

2024

Annual Portfolio Report Appendices

Building Data Capacity for the
Office of the Secretary Patient-
Centered Outcomes Research
Trust Fund

Appendix A: Key Impacts of OS-PCORTF

Projects Completed in FY 2024..... 2

Details the **eight projects** that concluded in FY 2024 and the **data and technical resources** developed by the projects.

- ▶ CURE ID: Aggregating and Analyzing COVID-19 Treatment from EHRs and Registries 3
- ▶ Data Linkage: Evaluating Privacy-Preserving Record Linkage Methodology and Augmenting the National Hospital Care Survey with Medicaid Administrative Records..... 5
- ▶ Developing a Multi-State Network of Linked Pregnancy Risk Assessment Monitoring System and Clinical Outcomes Data for Patient-Centered Outcomes Research 7
- ▶ Human-Centered Design Study on Federal Data Access and Acquisition Processes of CMS Data 9
- ▶ Linking State Medicaid and Child Welfare Data for Outcomes Research on Treatment for Opioid Use Disorder and Other Behavioral Health Issues 10
- ▶ Making Electronic Health Record Data More Available for Research and Public Health 12
- ▶ MAT-LINK2: Expansion of MATernaL and Infant NetworK to Understand Outcomes Associated with Treatment for Opioid Use Disorder during Pregnancy 14
- ▶ Severe Maternal Morbidity and Mortality EHR Data Infrastructure 16

Appendix B: New Projects Funded in FY 2024..... 18

Describes the **seven new projects** funded in FY 2024, including **project goals** and the patient-centered outcomes research (PCOR) data infrastructure enhancements addressed.

- ▶ Establishing the Governance, Legal and Analytical Framework for a Federated Linked Data System: Creating a New Data Research Environment 19
- ▶ Improving Maternal Health by Leveraging Existing Resources, Data Linkage and Standard Reporting for Patient-Centered Outcomes Research 20
- ▶ Improving the Availability of National Mortality Data by Accelerating Data Release and Laying the Groundwork for IT Modernization of the National Death Index..... 21
- ▶ Leveraging Data Standards for Patient-Centered Outcomes Research in Sickle Cell Disease under the United States Core Data for Interoperability Plus Program..... 22
- ▶ Linking the Boston Birth Cohort and Pregnancy to Early Life Longitudinal Data System and Social Determinants of Health Databases: A Longitudinal Cohort of Mother-Child Dyads..... 23
- ▶ Medical Expenditure Panel Survey Enhancements on Burdens and Economic Impacts of Medical Care..... 24
- ▶ N3C Renal Freeport: A Collaborative Data Analytics and Data Sharing Model for Patient-Centered Outcomes Research 25

Appendix C: In Progress Projects in FY 2024 26

Presents the **27 ongoing projects**.



Appendix A: Key Impacts of OS-PCORTF Projects Completed in FY 2024

This appendix highlights key impacts of the eight Office of the Secretary Patient-Centered Outcomes Research Trust Fund (OS-PCORTF) projects that concluded in Fiscal Year (FY) 2024.

The completed FY 2024 OS-PCORTF projects have made important contributions to improving data infrastructure for patient-centered outcomes research across four domains: 1) improving the quality of data; 2) providing more relevant, comprehensive data; 3) improving access and use of data; and 4) enhancing analytic resources. The following icons indicate how each of the completed projects described below are enhancing data capacity.

Improving the Quality of Data



Providing More Relevant, Comprehensive Data



Improving Access and Use of Data



Enhancing Analytic Resources



Completed Project Resources

Of the FY 2024 completed projects' 27 data and technical resources, **78 percent** are open-source and publicly available to researchers and others. The following icons indicate accessibility of key resources described below:



Icon indicates key resources that are *public/open-source*



Icon indicates key resources that are *restricted access* to the public

Ensuring Data Infrastructure Resources Meet End-User Needs

88 percent of completed FY 2024 projects engaged end users in the development and testing of resources.

End users were engaged in the design of web-based applications, developing use cases, refining and improving data collection methods, and informing electronic health record (EHR) data abstraction processes. They also participated in Delphi processes to reach consensus on data definitions, tested implementation guides at Health Level Seven (HL7®) Fast Healthcare Interoperability Resources (FHIR®) Connectathons and conducted human-centered design activities.

FY 2024 Completed Project Snapshots

CURE ID: Aggregating and Analyzing COVID-19 Treatment from EHRs and Registries

CURE ID, a collaboration between the Food and Drug Administration (FDA) and National Center for Advancing Translational Sciences at the National Institutes of Health (NIH/NCATS), is a web-based repository that allows clinicians to report and access case information on the use of existing drugs for difficult-to-treat infectious diseases. Recognizing the need to better understand the health outcomes of the different repurposed drugs used in the COVID-19 pandemic, this project expanded the CURE ID platform to include tens of thousands of COVID-19 case reports and expanded the platform to enable automated extraction and manual data collection from electronic health records (EHRs) and clinical disease registries.

Challenge Addressed

In the initial stages of the COVID-19 pandemic, the rapid emergence and spread of the disease left limited time for developing new, effective treatments.¹ A broad standardized platform to gather experiences with different drugs was necessary to facilitate both research on effective COVID-19 treatments and partnerships to obtain data and automate data extraction.

Project Impact on Expanding Data Capacity for Patient-Centered Outcomes Research



To improve the quality of available EHR data, the project transformed and harmonized EHR data from seven healthcare systems into a unified database

using the Observational Medical Outcomes Partnership (OMOP) common data model (CDM). The project also improved the representativeness of data captured in CURE ID by expanding data collection to gather patient-reported outcomes from vulnerable populations, including pregnant women and neonates.



To provide more relevant, comprehensive data, the project team enabled the CURE ID platform to capture new variables on COVID-19 and draw from new sources of data

including EHRs, clinical disease registries, and a patient-completed form on Long COVID treatment experiences. Added data elements include those on COVID-19, social determinants of health (SDOH) obtained from medical records, as well as more comprehensive information on demographics, characteristics of interventions and comparators, outcomes, and timing of interventions and outcomes up to 28 days, among other elements.





To improve clinician and researcher access to and use of timely data on COVID-19 treatments, the project made de-identified patient-level cases available to researchers via the CURE ID platform and made the platform accessible to researchers, as well as the broader community of patients, caregivers, clinicians, policymakers, and others.


Key Resources





CURE ID Database. The CURE ID Database now captures **over 115,000 acute COVID-19 treatment cases** from seven healthcare systems, and **more than 500 cases submitted by patients** with Long COVID. The capability for patients to report has continued to be refined for Long COVID, mpox, RASopathies, Sarcoma, and other diseases. The database is publicly accessible for researchers wishing to explore treatment used for all of these patients.

 **[CURE ID and EHR Data Dictionary](#)**. This resource provides guidance on the CURE ID data elements included in the EHR-extracted data, how they were defined, and the types of response values.

 **[CURE ID Researcher’s Guide](#)**. The guide summarizes data processing and pathways, data sources and characterization, data access, data security, data quality assurance, and frequently asked questions to inform research use and exploration of the data in CURE ID.

 **[CURE ID Data Extraction Toolkit/Guidance Documents](#)**. The open-source code for the automated data extraction tool is available on GitHub. The tool allows users to capture targeted fields from diverse EHRs and registries globally.


 **[CURE ID Report on COVID-19 Outcome Measures of Importance to Patients](#)**. This report captures findings from interviews conducted with four patients to identify COVID-19 outcome measures important to patients and that could be collected via CURE ID.

 **[CURE ID Patient Case Report Form](#)**. Patients with any infectious disease or their caregivers can submit a case report form for their condition in CURE ID. This includes a generic patient case report form that mimics the generic clinician case report form for most diseases, as well as disease-specific case report forms for mpox, Long COVID, Sarcoma, and RASopathies.

Long-Term Impacts on Interagency Partnerships and Sustainability

The CURE ID Database is already being leveraged for additional research initiatives. The data extraction infrastructure established for CURE ID is being repurposed and leveraged for the Society of Critical Care Medicine's (SCCM) initiative REDISCOVER-ICU, a clinical registry examining drug repurposing in critical care.² The CURE ID dataset was used at the 2024 SCCM Datathon, providing more than 100 attendees the opportunity to explore the dataset and learn about unique challenges and opportunities presented by the data. Two multi-disciplinary teams from this event submitted ancillary study proposals to use the CURE ID dataset for clinical and data science research projects. The SCCM’s newly launched Centers of Excellence program also submitted a letter of intent to the Patient-Centered Outcomes Research Institute (PCORI) for a longitudinal pediatric intensive care unit liberation study using CURE ID data and was invited for full study proposal submission. In addition, the CURE ID team is exploring whether a similar approach can be taken to identify rare causes of meningitis and their treatments in EHR data and to bring this data into the CURE ID platform.

