/ig/HL7/fhir-outcome-criteria-framework-ig/index.html>
- ¹³ Dolin, Gay. (2020). Outcome Criteria Implementation Guide. Retrieved December 13, 2021, from: <https://confluence.hl7.org/display/CQIWC/Outcome+Criteria+Framework+Implementation+Guide>
- ¹⁴ Agency for Healthcare Research. (n.d.). Outcome Criteria Framework Implementation Guide. Retrieved December 13, 2021 from: <https://github.com/HL7/fhir-outcome-criteria-framework-ig>
- ¹⁵ Gliklich RE, Leavy MB, Li F. Standardized Library of Depression Outcome Measures. Research White Paper. (Prepared by L&M Policy Research, LLC under Contract No. 290-2014-00004-C.) AHRQ Publication No. 19(20)-EHC005. Rockville, MD: Agency for Healthcare Research and Quality; May 2020. doi: <https://doi.org/10.23970/AHRQEPCLIBRARYDEPRESSION>
- ¹⁶ ISPOR Value and Outcomes Spotlight. Beyond Harmonization: Implementing Standardized Outcome Measures to Support Value-Based Care. (September 2020). Retrieved from: <https://www.ispor.org/publications/journals/value-outcomes-spotlight/vos-archives/issue/view/unlocking-the-promise-of-real-world-evidence/beyond-harmonization-implementing-standardized-outcome-measures-to-support-value-based-care>
- ¹⁷ Leavy MB, Cooke D, Hajjar S, Bickelman E, Egan B, Clarke D, Gibson D, Casanova B, Gliklich R. Outcome Measure Harmonization and Data Infrastructure for Patient-Centered Outcomes Research in Depression. Research White Paper. (Prepared by OM1, Inc., with subcontractors American Board of Family Medicine and American Psychiatric Association under Contract No. 75Q80119C00005.) AHRQ Publication No. 21-EHC003. Rockville, MD:

Agency for Healthcare Research and Quality; November 2020. doi:
<https://doi.org/10.23970/AHRQEPCREGISTRYOUTCOME>

¹⁸ Leavy MB, Swenson A, Gliklich R. A Prioritized Research Agenda for Using the Harmonized Outcome Measures To Support Patient-Centered Outcomes Research in Depression. (Prepared OM1, Inc. under Contract No. 75Q80119C00005.) AHRQ Publication No. 21(22)-EHC029. Rockville, MD: Agency for Healthcare Research and Quality; October 2021. DOI: 10.23970/AHRQEPCWHITEPAPERDEPRESSIONRESEARCH.

¹⁹ Friedler HS, Leavy MB, Bickelman E, Casanova B, Clarke D, Cooke D, DeMayo A, Egan B, Gibson D, Hajjar S, Gliklich R. Outcome Measure Harmonization and Data Infrastructure for Patient-Centered Outcomes Research in Depression: Data Use and Governance Toolkit. Research White Paper. (Prepared by OM1, Inc. and its subcontractors the American Board of Family Medicine and the American Psychiatric Association under Contract No. 75Q80119C00005.) AHRQ Publication No. 21(22)-EHC031. Rockville, MD: Agency for Healthcare Research and Quality; October 2021. DOI: 10.23970/AHRQEPCWHITEPAPERDEPRESSIONTOOLKIT.

²⁰ Leavy MB, Boussios C, Phillips RL, Clarke D, Sarvet B, Boxwala A, Gliklich R. Outcome Measure Harmonization and Data Infrastructure for Patient-Centered Outcomes Research in Depression: Final Report. (Prepared by OM1, Inc. and its subcontractors the American Board of Family Medicine, American Psychiatric Association, Baystate Health, and Elimu Informatics under Contract No. 75Q80119C00005.) AHRQ Publication No. TBD. Rockville, MD: Agency for Healthcare Research and Quality; Date TBD. Retrieved from:
<https://effectivehealthcare.ahrq.gov/sites/default/files/pdf/harmonization-depression-draft-report.pdf>

²¹ Overview: Outcome Measure Harmonization and Data Infrastructure for Patient-Centered Outcomes Research in Depression. Content last reviewed October 2021. Effective Healthcare Program, Agency for Healthcare Research and Quality, Rockville, MD. Retrieved from: <https://effectivehealthcare.ahrq.gov/products/outcomes-research-depression/overview>

²² B.G. Link, J. Phelan. Social conditions as fundamental causes of diseases. (1995). *J Health Soc Behav*, 36 (Extra Issue). pp. 80-94, 10.2307/2626958

²³ J.C. Phelan, B.G. Link, P. Tehranifar. Social conditions as fundamental causes of health inequalities: theory, evidence, and policy implications. (2010). *J Health Soc Behav*, 51 (Extra Issue). pp. S28-S40, 10.1177/0022146510383498

²⁴ Agency for Healthcare Research and Quality (AHRQ). (2020, December). Social Determinants of Health Database (Beta Version). Retrieved from: <https://www.ahrq.gov/sdoh/data-analytics/sdoh-data.html>

²⁵ Agency for Healthcare Research and Quality (AHRQ). (2020, January). AHRQ Resources to Address SDOH and Social Needs. Retrieved from: <https://www.ahrq.gov/sdoh/practice-improvement.html#resources>

²⁶ Assistant Secretary for Planning and Evaluation. Standardization and Querying of Data Quality Metrics and Characteristics for Electronic Health Data. Retrieved from: <https://aspe.hhs.gov/standardization-querying-data-quality-metrics-characteristics-electronic-health-data>

²⁷ Rubin R. Internet access as a social determinant of health. *JAMA* 326, no. 4 (2021): 298-298.
<https://jamanetwork.com/journals/jama/article-abstract/2782319>

²⁸ Mirel L. The NCHS Data Linkage Program: Leveraging the nation's health data for evidence-based decision making. (2020, October). <https://www.cdc.gov/nchs/data/datalinkage/Data-Linkage-Webinar.pdf>

²⁹ National Center for Health Statistics. Division of Analysis and Epidemiology. (2020, September). The Linkage of the 2016 National Hospital Care Survey to the 2016/2017 Centers for Medicare & Medicaid Services Medicare Enrollment, Claims/Encounters and Assessment Data: Matching Methodology and Analytic Considerations. Hyattsville, Maryland. Retrieved from: <https://www.cdc.gov/nchs/data-linkage/index.htm>

³⁰ National Center for Health Statistics. (2020, September). Restricted-Use Linked NHCS-CMS Medicare Data; 2016 NHCS – 2016/2017 Medicare Enrollment/Summary, Claims/Encounters, and Assessment Data. Retrieved from: <https://www.cdc.gov/nchs/data-linkage/CMS-Medicare-Restricted.htm>

³¹ National Center for Health Statistics. Division of Analysis and Epidemiology. (2021, August). Restricted-Use Linked NHCS – HUD Administrative Housing Data Hyattsville, Maryland. Retrieved from: <https://www.cdc.gov/nchs/data-linkage/index.htm>

