REPORT

Understanding the Impacts of OS-PCORTF Projects on Data Capacity: An Interim Qualitative Assessment

Prepared for
**The Office of the Assistant Secretary for Planning and Evaluation (ASPE)
at the U.S. Department of Health and Human Services**

by
NORC at the University of Chicago

September 2023

OFFICE OF THE ASSISTANT SECRETARY FOR PLANNING AND EVALUATION

The Assistant Secretary for Planning and Evaluation (ASPE) advises the Secretary of the U.S. Department of Health and Human Services (HHS) on policy development in health, disability, human services, data, and science; and provides advice and analysis on economic policy. ASPE leads special initiatives; coordinates the Department's evaluation, research, and demonstration activities; and manages cross-Department planning activities such as strategic planning, legislative planning, and review of regulations. Integral to this role, ASPE conducts research and evaluation studies; develops policy analyses; and estimates the cost and benefits of policy alternatives under consideration by the Department or Congress.

THE OFFICE OF HEALTH POLICY

The Office of Health Policy (HP) provides a cross-cutting policy perspective that bridges Departmental programs, public and private sector activities, and the research community, in order to develop, analyze, coordinate and provide leadership on health policy issues for the Secretary. HP carries out this mission by conducting policy, economic and budget analyses, assisting in the development and review of regulations, assisting in the development and formulation of budgets and legislation, and assisting in survey design efforts, as well as conducting and coordinating research, evaluation, and information dissemination on issues relating to health policy.

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

The Office of the Secretary Patient-Centered Outcomes Research Trust Fund (OS-PCORTF) was established as part of the 2010 Patient Protection and Affordable Care Act and is charged to build data capacity for patient-centered outcomes research. Coordinated by ASPE on behalf of the Department, OS-PCORTF has funded a rich portfolio of projects to meet emerging U.S. Department of Health and Human Services policy priorities and fill gaps in data infrastructure to enhance capabilities to collect, link, and analyze data for patient-centered outcomes research. For more information, visit https://aspe.hhs.gov/collaborations-committees-advisory-groups/os-pcortf

Suggested Citation: Evans, E., Srinivasan, M., Correa, K., Ryan, S., Harmsen, M., Lumsden, S., and Dullabh, P. Understanding the Impacts of OS-PCORTF Projects on Data Capacity: An Interim Qualitative Assessment. Office of the Assistant Secretary for Planning and Evaluation, U.S. Department of Health and Human Services. September 2023.

CONTRIBUTING AUTHORS

Emily Evans, PhD, MPH, Health Science Analyst, ASPE

Mithuna Srinivasan, PhD, Principal Research Scientist, NORC

Kiran Correa, MPP, Senior Research Associate II, NORC

Sofia Ryan, MSPH, Senior Research Associate I, NORC

Mikaela Harmsen, MPH, Research Associate II, NORC

Susan Lumsden, MS, Social Science Analyst, ASPE

Prashila Dullabh, MD, Vice President and Senior Fellow, NORC

KEY INFORMANTS[*]

Arlene Bierman (AHRQ)

Valeria Butler (ACF)

Keith Campbell (FDA)

Carol DeFrances (CDC)

Stephanie Garcia (ONC)

Robin Ghertner (ASPE)

Cordell Golden (CDC)

Cate Green (CDC)

Geoffrey Jackson (CDC)

Shin Kim (CDC)

Danica Marinac-Dabic (FDA)

Kate Miele (CDC)

Jenna Norton (NIH)

Pam Owens (AHRQ)

Gregory Pappas (FDA)

Papia Paul (ONC)

Marta Steliac (FDA)

PROJECT OFFICERS AND PROJECT LEADERSHIP

Sara Wei (ASPE)

Susan Lumsden (ASPE)

Emily Evans (ASPE)

Prashila Dullabh (NORC)

Rina Dhopeshwarkar (NORC)

Mithuna Srinivasan (NORC)

---

[*] We acknowledge the contributions of additional project team members during key informant interviews.

Table of Contents

Table of Exhibits

Executive Summary

Patient-centered outcomes research (PCOR) aims to generate evidence about the outcomes and effectiveness of treatments, services, and other health care interventions to support informed decisions by patients, caregivers, clinicians, and policymakers. Robust data capacity and infrastructure are integral to this evidence generation; conducting PCOR studies requires timely access to relevant, high-quality data that can be analyzed with rigorous and appropriate research methods. Within the U.S. Department of Health and Human Services (HHS), the Office of the Secretary Patient-Centered Outcomes Research Trust Fund (OS-PCORTF) works to address these critical data needs, funding a portfolio of projects that improve the capacity for collection, linkage, and analysis of data for patient-centered outcomes research.

To measure progress in advancing data capacity for PCOR studies, and ensure its efforts remain responsive to evolving data infrastructure needs, priorities, and other relevant developments, the OS-PCORTF conducts regular assessments of its portfolio. This report presents findings from an interim qualitative assessment that examined the ways in which a selected set of OS-PCORTF projects collectively advanced the ability of researchers to conduct PCOR studies by (1) improving the quality of data; (2) providing more relevant, comprehensive data; (3) enhancing analytical resources; and (4) reducing barriers to data access and use. Together, the contributions of OS-PCORTF projects across these four domains will enable PCOR researchers to address questions important to patients, caregivers, clinicians, and policymakers in a more robust and rigorous manner.
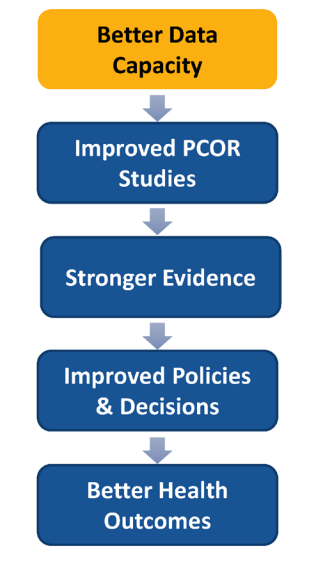
1. Introduction

Patients, caregivers, clinicians, and policymakers frequently seek valid scientific evidence to inform decisions about the health outcomes associated with different options in the use of health care.[1] In particular, strong evidence is needed about the outcomes and effectiveness of treatments, services, and other health care interventions across diverse patient groups and settings of care. Patient-centered outcomes research (PCOR) aims to generate this evidence, but researchers often lack timely access to relevant high-quality data that can be transformed into actionable evidence with rigorous and appropriate research methods.

Within the U.S. Department of Health and Human Services (HHS), the Office of the Secretary Patient-Centered Outcomes Research Trust Fund (OS-PCORTF) works to address these critical data needs, funding a portfolio of projects that improve the capacity for collecting, linking, and analyzing data for patient-centered outcomes research.[2,3] The OS-PCORTF also conducts regular assessments of its portfolio to understand the ways in which its efforts support the generation of evidence to address questions and inform decisions important to patients, caregivers, clinicians, and policymakers (Exhibit 1) and to ensure responsiveness to evolving data infrastructure needs, national health priorities, legislative or policy changes, and advances in health care and data science.