Implementing Agency FDA and NCATS/NIH	Point of Contact Heather Stone, heather.stone@fda.hhs.gov	Project Profile ASPE Website Project Profile
---	---	--

Strategic Plan Goals Addressed		Goal 1: Data Capacity for National Health Priorities
---------------------------------------	---	---

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

The National Center for Health Statistics (NCHS) Data Linkage Program creates new data resources that can be used to study the relationship between federal program participation and access to health care services and health outcomes. This project builds on two previously funded OS-PCORTF projects that conducted linkages to the National Hospital Care Survey (NHCS) under the NCHS Data Linkage Program: *Enhancing Data Resources for Researching Patterns of Mortality in Patient-Centered Outcomes Research* and *Augmenting the NHCS Data through Linkages with Administrative Records*. This project evaluated the accuracy of a privacy-preserving record linkage (PPRL) technique against the previously developed linkage algorithm. An initial assessment and a refined approach were evaluated. The project also developed an enhanced linkage algorithm used to link the 2016 NHCS with the Centers for Medicare & Medicaid Services (CMS) Transformed Medicaid Statistical Information System (T-MSIS) administrative records. The enhancement involved incorporating machine learning techniques to improve linkage accuracy and efficiency. The resulting linked NHCS-T-MSIS data expands the linked NHCS data resources developed under the prior OS-PCORTF projects. The linked NHCS data enables researchers to study patient-centered outcomes associated with means-tested government insurance programs, health care received, participation in federal housing programs, and management of chronic health conditions.

Challenge Addressed

Linkages conducted by the NCHS Data Linkage Program fill critical data gaps by integrating information from hospital patient visits with government health insurance program information and other health-related administrative data. Traditional linkage methods rely on the exchange of personal identifiable information, which poses threats to data privacy.³ PPRL is an alternative to traditional linkage approaches that may overcome this barrier. However, prior to implementing PPRL techniques it is important to understand their effect on data quality. The assessment of PPRL techniques is an important first step toward advancing the potential use of PPRL to integrate data without sharing direct identifiers and protecting health information.

Project Impact on Expanding Data Capacity for Patient-Centered Outcomes Research



To enhance analytic resources available for linking data while maintaining individual privacy, the project developed an enhanced machine

learning-based linkage algorithm (i.e., sequential coverage algorithm) which has improved the accuracy and efficiency of linking NHCS files to CMS data. The percentage of eligible records linked ranged from 43 percent to 71 percent, with samples stratified by age and sex.




To provide more relevant, comprehensive data, the project augmented the previously developed 2014 and 2016 NCHS-Medicare linked datasets with additional years of data,


including 2014 and 2015 Medicare claims, prescription drug and assessment data, and 2017 Medicare Advantage (MA) enrollment data. It also developed a new 2016 NHCS-T-MSIS linked dataset to support studies on Medicaid insurance and hospital utilization. These additional years of data will support trend analyses and longitudinal studies of utilization and expenditures over time.





To improve researchers' access and use of NHCS data for studying PCOR outcomes, the team has made the linked datasets available to researchers, through the NCHS and Federal Statistical Research Data Centers.


Key Resources and Final Report


- 


[Linked Dataset: 2016 NHCS and T-MSIS.](#) This dataset includes 2016 NHCS data linked with 2015-2017 Medicaid and Children’s Health Insurance Program (CHIP) claims data. The dataset includes five Medicaid/CHIP files: a demographic and eligibility file, an inpatient hospital file, a long-term care file, a pharmacy file, and other services file.
- 

[Data Dictionaries for Linked NHCS–Medicaid Dataset.](#) Six data dictionaries are available for the linked NHCS-CMS data and include information on match status and the five Medicaid/CHIP files.
- 

[The Linkage of the 2016 National Hospital Care Survey to 2015-2017 CMS T-MSIS Claims Data: Matching Methodology and Analytic Considerations.](#) This report describes the linkage of data from the 2016 NHCS to 2015-2017 CMS T-MSIS claims data. The report includes a brief overview of the data sources, a description of the linkage methodology used, and analytic considerations to assist researchers when using the files.
- 

[Linked Dataset: 2014 NCHS and 2014/2015 Medicare Data Files.](#) The dataset provides additional linked data files containing 2014 NHCS inpatient and emergency department claims and EHR data linked with 2014 and 2015 Medicare claims, prescription drug, and assessment data.
- 



[2017 Medicare Advantage Data Files.](#) The project team added 2017 MA encounter data to the set of linked 2016 NHCS-Medicare files to achieve consistency with data years for fee-for-service claims.
- 

[A Methodological Assessment of Privacy Preserving Record Linkage Using Survey and Administrative Data.](#) This journal article, published in *Statistical Journal of the IAOS*, describes the methodology used and results of the PPRL validation. The project compared a commercially available PPRL software to the results of a clear text matching approach to link survey and administrative records, using the OS-PCORTF funded NHCS-NDI linkage as the gold standard. This journal article serves as a resource for stakeholders interested in linking data without sharing direct identifiers.
- 

[Project Final Report.](#) The final report describes the project’s goals, activities, and key accomplishments as well as lessons learned and a summary of key resources and publications.

Long-Term Impacts on Interagency Partnerships and Sustainability

The linkage methods developed and subsequent improvements in linked NHCS-T-MSIS data quality accomplished through this project will be utilized in future NCHS data linkage projects. As more years of the NHCS become available, the algorithms developed through this project will continue to be used to link new sources of data and create new resources. As a result of this project, NCHS received funding for an OS-PCORTF FY 2021 project that explores the creation of synthetic linked data files and an OS-PCORTF FY 2022 project to continue the assessment of PPRL methodology with different sources of data and different PPRL tools.

<p>Implementing Agency CDC</p>	<p>Point of Contact Cordell Golden, cdg4@cdc.gov</p>	<p>Project Profile ASPE Website Project Profile</p>
<p>Strategic Plan Goals Addressed</p>	 <p>Goal 2: Data Standards and Linkages for Longitudinal Research</p>	 <p>Goal 3: Technology Solutions to Advance Research</p>

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

By establishing a Pregnancy Risk Assessment Monitoring System (PRAMS) Data Linkage Learning Community, the project assisted 11 states in successfully linking PRAMS survey data with clinical and/or administrative data, such as birth certificates, hospital discharge data, Medicaid claims, and all-payer claims databases (APCDs). States can use the resulting data resources for future analyses on maternal and child health outcomes. This project aligns with the federal [White House Blueprint for Addressing the Maternal Health Crisis](#), which highlights PRAMS data and PRAMS data linkages as opportunities to improve data for better surveillance, quality measurement, and ultimately to improve outcomes.

Challenge Addressed

Since 1987, the Centers for Disease Control and Prevention (CDC) has worked with state, territorial, and local health departments to implement PRAMS, a population-based survey capturing the lived experiences of persons with a recent live birth.⁴ There is considerable interest in linking PRAMS data with clinical and/or administrative data to conduct maternal and infant patient-centered outcomes research, but linkage efforts vary greatly by state and there is not a consistent set of standard methodologies to make these data available for research at the national level.⁵

Project Impact on Expanding Data Capacity for Patient-Centered Outcomes Research



The project team produced **analytic resources** to ensure replicable and sustainable PRAMS linkages. The PRAMS Data Linkage Framework includes four key resources to support standardized PRAMS linkages: 1) template data use agreement for public health data linkage projects, 2) data linkage readiness assessment tool, 3) data linkage process list that outlines five phases of a linkage project, and 4) data linkage process map that visually depicts each phase.



The project supported 11 states in linking PRAMS to clinical outcomes data and/or administrative data (e.g., hospital discharges, Medicaid claims, APCDs) to **support more relevant, comprehensive data** for states' assessment of healthcare factors that influence maternal and child health clinical or health-related outcomes. Depending on the state, external researchers may be able to access the linked data.

Key Resources and Final Report



PRAMS Data Linkage Framework. The framework provides linkage tools and resources to assist states to develop a comprehensive systematic linkage process for linking PRAMS with administrative data sources, regardless of their data source or outcomes of interest.





Project Final Report. The final report provides an overview of project activities, including participating states' data linkage activities and capacity building and technical assistance offered to states, as well as lessons learned and future recommendations.