³² National Center for Health Statistics. Division of Analysis and Epidemiology. (2021, August). The Linkage of the 2014 and 2016 National Hospital Care Survey to U.S. Department of Housing and Urban Development Administrative

Data: Matching Methodology and Analytic Considerations. Hyattsville, Maryland. Retrieved from:
<https://www.cdc.gov/nchs/data-linkage/index.htm>

³³ Adams, N. (2022, February 28). Opioid Involvement NLP. Retrieved from:
https://github.com/CDCgov/Opioid_Involvement_NLP

³⁴ Adams, N. (2022, March 1). Opioid_SUD_MHI_MedCodes. Retrieved from:
https://github.com/CDCgov/Opioid_SUD_MHI_MedCodes

³⁵ Adams, N. (2022, March 1). SUD_MHI_NLP. Retrieved from: https://github.com/CDCgov/SUD_MHI_NLP

³⁶ Campbell SR, Resnick DM, Cox CS, and Mirel LB. Using supervised machine learning to identify efficient blocking schemes for record linkage. *Statistical Journal of the IAOS*, (Preprint), 1-8. Retrieved from:
<https://stacks.cdc.gov/view/cdc/109539>

³⁷ Mirel, L. (October, 2020). An Overview of the NCHS Data Linkage Program: Leveraging the nation's health data for evidence-based decision making. Retrieved from: <https://www.cdc.gov/nchs/data-linkage/datalinkage-webinar.htm>

³⁸ National Center for Health Statistics. Linked National Hospital Care Survey Data Citation List as of 9/2/2021. (2021, September). Retrieved from: <https://www.cdc.gov/nchs/data/datalinkage/Linked-NHCS-Data-Citation-List-20210902-508.pdf>

³⁹ World Health Organization. Global strategy on diet, physical activity, and health: childhood overweight and obesity. Accessed March 9, 2012.

⁴⁰ U.S. Preventive Services Task Force. (2017). Screening for Obesity in Children and Adolescents: U.S. Preventive Services Task Force Recommendation Statement. *JAMA*, 317(23), 2417–2426. doi:10.1001/jama.2017.6803

⁴¹ U.S. Preventive Services Task Force. (2017). Screening for Obesity in Children and Adolescents: U.S. Preventive Services Task Force Recommendation Statement. *JAMA*, 317(23), 2417–2426. doi:10.1001/jama.2017.6803

⁴² MITRE. (n.d.). MITRE/CODI. Retrieved December 13, 2021, from
<https://github.com/mitre/codi/blob/main/CODI%20Data%20Model%20Implementation%20Guide.pdf>

⁴³ MITRE. (n.d.). MITRE/CODI. Retrieved December 13, 2021, from
<https://github.com/mitre/codi/blob/main/CODI%20PPRL%20Implementation%20Guide.pdf>

⁴⁴ Kochanek KD, Xu JQ, Arias E. Mortality in the United States, 2019. NCHS Data Brief, no 395. Hyattsville, MD: National Center for Health Statistics. 2020.

⁴⁵ Johnson NB, Hayes LD, Brown K, Hoo EC, Ethier KA, & Centers for Disease Control and Prevention (CDC) (2014). CDC National Health Report: leading causes of morbidity and mortality and associated behavioral risk and protective factors--United States, 2005-2013. *MMWR supplements*, 63(4), 3-27.

⁴⁶ Centers for Disease Control and Prevention. (June 2021). Health and Economic Costs of Chronic Diseases. Retrieved from: <https://www.cdc.gov/chronicdisease/about/costs/index.htm>

⁴⁷ The Centers for Disease Control and Prevention. About PRAMS. Retrieved from:
<https://www.cdc.gov/prams/about/prams-faq.htm>

⁴⁸ The Centers for Disease Control and Prevention. PRAMS Special Projects. Retrieved from:
<https://www.cdc.gov/prams/special-projects/index.htm#link>

⁴⁹ The Centers for Disease Control and Prevention. PRAMS Questionnaires. Retrieved from:
<https://www.cdc.gov/prams/questionnaire.htm>

⁵⁰ The Centers for Disease Control and Prevention. (2022, January 21). Participating PRAMS Sites. Retrieved from:
<https://www.cdc.gov/prams/states.htm>

⁵¹ Whitney P, Witt, W. P. (2018). The future of maternal and child health data in the United States. *Am J Public Health*, 108(10), 1277-1279. doi:10.2105/ajph.2018.304663,

⁵² The Association of State and Territorial Health Officials (ASTHO). About Us. Retrieved from:
<https://www.astho.org/About/>

⁵³ The Centers for Disease Control and Prevention. PRAMS Special Projects. Retrieved from:
<https://www.cdc.gov/prams/special-projects/index.htm#link>