For this report, we conducted an interim qualitative assessment of the contributions made by a set of nine OS-PCORTF projects across four core areas (hereafter, "domains") of expected impact on PCOR data capacity: [a]

- **Domain 1: Improving the quality of data** for PCOR studies by addressing key aspects of data quality that affect research conclusions (for example, completeness, accuracy, and consistency).

- **Domain 2: Providing more relevant, comprehensive data** to address PCOR questions, as reflected in improved availability of data types and sources (for example, new variables and linked and/or longitudinal data).

- **Domain 3: Enhancing analytic resources** for PCOR studies, as reflected in the development of new methods for linking and analysis, and/or improved understanding of the strengths and limitations of such approaches.

- **Domain 4: Reducing barriers to data access and use** to promote timelier PCOR studies, as reflected in lower costs, non-duplication of efforts (for example, making linked datasets available), and more efficient data sharing (for example, interoperability) and access (for example, governance).

Exhibit 1. Pipeline from Improving Data Capacity to Improving Health Outcomes (simplified)

Better Data Capacity → Improved PCOR Studies → Stronger Evidence → Improved Policies & Decisions → Better Health Outcomes

The goal of this assessment was to better understand the more direct effects—as captured in the four domains—of projects funded in Fiscal Year (FY) 2018 and FY 2019 (and thus recently completed or nearly complete) as well as to inform future assessment efforts. Generally, impacts in these domains are what can be expected and observed in this timeframe, given the upstream position of efforts to improve data capacity and infrastructure. A more comprehensive assessment, including ascertaining longer-term impacts, such as uptake of data products in PCOR studies or incorporation of evidence into policies, requires additional time for observation and analysis.

---

[a] The nine projects were chosen to represent a cross-section of the OS-PCORTF portfolio (HHS agencies, impact domains, and range of work funded). Projects not selected for inclusion in the report also made important contributions across the four domains. Additional information on the selected projects (including the data products produced) is available in Appendix A, and additional discussion of the assessment approach is available in Appendix B.

2. Findings

We conducted a qualitative impact assessment, comprising a targeted review of project materials and interviews with project leads and other team members. This section presents selected findings, organized by impact domain.

Domain 1: Improving Quality of Data Available for PCOR Studies

Data quality is a multidimensional concept,[4,5] broadly defined as the "degree to which the data capture the desired information using appropriate methodology in a manner that sustains public trust."[6] Dimensions of data quality that affect research conclusions significantly include *completeness* (i.e., presence of necessary data), *accuracy* (i.e., closeness between the data and true value), and *consistency* (i.e., uniformity in data across settings).[4] Other key dimensions include *utility* (i.e., usefulness of data for the end users), *objectivity* (i.e., data are reliable and unbiased), and *integrity* (i.e., data adhere to appropriate scientific standards and are protected from manipulation and unauthorized access).[6]

Generating evidence to inform decisions and improve health outcomes depends on the ability of PCOR researchers to extract and use high-quality data from a range of sources, including electronic health records (EHRs), medical billing claims, and patient-provided information. Data collected for purposes other than research may not be suitable for PCOR studies (or other types of health-related research) without significant assessment, cleaning, and transformation efforts. Moreover, even the most advanced study design and analytical approaches cannot overcome the limitations of low-quality data.

OS-PCORTF projects have improved the quality of data available for research, which is critical to the credibility of findings generated by PCOR studies:

- The ***Medical Needs in a Disaster* project**[7] improved the utility of the Agency for Healthcare Research and Quality's (AHRQ) Healthcare Cost and Utilization Project (HCUP) databases by: (1) acquiring and processing more timely (quarterly rather than annual) inpatient and outpatient data from eight states prioritized as being at most risk for natural disasters; (2) generating standardized and uniformly formatted quarterly data files for analytic and research purposes; and (3) combining encounter-level HCUP data on hospital inpatient and emergency department utilization with data from other federal data sources, including hurricane and weather-related data from the National Oceanic and Atmospheric Administration (NOAA) and community vulnerability data from the Centers for Disease Control and Prevention (CDC). The resulting improvements in data quality—timeliness, consistency (standardization), and utility (including linkage)—have supported informed disaster response and recovery operations, including more effective deployment of resources in advance of Hurricane Florence.[8] The data were also provided to partners early during the COVID-19 pandemic to inform projections about hospital utilization and surges in highly affected states.

- The ***eCare Planning* project**[9] developed an electronic care (eCare) plan that will improve the quality of data for PCOR studies by facilitating aggregation and sharing of data across home, community, clinic, and research settings that do not typically share an EHR system. Project staff worked to address issues related to data standardization, completeness, comprehensiveness, and accuracy of care plan data elements. Value sets were created to organize and standardize key data elements for specific conditions (for example, chronic kidney disease and long COVID), which can be used as is or updated as needed by PCOR researchers to extract and aggregate EHR data for research.

- The ***SHIELD* project**[10] developed mapping manuals to support consistency and interoperability in the coding of laboratory in vitro diagnostic (IVD) tests. IVD tests—tests done on samples taken as part of the provision of health services—are important sources of patient information for both clinical and research purposes. However, test results are often coded according to different terminology standards,

both within and across health care organizations, making data difficult to exchange or analyze. The open-access manuals created by this project use the Logical Observation Identifiers Names and Codes (LOINC) terminology standard to facilitate consistent mapping of LOINC terms for laboratory tests and orders to IVD tests.

Domain 2: Providing More Relevant, Comprehensive Data for PCOR Studies

The core elements of PCOR questions are reflected in a framework commonly referred to as PICOTS (Patient population, Intervention, Comparator, Outcome, Timing, Setting).[11] Addressing PCOR questions requires sufficient data on the following elements: (1) patient populations (including sociodemographic, health, and health care information); (2) characteristics of interventions (and comparators); (3) outcomes, particularly those important to patients and caregivers; (4) timing, which may include longer-term follow up; and (5) settings where the intervention is delivered. Data may also be needed on other relevant variables, depending on the underlying causal model (for example, potential confounders, moderators, and mediators).

However, researchers often lack sufficient data on these elements and related variables, limiting the ability to address PCOR questions important to patients, caregivers, clinicians, and policymakers. In some cases, new data are needed, such as data on patient-reported outcomes. In other cases, the data may exist in different databases but are not linked or otherwise accessible in research-ready formats.