Long-Term Impacts on Interagency Partnerships and Sustainability

Most of the 12 states who entered the PRAMS Data Linkage Learning Community indicated that they planned to use the data internally for future program implementation or other surveillance activities. For example, the Nebraska Health Department will sustain data linkages to better incorporate race/ethnicity and other structural factors into the

analysis of hospital discharge data. The Alaska Department of Health will sustain PRAMS data linkages to continue to assess risk factors for child maltreatment. States have also presented findings from analyses using linked PRAMS data to multiple audiences, including state legislators.

Implementing Agency CDC	Point of Contact Shanna Cox, cio8@cdc.gov	Project Profile ASPE Website Project Profile
Strategic Plan Goals Addressed	 Goal 1: Data Capacity for National Health Priorities	 Goal 2: Data Standards and Linkages for Longitudinal Research

Human-Centered Design Study on Federal Data Access and Acquisition Processes of CMS Data

CMS conducted a human-centered design study to understand and identify current gaps in the process used by federal agencies to request access to CMS data. This human-centered design study sought to address root causes of the challenges and limitations federal agencies experience. Using insights from federal end users, the project developed four prototype resources to support a more efficient federal agency data request process and provided recommendations to inform future updates to research request workflows, the CMS website, and resources at CMS. The goal of the project was to create a more streamlined, transparent, and user-friendly process that meets federal agency partners’ needs, while ensuring compliance with relevant laws and CMS policies.

Challenge Addressed

To access CMS data containing personally identifiable information and protected health information, researchers must comply with a standard process to request the data.⁶ Federal agencies may request special terms to accessing the data, however, this has required additional processes that can be difficult and resource-intensive to navigate.

Project Impact on Expanding Data Capacity for Patient-Centered Outcomes Research



To improve researchers’ access to and use of CMS data, the CMS team developed four prototype resources to support a more efficient federal agency data request process. These resources aim to narrow information gaps for requestors. They include a landing page highlighting key information, a visualization of the data request process phases, a form to help requestors determine key variables to include in their request, and a Frequently Asked Questions (FAQ) page. These resources will improve federal accessibility of CMS data and may lead to new linked datasets that may ultimately be made available to extramural researchers.


Key Resources and Final Report



Project Final Report. This final report describes the project objectives, background, methods, lessons learned, and additional recommendations, as well as opportunities to improve the federal requestor experience.

Long-Term Impacts on Interagency Partnerships and Sustainability

Streamlining and improving the process for federal requestors of CMS data will allow for federal researchers to have a clearer and more efficient experience with requesting CMS data, ultimately supporting the use of CMS data for patient-centered outcomes research. By addressing root causes of data request challenges and developing user-friendly resources, the project will reduce burden on other federal agencies in requesting data while helping them understand what data are available and for what purposes. CMS will continue to use the recommendations identified through the project to refine and implement the prototype process changes and inform future updates to research request workflows, the CMS website, and resources at CMS.

<p>Implementing Agency CMS</p>	<p>Point of Contact Kari Gaare, kari.gaare@cms.hhs.gov</p>	<p>Project Profile ASPE Website Project Profile</p>
<p>Strategic Plan Goals Addressed</p>	<p> Goal 2: Data Standards and Linkages for Longitudinal Research</p>	

Linking State Medicaid and Child Welfare Data for Outcomes Research on Treatment for Opioid Use Disorder and Other Behavioral Health Issues

This project linked records from state Medicaid and child welfare systems for two states to develop the Child and Caregiver Outcomes Using Linked Data (CCOULD) dataset. The project also developed a data linkage methodology and other resources to illustrate key lessons to inform similar state efforts in the future.

Challenge Addressed

Child well-being is linked to family well-being and parents' capacity to care for their children.⁷ Parental substance use disorder (SUD) is a major driver of child welfare system involvement and can influence a child's likelihood of developing SUD in the future.⁸ However, there is a lack of comprehensive data about the use of Medicaid services for SUD or mental health treatment by both parents and children involved with child welfare systems, and how that treatment might be associated with child outcomes.

Project Impact on Expanding Data Capacity for Patient-Centered Outcomes Research



To improve the quality of data, the Office of the Assistant Secretary for Planning and Evaluation/Administration for Children and Families (ASPE/ACF) team created a CDM for the linked CCOULD dataset. The CDM harmonized records across child welfare and Medicaid data systems, making them easier for researchers to use for analyses.



To provide more relevant, comprehensive data, the team linked records from child welfare data systems with Medicaid enrollment and claims data for two states (Florida and Kentucky), which allows for a more complete picture of the services families receive that are funded by child welfare agencies and Medicaid.



To improve researchers' access and use of the linked CCOULD dataset, a final research use (i.e., all identifiers removed) dataset was made available free of charge. The team also developed a linkage road map for other states interested in developing their own research use linked data across child welfare and Medicaid systems along with a toolkit to facilitate bi-direction data exchange between systems.



To enhance analytic resources, this project developed a data linkage methodology that allowed for linkage of records that lacked Social Security numbers, which helped address prior data linkage challenges.

Key Resources and Final Report



CCOULD Dataset and User Guide. The de-identified, general research use CCOULD dataset is available to researchers free of charge at the National Data Archive on Child Abuse and Neglect at Cornell University. The CCOULD dataset contains linked records on **1,087,763 children** and **89,871 caregivers** and covers the periods of January 2017 through June 2021 in Florida, and January 2017 through 2020 in Kentucky.



Toolkit: Data Sharing for Child Welfare Agencies and Medicaid. This toolkit provides guidance to Medicaid and child welfare agencies to create automated, bi-directional data exchanges between their respective information systems.



Linking Child Welfare and Medicaid Data: Lessons Learned from Two States. This report provides lessons learned on recruitment, data linkage and harmonization, data privacy, data governance, and sustainability to state, tribal, and local child welfare and behavioral health agencies interested in linking their Medicaid and child welfare data.



Project Final Report. The final report summarizes critical background information, major project accomplishments, lessons learned, and key products and publications resulting from the project activities.





Linked Datasets for Florida and Kentucky. These datasets contain de-identified linked Medicaid eligibility, enrollment, and claims data with child welfare records developed by Florida and Kentucky, which linked 80 and 81 percent, respectively, of the children in the child welfare system with a report to a Medicaid record. States plan to continue performing these linkages on an annual basis in order to generate a more complete understanding of the services families receive that are funded by child welfare agencies and Medicaid. Data are not publicly available to researchers.

Long-Term Impacts on Interagency Partnerships and Sustainability

ASPE/ACF worked with state agencies to develop best practices and technical assistance for ongoing, annual linkage of the state child welfare and Medicaid data systems to support the states’ needs.

The success of the CCOULD project spurred a new, larger initiative to link data in more states, thus developing a sustainable data infrastructure. Initiated in 2022, the Child Welfare and Health Infrastructure for Linking and Data Analysis of Resources, Effectiveness, and Needs (CHILDREN) project builds directly on the lessons learned from CCOULD.⁹ The CHILDREN project will assess states’ readiness to link Medicaid and child welfare data. Building on the success of the Florida and Kentucky datasets, CHILDREN will also develop datasets that link Medicaid and child welfare records for new states.

Additionally, as of October 2024, 69 external researchers have been granted access to the CCOULD dataset and 50 requests for information about the data have been submitted by 24 unique individuals to ACF. The CCOULD dataset was selected to be the focus of the Institute for Research on Poverty’s research grant in 2024. Recipients of this grant and other selected researchers are participating in a research consortium focused on child welfare and Medicaid research currently run by ASPE staff.

<p>Implementing Agency ASPE/ACF</p>	<p>Point of Contact Robin Ghertner, Robin.Ghertner@hhs.gov</p>	<p>Project Profile ASPE Website Project Profile</p>
<p>Strategic Plan Goals Addressed</p>	 <p>Goal 1: Data Capacity for National Health Priorities</p>	 <p>Goal 2: Data Standards and Linkages for Longitudinal Research</p>

Making Electronic Health Record Data More Available for Research and Public Health

This project developed a reference architecture and Substitutable Medical Apps, Reusable Technologies on Fast Healthcare Interoperability Resources application (SMART-on-FHIR® app) for a real-time data exchange system (MedMorph) to improve the availability of clinical EHR data for research and public health in four use cases that represent infectious disease, chronic disease, healthcare services, and research, respectively: hepatitis C, cancer, health care surveys, and research data exchange.

Challenge Addressed

The lack of data interoperability between EHR systems leads to reliance on manual, time-consuming methods for researchers and public health officials to access patient clinical data that reside across different EHR systems.¹⁰ This leads to unreliable and insufficient data exchange processes that inhibit patient-centered outcomes research and public health surveillance.

Project Impact on Expanding Data Capacity for Patient-Centered Outcomes Research



To improve the quality of EHR data used by researchers and public health professionals,

MedMorph improved the completeness and timeliness of EHR data

available, compared to traditional approaches that public health agencies utilize to acquire EHR data.