-
- ⁵⁴ Owens PL, Weiss AJ, Barrett ML, Reid LD. (June 2020). Social Determinants of Health and County Population Rates of Opioid-Related Inpatient Stays and Emergency Department Visits, 2016. HCUP Statistical Brief #260. Agency for Healthcare Research and Quality, Rockville, MD. Retrieved from: www.hcup-us.ahrq.gov/reports/statbriefs/sb260-Social-Determinants-County-Opioid-Rates-Hospital-Use-2016.pdf.
- ⁵⁵ Vivolo-Kantor AM, Seth P, Gladden RM, et al. (2018). Vital Signs: Trends in Emergency Department Visits for Suspected Opioid Overdoses — United States, July 2016–September 2017. *Morbidity and Mortality Weekly Report*, 67, 279–285. <http://dx.doi.org/10.15585/mmwr.mm6709e1>.
- ⁵⁶ National Center for Health Statistics. Office of Analysis and Epidemiology. (2018, April). The Linkage of the 2014 National Hospital Care Survey to the 2014/2015 National Death Index: Methodology Overview and Analytic Considerations. Hyattsville, Maryland. Retrieved from: https://www.cdc.gov/nchs/data/datalinkage/NHCS14_NDI14_15_Methodology_Analytic_Consider.pdf
- ⁵⁷ National Center for Health Statistics. Office of Analysis and Epidemiology. (2019, May). Enhancing Identification of Opioid-involved Health Outcomes using Linked Hospital Care and Mortality Data: Summary Report on Task 1. Retrieved from https://www.cdc.gov/nchs/data/nhcs/opioid_health_outcomes.pdf
- ⁵⁸ National Center for Health Statistics Research Data Center (NCHS RDC). (2020, October). Linked Data on Hospitalizations, Mortality, and Drugs: Data from the National Hospital Care Survey 2016, National Death Index 2016-2017, and the Drug-Involved Mortality 2016-2017. Retrieved from: <https://www.cdc.gov/nchs/data/nhcs/Task-3-Doc-508.pdf>
- ⁵⁹ White DG, Adams NB, Brown AM, O’Jiaku-Okorie A, Badwe R, Shaikh S, Adegbeye A. Enhancing identification of opioid-involved health outcomes using National Hospital Care Survey data. National Center for Health Statistics. Vital Health Stat 2(188). 2021. DOI: <https://dx.doi.org/10.15620/cdc:108998>
- ⁶⁰ Jackson G, Brown AM, DeFrances C. (2020, August 4). Opioid-involved Emergency Department Visits in the National Hospital Care Survey and the National Hospital Ambulatory Medical Care Survey. National Health Statistics Reports; no 149. Hyattsville, MD: National Center for Health Statistics. 2020. Retrieved January 12, 2022, from <https://www.cdc.gov/nchs/data/nhsr/nhsr149-508.pdf>
- ⁶¹ Jackson G, DeFrances C, Ashman J. Assessing National Hospital Care Survey and National Hospital Ambulatory Medical Care Survey Data: A Comparison of Opioid and Respiratory Disease Encounters. <https://www.cdc.gov/nchs/data/nhcs/JSM-508.pdf>
- ⁶² White DG, Adams NB, Brown AM, O’Jiaku-Okorie A, Badwe R, Shaikh S, Adegbeye A. Enhancing identification of opioid-involved health outcomes using National Hospital Care Survey data. National Center for Health Statistics. Vital Health Stat 2(188). 2021. DOI: <https://dx.doi.org/10.15620/cdc:108998>
- ⁶³ Health Resources and Services Administration. National Health Data Center: Expanded summary for 2019 UDS tables 3A-9E. Retrieved: <https://bphc.hrsa.gov/uds/datacenter.aspx?q=t6b&year=2018&state=>.
- ⁶⁴ National Center for Health Statistics. National Ambulatory Medical Care Survey – Community Health Centers: 2014 State and National Summary Tables. Hyattsville, MD: National Center for Health Statistics; 2019. Retrieved: https://www.cdc.gov/nchs/data/ahcd/namcs_summary/2014_namcs_chc_web_tables-508.pdf.
- ⁶⁵ Substance Abuse and Mental Health Services Administration (SAMHSA). (2018, September). 2017 National Survey of Drug Use and Health (NSDUH) Releases: NSDUH data. Retrieved from: <https://www.samhsa.gov/data/release/2017-national-survey-drug-use-and-health-nsduh-releases>
- ⁶⁶ Substance Abuse and Mental Health Services Administration (SAMHSA). (2016, September). 2015 national survey of drug use and Health (NSDUH) Releases. Retrieved from: <http://www.samhsa.gov/data/release/2015-national-survey-drug-use-and-health-nsduh-releases>
- ⁶⁷ National Center for Health Statistics Research Data Center (NCHS RDC). (2021, June). Identifying Co-Occurring Disorders among Opioid Users Using Linked Hospital Care and Mortality Data: Capstone to an Existing FY18 PCORTF Project. Retrieved from: <https://www.cdc.gov/nchs/data/nhcs/FY19-RDC-2021-06-01-508.pdf>.
- ⁶⁸ Centers for Disease Control and Prevention. (2021, November). *Making Rapid Data Exchange a Reality*. Retrieved from: https://www.cdc.gov/csels/phio/making_rapid_data_exchange_reality.html
- ⁶⁹ Github. (n.d.) Making EHR Data More Available for Research and Public Health (MedMorph) Reference Architecture HL7 FHIR Implementation Guide. Retrieved from: <https://github.com/HL7/fhir-medmorph>

⁷⁰ HL7 International – Public Health Work Group. (2021, December). Making EHR Data More available for Research and Public Health (MedMorph). Retrieved from: <http://build.fhir.org/ig/HL7/fhir-medmorph/>

⁷¹ HL7 International – Public Health Work Group. (2021, December). Central Cancer Registry Reporting Content IG. Version 0.1.0. Retrieved from: <https://build.fhir.org/ig/HL7/fhir-central-cancer-registry-reporting-ig/index.html>

⁷² HL7 International – Public Health Work Group. (2021, December). Healthcare Surveys Content Implementation Guide. Version 0.1.0. Retrieved from: <https://build.fhir.org/ig/HL7/fhir-health-care-surveys-reporting-ig/index.html>

⁷³ HL7 International – Public Health Work Group. (2021, December). MedMorph Research Data Exchange Content I. Version 0.1.0. Retrieved from: <http://build.fhir.org/ig/HL7/fhir-medmorph-research-dex-ig/branches/master/index.html>

⁷⁴ Centers for Disease Control and Prevention. (2021, November). *Making Rapid Data Exchange a Reality*. Retrieved from: https://www.cdc.gov/csels/phio/making_rapid_data_exchange_reality.html

⁷⁵ Maria Michaels, Sameemuddin Syed, William B Lober, Blueprint for aligned data exchange for research and public health, Journal of the American Medical Informatics Association, Volume 28, Issue 12, December 2021, Pages 2702–2706, <https://doi.org/10.1093/jamia/ocab210>

⁷⁶ Haight SC, Ko JY, Tong VT, Bohm MK, Callaghan WM. (2018). Opioid use disorder documented at delivery hospitalization – United States, 1999–2014. *Morbidity and Mortality Weekly Report*, 67, 845-849.

⁷⁷ Fill MA, Miller AM, Wilkinson RH, Warren MD, Dunn JR, Schaffner W, et al. (2018). Educational Disabilities among Children Born with Neonatal Abstinence Syndrome. *Pediatrics*, 124(3).

⁷⁸ Oei JL, Melhuish E, Uebel H, Azzam N, Breen C, Burns L, et al. (2017). Neonatal Abstinence Syndrome and High School Performance. *Pediatrics*, 139(2).

⁷⁹ Centers for Disease Control and Prevention. (2020, July 15). Substance use during pregnancy. Retrieved from: <https://www.cdc.gov/reproductivehealth/maternalinfanthealth/substance-abuse/substance-abuse-during-pregnancy.htm>

⁸⁰ Centers for Disease Control and Prevention. (2020, December). MAT-LINK: MATernal and Infant NetworK to Understand Outcomes Associated with Treatment for Opioid Use Disorder during Pregnancy. Retrieved from: <https://www.cdc.gov/ncbddd/aboutus/mat-link.html>

⁸¹ Centers for Disease Control and Prevention (2020, October) Reducing Polysubstance Pregnancy. Retrieved from: <https://www.cdc.gov/grand-rounds/pp/2020/20200729-reducing-polysubstance-pregnancy.html>

⁸² Tran, E.L., Kim, S.Y., England, L.J., Green, C., Dang, E.P., Broussard, C.S., Fehrenbach, N., Hudson, A., Yow-Conley, T., Gilboa, S.M., & Meaney-Delman, D. (2020). The MATernal and Infant NetworK to understand outcomes associated with treatment of opioid use disorder during pregnancy (MAT-LINK): Surveillance opportunity. *Journal of Women's Health*, 29(12), 1491-1499. <http://doi.org/10.1089/jwh.2020.8848>

⁸³ Maeda, A., et al., *Opioid abuse and dependence during pregnancy: temporal trends and obstetrical outcomes*. Anesthesiology, 2014. 121(6): p. 1158-65.