OS-PCORTF projects have improved the availability of the types and sources of data for PCOR studies through development of new datasets (for example, through data collection or linkage) or augmentation of existing datasets (for example, through addition of new variables):

- The *MAT-LINK* **project**,[12] launched in 2019 and expanded in 2021, is designed to provide comprehensive data to assess and address the longer-term effects of prenatal opioid exposure on children. The project developed a surveillance network that collects and standardizes EHR data from seven geographically diverse clinical sites across the U.S. on maternal, infant, and child health outcomes (through age six) associated with medication for opioid use disorder (MOUD) during pregnancy. The resulting dataset will have clinical and health outcomes data for over 5,000 pregnant person-infant pairs, which will support PCOR studies to assess the effects of different MOUD regimens on infant and child development and to understand the role of mediating and moderating factors (including exposure to multiple substances, maternal comorbidities, and other psychosocial factors) on maternal and infant outcomes.

- Several OS-PCORTF projects have enhanced the relevance and comprehensiveness of data available for PCOR studies through linkage with the National Hospital Care Survey (NHCS).[13] The NHCS collects data related to health care delivery in hospital-based settings, including data on demographics, diagnoses, procedures, laboratory tests, and medications, as well as patient-level identifiers that allow for linkage with other data sources. Under the *Opioid Hospital Care and Mortality Data Linkage* **project**,[14] NHCS data were linked to data from the National Death Index to improve the availability of data on hospital care and deaths related to opioid-involved drug overdose. The linked dataset allows researchers to follow patients with an opioid event from presentation at a hospital to death (if applicable) and to analyze previous encounters retroactively for more information. The project later augmented the dataset with information on additional variables, including co-occurring substance use disorders (SUDs) and mental health conditions. Separately, the *NHCS-Administrative Data Linkage* **project**[15] linked NHCS data with: (1) Medicare fee-for-service claims data from the Centers for Medicare & Medicaid Services (CMS); and (2) federal housing data from the U.S. Department of Housing and Urban Development (HUD). Researchers can use the linked data to investigate PCOR questions related to initiatives focusing on opioid use and mental health care, health outcomes associated with distinct types of post-acute care services, and the roles of federal social support programs and access to stable housing in health outcomes.

- The *Linking State Medicaid and Child Welfare Data* project[16] linked administrative Medicaid and child welfare data in two states (Kentucky and Florida) to produce a first-of-its-kind dataset focused on understanding outcomes among families in which a parent has a SUD and co-occurring behavioral health issues and children in the child welfare system. The dataset also includes a sample of children and adults receiving Medicaid who were not child-welfare involved for comparison purposes. Researchers can leverage this dataset to investigate PCOR questions related to health and family welfare outcomes, the effectiveness of interventions, and the role of social determinants of health (SDOH). As a result of this project, ASPE launched the Child Welfare and Health Infrastructure for Linking and Data Analysis of Resources, Effectiveness, and Needs (CHILDREN) initiative, which is linking Medicaid and child welfare data for five additional states.[17]

Domain 3: Enhancing Analytic Resources for PCOR Studies

Improving data capacity for PCOR studies also requires enhancements to analytical resources—including methodologies and tools to extract, link, and analyze data, as well as improved understanding of the strengths and limitations of these approaches. PCOR researchers and other stakeholders also benefit from the development of frameworks and toolkits that aid in implementing new methods and approaches, facilitate data queries, and support data visualization. Enhanced resources can assist PCOR researchers throughout the research process, enabling better formulation and prioritization of research questions; more rigorous research design, conduct, and analysis; and more robust dissemination and uptake of findings and products.

OS-PCORTF projects have strengthened the analytic resources available for PCOR studies and supported their appropriate use (through documentation and dissemination of their methodological work), thereby improving the capacity of researchers to generate evidence to inform decisions and improve health outcomes:

- The *Opioid Hospital Care and Mortality Data Linkage* project[14] leveraged artificial intelligence approaches to extract data from EHR clinical notes, a data source that is typically difficult for researchers to access. The project developed a novel two-part algorithm to improve identification of opioid-involved encounters by: (1) searching diagnosis and procedure codes for opioid-related encounters; and (2) using natural language processing (NLP) methods to sift efficiently through large amounts of unstructured, text-based EHR data to identify indicators of opioid use. The NLP application enabled identification of additional cases of opioid use—as well as the specific opioid drug(s) involved—that were not detected with previous approaches.

- The *NHCS-Administrative Data Linkage* project[15] assessed standard linkage algorithms and developed a machine learning algorithm to improve the quality of patient-level record linkages. Typically, in the absence of unique identifiers, each record in a given dataset must be compared with all records in a second dataset; this approach can be impractical and prohibitive in the case of datasets with a large volume of records (for example, CMS claims data). To avoid such logistical challenges, as well limitations of deterministic matching approaches that require exact matches between specific fields, the project developed a machine learning algorithm that used probabilistic matching techniques to improve the accuracy and efficiency of linkages between EHR and administrative data.

- The *CRN Community of Practice* project[18] developed a framework to assess the performance and maturity of coordinated registry networks (CRNs). CRNs build on clinical patient registries, which store patient health data, and strategically harmonize and link data from multiple sources.[19,20] The maturity framework helps to identify areas that need investment to improve the utility and value of a CRN for patient-centered outcomes research.[21]

- The *Synthetic Health Data Generation Engine* project[22] developed new modules within Synthea™, a data generation engine that uses publicly available data to create synthetic health records (i.e., health data on fictitious patients). Synthetic data provide a useful option for mitigating privacy risks by

allowing researchers to initiate, refine, or test methods without accessing real patient data.[23] This project produced five new clinical modules—for cerebral palsy, prescribing opioids for chronic pain and treatment of opioid use disorder, sepsis, spina bifida, and acute myeloid leukemia—along with companion guides to support their use.

Domain 4: Reducing Barriers to Data Access and Use for PCOR Studies

Researchers need timely and efficient access to data to ensure that patients, caregivers, clinicians, and policymakers have relevant evidence to inform decision making. Strategies to reduce barriers to data access and use include, but are not limited to, interoperability standards for the exchange of data; de-identification processes to remove personal identifiers or other protected information; tiered access to datasets (for example, restricted-use files versus public-use files); and making documentation publicly available (for example, data extraction codes, implementation guides, and data documentation).

OS-PCORTF projects have developed both technical and non-technical resources to facilitate researchers' timely access to (and use of) data for PCOR studies:

- The *Linking State Medicaid and Child Welfare Data* **project**[16] published technical guidance and implementation resources to support states in: (1) establishing data sharing agreements between state Medicaid and child welfare agencies and (2) conducting data linkages. The resources include the common data model and extraction code used for the two states involved in the project and best practices for establishing data infrastructure to promote responsible data stewardship. In addition, a restricted use, de-identified version of the linked Medicaid claims and child welfare dataset has been made available to researchers at no cost, along with an accompanying user guide containing codebooks and information on the background of the study, de-identification processes, data limitations, and analytic guidance.