To improve researchers' and public health systems' access and use of EHR data,

MedMorph leveraged FHIR® standards to improve interoperability and efficiency of

data access, combining different FHIR functions to facilitate data exchange between data senders and authorized data receivers.

Key Resources



[Health Data Exchange App \(HDEA\)](#). The source code and other materials for MedMorph's HDEA FHIR app are publicly available. The HDEA enhanced the existing eCR Now app to accommodate additional types of use cases (the eCR Now app was initially designed for COVID-19 reporting and expanded to most reportable conditions), resulting in one harmonized FHIR app. The app involves the automated identification of health events in EHRs and their transmission to an appropriately authorized entity (e.g., state and local public health authorities, CDC, PCORnet clinical research networks). A harmonized and reusable infrastructure helps reduce implementation burden while the enhanced functionality supports more types of use cases.



[MedMorph Reference Architecture FHIR® Implementation Guide](#). This resource provides a common method for obtaining data for research and public health, which minimizes the burden on both data senders and receivers. This reference architecture Implementation Guide serves as a common framework that can be leveraged for different use cases, which are provided in separate content implementation guides (see below).



[Open-Source Code for MedMorph Reference Architecture HL7 FHIR® Implementation Guide](#). This open-source code is for products and components that inform the MedMorph FHIR Reference Architecture Implementation Guide and related documents, which were balloted and published through HL7.



[MedMorph Central Cancer Registry Reporting Content HL7 FHIR® Implementation Guide](#). This content implementation guide is based on the MedMorph Reference Architecture and supports the cancer reporting use case. This resource is for external organizations interested in adopting the reference architecture for reporting cancer cases to central cancer registries.



[MedMorph Healthcare Surveys Content HL7 FHIR® Implementation Guide](#). This content implementation guide is based on the MedMorph Reference Architecture and supports the healthcare surveys use case. This resource is for external organizations interested in adopting the reference architecture to report healthcare surveys to CDC.



[MedMorph Research Data Exchange Content HL7 FHIR® Implementation Guide](#). This content implementation guide is based on the MedMorph Reference Architecture and supports research use cases. This resource is for external organizations interested in adopting the reference architecture to collect data from research partners.


Long-Term Impacts on Interagency Partnerships and Sustainability

The MedMorph application was tested and implemented in a real-world setting. As part of the OS-PCORTF project activities, the project team implemented the MedMorph Reference Architecture and leveraged the existing eCR FHIR Implementation Guide for the content to report hepatitis C cases to a public health authority (New York City Department of Health and Mental Hygiene, [NYC DOHMH]) and the MedMorph Research Data Exchange Implementation Guide to report research data to a research organization (Westat). This tested end-to-end and demonstrated reuse of the MedMorph Reference Architecture for automated data exchange from the same clinical sites’ EHR (MDLand) to two different data receivers for two different purposes (i.e., public health reporting and research, respectively). Going forward, the MedMorph Reference Architecture infrastructure can be reused by the pilot partners: NYC DOHMH for future FHIR-based public health data collection efforts, Westat for future research data collection efforts, and MDLand and its clinical customers for automating the sending of required data for multiple purposes (e.g., public health reporting, research) to multiple entities.

Multiple federal programs have expressed strong interest in adopting the MedMorph Reference Architecture for their data exchange use cases. Additional pilots in cancer registry reporting, health care surveys, and other use cases have begun. For example, the CDC Respiratory Virus Hospitalization Surveillance Network (RESP-NET) program will develop and pilot a content Implementation Guide for surveillance of hospitalizations from respiratory viruses.

Recognition and Awards

The MedMorph team won the CDC Honor Award for Excellence in Surveillance and Health Monitoring (Domestic) in 2020. MedMorph was also nominated for the CDC Center for Surveillance, Epidemiology, and Laboratory Services’ Honor Award for Excellence in Surveillance and Health Monitoring.

<p>Implementing Agency CDC</p>	<p>Point of Contact Maria Michaels, maria.michaels@ahrq.hhs.gov</p>	<p>Project Profile ASPE Website Project Profile</p>
<p>Strategic Plan Goals Addressed  Goal 3: Technology Solutions to Advance Research</p>		

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

This project builds off the OS-PCORTF **MAT-LINK 1.0** project, which established MAT-LINK, a surveillance system to monitor maternal, infant, and child health outcomes associated with medication for opioid use disorder (MOUD) during pregnancy. The project developed a data abstraction and extraction methodology to gather MOUD and outcome data for pregnant people, infants, and children from EHRs. In the second project iteration, the CDC expanded the sample population and collected additional exposure, risk factor, and outcome variables to improve the quality and comprehensiveness of data for research on long-term outcomes of children with prenatal opioid exposure.

Challenge Addressed

Improving maternal and child outcomes requires strong, high-quality evidence on the long-term outcomes for children with prenatal opioid exposure as well as differences in exposures, risk factors, and outcomes among pregnant people with opioid use disorder (OUD).¹¹ Understanding the causes and contributors to rising maternal mortality rates in the United States (U.S.) has been inhibited by data infrastructure limitations and a lack of data standards for maternal and infant data. Researchers lack a consistent way to collect, share, and link maternal and infant health data to analyze how one's medical history, as well as basic socioeconomic and demographic characteristics, impact pregnancy outcomes for the pregnant person and child.¹²

Project Impact on Expanding Data Capacity for Patient-Centered Outcomes Research



MAT-LINK2 improved the overall quality of the MAT-LINK dataset through several activities, including increasing EHR data completeness and correctness among

participating clinical sites and improving data collection consistency across participating sites. The project improved the representativeness of MAT-LINK data by capturing more diverse populations in terms of geography, race and ethnicity, and insurance status in the dataset.



To provide more relevant, comprehensive data for research on outcomes associated with OUD during pregnancy, MAT-LINK2 expanded MAT-LINK to three new clinical

sites, extended child follow-up through six years of age, and added eight new child follow-up variables capturing health care encounters and referrals, vaccinations, school-readiness information, and screening and evaluation results for existing and new MAT-LINK sites.



To improve researchers' access and use of MAT-LINK data, the project improved the accessibility of MAT-LINK data for a variety of research purposes by creating

multiple tiers of data to control the accessibility, granularity, and quality of the data, and improved data processing and data exchange efficiency.



To enhance analytic resources, MAT-LINK2 implemented machine learning techniques to analyze MAT-LINK data and improve the accuracy of diagnostic data. The project

utilized clustering analysis to more efficiently identify patterns in diagnostic codes within the MAT-LINK dataset. These efforts contribute to CDC's overall data modernization efforts and will be leveraged for future projects.

Key Resources



Expanded MATernal and Infant Clinical Network (MAT-LINK) Dataset. This dataset captures almost **6,000 pregnant woman-infant dyads from 11 clinical sites** across the U.S. The expanded dataset increased the number of treatment options studied, improved the racial/ethnic diversity of the sample population, and added new child follow-up variables.ⁱ The restricted-use dataset is available to researchers through the CDC Research Data Center upon application.



Final Child Follow-Up Data Dictionary. The eight child follow-up variables were added to the MAT-LINK data dictionary. The data dictionary will be provided upon request. Please contact rdca@cdc.gov.

Long-Term Impacts on Interagency Partnerships and Sustainability

Ongoing federal initiatives and OS-PCORTF projects are leveraging or improving the expanded MAT-LINK data. The National Center on Birth Defects and Developmental Disabilities (NCBDD), where MAT-LINK is housed, released a four-year Notice of Funding Opportunity titled, “Pregnant Person—Infant Linked Longitudinal Surveillance” in Spring 2023, that included an expansion for MAT-LINK. This effort aimed to sustain, improve, and expand surveillance efforts from clinical sites and health departments to collect data captured in MAT-LINK.

Another OS-PCORTF project from the NCHS, *Utilizing Natural Language Processing and Machine Learning to Enhance the Identification of Stimulant and Opioid-Involved Health Outcomes in the NHCS*, is using MAT-LINK data as a use case to test machine learning and natural language processing techniques to identify hospital encounters involving the use of illicit stimulants, misuse of prescription stimulants, and the co-involvement of stimulants and illicit or misused prescription opioids. Additionally, external researchers are conducting four ongoing analyses of the MAT-LINK data focusing on three topics: homelessness, post-traumatic stress disorder, and hepatitis C among both mothers and infants.

Awards and Recognition

For their efforts to develop the first surveillance system to collect data on MOUD during pregnancy for pregnant person-infant dyads across clinical sites, MAT-LINK received the Information Technology award in 2022 from the CDC NCBDD for the project’s accomplishments in data pipelining, optimization, and transformation.