⁸⁴ Hudak, M.L., et al., *Neonatal drug withdrawal*. Pediatrics, 2012. 129(2): p. e540-60.

⁸⁵ Department of Health and Human Services. (2017, October 26). HHS Acting Secretary Declares Public Health Emergency to Address National Opioid Crisis. Retrieved from: <https://public3.pagefreezer.com/browse/HHS.gov/31-12-2020T08:51/https://www.hhs.gov/about/news/2017/10/26/hhs-acting-secretary-declares-public-health-emergency-address-national-opioid-crisis.html>

⁸⁶ Ventura SJ. The U.S. National Vital Statistics System: Transitioning into the 21st century, 1990–2017. National Center for Health Statistics. *Vital Health Stat* 1(62), 2018. Retrieved from: https://www.cdc.gov/nchs/data/series/sr_01/sr01_062.pdf

⁸⁷ Warner M, Paulozzi L, Nolte K, Davis G, Nelson L. State variation in certifying manner of death and drugs involved in drug intoxication deaths. *Acad Forensic Pathol*. 2013;3(2):231–237.

⁸⁸ Friedman DJ, Parrish RG. Enhancing Electronic Health Record Systems to Generate and Exchange Data with Electronic Vital Registration Systems. Public Health Informatics Institute. 2015. Retrieved from: <https://www.cdc.gov/nchs/data/dvs/evital/12-PHI-EHR-eVital-Records-Standards.pdf>

⁸⁹ Vital Statistics Rapid Release. National Center for Health Statistics. Retrieved from: <https://www.cdc.gov/nchs/nvss/vsrr.htm>.

-
- ⁹⁰ HL7 International, Kahn K, Shakir A. (2020). Vital Records Mortality and Morbidity Reporting FHIR Implementation Guide. HL7 International. Retrieved from: <http://hl7.org/fhir/us/vrdr/index.html>
- ⁹¹ Centers for Disease Control and Prevention. (2021). NVSS Modernization Tool Kit. Retrieved from: <https://www.cdc.gov/nchs/nvss/modernization/tool-kit.htm>
- ⁹² Ahmad FB, Cisewski JA. Quarterly provisional estimates for selected indicators of mortality, 2019-Quarter 1, 2021. National Center for Health Statistics. National Vital Statistics System, Vital Statistics Rapid Release Program. 2021. Retrieved from: <https://www.cdc.gov/nchs/nvss/vsrr/mortality-dashboard.htm>
- ⁹³ Ahmad FB, Rossen LM, Sutton P. Provisional drug overdose death counts. National Center for Health Statistics. 2021. Retrieved from: <https://www.cdc.gov/nchs/nvss/vsrr/drug-overdose-data.htm>
- ⁹⁴ Jackson G, Brown AM, DeFrances C. Opioid-involved emergency department visits in the National Hospital Care Survey and the National Hospital Ambulatory Medical Care Survey. National Health Statistics Reports; no 149. Hyattsville, MD: National Center for Health Statistics. 2020. Retrieved from: <https://www.cdc.gov/nchs/data/nhsr/nhsr149-508.pdf>
- ⁹⁵ Ahmad FB, Dokpesi P, Escobedo L, Rossen L. Timeliness of death certificate data by sex, age, and geography. Vital Statistics Rapid Release; no 9. Hyattsville, MD: National Center for Health Statistics. June 2020. Retrieved from: <https://www.cdc.gov/nchs/data/vsrr/vsrr009-508.pdf>
- ⁹⁶ Spencer MR, Warner M, Bastian BA, Trinidad JP, Hedegaard H. Drug overdose deaths involving fentanyl, 2011–2016. National Vital Statistics Reports; vol 68 no 3. Hyattsville, MD: National Center for Health Statistics. 2019. Retrieved from: https://www.cdc.gov/nchs/data/nvsr/nvsr68/nvsr68_03-508.pdf
- ⁹⁷ Ely DM, Martin JA, Hoyert DL, Rossen LM, Drake P. Drug-involved infant deaths in the United States, 2015–2017. National Vital Statistics Reports; vol 70 no 7. Hyattsville, MD: National Center for Health Statistics. 2021. DOI: <https://dx.doi.org/10.15620/cdc:105508>
- ⁹⁸ Black JC, Bau GE, Iwanicki JL, Dart RC. Association of Medical Stimulants With Mortality in the U.S. From 2010 to 2017. *JAMA Intern Med.* 2021;181(5):707–709. doi:10.1001/jamainternmed.2020.7850
- ⁹⁹ MDEpiNet. (n.d.). Tools. Retrieved from: <https://www.mdepinet.net/tools>
- ¹⁰⁰ Stoddard MD, Zheng X, Mao J, Te A, Sedrakyan A, Chughtai B. Safety and Efficacy of Outpatient Surgical Procedures for the Treatment of Benign Prostatic Enlargement in New York State and California (2005-2016). *J Urol.* 2021 Mar;205(3):848-854. doi: 10.1097/JU.0000000000001401. Epub 2020 Oct 7. PMID: 33026907.
- ¹⁰¹ Chughtai B, Mao J, Matheny ME, Mauer E, Banerjee S, Sedrakyan A. Long-Term Safety with Sling Mesh Implants for Stress Incontinence. *J Urol.* 2021 Jan;205(1):183-190. doi: 10.1097/JU.0000000000001312. Epub 2020 Aug 4. PMID: 32749936.
- ¹⁰² Bouhadana D, Nguyen DD, Zhang X, Mao J, Elterman DS, Te AE, Zorn KC, Bhojani N, Sedrakyan A, Chughtai B. Safety and efficacy of TURP vs. laser prostatectomy for the treatment of benign prostatic hyperplasia in multi-morbid and elderly individuals aged ≥ 75. *World J Urol.* 2021 Dec;39(12):4405-4412.
- ¹⁰³ Bertges DJ, Eldrup-Jorgensen J, Robbins S, Ssemaganda H, Malone M, Marinac-Dabic D, Smale J, Lottes AE, Majithia A, Resnic FS; Society for Vascular Surgery Vascular Quality Initiative. Vascular Quality Initiative Surveillance of Femoropopliteal Artery Paclitaxel Devices. *JACC Cardiovasc Interv.* 2021 Dec 13;14(23):2598-2609. doi: 10.1016/j.jcin.2021.08.058. PMID: 34887051
- ¹⁰⁴ Mao J, Moore KO, Columbo JA, Mehta KS, Goodney PP, Sedrakyan A. Validation of an indirect linkage algorithm to combine registry data with Medicare claims. *J Vasc Surg.* 2022 Feb 15;. doi: 10.1016/j.jvs.2022.01.132. [Epub ahead of print] PubMed PMID: 35181518.
- ¹⁰⁵ Chughtai B, Cho A, Simonyan V, Mao J, Viviano C, Marinac-Dabic D, Menafee S, Rardin C, Sedrakyan A. Development and Usability Testing of a Mobile Application to Monitor Patient-Reported Outcomes after Stress Urinary Incontinence Surgery. *Urology.* 2021 Oct 24:S0090-4295(21)00984-5. doi: 10.1016/j.urology.2021.10.011. Epub ahead of print. PMID: 34706249.
- ¹⁰⁶ Behrendt CA, Sedrakyan A, Katsanos K, Nordanstig J, Kuchenbecker J, Kreutzburg T, Secemsky EA, Debus ES, Marschall U, Peters F. Sex Disparities in Long-Term Mortality after Paclitaxel Exposure in Patients with Peripheral Artery Disease: A Nationwide Claims-Based Cohort Study. *J Clin Med.* 2021 Jul 2;10(13):2978.