- The *CRN Community of Practice* **project**[18] developed core data elements and Fast Healthcare Interoperability Resources® (FHIR) standards to allow researchers to identify, encode, and organize data related to medical devices in a consistent and standardized way across CRNs. The core data elements and FHIR standards reduce the time and resources required for researchers to access and use data for PCOR studies.

Limitations

The findings presented in this report are based on an interim qualitative assessment of a subset of OS-PCORTF projects and their impacts across a selected set of domains. As such, they do not reflect a comprehensive evaluation of the portfolio or its impacts (short- or long-term). Despite the somewhat narrow focus, this assessment provided the OS-PCORTF with an opportunity to better understand recent collective contributions of, as well as challenges with, projects; strengthen project funding and award management practices based on these insights; and pilot an approach (impact domains) for informing future, more comprehensive assessments.

3. Conclusion

To build data capacity and infrastructure for patient-centered outcomes research, OS-PCORTF funds HHS agency projects that (1) improve the quality of data; (2) provide more relevant, comprehensive data; (3) enhance analytical resources; and (4) reduce barriers to data access and use. The collective contributions of these projects, as illustrated in this interim qualitative assessment, will enable PCOR researchers to address questions important to patients, caregivers, clinicians, and policymakers in a more robust and rigorous manner.

The path from improving data capacity to better health outcomes can be long, indirect, and affected by a host of other factors, however. PCOR researchers' uptake and implementation of data products developed and

made available by OS-PCORTF projects (and others) is only the first step. Findings from PCOR studies must then be communicated and disseminated to relevant decision makers in a way that meets their informational needs, and decision makers must be empowered to make informed choices that are able to be fully implemented within the relevant environment and systems.

Nevertheless, data capacity and infrastructure are the essential foundation on which research, informed decisions, and improved outcomes rest. Efforts to improve the collection, linkage, and analysis of data must therefore be shaped by, and remain responsive to, evolving data infrastructure needs, national health priorities, legislative or policy changes, and advances in health care and data science. The OS-PCORTF, guided by its new Strategic Plan[24] and continued collaboration with agency partners, is well-positioned to build the robust data capacity and infrastructure needed to realize its vision of delivering better data to improve evidence generation, decision making, and health outcomes for all Americans.

Appendix A: Descriptions of Featured OS-PCORTF Projects

	Project Title	Agencies	Project Duration	High-Level Impact Summary
1	**Assessing and Predicting Medical Needs in a Disaster**[7] *Short Title as Referenced in the Report*: Medical Needs in a Disaster	AHRQ, ASPR	2018-2020	*What is the challenge?* Coordinated medical and public health efforts are critical for responding to weather-related disasters. However, effective responses require timely, tailored data to address specific needs (for example, disaster type, geography, population). *How did the project address the challenge?* This project brought together quarterly Healthcare Cost and Utilization Project (HCUP) data on hospital utilization, disaster declaration data from the Federal Emergency Management Agency (FEMA), weather-related data from the National Oceanic and Atmospheric Administration (NOAA), and community characteristics data from the CDC to help researchers understand how to tailor disaster medical response to local needs. *Key Data Products:* • The HCUP Hurricane Data Resource and data dictionary provide county-level hospital utilization data, storm and community characteristics data, and hospital-level characteristics data for use by federal researchers in disaster response. • The **HCUP Fast Stats Hurricane Impact on Hospital Use**[25] is a public-facing statistical query system that allows users to download data files on hospital utilization by hurricane, age, condition, and setting.
2	**Surveillance Network: Maternal, Infant, and Child Health Outcomes Associated with Treatment of Opioid Use Disorder During Pregnancy**[12] *Short Title as Referenced in Report:* MAT-LINK	CDC	2019-2022	*What is the challenge?* 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). *How did the project address the challenge?* MAT-LINK is 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 exposure and outcome data for pregnant people, infants, and children from electronic health records (EHRs). *Key Data Products:* • **MAT-LINK**[26] has collected, linked, and made available data on over 5,000 pregnant person-child dyads to support research on maternal, infant, and child health outcomes among those who had an OUD diagnosis during pregnancy (and either received or did not receive MOUD during pregnancy).

	Project Title	Agencies	Project Duration	High-Level Impact Summary
3	**Linking State Medicaid and Child Welfare Data for Outcomes Research on Treatment for Opioid Use Disorder and Other Behavioral Health Issues**[16] *Short Title as Referenced in Report:* Linking State Medicaid and Child Welfare Data	ACF, ASPE	2019-2022	***What is the challenge?*** Little is known about the relationship(s) between Medicaid treatments for parental substance use disorder or mental health conditions and outcomes for children in the child welfare system. Understanding how services provided by state child welfare agencies and Medicaid affect outcomes can enable agencies to serve these populations more efficiently and effectively. ***How did the project address the challenge?*** The project developed a dataset of linked child welfare and Medicaid data for researchers in two states as well as a roadmap and toolkit to support other states interested in conducting such linkages. ***Key Data Products:*** • A restricted use dataset containing the linked data from Kentucky and Florida child welfare and Medicaid agencies • A **de-identified general-use dataset**[27] and an **accompanying users' guide**[28] • A **lessons learned document**[29] detailing the requirements to conduct linkages and guidance on how to overcome barriers and challenges
4	**SHIELD - Standardization of Lab Data to Enhance Patient-Centered Outcomes Research and Value-Based Care**[10] *Short Title as Referenced in Report:* SHIELD	FDA	2019-2022	***What is the challenge?*** Electronic encoding of in vitro diagnostic (IVD) testing data is often different within and across institutions. The lack of semantic interoperability leads to challenges in using and aggregating data for clinical, public health, and research purposes. ***How did the project address the challenge?*** The project developed several tools and guides to support standardized encoding of laboratory data in key domains as well as standardized reporting of those data. The project also conducted a pilot in five clinical sites to better understand the current implementation of standard codes and remaining gaps. ***Key Data Products:*** • **Pilot-tested and vetted resources**[30] that provide guidance for using Logical Observation Identifiers Names and Codes (LOINC) terms across six laboratory domains: allergy, cell markers, chemistry, drug and toxicology, hematology and serology, and molecular pathology