Implementing Agency CDC	Point of Contact Shin Kim, dgx5@cdc.gov	Project Profile ASPE Website Project Profile
Strategic Plan Goals Addressed	 Goal 1: Data Capacity for National Health Priorities	 Goal 2: Data Standards and Linkages for Longitudinal Research

ⁱ The expanded MAT-LINK dataset was funded through several mechanisms, including the OS-PCORTF.

Severe Maternal Morbidity and Mortality EHR Data Infrastructure

This project developed the MaternalHealthLink application programming interface (API) and the Longitudinal Maternal & Infant Information for Research HL7 FHIR standard and Implementation Guide (IG) to enable the extraction and exchange of maternal and infant EHR data and vital records data for research on maternal mortality. These data can be used to examine the effect of medical conditions, demographic characteristics, and interventions on health outcomes for pregnant, postpartum or lactating women and their infants, including morbidity, mortality, and health disparities.

Challenge Addressed

Understanding the causes and contributors to rising maternal mortality rates in the U.S. has been hindered by a limited data infrastructure and standards for maternal and infant data. Researchers lack a consistent way to collect, share, and link maternal and infant health data to analyze how one's medical history, as well as socioeconomic and demographic characteristics, impact pregnancy outcomes for mothers and infants.¹³

Project Impact on Expanding Data Capacity for Patient-Centered Outcomes Research



The project developed a new analytic resource, the MaternalHealthLink API to query EHRs for clinical data from patients in the target cohort and subsequently query for the patient's child's data to make these data available for analysis and visualization. The tool was piloted within a health information exchange system as part of project activities.



To provide more relevant, comprehensive data, the project linked clinical data and socioeconomic and demographic information with vital records, enabling researchers to identify the full range of risk factors, pre-existing conditions, and causes of severe maternal morbidity and pregnancy-related mortality, as well as their association with infant outcomes.



To improve researchers' access to and use of maternal and infant health data, the project has made clinical data available through MaternalHealthLink. The Longitudinal Maternal & Infant Information for Research FHIR IG enables researchers to apply the requisite workflows for automating the extraction and exchange of clinical data for maternal and infant health within their own institutions systems.

Key Resources and Final Report



[Longitudinal Maternal & Infant Information for Research FHIR Implementation Guide](#). This IG defines a framework to enable maternal health researchers to aggregate, calculate, and analyze clinical information of research populations to explore the root causes for maternal and child morbidity and mortality.



[MaternalHealthLink](#). This application is an open-source FHIR-based web application for automating the extraction and exchange of clinical data for maternal and infant health, for researcher use.




[Project Final Report](#). This report describes project objectives and corresponding activities and accomplishments, as well as noteworthy dissemination products produced by the project team.

Long-Term Impacts on Interagency Partnerships and Sustainability

Lessons learned from the project have been leveraged in other OS-PCORTF projects and federal initiatives. As part of the project's funded activities, NIH shared lessons learned from the Longitudinal Maternal & Infant Health Information

for Research IG pilot testing in the Kids First Data Resource with the Assistant Secretary for Technology Policy/Office of the National Coordinator for Health Information Technology (ASTP/ONC), which will support their integration of the IG in the Kids First Data Resource platform. Additionally, the OS-PCORTF project, ***Improving Maternal Health by Leveraging Existing Resources, Data Linkage and Standard Reporting for Patient-Centered Outcomes Research***, is testing the IG and will provide additional feedback on the standards.

The project also facilitated the OS-PCORTF Maternal Health Consortium until 2023, which has been extended under ASPE facilitation as Maternal Health 2.0 Consortium and provides a forum for leads of OS-PCORTF projects focused on maternal health to collaborate and provide input on project activities.

<p>Implementing Agency NIH</p>	<p>Point of Contact Nahida Chakhtoura, nahida.chakhtoura@nih.gov</p>	<p>Project Profile ASPE Website Project Profile</p>
<p>Strategic Plan Goals Addressed</p>	<p> Goal 1: Data Capacity for National Health Priorities</p> <p> Goal 3: Technology Solutions to Advance Research</p>	<p> Goal 2: Data Standards and Linkages for Longitudinal Research</p> <p> Goal 4: Person-Centeredness, Inclusion, and Equity</p>



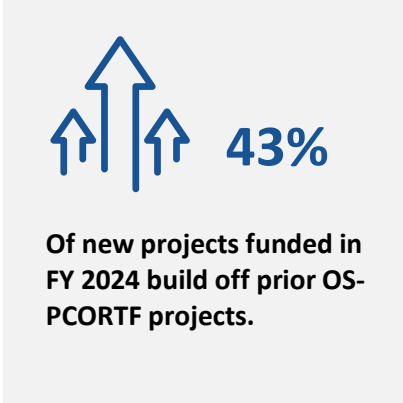
Appendix B: New Projects Funded in FY 2024

This appendix provides an overview of the seven new Office of the Secretary Patient-Centered Outcomes Research Trust Fund (OS-PCORTF) projects funded in Fiscal Year (FY) 2024.

Iterative Data Infrastructure

The accomplishments of past OS-PCORTF products continue to be leveraged in new work. This underscores the impact of the OS-PCORTF portfolio and the consistency and interconnection between projects. In FY 2024, **three of the seven** newly funded projects build off previous OS-PCORTF products and findings to further advance data infrastructure.

- Two projects will link additional data to the [Agency for Healthcare Research and Quality \(AHRQ\) Social Determinants of Health \(SDOH\) Database](#).
- One project will test and provide feedback on the National Institute of Health (NIH) National Institute of Child Health and Human Development (NICHD) [Longitudinal Maternal & Infant Health Information for Research Implementation Guide](#).



The newly funded FY 2024 OS-PCORTF projects will make important contributions to improving data infrastructure for patient-centered outcomes research across four domains: 1) improving the quality of data; 2) providing more relevant, comprehensive data; 3) improving access and use of data; and 4) enhancing analytic resources. The following icons indicate how each of the new projects described below are enhancing data capacity.

Improving the Quality of Data



Providing More Relevant, Comprehensive Data



Improving Access and Use of Data



Enhancing Analytic Resources



New Project Profiles

Establishing the Governance, Legal and Analytical Framework for a Federated Linked Data System: Creating a New Data Research Environment

This project will develop a prototype legal structure and governance framework for a federated linked data system that supports cancer research. This project is one of the four Cancer Moonshot use case workstreams specified in the U.S. Department of Health and Human Services (HHS) Data Strategy.¹⁴

Linked datasets are often required to conduct longitudinal research on cancer and other diseases, but the fragmented nature of data collection poses challenges.¹⁵ A legal, privacy, and data governance framework is needed to facilitate the secure sharing of patient data across databases. However, datasets have varying legal, regulatory, and policy restrictions that govern when and how privacy-protected health data can be used.



To enhance analytic resources, the team will partner with Advanced Research Projects Agency for Health (ARPA-H), to develop a privacy-preserving record linkage (PPRL) methodology to create and store longitudinal records.



The prototype legal structure and governance framework will include a framework for governing data access and use, security and privacy controls, and a sustainable operating model. In conjunction with linkage infrastructure developed by the ARPA-H project, this framework will create a scalable federated data system that makes data accessible to researchers.

<p>Implementing Agency NIH/NCI</p>	<p>Point of Contact Allison Dennis, Allison.Dennis@nih.gov</p>	<p>Project Profile ASPE Website Project Profile</p>
<p>Strategic Plan Goals Addressed</p>	 <p>Goal 1: Data Capacity for National Health Priorities</p>	 <p>Goal 2: Data Standards and Linkages for Longitudinal Research</p>

Improving Maternal Health by Leveraging Existing Resources, Data Linkage and Standard Reporting for Patient-Centered Outcomes Research

This project will expand data elements included in the United States Core Data for Interoperability Plus (USCDI+) Maternal Health initiative, which guides the development and adoption of data standards necessary for maternal health research at the national level. The resulting USCDI+ maternal health dataset will provide essential guidance for data standardization for maternal health data and is part of a larger health information technology (IT) alignment effort led by the Assistant Secretary for Technology Policy/Office of the National Coordinator for Health Information Technology (ASTP/ONC) that aims to improve the efficacy and efficiency of harmonizing data across HHS programs.¹⁶




In the United States (U.S.), 80 percent of pregnancy-related deaths are preventable,¹⁷ highlighting the need for continued research on maternal health treatments and outcomes. However, the data currently available for maternal health research, surveillance, and program reporting are not standardized, which hinders the research necessary to update clinical guidelines and improve maternal health outcomes.¹⁸



This project will expand the USCDI+ MH dataset to include 28 new data elements related to postpartum, labor and delivery, behavioral health, interventions, mortality, and lactation.