-
- ¹⁰⁷ Chen A, Paxton L, Zheng X. (2021). Association of Sex With Risk of 2-Year Revision Among Patients Undergoing Total Hip Arthroplasty. *JAMA network open*, 4(6), e2110687. <https://doi.org/10.1001/jamanetworkopen.2021.10687>
- ¹⁰⁸ Garland SK, Falster MO, Beiles CB, Freeman AJ, Jorm LR, Sedrakyan A, Sotade O, Varcoe RL. Long Term Outcomes Following Elective Repair of Intact Abdominal Aortic Aneurysms: A Comparison Between Open Surgical and Endovascular Repair Using Linked Administrative and Clinical Registry Data. *Ann Surg.* 2021 Oct 20;. doi: 10.1097/SLA.0000000000005259. [Epub ahead of print] PubMed PMID: 35129507.
- ¹⁰⁹ Public Law 116-136, § 18115(a), the Coronavirus Aid, Relief, and Economic Security (CARES) Act. The U.S. Department of Health and Human Service. (2020). COVID-19 Pandemic Response, Laboratory Data Reporting: CARES Act Section 18115. Retrieved from: <https://www.congress.gov/116/bills/hr748/BILLS-116hr748enr.pdf>
- ¹¹⁰ The U.S. Department of Health and Human Services. (2020). COVID-19 Pandemic Response, Laboratory Data Reporting: CARES Act Section 18115. Retrieved from: <https://www.hhs.gov/sites/default/files/covid-19-laboratory-data-reporting-guidance.pdf>
- ¹¹¹ The U.S. Department of Health and Human Services. (2020). COVID-19 Pandemic Response, Laboratory Data Reporting: CARES Act Section 18115. Retrieved from: <https://www.hhs.gov/sites/default/files/covid-19-laboratory-data-reporting-guidance.pdf>
- ¹¹² Centers for Diseases Control and Prevention. (2020, June 4). LOINC In Vitro Diagnostic (LIVD) Test Code Mapping for SARS-CoV-2 Tests. Retrieved from: <https://www.cdc.gov/cseis/dls/sars-cov-2-livd-codes.html>
- ¹¹³ Medical Device Innovation Consortium (MDIC). (2020, June 2). Systemic Harmonization and Interoperability Enhancement for Lab Data (SHIELD). Retrieved from: <https://mdic.org/program/systemic-harmonization-and-interoperability-enhancement-for-lab-data-shield/>
- ¹¹⁴ The U. S. Food and Drug Administration. (2020, July 21). FDA's Ongoing Work to Support and Advance COVID-19 Diagnostic Test Accuracy and Availability. Retrieved from: <https://www.fda.gov/news-events/fda-voices/fdas-ongoing-work-support-and-advance-covid-19-diagnostic-test-accuracy-and-availability>
- ¹¹⁵ MDEpinet. (n.d.). SHIELD. Retrieved from: <https://www.mdepinet.net/shield>
- ¹¹⁶ National Institute on Drug Abuse. (2018, June). Medications to Treat Opioid Use Disorder. Retrieved from: <https://www.drugabuse.gov/publications/research-reports/medications-to-treat-opioid-addiction/overview>
- ¹¹⁷ Substance Abuse and Mental Health Services Administration (n.d.) Certification of Opioid Treatment Programs. Retrieved from: <https://www.samhsa.gov/medication-assisted-treatment/become-accredited-opioid-treatment-program>
- ¹¹⁸ U.S. Food and Drug Administration. (2019, February 14). Information about Medication-Assisted Treatment (MAT). <https://www.fda.gov/drugs/information-drug-class/information-about-medication-assisted-treatment-mat>
- ¹¹⁹ Schwartz RP, Gibson D, Pagano A, et al. Addiction medicine practice-based research network (AMNet): Building partnerships. *Psychiatric Services*. 2021; 72:7, 845-847. <https://doi.org/10.1176/appi.ps.202000390>
- ¹²⁰ Clarke DE, Ibrahim A, Doty B, et al. Addiction Medicine Practice-Based Research Network (AMNet): Assessment Tools and Quality Measures. *Subst Abuse Rehabil.* 2021 Jun 25;12:27-39. doi: 10.2147/SAR.S305972.
- ¹²¹ Centers for Disease Control and Prevention. Pregnancy Mortality Surveillance System. Retrieved from: <https://www.cdc.gov/reproductivehealth/maternal-mortality/pregnancy-mortality-surveillance-system.htm#about-pmss>
- ¹²² Creanga AA, Berg CJ, Ko JY, et al. Maternal mortality and morbidity in the United States: Where are we now? *J Womens Health (Larchmt)* 2014;23:3–9.
- ¹²³ Creanga AA, Syverson C, Seed K, Callaghan WM. Pregnancy-related mortality in the United States, 2011e2013. *Obstet Gynecol* 2017;130:366–73.
- ¹²⁴ Petersen EE, Davis NL, Goodman D, et al. Racial/ethnic disparities in pregnancy-related deaths—United States, 2007 e2016. *MMWR Morb Mortal Wkly Rep* 2019; 68:762–5.
- ¹²⁵ National Public Radio. For Every Woman Who Dies In Childbirth In The U.S., 70 More Come Close. May 2018. Retrieved from: <https://www.npr.org/2018/05/10/607782992/for-every-woman-who-dies-in-childbirth-in-the-u-s-70-more-come-close>
- ¹²⁶ HL7. FHIR Implementation Guide. Retrieved from: <https://www.hl7.org/fhir/implementationguide.html>