	Project Title	Agencies	Project Duration	High-Level Impact Summary
5	**Data Capacity for PCOR through eCare Planning for People with Multiple Chronic Conditions & Disability**[9] *Short Title as Referenced in Report:* eCare Planning	AHRQ, NIH	2019-2024	*What is the challenge?* Patients with multiple chronic conditions often see multiple providers in a variety of care settings, and their data are likely to be dispersed or missing critical elements in the EHR (for example, data on patient-reported outcomes and social needs), resulting in fragmented care and difficulty for researchers trying to understand and improve outcomes for the population. *How did the project address the challenge?* This project developed an electronic care (eCare) Plan application and accompanying Health Level Seven (HL7) Fast Healthcare Interoperability Resources® (FHIR) implementation guide, including value sets and data standards, clinical information models, and FHIR profiles. The eCare Plan application for patients and clinicians aggregates patient data from a variety of sources to improve clinical care and care coordination and to enable more research on the patient population. *Key Data Products:* • eCare Plan Provider Facing software application and **HL7 FHIR Implementation Guide**[31] • Data elements and value sets for chronic kidney disease, cardiovascular disease, chronic pain, pain associated with opioid use disorder, type 2 diabetes, and social determinants of health (SDOH)
6	**Capstone: Augmenting the National Hospital Care Survey (NHCS) Data through Linkages with Administrative Records**[15] *Short Title as Referenced in Report:* NHCS-Administrative Data Linkage	CDC	2019-2022	*What is the challenge?* Data on hospital utilization are often siloed from SDOH data, such as health insurance status and housing access, making it difficult for researchers to understand the relationships between SDOH and patient outcomes. *How did the project address the challenge?* The project linked data from the NHCS, which collects inpatient, outpatient, and emergency department data from U.S. hospitals, to Medicare claims and utilization data and administrative data from the U.S. Department of Housing and Urban Development (HUD). *Key Data Products:* • Restricted use **linked 2016 NHCS/2016-2017 CMS Medicare data files**[32] and **linked 2014 NHCS-2013/2015 HUD and 2016 NHCS-2015/2017 HUD data files**,[33] both with associated resources including code books, analytic reports, and considerations • A **case study**[34] assessing how the Sequential Coverage Algorithm, a supervised machine learning algorithm developed to link NHCS and Medicare data, improved linkage efficiency

	Project Title	Agencies	Project Duration	High-Level Impact Summary
7	**WHT-CRN Project: Bridging the PCOR Infrastructure and Innovation through Coordinated Registry Network (CRN) Community of Practice**[18] *Short Title as Referenced in Report:* CRN Community of Practice	FDA	2019-2022	*What is the challenge?* Coordinated registry networks (CRNs) are an important source of data collected in routine care settings ("real-world data") but often do not collect data in a standardized format or link the data with other types of data needed for PCOR studies. *How did the project address the challenge?* The project developed a common minimum dataset for CRNs and a framework to assess the maturity of CRNs. The project also developed tools to strengthen the ability of CRNs to collect additional patient-related data (for example, patient goals, patient-reported outcomes); improve interoperable data sharing; and link CRN data to claims data and clinical trial data. *Key Data Products:* • **Framework to assess CRN maturity**[21] • Common minimum dataset for CRNs • **Set of analytic tools and resources**[35]
8	**A Synthetic Health Data Generation Engine to Accelerate Patient-Centered Outcomes Research**[22] *Short Title as Referenced in Report:* Synthetic Health Data Generation Engine	ONC	2019-2022	*What is the challenge?* Researchers often have difficulty accessing health-related data due to privacy and security concerns. *How did the project address the challenge?* Synthea™ provides open-source, freely available synthetic data that researchers can use to initiate, refine, or test research questions and approaches. The project developed and validated Synthea™ modules to develop synthetic data for five clinical use cases, whose source code can be modified and built on by external researchers for their own use cases. *Key Data Products:* • Synthea™ clinical modules for five use cases: **cerebral palsy**,[36] **prescribing opioids for chronic pain and treatment of opioid use disorder**,[37] **sepsis**,[38] **spina bifida**,[39] and **acute myeloid leukemia**[40] • Accompanying **companion guides**[41] that provide information on how the modules were developed and can be used or modified for each of the five use cases

	Project Title	Agencies	Project Duration	High-Level Impact Summary
9	**Enhancing Identification of Opioid-Involved Health Outcomes Using Linked Hospital Care and Mortality Data**[42] **– Phase 1** **Identifying Co-Occurring Disorders among Opioid Users Using Linked Hospital Care and Mortality Data**[14] **– Phase 2** *Short Title as Referenced in Report:* Opioid Hospital Care and Mortality Data Linkage	CDC	2018-2022 – Phase 1 2019-2021 – Phase 2	*What is the challenge?* Data needed to provide a more complete picture of patient care are often in distinct locations and/or are difficult to extract. Leveraging the full potential of such data, particularly the additional detail that can be found in clinical notes and narrative text from death certificates, offers a chance to better understand opioid-involved hospital encounters, co-occurring mental health conditions, and overdose deaths. *How did the project address the challenge?* The project linked data from the National Death Index (NDI), Drug-Involved Mortality file (DIM), and NHCS and developed methods using natural language processing (NLP) techniques to extract more granular data from clinical notes. The project produced a new dataset that enables researchers to study risk factors identified in opioid-involved hospital encounters and opioid overdoses. *Key Data Products:* • Linked NHCS/NDI/DIM datasets **with enhanced opioid identification**[43] and with **enhanced identification of co-occurring disorders**[44] and a summary report detailing linkage processes, variables, and potential analyses • Benchmarking data portal for participating hospitals • Publicly available code for NLP algorithms

NOTES: ACF=Administration for Children and Families; AHRQ=Agency for Healthcare Research and Quality; ASPR=Administration for Strategic Preparedness and Response; CDC=U.S. Centers for Disease Control and Prevention; FDA=Food and Drug Administration; NIH=National Institutes of Health; ONC=Office of the National Coordinator for Health Information Technology.

Appendix B: Assessment Methods

The report presents the findings of a three-stage assessment approach: (1) identifying impact assessment domains, (2) selecting projects for inclusion, and (3) conducting and analyzing key informant interviews (KIIs) with project leads and other team members.

Identification of Impact Assessment Domains

An iterative process was used to identify four domains that captured key areas of direct impact for OS-PCORTF projects. The four impact assessment domains and their corresponding subdomains are presented below.

Domain	Subdomains
Improving the quality of data available for PCOR studies	• Data utility: relevance, accessibility, timeliness, punctuality, and granularity of data available • Data objectivity: data accuracy, reliability, and coherence • Data integrity: data credibility, security, and confidentiality
Providing more relevant, comprehensive data for PCOR studies	• New datasets (individual and/or linked) • New variables (for example, PCOR-relevant outcomes) within existing datasets
Enhancing analytic resources for PCOR studies	• New methods of data collection or extraction • New methods and/or improved understanding of methods for using, linking, analyzing, and/or validating data • New or improved analytical tools and resources (for example, dashboards)
Reducing barriers to PCOR data access and use	• Timelier PCOR studies by means of more efficient data sharing (for example, interoperability) • Improved data access and governance (for example, making certain types of data more widely available to researchers)

Selection of Projects for Impact Assessment

The nine projects featured in this report were selected from a set of nineteen OS-PCORTF projects awarded funding in Fiscal Year (FY) 2018 or FY 2019 (and for this reason, either completed or nearing completion) and not completed at the time of a previous impact report.[45] For each of the 19 eligible projects, the NORC team reviewed materials that included project products, progress reports, **OS-PCORTF Portfolio Reports**[46] that featured the project, and project final reports (when available). The team then mapped the information available for each project to the four domains described above.