The standards created from the expanded USCDI+ MH dataset will enhance data interoperability and availability to support decision-making related to maternal health care. The project will also test and provide feedback on the Longitudinal Maternal & Infant Health Information for Research IG developed by the OS-PCORTF project ***Severe Maternal Morbidity and Mortality EHR Data Infrastructure***.

<p>Implementing Agency ASTP/ONC</p>	<p>Point of Contact Kathleen Tully, Kathleen.Tully@hhs.gov</p>	<p>Project Profile ASPE Website Project Profile</p>
<p>Strategic Plan Goals Addressed</p>	<p> Goal 1: Data Capacity for National Health Priorities</p> <p> Goal 4: Person-Centeredness, Inclusion, and Equity</p>	<p> Goal 2: Data Standards and Linkages for Longitudinal Research</p>

Improving the Availability of National Mortality Data by Accelerating Data Release and Laying the Groundwork for IT Modernization of the National Death Index

This project will improve the timeliness and quality of mortality data by increasing the frequency of National Death Index (NDI) data releases and laying the groundwork for modernization and potential cloud migration of the NDI.

PCOR investigators need to be able to efficiently link data with an accurate, reliable source of mortality data.¹⁹ The NDI is the only national data source containing cause of death information that can be used for linkages with research datasets.²⁰ However, data available through the NDI may be up to 24 months out of date, and the NDI does not yet leverage modern cloud-based IT infrastructure, which creates additional barriers to data access and use.²¹



This project will improve the timeliness of mortality data available for research by increasing the frequency of NDI data releases from annual to monthly. The project will also develop an automated process for creating tables of estimated NDI completeness for end users.



To support the increased frequency of NDI releases, and to make NDI data more accessible to researchers, the project will develop a cloud-based migration strategy and roadmap for the NDI IT infrastructure. Migrating the NDI IT infrastructure to the cloud will allow for modernization of the database and enable a more secure, flexible, and scalable technology infrastructure

Implementing Agency

CDC

Point of Contact

Paul Sutton, PV59@cdc.gov

Strategic Plan Goals Addressed



Goal 2: Data Standards and Linkages for Longitudinal Research

Leveraging Data Standards for Patient-Centered Outcomes Research in Sickle Cell Disease under the United States Core Data for Interoperability Plus Program

This project will strengthen data infrastructure for patient-centered outcomes research in sickle cell disease by resolving gaps in the USCDI+ standards and facilitating the harmonization of shared data elements across agencies.




Sickle cell disease is an inherited blood disorder that affects about 100,000 people in the U.S. and is associated with substantial morbidity and decreased quality of life.^{22,23} Additional research is needed to understand the full scope of treatments, medical services use, and outcomes among people living with sickle cell disease. However, data interoperability poses a challenge for research as there is not currently a standard for data collection and exchange across HHS sickle cell disease programs.²⁴



To improve the quality of data available for research, this project will develop a minimum set of USCDI+ data elements for sickle cell disease. These data elements will be standardized with vocabulary standards, data formats, and value sets, as needed.



The project intends to capture the data needs specific to the sickle cell disease population to support standardized sickle cell disease data exchange and harmonization of sickle cell disease data elements. The steps involved include identifying the use case, listing the sickle cell disease data elements, and developing the FHIR implementation guide. The FHIR implementation guide will support the interoperability of data elements across systems and provide researchers the ability to identify the full range of factors that affect patients with sickle cell disease. Additionally, the sickle cell disease dataset will undergo testing.

Implementing Agency ASTP/ONC		Point of Contact LaVerne Perlie, Laverne.Perlie@hhs.gov	
Strategic Plan Goals Addressed		Goal 1: Data Capacity for National Health Priorities	
		Goal 4: Person-Centeredness, Inclusion, and Equity	Goal 2: Data Standards and Linkages for Longitudinal Research

Linking the Boston Birth Cohort and Pregnancy to Early Life Longitudinal Data System and Social Determinants of Health Databases: A Longitudinal Cohort of Mother-Child Dyads

This project will link data among a cohort of underserved and underrepresented populations in Boston, Massachusetts to generate a longitudinal data resource for assessing patient-centered maternal and infant health outcomes and health equity from a life course perspective.²⁵




Maternal mortality and severe maternal morbidity rates in the U.S. remain high,²⁶ and multiple factors contribute to racial and ethnic disparities.²⁷ Researchers need to better understand the factors that impact the health of mothers and children across the lifespan. However, there is a lack of longitudinal data in the U.S. linking mother-child dyads with multi-sectoral data to assess maternal health from a life course perspective, especially among vulnerable populations.²⁸



This project will link data from the Boston Birth Cohort study with Massachusetts’ Pregnancy to Early Life Longitudinal Data System and AHRQ’s SDOH Database. This linked longitudinal dataset will provide more comprehensive information for studying the factors across the life course that affect maternal and infant health outcomes.



A de-identified version of the linked dataset along with a codebook will be available to researchers using a tiered access approach. The project will also develop a technical report describing the data linkage approach, available data elements, and potential issues with the linked data.

Implementing Agency HRSA/MCHB	Point of Contact Rui Li, rli@hrsa.gov	Project Profile ASPE Website Project Profile
Strategic Plan Goals Addressed	 <p>Goal 1: Data Capacity for National Health Priorities</p>  <p>Goal 4: Person-Centeredness, Inclusion, and Equity</p>	 <p>Goal 2: Data Standards and Linkages for Longitudinal Research</p>

Medical Expenditure Panel Survey Enhancements on Burdens and Economic Impacts of Medical Care

This project will expand the Medical Expenditure Panel Survey (MEPS) to promote data capacity for patient-centered outcomes research on economic burdens and outcomes. The project will enhance burden and economic impact data already available in MEPS by creating a supplemental self-administered questionnaire from the patient and caregiver perspective.


Economic impacts of care are an important equity issue, and it is important to understand economic cost impacts of health care in the context of other social risk factors.²⁹ However, more data are needed to fully examine health care economic impacts and equity in health care access, utilization, and outcomes.³⁰



This project will provide more relevant, comprehensive economic outcomes data by creating and implementing a self-administered questionnaire that collects new data on economic burdens related to seeking and receiving healthcare (e.g., indirect costs of care) from the patient and caregiver perspective.



This project will also facilitate researchers' ability to study associations between community characteristics and economic impacts by making the OS-PCORTF-funded AHRQ SDOH Database files available directly in the MEPS Data Center.

<p>Implementing Agency AHRQ</p>	<p>Point of Contact Patricia Keenan, patricia.keenan@ahrq.hhs.gov</p>	<p>Project Profile ASPE Website Project Profile</p>
<p>Strategic Plan Goals Addressed</p>	 <p>Goal 4: Person-Centeredness, Inclusion, and Equity</p>	

N3C Renal Freeport: A Collaborative Data Analytics and Data Sharing Model for Patient-Centered Outcomes Research

This project will develop and evaluate the National Clinical Cohort Collaborative (N3C) Renal Freeport, a secure data sharing infrastructure containing linked datasets of real-world evidence related to kidney disease. This project builds on the N3C renal tenant pilot conducted by NIH National Center for Advancing Translational Sciences (NIH/NCATS) and aligns with a 2019 Executive Order on Advancing American Kidney Health³¹ by enabling patient-centered outcomes research on health disparities for patients with chronic kidney disease (CKD), end-stage renal disease (ESRD), and kidney transplants.


In the U.S., CKD is a common disorder with significant disparities in treatments and outcomes, including progression to ESRD and kidney transplantation.³² Previous HHS efforts to compile data related to CKD have been limited to claims and cost of care data, but additional EHR and registry data are needed to enable better longitudinal tracking of individuals.



The project will link de-identified EHR data, CMS claims data, and the Scientific Registry for Transplant Recipients data, using linkage processes developed for the N3C COVID Enclave. The linked data will address challenges with data fragmentation to provide more comprehensive data on the health care journey for CKD, ESRD, and kidney transplant patients.



The N3C Renal Freeport will develop a governance structure and technical requirements to facilitate data sharing across federal and non-federal organizations. The resulting data enclave will provide a secure workplace for researchers to access longitudinal data, perform analyses, and share results.

<p>Implementing Agency NIH/NCATS</p>	<p>Point of Contact Sam Michael, michaelsg@mail.nih.gov</p>	<p>Project Profile ASPE Website Project Profile</p>
<p>Strategic Plan Goals Addressed  Goal 2: Data Standards and Linkages for Longitudinal Research</p>		



Appendix C: In Progress Projects in FY 2024

This appendix provides a brief overview of the 27 Office of the Secretary Patient-Centered Outcomes Research Trust Fund projects in progress during Fiscal Year (FY) 2024. Projects are organized by primary Strategic Plan goal alignment.