-
- ¹²⁷ Larson, S.A., Eschenbacher, H.J., Taylor, B., Pettingell, S., Sowers, M., & Bourne, M.L. (2020). In-home and residential long-term supports and services for persons with intellectual or developmental disabilities: Status and trends through 2017. Minneapolis: University of Minnesota, Research and Training Center on Community Living, Institute on Community Integration. https://ici-s.umn.edu/files/aCHyYaFjMi/risp_2017
- ¹²⁸ Braddock D, Hemp R, Tanis S, Lulinski A, Scotti S. Public Support for People with Intellectual and Developmental Disabilities. National Conference of State Legislatures. Published January 2018. Accessed January 13, 2022. <https://www.ncsl.org/research/labor-and-employment/public-support-for-people-with-intellectual-and-developmental-disabilities.aspx>
- ¹²⁹ Larson, S.A., Eschenbacher, H.J., Taylor, B., Pettingell, S., Sowers, M., & Bourne, M.L. (2020). In-home and residential long-term supports and services for persons with intellectual or developmental disabilities: Status and trends through 2017. Minneapolis: University of Minnesota, Research and Training Center on Community Living, Institute on Community Integration. https://ici-s.umn.edu/files/aCHyYaFjMi/risp_2017
- ¹³⁰ Larson, S.A., Eschenbacher, H.J., Taylor, B., Pettingell, S., Sowers, M., & Bourne, M.L. (2020). In-home and residential long-term supports and services for persons with intellectual or developmental disabilities: Status and trends through 2017. Minneapolis: University of Minnesota, Research and Training Center on Community Living, Institute on Community Integration. https://ici-s.umn.edu/files/aCHyYaFjMi/risp_2017
- ¹³¹ Kepple NJ. (2018). Does Parental Substance Use Always Engender Risk for Children? Comparing Incidence Rate Ratios of Abusive and Neglectful Behaviors across Substance Use Behavior Patterns. *Child Abuse & Neglect*, 76, 44–55. Retrieved from: <https://doi.org/10.1016/j.chabu.2017.09.015>
- ¹³² Smith VC, Wilson CR. (2016). Families Affected by Parental Substance Use. *Pediatrics*, 138(2), e20161575–e20161575. Retrieved from: <https://doi.org/10.1542/peds.2016-1575>
- ¹³³ Peleg-Oren N, Teichman M. (2006). Young Children of Parents with Substance Use Disorders (SUD): A Review of the Literature and Implications for Social Work Practice. *Journal of Social Work Practice in the Addictions*, 6(1-2), 49-61. Retrieved from: http://dx.doi.org/10.1300/J160v06n01_03
- ¹³⁴ Kepple NJ. (2018). Does parental substance use always engender risk for children? Comparing incidence rate ratios of abusive and neglectful behaviors across substance use behavior patterns. *Child Abuse & Neglect*, 76, 44–55. Retrieved from: <https://doi.org/10.1016/j.chabu.2017.09.015>
- ¹³⁵ Office of the Administration for Children & Families. State Medicaid and Child Welfare Data Linkages for Outcomes Research, 2019 – 2021. Retrieved from: <https://devd8.acf.hhs.gov/opre/project/state-medicaid-and-child-welfare-data-linkages-outcomes-research-2019-2021>
- ¹³⁶ Office of Planning, Research and Evaluation & the Office of the Assistant Secretary for Planning and Evaluation . (n.d.). Child and Caregiver Outcomes Using Linked Data, 2019-2022: Project Overview. Retrieved from: <https://aspe.hhs.gov/sites/default/files/documents/d8f17022f085efd79013d33d81994e91/ccould-project-brief.pdf>
- ¹³⁷ Clinovations Government + Health. (2020, October 28). Treatment of Sialorrhea in Cerebral Palsy. Version 1.0. Retrieved from: https://synthetichealth.github.io/module-builder/#cerebral_palsy
- ¹³⁸ Clinovations Government + Health. (2020, November 2). Prescribing Opioids for Chronic Pain and Treatment of Opioids Use Disorder. Version 1.0. Retrieved from: https://synthetichealth.github.io/module-builder/#prescribing_opioids_for_chronic_pain_and_treatment_of_oud
- ¹³⁹ Clinovations Government + Health. (2020, December 15). Sepsis. Version 1.0. Retrieved from: <https://synthetichealth.github.io/module-builder/#sepsis>
- ¹⁴⁰ Clinovations Government + Health. (2020, November 30). Spina Bifida. Version 1.0. Retrieved from: https://synthetichealth.github.io/module-builder/#spina_bifida
- ¹⁴¹ Clinovations Government + Health. (2021, January). Synthea™ Module Companion Guide: Cerebral Palsy. Retrieved from: http://hdl.mitre.org/downloads/module_guides/Module_Companion_Guide_Cerebral_Palsy_508.pdf
- ¹⁴² Clinovations Government + Health. (2021, January). Synthea™ Module Companion Guide: Prescribing Opioids for Chronic Pain and Treatment of Opioid Use Disorder. Retrieved from: http://hdl.mitre.org/downloads/module_guides/Module_Companion_Guide_Opioids_508.pdf
- ¹⁴³ Clinovations Government + Health. (2021, January). Synthea™ Module Companion Guide: Sepsis. Retrieved from: http://hdl.mitre.org/downloads/module_guides/Module_Companion_Guide_Sepsis_508.pdf