After information about the nineteen projects had been mapped to the four assessment domains, the NORC team worked with ASPE to select nine projects (see Appendix A) for the project. Three criteria guided project selection: (1) representation of a cross-section of U.S. Department of Health and Human Services (HHS) agencies; (2) representation of a cross-section of impact domains (and subdomains); and (3) representation of the range of work that the OS-PCORTF has funded. *Note: The purposive sample of nine projects was intended to ensure representation across the domains; projects not selected for inclusion in the report also made important contributions across the four domains.*

Key Informant Interviews with Project Leads

To gain a deeper understanding of each project's impacts, the NORC team and ASPE conducted virtual semi-structured KIIs (in some cases small group discussions), each lasting an hour, with the leads (and other project team members, if applicable) of the nine projects. We developed a general interview protocol (Appendix C) structured around the four direct impact assessment domains. Within the general protocol, we tailored each interview to: (1) prioritize questions on the specific impact domains mapped to the project, and (2) gather additional information on specific products the project produced.

During April–May 2023, we conducted nine interviews, speaking with twenty-one key informants. A NORC interviewer and notetaker were present during each call, and the interviews were recorded for subsequent transcription with respondent consent. The KII findings were coded thematically to identify contributions across the four domains of interest.

Appendix C: Key Informant Interview Discussion Questions

The discussion questions below were used to guide the semi-structured KIIs. Questions were tailored for each interview based on the specific project.

Introduction

1. I want to start by having you do a brief introduction of yourself. Can you please introduce yourself and briefly tell us about your role in [Project Name]?

2. How would you describe the significance of your project to a non-technical audience? That is, what are the problems/challenges in PCOR data capacity that the project is trying to address and for whom?

Domain 1: Improving the Quality of Data Available for PCOR Studies

We'd like to turn now to asking you a few questions about how your project may have improved the quality of data available for PCOR studies. By quality, we are referring to some specific dimensions related to data utility, data objectivity, and data integrity, which were defined in the background document we sent prior to the interview.

3. Keeping these dimensions in mind, can you describe the specific contributions made by your project to the quality of data available for PCOR studies? This could be contributions via enhancements to the data quality of existing datasets, or via the creation of new, high-quality datasets.
    a. *How has it resulted in more relevant, accessible, timely, or granular data [data utility]? You can talk about both resulting datasets and documentation.*
    b. *What about any impacts to the availability of more accurate or reliable data for PCOR studies [data objectivity]?*
    c. *Has the project impacted the scientific integrity of the data, or security or confidentiality of data [data integrity], available for PCOR studies?*

4. Of the contributions described, which do you think has/had the most significant impact on improving the quality of data available for PCOR studies?

Domain 2: Providing more Relevant, Comprehensive Data for PCOR Studies

I would now like to ask you how your project has contributed to providing more relevant, comprehensive data for PCOR studies. This might include making new variables available within an existing dataset or making available new datasets. I want to note that we are interested in both new datasets that represent new individual data sources (cross-sectional, longitudinal, nationally representative, etc.), as well as new linked data sources.

5. How would you describe your project's contributions to enhancing the relevance and comprehensiveness of data for PCOR studies in existing datasets?
    a. *Can you describe/discuss the variables that were chosen and why they were chosen? Is the resulting dataset (inclusive of these variables) easily accessible/widely available to PCOR researchers?*
    b. *Can you describe any particular methods that your project utilized? Did your project utilize any techniques to capture and incorporate new variables (for example, AI, machine learning, natural language processing)? If so, how was this achieved? What lessons were learned?*
        i. *What impact do you think these techniques may have on advancing PCOR more generally?*

  c. Can your project inform the future addition of new variables into existing PCOR-relevant datasets? If yes, please discuss.

6. What about contributions related to new or unique data sources for PCOR studies, stemming from your project? For example, did the project improve the availability of longitudinal or nationally representative data that can be used for PCOR studies?

  a. Are these new data publicly available, or can they be made publicly available in the future even if not now?

  b. [For longitudinal data]: Can you describe plans in place for additional waves of data collection? For example, how frequently will data be collected?

  c. Based on your general expertise or lessons learned from your project, what recommendations do you have for improving the availability of data to promote more rigorous PCOR (for example, to result in more valid and actionable findings)? Are there specific subpopulations for which data gaps are most pronounced? What incentives need to be in place for promoting such data collection?

7. Of the contributions described, which do you think has/had the most significant impact on making available more relevant and comprehensive for PCOR studies?

Domain 3: Enhancing Analytic Resources for PCOR Studies

Next, I'd like to turn to asking you a few questions about how your project may have enhanced analytic know-how for PCOR studies. We're interested in both new methods that might have been developed or tested within your project, or even if your project did not develop a new method, any methodological reports or papers aimed at improving understanding of existing methods. New methods can include techniques for collecting or extracting PCOR-relevant data, as well as any analytic techniques for using, linking, analyzing, or validating data.

8. Can you talk about the noteworthy contributions of your project to the collection or extraction of PCOR-relevant data?

  a. Can you explain resulting unique or improved ways to collect PCOR-relevant primary data (for example, by reducing data collector or respondent burden, improving accuracy, etc.)?

  b. Can you talk about enhancements to existing data collection instruments or tools?

  c. Can you talk about the contributions made to extracting and incorporating data from historically less-used sources such as text fields or clinical notes?

  d. In the course of your project, did you identify gaps or areas for improvement for PCOR data collection and extraction? Are there any plans to address them in future work?

9. Can you speak to contributions stemming from your project that are related to generating or enhancing methodological know-how for conducting PCOR studies?

  a. For example, can you talk about new resulting methods to link or analyze PCOR-relevant data, or validate existing data?

  b. Did the project improve understanding of existing methods to link or analyze PCOR-relevant data, or validate existing data, via methodological reports or other publications?

    i. What is your sense of how widely these reports/publications are being used by the research community?

10. We are interested in better understanding other kinds of analytical resources produced by your project that PCOR researchers or other stakeholders might find useful. For example, this might be a new data visualization tool like a dashboard or an implementation guide. Can you talk about these tools or resources?

        a. *What is the main value-add of this tool for PCOR? Which stakeholder groups are most likely to benefit from this?*

        b. *How will the implementation guides you developed assist researchers in using your new datasets?*

11. Of the contributions described, which do you think has/had the most significant impact related to producing analytic methods or outputs for PCOR studies?

Domain 4: Reducing Barriers to PCOR Data Access and Use

Now I'd like to discuss your project's contributions to reduce barriers to PCOR data access and use. This is a broad category that might encompass improving interoperability to make data sharing more efficient or efforts to make it easier for stakeholders to access datasets.