Projects in Progress in FY 2024



Goal 1: Improved Data Capacity for National Health Priorities

Building Infrastructure and Evidence for COVID-19 Related Research, Using Integrated Data (CDC/NCHS). The project aims to create linked synthetic data files that integrate social determinants of health (SDOH) data and other health-related data collected through National Center for Health Statistics (NCHS) surveys, with a focus on addressing key concerns related to COVID-19.

Capturing the Full Trajectory of Patient-Centered Cancer Care Via Enhanced Data Linkages (NIH/NCI and NIH/NCATS). This project will create and evaluate longitudinal, linked datasets from cancer registry and electronic health record (EHR) data that capture the full trajectory of cancer patients' care and that can be used by researchers to understand the impacts of treatments on patient outcomes.

Enhancing Surveillance of Maternal Health Clinical Practices and Outcomes with Federally Qualified Health Centers' Electronic Health Records Visit Data (CDC/NCHS). This project will improve the collection and utility of EHR data from Federally Qualified Health Centers (FQHCs) and link FQHC EHR data to federal data sources on mortality and SDOH data to enable the study of health care utilization on maternal health and COVID-19-related outcomes.

Improving Quality of Care & Outcomes Data for Pregnant Medicaid Beneficiaries & Newborn Infants (CMS). This project will increase the amount of comprehensive data available on the impact of prenatal care on birth outcomes for mothers and infants by linking vital records to Transformed Medicaid Statistical Information System (T-MSIS) data.

Linking Federally Qualified Health Center EHR and Medicaid Data for Increased Data Capacity to Understand Maternal Health Care (CDC/NCHS). The project will link FQHC EHR data to T-MSIS to enhance researchers' ability to analyze data on maternal health care provided at FQHCs and subsequently gain insight into the relationships among patient characteristics, service utilization, and outcomes for maternal health visits.

Understanding Long-term Outcomes in COVID-19 Survivors with Multiple Chronic Conditions (MCC) through eCare Plan Development (AHRQ and NIH/NIDDK). This project will improve the availability of comprehensive data for people living with MCC who experience COVID-19 by developing and testing an expanded electronic care (eCare) Plan app that capture COVID-19 data elements.



Goal 2: Enhancing Data Standards and Linkages for Longitudinal Research

[Code Map Services for Interoperability of Common Data Models and Data Standards \(NIH/NCATS\)](#). This project will create an automated set of code mapping services of common data models (CDMs) and tools to promote interoperability and data sharing for patient-centered outcomes research (PCOR).

[Compendium of Policies on Use of Linkable Mortality Data for PCOR in the Nation's 57 Mortality-Reporting Jurisdictions \(ASPE/BHDAP\)](#). This project will conduct a systematic inventory of the formal policies on the use of linkable mortality data for patient-centered outcomes research for each of the 57 mortality-reporting jurisdictions, making it easier and less expensive for PCOR investigators to determine the best ways to obtain data on survival and mortality for their studies.

[Creating a Federal COVID-19 Longitudinal Patient Outcomes Research Database Linked to Health Systems and Clinical Data \(NIH/NCATS\)](#). This project will build a national longitudinal data infrastructure on COVID-19 and improve comprehensiveness of the National Clinical Cohort Collaborative (N3C) dataset through linkages between Centers for Medicare & Medicaid Services (CMS) claims data and EHR data in the N3C Data Enclave.

[Digitizing Consent and Regulatory Metadata Towards Streamlining Governance of Pediatric COVID-19 Research Data Linkages \(NIH/NICHHD\)](#). Based on pediatric COVID-19 studies, this project will develop a data governance approach that standardizes and digitizes consent metadata to help researchers determine dataset appropriate use of datasets.

[Evaluation of Privacy-Preserving Record Linkage Solutions to Broaden Linkage Capabilities in Support of Patient-Centered Outcomes Research Objectives \(CDC/NCHS\)](#). To support future linkages between administrative datasets and diverse data sources, this project will evaluate the effectiveness of three privacy-preserving record linkage tools.

[Linking the National Survey of Children's Health Longitudinal Cohort and Medicaid Claims Data to Explore Pediatric COVID-19 \(HRSA/MCHB\)](#). This project will link National Survey of Children's Health Longitudinal Cohort Study data to T-MSIS to provide more comprehensive data on the long-term impacts of COVID-19 on children and families.

[Making Medicaid Data More Accessible Through Common Data Models and FHIR APIs \(NIH/NLM and FDA\)](#). This project will improve data quality and enable multi-database studies through the development of open-source code to format T-MSIS data into a standardized CDM format.

[Multistate Emergency Medical Services \(EMS\) and Medicaid Dataset \(MEMD\): A Linked Dataset for PCOR \(ASPE/BHDAP\)](#). This project will create a publicly available dataset linking Medicaid and emergency medical services (EMS) records in five states, providing more comprehensive data regarding health outcomes among individuals receiving EMS.

[Panoramic View of Patient Care Through Data Innovations and Linkages Implementing \(AHRQ\)](#). This project will link data across multiple sources, including social determinants of health (SDOH) data and community information, to capture a panoramic view of patient care in at least four states, providing richer, more complete data to answer PCOR questions related to health care coverage and health outcomes.



Goal 3: Technology Solutions to Advance Research

[Using Machine Learning Techniques to Enable Health Information Exchange \(HIE\) Data Sharing to Support COVID-19-focused PCOR \(ASTP/ONC\)](#). This project will leverage data standardization, application programming interfaces,

and privacy-preserving machine learning techniques to facilitate the use of EHR data from health information exchanges for research, leveraging COVID-19 as a use case.

Utilizing Natural Language Processing and Machine Learning to Enhance the Identification of Stimulant and Opioid-Involved Health Outcomes in the National Hospital Care Survey (CDC/NCHS). To provide researchers access to more relevant, high-quality data on stimulant-involved health encounters, this project will develop an algorithm using natural language processing and machine learning to identify hospital encounters involving stimulants in both structured and unstructured hospital data.



Goal 4: Person-Centeredness, Inclusion, and Equity

Childhood Obesity Data Initiative (CODI): Integrated Data for Patient-Centered Outcomes Research Project 2.0 (CDC). This project will refine CODI's existing data tools and services to promote data standardization, improving researchers' ability to combine pediatric patient-level EHR, weight management intervention, and community (census) data.

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 (NIH/NIDDK and AHRQ). Through development of a patient-facing eCare Plan app and expansion of an existing eCare Plan Fast Healthcare Interoperability Resources implementation guide, this project will enhance the availability of patient-centric information on multiple chronic conditions across sources and enhance researchers' ability to access data from EHRs.

Data Infrastructure Supporting Research on Refugee Medicaid Service Utilization and Outcomes (ACF). This project will create a linked dataset of Refugee Arrivals Data System and T-MSIS data from 2016-2019 that provides more comprehensive data on refugee service utilization and health outcomes, strengthening the data infrastructure for conducting PCOR studies with refugee populations.

Dataset on Intellectual and Developmental Disabilities: Linking Data to Enhance Person-Centered Outcome Research (ASPE/BHDAP). To provide more relevant comprehensive data for research on person-centered outcomes for the intellectual and developmental disabilities (ID/DD) population, this project will link Support Intensity Scale scores, Medicaid claims, National Core Indicators survey data, National Core Indicators COVID-19 supplement survey data, and other relevant state-level data for four to six states.

Engaging People with Intellectual and Developmental Disabilities to Enhance Functional Disability Representation in Point of Care Settings (CDC/NCBDDD). Through development of a functional disability demographic checklist, this project will facilitate identification of people with intellectual and developmental disabilities (ID/DD) in administrative data systems and allow for collection of ID/DD data that were not previously available.

Expanding Climate Change and Health Data Infrastructure to Advance Health Interventions: Linking Health and Environmental Data to Improve Patient and Community Health (NIH/NIEHS). This project will provide researchers with accessible, timely, harmonized, interoperable, and linkable data by developing a publicly accessible catalog of curated data sources, tools, methods, and educational resources related to climate change and health.

Small-Area Community SDOH Data: Enhancements and Linkages to Inform Action (AHRQ). This project will improve the granularity and comprehensiveness of data available in the Agency for Healthcare Research and Quality SDOH Database by adding more recent years of data and adding new SDOH variables, as well as standardize data in the database to enhance its usability for researchers.