-
- ¹⁴⁴ Clinovations Government + Health. (2021, January). Synthea™ Module Companion Guide: Spina Bifida. Retrieved from: http://hdl.mitre.org/downloads/module_guides/Module_Companion_Guide_Spina_Bifida_508.pdf
- ¹⁴⁵ The Office of the National Coordinator for Health Information Technology. (2021). Synthetic Health Data Generation to Accelerate Patient-Centered Outcomes Research. Retrieved from: <https://www.healthit.gov/topic/scientific-initiatives/pcor/synthetic-health-data-generation-accelerate-patient-centered-outcomes>
- ¹⁴⁶ Assistant Secretary for Planning and Evaluation (ASPE). (2020, April 2). Strategic Plan: Strategic Goal 1: Reform, Strengthen, and Modernize the Nation's Healthcare System. Retrieved from: <https://www.hhs.gov/about/strategic-plan/strategic-goal-1/index.html>
- ¹⁴⁷ HHS Press Office. (2021, January 19). HHS Awards Funds to Expand Immunization Information Sharing Collaboration. Retrieved from: <https://www.hhs.gov/about/news/2021/01/19/hhs-awards-funds-to-expand-immunization-sharing-collaboration.html>
- ¹⁴⁸ Agency for Healthcare Research and Quality. Implementation of an electronic care plan for people with multiple chronic conditions (North Carolina). Retrieved from: <https://digital.ahrq.gov/ahrq-funded-projects/implementation-electronic-care-plan-people-multiple-chronic-conditions>
- ¹⁴⁹ HL7 International – Patient Care Workgroup. (2021, December 14). MCC eCare Plan Draft Implementation Guide Version 1.0. Retrieved from: https://trifolia-fhir.lantanagroup.com/igs/lantana_hapi_r4/MCC-IG/index.html
- ¹⁵⁰ GitHub. (2021, February 25). Multiple Chronic Condition Careplan Version 1.0.5. Retrieved from: <https://github.com/MccCareplan> <https://github.com/MccCareplan/providersmartapp>
- ¹⁵¹ McKenzie, L. (2021, October 29). HL7 FHIR Connectathons. Retrieved from: <https://confluence.hl7.org/display/FHIR/Connectathons>
- ¹⁵² Agency for Healthcare Research and Quality. About the eCare Plan Joint NIH/NIDDK AHRQ Project. Retrieved from: <https://ecareplan.ahrq.gov/>
- ¹⁵³ GitHub. (2021, February 3). Multiple Chronic Condition Careplan Version 1.0. Retrieved from: <https://github.com/MccCareplan/patientsmartapp>
- ¹⁵⁴ HL7 International – Patient Care Workgroup. (2020, November 2). MCC eCare Plan Draft Implementation Guide Version 1.0. Retrieved from: https://trifolia-fhir-dev.lantanagroup.com/igs/lantana_prod_hapi_r4/MCC-IG/index.html
- ¹⁵⁵ National Center for Health Statistics and Centers for Disease Control and Prevention. (2021, February). Enhancing Data Resources for Studying Patterns and Correlates of Mortality in Patient-Centered Outcomes Research: Project 1 - Adding Cause-Specific Mortality to NCHS's National Hospital Care Survey by Linking to the National Death Index. Retrieved from <https://aspe.hhs.gov/system/files/pdf/259016/cdc-project-1-final-report.pdf>
- ¹⁵⁶ National Center for Health Statistics. (2021). Restricted-Use Linked NHCS-NDI Data. Retrieved from: <https://www.cdc.gov/nchs/data-linkage/nhcs-ndi.htm>
- ¹⁵⁷ Restricted-Use Linked NCHS-Mortality NDI Data, codebook for the 2015 Linked Mortality File (LMF). Retrieved from: https://www.cdc.gov/nchs/data/datalinkage/LMF2015_DataDictionary.pdf
- ¹⁵⁸ National Center for Health Statistics. (2019, August 6). The Linkage of the 2014 National Hospital Care Survey to the 2014/2015 National Death Index: Methodology Overview and Analytic Considerations. Retrieved from: https://www.cdc.gov/nchs/data/datalinkage/NHCS14_NDI14_15_Methodology_Analytic_Consider.pdf
- ¹⁵⁹ National Center for Health Statistics. (2021, March 5). Restricted-Use Linked NHCS-CMS Medicare Data. Retrieved from: <https://www.cdc.gov/nchs/data-linkage/CMS-Medicare-Restricted.htm>
- ¹⁶⁰ National Center for Health Statistics. (2019, August 6). The Linkage of the 2014 National Hospital Care Survey to the 2014/2015 Centers for Medicare & Medicaid Services Master Beneficiary Summary File: Methodology Overview and Analytic Consideration. Retrieved from: <https://www.cdc.gov/nchs/data/datalinkage/NHCS-CMS-Medicare-Linkage-Methods-and-Analytic-Considerations.pdf>
- ¹⁶¹ National Center for Health Statistics. (2021, January 8). Restricted-Use Linked NHCS-NDI Data. Retrieved from: <https://www.cdc.gov/nchs/data-linkage/nhcs-ndi.htm>
- ¹⁶² National Center for Health Statistics. (2019, August 6). The Linkage of the 2014 National Hospital Care Survey to the 2014/2015 Centers for Medicare & Medicaid Services Master Beneficiary Summary FileMethodology Overview

and Analytic Consideration. Retrieved from: <https://www.cdc.gov/nchs/data/datalinkage/NHCS-CMS-Medicare-Linkage-Methods-and-Analytic-Considerations.pdf>

¹⁶³ Bercovitz A, Jamoom E, Lau DT. (2018). National Hospital Care Survey Demonstration Projects: Characteristics of Inpatient and Emergency Department Encounters among Patients with any Listed Diagnosis of Alzheimer disease. National Health Statistics Reports; no 121. Hyattsville, MD: National Center for Health Statistics Reports. Retrieved from: <https://www.cdc.gov/nchs/data/nhsr/nhsr121-508.pdf>

¹⁶⁴ Ashman JJ, Cairns C, DeFrances CJ, Schwartzman A. (2021). Respiratory illness emergency department visits in the National Hospital Care Survey and the National Hospital Ambulatory Medical Care Survey. National Health Statistics Reports; no 151. Hyattsville, MD: National Center for Health Statistics. 2021. Retrieved from: <https://www.cdc.gov/nchs/data/nhsr/nhsr151-508.pdf>

¹⁶⁵ Hoyert DL, Miniño AM. (2020). Maternal mortality in the United States: Changes in coding, publication, and data release, 2018. National Vital Statistics Reports; vol 69 no 2. Hyattsville, MD: National Center for Health Statistics. Retrieved from: <https://www.cdc.gov/nchs/data/nvsr/nvsr69/nvsr69-02-508.pdf>

¹⁶⁶ Jackson G, Chari K. (2019). National Hospital Care Survey demonstration projects: Stroke inpatient hospitalizations. National Health Statistics Reports; no 132. Hyattsville, MD: National Center for Health Statistics. Retrieved from: <https://www.cdc.gov/nchs/data/nhsr/nhsr132-508.pdf>

¹⁶⁷ Jackson G, Brown AM, DeFrances C. (2020). Opioid-involved emergency department visits in the National Hospital Care Survey and the National Hospital Ambulatory Medical Care Survey. National Health Statistics Reports; no 149. Hyattsville, MD: National Center for Health Statistics. Retrieved from: <https://www.cdc.gov/nchs/data/nhsr/nhsr149-508.pdf>

¹⁶⁸ Spencer MR, Flagg LA, Jackson G, DeFrances C, Hedegaard H. (2020). National Hospital Care Survey demonstration projects: Opioid-involved emergency department visits, hospitalizations, and deaths. National Health Statistics Reports; no 141. Hyattsville, MD: National Center for Health Statistics. Retrieved from: <https://www.cdc.gov/nchs/data/nhsr/nhsr141-508.pdf>

¹⁶⁹ Fuller CC, Hua W, Leonard CE, et al. (2020, November 2). Developing a Standardized and Reusable Method to Link Distributed Health Plan Databases to the National Death Index: Methods Development Study Protocol. *JMIR Res Protoc.* 9(11):e21811. doi: 10.2196/21811.

¹⁷⁰ Office of the National Coordinator for Health IT. National Health IT Priorities. Retrieved from: <https://www.healthit.gov/topic/scientific-initiatives/national-health-it-priorities-research-policy-and-development-agenda>

¹⁷¹ Assistant Secretary for Planning and Evaluation. Standardization and Querying of Data Quality Metrics and Characteristics for Electronic Health Data. Retrieved from: <https://aspe.hhs.gov/standardization-and-querying-data-quality-metrics-and-characteristics-electronic-health-data>

¹⁷² DataBricks. Extract Transform Load. Retrieved from: <https://databricks.com/glossary/extract-transform-load>

¹⁷³ National Institutes of Health. (2021). TB Portals. Retrieved from: <https://tbportals.niaid.nih.gov/>

¹⁷⁴ Zheng Q, Lu Y, Lure F, Jaeger S, Lu P. (2020). Clinical and radiological features of novel coronavirus pneumonia. Journal of X-Ray Science and Technology, 28(3), 391-404. doi:10.3233/XST-200687.