12. Can you discuss your project's impacts on PCOR data sharing and interoperability?
        a. *Were Fast Healthcare Interoperability Resources application programming interfaces (or FHIR APIs) used within your project to facilitate data exchange? If so, how did they facilitate data exchange?*
        b. *What lessons learned on data exchange/interoperability have resulted from your project, that might be informative for (other) data stewards?*

13. What challenges in governance for PCOR-relevant data did your project address? By governance, we mean the processes/models that are used to manage data access and sharing.
        a. *How has access to the project's resulting datasets been improved (for example, streamlined application process, lower/no costs, available to more users)?*

14. Of the contributions described, which do you think has/had the most significant impact on improving access and/or use of data for PCOR studies?

Broader Impact

15. To your knowledge, how are the products produced by your project being used by end users to result in improved PCOR studies/stronger evidence (for example, addressing new PCOR questions) or contributing to decision- making at your agencies?

16. We have discussed your project's impacts related to various data and analytic domains. Are there any other impacts emanating from this project that you think tell a stronger story of impact than what we have discussed so far?
        a. *What do you feel is your project's biggest contribution to PCOR?*

17. Based on the work done in your project, what future work do you think the OS-PCORTF can support in this area?
        a. *In your opinion, what should be the top priorities for the Trust Fund in this area over the short-term? In the long-term?*

References

[1] Institute of Medicine (US) Committee on Quality of Health Care in America. Crossing the Quality Chasm: A New Health System for the 21st Century. Washington (DC): National Academies Press (US); 2001. PMID: 25057539.

[2] Office of the Assistant Secretary for Planning and Evaluation. Explore the OS-PCORTF Project Profiles. Accessed June 29, 2023. https://aspe.hhs.gov/collaborations-committees-advisory-groups/os-pcortf/explore-portfolio

[3] Patient Protection and Affordable Care Act, Publ. L. No. 111-148, 124 Stat. 119 (2010). https://www.govinfo.gov/content/pkg/PLAW-111publ148/html/PLAW-111publ148.htm

[4] Zozus MN, Hammond WE, Green BB, Kahn MG, Rachel L, Richesson RL, Rusincovitch SA, Simon GE, Smerek MM. Assessing Data Quality for Healthcare Systems Data Used in Clinical Research (Version 1.0). https://dcricollab.dcri.duke.edu/sites/NIHKR/KR/Assessing-data-quality_V1%200.pdf

[5] Kahn MG, Callahan TJ, Barnard J, Bauck AE, Brown J, Davidson BN, Estiri H, Goerg C, Holve E, Johnson SG, Liaw ST. A harmonized data quality assessment terminology and framework for the secondary use of electronic health record data. EGEMS. 2016;4(1). doi:10.13063/2327-9214.1244

[6] Federal Committee on Statistical Methodology. 2020. A Framework for Data Quality. FCSM 20-04. Federal Committee on Statistical Methodology. September 2020.

[7] Assessing and Predicting Medical Needs in a Disaster. Project Profile. Accessed August 14, 2023. https://www.aspe.hhs.gov/assessing-predicting-medical-needs-disaster

[8] Building the Data Capacity for Patient-Centered Outcomes Research: 2019 Vignettes. Accessed August 18, 2023. https://aspe.hhs.gov/sites/default/files/private/pdf/259016/OS-PCORTF_Project_Highlights.pdf

[9] Data Capacity for Patient-Centered Outcomes Research through Creation of an Electronic Care Plan for People with Multiple Chronic Conditions. Project Profile. Accessed August 14, 2023. https://aspe.hhs.gov/data-capacity-patient-centered-outcomes-research-through-creation-electronic-care-plan-people-0

[10] SHIELD - Standardization of Lab Data to Enhance Patient-Centered Outcomes Research and Value-Based Care. Project Profile. Accessed August 14, 2023. https://aspe.hhs.gov/shield-standardization-lab-data-enhance-patient-centered-outcomes-research-value-based-care

[11] Matchar DB. Introduction to the Methods Guide for Medical Test Reviews. AHRQ Publication No. 12-EHC073-EF. Chapter 1 of Methods Guide for Medical Test Reviews (AHRQ Publication No. 12-EHC017). Rockville, MD: Agency for Healthcare Research and Quality; June 2012. https://effectivehealthcare.ahrq.gov/products/methods-guidance-tests-introduction/methods#toc-7

[12] Surveillance Network: Maternal, Infant, and Child Health Outcomes Following Treatment of Opioid Use Disorder (OUD) During Pregnancy. Project Profile. Accessed August 14, 2023. https://aspe.hhs.gov/surveillance-network-maternal-infant-child-health-outcomes-following-treatment-opioid-use-disorder

[13] National Hospital Care Survey. National Center for Health Statistics, Centers for Disease Control and Prevention. Last updated May 23, 2023. https://www.cdc.gov/nchs/nhcs/index.htm

[14] National Center for Health Statistics. Identifying Co-Occurring Disorders among Opioid Users Using Linked Hospital Care and Mortality Data: Capstone to an Existing FY18 PCORTF Project (Final Report). October 2021. Accessed August 14, 2023. https://aspe.hhs.gov/sites/default/files/documents/6b78783ff57c2d5baccce2e71e148f4b/fy19-nhcs-aspe-final-report-10-18-21.pdf

[15] Augmenting the National Hospital Care Survey (NHCS) Data through Linkages with Administrative Records: A Project. Project Profile. Accessed August 14, 2023. https://aspe.hhs.gov/augmenting-national-hospital-care-survey-nhcs-data-through-linkages-administrative-records-project

[16] Linking State Medicaid and Child Welfare Data for Outcomes Research on Treatment for Opioid Use Disorder and Other Behavioral Health Issues. Project Profile. Accessed August 14, 2023. https://aspe.hhs.gov/linking-state-medicaid-child-welfare-data-outcomes-research-treatment-opioid-use-disorder-other

[17] Mathematica. Project Overview: CHILDREN INITIATIVE (2022-2027). Accessed August 18, 2023. https://www.mathematica.org/projects/child-welfare-and-health-infrastructure-for-linking-and-data-analysis-of-resources

[18] Bridging the PCOR Infrastructure and Technology Innovation through Coordinated Registry Networks (CRN) Community of Practice. Project Profile. Accessed August 14, 2023. https://aspe.hhs.gov/bridging-pcor-infrastructure-technology-innovation-through-coordinated-registry-networks-crn

[19] Coordinated Registry Network for Women's Health Technologies. HealthIT.gov. Updated June 14, 2020. Accessed June 16, 2023. https://www.healthit.gov/topic/scientific-initiatives/pcor/coordinated-registry-network-womens-health-technologies-crn

[20] Coordinated Registry Networks. Medical Device Epidemiology Network (MDEpiNet). Accessed June 16, 2023. https://www.mdepinet.net/coordinated-registry-networks