References

- ¹ Robinson PC, Liew DFL, Tanner HL, et al. COVID-19 therapeutics: Challenges and directions for the future. *Proc Natl Acad Sci U S A*. 2022;119(15):e2119893119. doi:10.1073/pnas.2119893119
- ² The Society of Critical Care Medicine. REDISCOVER-ICU. Accessed November 4, 2024. <https://www.sccm.org/research/discovery-research-network/discovery-research-collaboratives/rediscover-icu>
- ³ Harron K, Dibben C, Boyd J, et al. Challenges in administrative data linkage for research. *Big Data Soc*. 2017;4(2):2053951717745678. doi:10.1177/2053951717745678
- ⁴ Centers for Disease Control and Prevention. Pregnancy Risk Assessment Monitoring System. Accessed October 31, 2024. <https://www.cdc.gov/prams/index.html>.
- ⁵ Heins S, Grigorescu V, Predmore Z, Zhou A, Hoch E., Smith S. *Linking State Medicaid Data and Birth Certificates For Maternal Health Research*. Office of the Assistant Secretary for Planning and Evaluation, U.S. Department of Health and Human Services. October 2023. <https://aspe.hhs.gov/sites/default/files/documents/147eac15e9e1a004d5650a400d4fcf51/state-medicaid-data-linkages.pdf>
- ⁶ Research Data Assistance Center. Research Identifiable File (RIF) Requests. Accessed October 31, 2024. <https://resdac.org/research-identifiable-files-rif-requests>.
- ⁷ Ahrens KR, Garrison MM, Courtney ME. Health outcomes in young adults from foster care and economically diverse backgrounds. *Pediatrics*. 2014;134(6):1067-1074. doi:10.1542/peds.2014-1150
- ⁸ Radel L, Baldwin M, Crouse G, Ghertner R, Waters A. *Substance use, the opioid epidemic, and the child welfare system: Key findings from a mixed methods study*. Office of the Assistant Secretary for Planning and Evaluation; March 2018. Accessed October 7, 2024. https://aspe.hhs.gov/sites/default/files/migrated_legacy_files//179966/SubstanceUseChildWelfareOverview.pdf
- ⁹ Mathematica. Child Welfare and Health Infrastructure for Linking and Data Analysis of Resources, Effectiveness, and Needs (CHILDREN) Initiative. Accessed November 4, 2024. <https://www.mathematica.org/projects/child-welfare-and-health-infrastructure-for-linking-and-data-analysis-of-resources>
- ¹⁰ Roman LC, Ancker JS, Johnson SB, Senathirajah Y. Navigation in the electronic health record: A review of the safety and usability literature. *J Biomed Inform*. 2017;67:69-79. doi:10.1016/j.jbi.2017.01.005
- ¹¹ Mascola MA, Borders AE, Terplan M. Opioid use and opioid use disorder in pregnancy. *Obstetrics and Gynecology*. 2017;130(2):488-489.
- ¹² MacDorman MF, Declercq E. The failure of United States maternal mortality reporting and its impact on women's lives. *Birth*. 2018;45(2):105-108. doi:10.1111/birt.12333
- ¹³ National Institute of Child Health and Human Development. Item of Interest: NICHD-Developed Guide Sets Framework to Link Maternal, Infant Health Data. National Institute of Child Health and Human Development; June 2023. <https://www.nichd.nih.gov/newsroom/news/062623-maternal-guide#:~:text=Research%20increasingly%20shows%20that%20maternal,share%20knowledge%20about%20their%20efforts>. Accessed October 31, 2024.
- ¹⁴ U.S. Department of Health and Human Services Office of the Chief Data Officer. Data Strategy 2023-2028. Accessed October 30, 2024. <https://cdo.hhs.gov/s/hhs-data-strategy>
- ¹⁵ National Institutes of Health National Cancer Institute. National Cancer Plan. April 2023. <https://nationalcancerplan.cancer.gov/national-cancer-plan.pdf>
- ¹⁶ Office of the National Coordinator for Health Information Technology, HHS Office of the Secretary. HHS Health IT Alignment Policy. Accessed October 30, 2024. <https://www.healthit.gov/topic/hhs-health-it-alignment-policy>
- ¹⁷ Centers for Disease Control and Prevention Newsroom. Four in 5 pregnancy-related deaths in the U.S. are preventable. Available at: <https://www.cdc.gov/media/releases/2022/p0919-pregnancy-related-deaths.html>. Accessed October 30, 2024.

- ¹⁸ Declercq E, Thoma M. Measuring US Maternal Mortality. *JAMA*. 2023;330(18):1731-1732. doi:10.1001/jama.2023.19945
- ¹⁹ Eisenstein EL, Prather K, Greene SJ, et al. Death: The Simple Clinical Trial Endpoint. *Stud Health Technol Inform*. 2019;257:86-91.
- ²⁰ National Center for Health Statistics. National Death Index (NDI) Non-Economic Barriers Report. National Center for Health Statistics, Centers for Disease Control and Prevention; September 2019.
<https://aspe.hhs.gov/sites/default/files/private/pdf/259016/NDI-Non-Economic-Barriers-Report-508.pdf>
- ²¹ NIH Collaboratory on Pragmatic Trials Textbook, Section 4: Using Death as an Endpoint. NIH Collaboratory, National Institutes of Health. <https://rethinkingclinicaltrials.org/chapters/design/choosing-specifying-end-points-outcomes/using-death-as-an-endpoint/>
- ²² Lee L, Smith-Whitley K, Banks S, Puckrein G. Reducing Health Care Disparities in Sickle Cell Disease: A Review. *Public Health Rep*. 2019 Nov/Dec;134(6):599-607. doi: 10.1177/0033354919881438.
- ²³ Krishnamurti L. Gaps in Medicaid Coverage and Financial Toxicity: A Health Disparity in Hematopoietic Cell Transplantation for Sickle Cell Disease. *Transplant Cell Ther*. 2021;27(4):284-285. doi:10.1016/j.jtct.2021.03.016
- ²⁴ Perlie L, Wong D. Patient Centered Outcomes Research Health Equity Consortium: Sickle Cell Disease under the United States Core Data for Interoperability Program (USCDI+). Presented May 28 2024.
- ²⁵ Fine A, Kotelchuck M. Rethinking MCH: The life course model as an organizing framework. US Department of Health and Human Services, Health Resources and Services Administration, Maternal and Child Health Bureau. 2010 Nov:1-20.
- ²⁶ The White House. White House Blueprint for Addressing the Maternal Health Crisis. The White House; June 2022.
<https://www.whitehouse.gov/wp-content/uploads/2022/06/Maternal-Health-Blueprint.pdf>
- ²⁷ Crear-Perry J, Correa-de-Araujo R, Lewis Johnson T, McLemore MR, Neilson E, Wallace M. Social and Structural Determinants of Health Inequities in Maternal Health. *J Womens Health (Larchmt)*. 2021;30(2):230-235. doi:10.1089/jwh.2020.8882
- ²⁸ Assistant Secretary for Planning and Evaluation. *Enhancing Data Infrastructure to Improve Women's Health Outcomes*. Office of the Assistant Secretary for Planning and Evaluation, U.S. Dept of Health and Human Services October 2022.
<https://aspe.hhs.gov/sites/default/files/documents/Ofce1126682bfe8e00e40b8b198671e3/Women-Health-Vignette.pdf>
- ²⁹ Daras LC, Bozzolo C, Boyer N, et al. *Patient-Centered Outcomes Research Institute's Public Comment Period on the Full Range of Outcomes: Summary and Potential Considerations*. Patient-Centered Outcomes Research Institute. February 2021.
<https://www.pcori.org/sites/default/files/PCORI-Principles-for-Consideration-of-Full-Range-of-Outcomes-Data-in-PCORI-Funded-Research-Public-Comment-Report.pdf>
- ³⁰ Brown D, Srinivasan M, Arbulu L, Dullabh P, Curtis K, Zott C, and Smith SR. *Federal Data for Conducting Patient-Centered Outcomes Research on Economic Outcomes*. Office of the Assistant Secretary for Planning and Evaluation, U.S. Department of Health and Human Services. September 2021.
<https://aspe.hhs.gov/sites/default/files/documents/701038d7cdc48c98daf05ddef163cd86/EconomicVariablesWhitePaper.pdf>
- ³¹ Exec. Order No. 13879. 2019. <https://www.federalregister.gov/d/2019-15159>
- ³² Crews DC, Liu Y, Boulware LE. Disparities in the burden, outcomes, and care of chronic kidney disease. *Curr Opin Nephrol Hypertens*. 2014;23(3):298-305. doi: 10.1097/01.mnh.0000444822.25991.f6.