¹⁷⁵ Ma L, Wang Y, Guo L, et al. (2020). Developing and verifying automatic detection of active pulmonary tuberculosis from multi-slice spiral CT images based on deep learning. Journal of X-Ray Science and Technology, vol. 28, no. 5, pp. 939-951. doi: 10.3233/XST-200662

¹⁷⁶ Yang F, Yu H, Kantipudi K, et al. Differentiating between Drug-Sensitive and Drug-Resistant Tuberculosis with Machine Learning for Clinical and Radiological Features. Quantitative Imaging in Medicine and Surgery, 2021. <https://qims.amegroups.com/article/view/77560/html>

¹⁷⁷ Karki M, Kantipudi K, Yu H, Yang F, Kassim YM, Yaniv Z, Jaeger S. Identifying Drug-Resistant Tuberculosis in Chest Radiographs: Evaluation of CNN Architectures and Training Strategies. Annual International Conference of the IEEE Engineering in Medicine and Biology Society (EMBC), 2021. <https://pubmed.ncbi.nlm.nih.gov/34891867/>

¹⁷⁸ Karki M, Kantipudi K, Yang F, Yu H, Wang YXJ, Yaniv Z, Jaeger S. Generalization Challenges in Drug-Resistant Tuberculosis Detection from Chest X-rays. Diagnostics 2022, 12, 188. <https://doi.org/10.3390/diagnostics12010188>

-
- ¹⁷⁹ Rankin SK. (2021). onc-healthit/2021PCOR-ML-AI: ONC-BAH-1 (v1.1). Zenodo. <https://doi.org/10.5281/zenodo.5514290>
- ¹⁸⁰ Rankin SK. (2021). onc-healthit/2021PCOR-ML-AI: ONC-BAH-1 (v1.1). Zenodo. <https://doi.org/10.5281/zenodo.5514290>
- ¹⁸¹ The Office of the National Coordinator for Health Information Technology. (2021). Implementation Guidance and Code. Retrieved from: <https://www.healthit.gov/topic/scientific-initiatives/pcor/machine-learning>
- ¹⁸² The Office of the National Coordinator for Health Information Technology. (2020). Training Data for Machine Learning to Enhance Patient-Centered Outcomes Research Data Infrastructure. Retrieved from: <https://www.healthit.gov/topic/scientific-initiatives/pcor/machine-learning>
- ¹⁸³ The Office of the National Coordinator for Health Information Technology. (2020). Training Data for Machine Learning to Enhance Patient-Centered Outcomes Research Data Infrastructure. Retrieved from: [https://www.healthit.gov/sites/default/files/page/2021-09/ONC%20Training%20Data%20for%20ML_Project_Infographic-FINAL.pdf](https://www.healthit.gov/sites/default/files/page/2021-09/ONC%20Training%20Data%20Project_Infographic-FINAL.pdf)
- ¹⁸⁴ The Office of the National Coordinator for Health Information Technology. (2021, September 15). Training Data for Machine Learning to Enhance Patient-Centered Outcomes Research (PCOR) Data Infrastructure: Project Webinar. Retrieved from: https://www.healthit.gov/sites/default/files/page/2021-09/ONC_Training%20Data%20for%20ML_Webinar_508.pdf
- ¹⁸⁵ Garcia S. (2020, March 31). The Application of Machine Learning to Address Kidney Disease. HealthIT Buzz. Retrieved from: <https://www.healthit.gov/buzz-blog/health-it/the-application-of-machine-learning-to-address-kidney-disease>
- ¹⁸⁶ Rahn M, Hua J, Wong A, et al. (2021). Training Data for Machine Learning to Enhance Patient-Centered Outcomes Research Data Infrastructure Final Report. The Office of the National Coordinator for Health Technology. Retrieved from: https://www.healthit.gov/sites/default/files/page/2021-09/ONC_Training%20Data%20for%20ML_PCOR_Final%20Report.pdf
- ¹⁸⁷ Hao Q, Zhou L, Dong B, Yang M, Dong B, Weil Y. The role of frailty in predicting mortality and readmission in older adults in acute care wards: a prospective study. *Scientific reports*. 2019; 9(1), 1207. <https://doi.org/10.1038/s41598-018-38072-7>
- ¹⁸⁸ Charlson ME, Pompei P, Ales KL, MacKenzie CR. A new method of classifying prognostic comorbidity in longitudinal studies: development and validation. *Journal of chronic diseases*. 1987; 40(5), 373–383. [https://doi.org/10.1016/0021-9681\(87\)90171-8](https://doi.org/10.1016/0021-9681(87)90171-8)
- ¹⁸⁹ Elixhauser Comorbidity Software Refined for ICD-10-CM Healthcare Cost and Utilization Project (HCUP). October 2021. Agency for Healthcare Research and Quality, Rockville, MD. Retrieved from: www.hcup-us.ahrq.gov/toolssoftware/comorbidityicd10/comorbidity_icd10.jsp
- ¹⁹⁰ Kim DH , Schneeweiss S, Glynn RJ, Lipsitz LA, Rockwood K, Avorn J. Measuring Frailty in Medicare Data: Development and Validation of a Claims-Based Frailty Index. *The journals of gerontology. Series A, Biological sciences and medical sciences*. 2018; 73(7), 980–987. <https://doi.org/10.1093/gerona/glx229>
- ¹⁹¹ Faurot KR, Jonsson Funk M, Pate V, et al. Using claims data to predict dependency in activities of daily living as a proxy for frailty. *Pharmacoepidemiol Drug Saf*. 2015;24(1):59–66. doi:10.1002/pds.3719
- ¹⁹² Office of the Assistant Secretary for Planning and Evaluation (ASPE). (2020, August 5). Electronic Health record (EHR) Learning Network on Algorithms to Identify Frailty and Functional Disability. Retrieved from: <https://aspe.hhs.gov/frailty-EHR-learning-network>
- ¹⁹³ NIH U.S. National Library of Medicine. (2020). Common Data Element (CDE) Resource Portal Glossary. Retrieved from: <https://www.nlm.nih.gov/cde/glossary.html>
- ¹⁹⁴ Patient-Centered Outcomes Research Institute. (2013, December). PCORnet: The National Patient-Centered Clinical Research Network. Retrieved from: <https://www.pcori.org/research-results/pcornet-national-patient-centered-clinical-research-network>
- ¹⁹⁵ SMART®. (2018, March 8). What is smart? SMART—An App Platform for Healthcare. Retrieved from: <https://smarthealthit.org/an-app-platform-for-healthcare/about/>
- ¹⁹⁶ NIH U.S. National Library of Medicine. (2019, September 18). SNOMED-CT. Retrieved from: <https://www.nlm.nih.gov/healthit/snomedct/>