[21] Sedrakyan A, Aryal S. Maturity framework and select approaches for developing Coordinated Registry Networks (CRNs): Medical Device Epidemiology Network (MDEpiNet) supplement. BMJ Surg Interv Health Technol. 2022;4(Suppl 1):e000148. Published 2022 Nov 11. doi:10.1136/bmjsit-2022-000148

[22] Office of the National Coordinator for Health Information Technology. Final Report: Synthetic Health Data Generation to Accelerate Patient-Centered Outcomes Research (PCOR). March 2022. Accessed August 14, 2023. https://www.healthit.gov/sites/default/files/page/2022-03/20220314_Synthetic%20Data%20Final%20Report_508.pdf

[23] Synthetic Health Data Generation to Accelerate Patient-Centered Outcomes Research. HealthIT.gov. Updated April 28, 2022. Accessed June 27, 2023. https://www.healthit.gov/topic/scientific-initiatives/pcor/synthetic-health-data-generation-accelerate-patient-centered-outcomes

[24] Office of the Assistant Secretary for Planning and Evaluation. OS-PCORTF Strategic Plan for 2020-2029. Accessed June 29, 2023. https://aspe.hhs.gov/os-pcortf-strategic-plan-2020-2029

[25] Agency for Healthcare Research and Quality. Healthcare Cost and Utilization Project (HCUP) Fast Stats: Hurricane Impact on Hospital Use. Accessed August 14, 2023. **https://datatools.ahrq.gov/hcup-fast-stats?tab=special-emphasis&dash=102**

[26] National Center on Birth Defects and Developmental Disabilities, Centers for Disease Control and Prevention. MAT-LINK: MATernaL and Infant NetworK to Understand Outcomes Associated with Medication for Opioid Use Disorder during Pregnancy. May 2023. Accessed August 14, 2023. **https://www.cdc.gov/ncbddd/aboutus/mat-link.html**

[27] Mark, TL., Dolan, M., Allaire, B., Smith, K., Parish, W., Bradley, C., Madden, E., & Butler, V. Child and Caregiver Outcomes Using Linked Data [dataset]. National Data Archive on Child Abuse and Neglect. 2022. doi: https://doi.org/10.34681/33ge-9z29

[28] Sernaker, S., Smith, K. User's Guide: Child and Caregiver Outcomes Using Linked Data (CCOULD) Dataset (NDACAN Dataset Number 272). Accessed August 14, 2023. **https://www.ndacan.acf.hhs.gov/datasets/pdfs_user_guides/dataset272codebook.pdf**

[29] Mark, TL., Dolan, M., Allaire, B., Bradley, C. Linking Child Welfare and Medicaid Data: Lessons Learned from Two States. October 2022. Accessed August 14, 2023. **https://aspe.hhs.gov/sites/default/files/documents/be8e500260bc04552f9b46f64391de9b/ccould-lessons-learned.pdf**

[30] Regenstrief Institute. LOINC Mapping Guides. Accessed August 14, 2023. **https://loinc.org/guides/#in-development**

[31] HL7 International - Patient Care. MCC eCare Plan Implementation Guide. Accessed August 14, 2023. **http://build.fhir.org/ig/HL7/fhir-us-mcc/**

[32] National Center for Health Statistics. Restricted-Use Linked NHCS-CMS Medicare Data. Accessed August 14, 2023. **https://www.cdc.gov/nchs/data-linkage/CMS-Medicare-Restricted.htm**

[33] National Center for Health Statistics. Restricted-Use Linked NHCS-HUD Administrative Housing Data. Accessed August 14, 2023. **https://www.cdc.gov/nchs/data-linkage/nhcs-hud.htm**

[34] Campbell SR, Resnick DM, Cox CS, Mirel LB. Using supervised machine learning to identify efficient blocking schemes for record linkage. Stat J IAOS. 2021 Jun 3;37(2):673-680. doi: 10.3233/sji-200779. PMID: 34413910; PMCID: PMC8371678.

[35] Resources. Medical Device Epidemiology Network (MDEpiNet). Accessed August 14, 2023. **https://www.mdepinet.net/resources**

[36] Synthea Module Builder. Cerebral Palsy. Accessed August 14, 2023. **https://synthetichealth.github.io/module-builder/#cerebral_palsy**

[37] Synthea Module Builder. Prescribing Opioids for Chronic Pain and Treatment of OUD. Accessed August 14, 2023. **https://synthetichealth.github.io/module-builder/#prescribing_opioids_for_chronic_pain_and_treatment_of_oud**

[38] Synthea Module Builder. Sepsis. Accessed August 14, 2023. **https://synthetichealth.github.io/module-builder/#sepsis**

[39] Synthea Module Builder. Spina Bifida. Accessed August 14, 2023**. https://synthetichealth.github.io/module-builder/#spina_bifida**

[40] Synthea Module Builder. Acute Myeloid Leukemia for PCOR Research. Accessed August 14, 2023**.
https://synthetichealth.github.io/module-builder/#acute_myeloid_leukemia**

[41] Synthea Module Companion Guides. Accessed August 14, 2023.
https://github.com/synthetichealth/synthea/wiki/Module-Companion-Guides

[42] Enhancing Identification of Opioid-Involved Health Outcomes Using Linked Hospital Care and Mortality Data. Project Profile. Accessed August 14, 2023. **https://aspe.hhs.gov/enhancing-identification-opioid-involved-health-outcomes-using-linked-hospital-care-mortality-data**

[43] National Center for Health Statistics Research Data Center (NCHS RDC)**.** 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. October 2020. Accessed August 14, 2023. **https://www.cdc.gov/nchs/data/nhcs/Task-3-Doc-508.pdf**

[44] National Center for Health Statistics Research Data Center (NCHS RDC)**.** Identifying Co-Occurring Disorders among Opioid Users Using Linked Hospital Care and Mortality Data: Capstone to an Existing FY18 PCORTF Project. June 2021. Accessed August 14, 2023. **https://www.cdc.gov/nchs/data/nhcs/FY19-RDC-2021-06-01-508.pdf**

[45] Dhopeshwarkar R, Dullabh P, Dungan R, et al. *Building Data Capacity for Patient-Centered Outcomes Research: Portfolio Highlights (2016-2019) Impact, Opportunities, and Case Studies.* Office of the Assistant Secretary for Planning and Evaluation. Accessed June 29, 2023. **https://aspe.hhs.gov/sites/default/files/private/pdf/259016/OS-PCORTFImpactReport508.pdf**

[46] Office of the Assistant Secretary for Planning and Evaluation. OS-PCORTF Annual Portfolio Reports. Accessed August 14, 2023. **https://aspe.hhs.gov/collaborations-committees-advisory-groups/os-pcortf/about-os-pcortf/os-pcortf-annual-reports**