Building Data Capacity for Patient-Centered Outcomes Research

2020 2029

OS-PCORTF Strategic Plan

September 2022

Department of Health and Human Services

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. For more information, visit https://aspe.hhs.gov/.

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 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. For additional information, visit https://aspe.hhs.gov/office-health-policy-hp.

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

Building Data Capacity for Patient-Centered Outcomes Research. Office of the Secretary Patient-Centered Outcomes Research Trust Fund Strategic Plan: 2020-2029. Office of the Assistant Secretary for Planning and Evaluation, U.S. Department of Health and Human Services. September 2022.

Executive Summary

The mission of the U.S. Department of Health and Human Services (HHS) is to "enhance the health and well-being of all Americans, by providing for effective health and human services and by fostering sound, sustained advances in the sciences underlying medicine, public health, and social services."[1] HHS relies on data for fulfilling its mission and advancing science to improve health and human services. As such, HHS's coordination of federal efforts to enhance capacity to collect, link, and analyze data is critical to strengthening research and supporting the respective missions, statutory authorities, and annual priorities for HHS agencies and the Department as a whole.

The Office of the Assistant Secretary for Planning and Evaluation (ASPE), under delegation of authority by the Secretary of HHS, coordinates across relevant federal health programs to build data capacity for patient-centered outcomes research (PCOR), including administering the Office of the Secretary's Patient-Centered Outcomes Research Trust Fund (OS-PCORTF).

Patient-centered outcomes research aims to generate high-quality evidence about the effectiveness of treatments, services, and other health care interventions on the full range of outcomes that patients, caregivers, clinicians, policymakers, and other stakeholders have identified as important. Conducting PCOR studies requires timely access to relevant, high-quality data and the use of rigorous and appropriate research methods. Robust data capacity and infrastructure are therefore integral to the success of patient-centered outcomes research.

This document represents HHS's Strategic Plan for building data capacity for patient-centered outcomes research through the OS-PCORTF (2020–2029). The plan presents a vision of better data to generate stronger evidence and foster informed decisions by improving data infrastructure to conduct patient-centered outcomes research. The Strategic Plan was developed through a comprehensive review of literature; interviews with agency leaders, program officials, and data stewards; a series of public meetings; and close collaboration with a committee of HHS agency representatives. Building on a portfolio of PCOR projects that began in 2011, the Strategic Plan charts a course for continuously improving the capacity for collecting, linking, and analyzing data for PCOR studies.

The Strategic Plan provides a framework for prioritizing, supporting, and evaluating efforts to build data capacity with respect to four goals that target areas of critical importance. Figure ES-1 summarizes the core elements of the Strategic Plan.

Figure ES-1. Summary of the OS-PCORTF Strategic Plan

Goal 1. Data Capacity for National Health Priorities

Build data capacity for patient-centered outcomes research that informs the needs of federal health programs, providers, and the people served by these programs.

Goal 1 focuses on building and strengthening data capacity related to national health priorities, covering a range of health conditions, populations, and communities, which may evolve over time. Efforts to build and strengthen data capacity for these priorities, and by extension, patient-centered outcomes research more broadly, will also require sustained, successful collaborations and robust dissemination efforts. Collective learning and problem-solving help to ensure that the data, tools, and services developed will advance patient-centered outcomes research by meeting the needs of a variety of users and increasing knowledge and awareness of resources.

Goal 2. Data Standards and Linkages for Longitudinal Research

Expand longitudinal data resources that enable patient-centered outcomes research to advance evidence generation.

Data about people over time, across geographic boundaries and care settings, are often available in numerous disparate databases. Enabling data linkages, while preserving the privacy and security of the data, greatly increases the usefulness of these data for PCOR studies, particularly the long-term outcomes that matter to people and about which we often lack the most evidence. Goal 2 supports efforts in data harmonization, privacy-preserving linkage methods, and other innovative approaches to enhance the availability, quality, accessibility, and suitability of linked data for longitudinal research in PCOR studies.

Goal 3. Technology Solutions to Advance Research

Leverage leading technology solutions to improve data capacity for patient-centered outcomes and comparative clinical effectiveness research.

> Advances in analytic methods and leading technologies, such as artificial intelligence solutions, can support increased integration and use of real-world data to further patient-centered outcomes research and the generation of real-world evidence. Goal 3 focuses on leveraging advanced technology solutions to improve the utilization of large volumes of data as well as the variety and timeliness of data availability for PCOR studies, thereby increasing the richness and robustness of the evidence generated.

Goal 4. Person-Centeredness, Inclusion, and Equity

Expand the collection and analysis of socioeconomic, environmental, and other data so all people making health care decisions have the evidence they value about the outcomes and effectiveness of health care.

> Socioeconomic, environmental, and other non-health care data assets enhance the understanding of the whole person (person-centeredness), including what is important to the person, and communities (inclusion). Improving the availability and suitability of these data to complement health care data will contribute to better informed health care and health policy decisions for individuals and people who are medically underserved, underrepresented in biomedical research, and disproportionately affected (equity and inclusion). Goal 4 focuses on addressing data capacity limitations to support a more comprehensive view of health outcomes, identify and address disparities, and examine economic outcomes.

In addition to providing direction and accountability, the Strategic Plan allows for responsiveness to evolving data infrastructure needs, priorities, and relevant developments, including legislative or policy changes and advances in health care, data science, and the needs of decision makers. Guided by the Strategic Plan and continued collaboration with agency partners, HHS is positioned to fulfill the vision of the OS-PCORTF to deliver better data to improve evidence generation, decision making, and health outcomes for all Americans.

Table of Contents

List of Figures

List of Tables

1 Introduction

The mission of the U.S. Department of Health and Human Services (HHS) is to "enhance the health and well-being of all Americans, by providing for effective health and human services and by fostering sound, sustained advances in the sciences underlying medicine, public health, and social services."[1] Data are essential to this mission. Agencies within HHS routinely collect, link, and analyze data for research and policymaking. Collectively, these activities generate new data and work to improve the usability of existing and future data. Coordination of these efforts to enhance data capacity strengthens research and supports the respective missions, statutory authorities, and annual priorities for HHS agencies and the Department as a whole.

The Office of the Assistant Secretary for Planning and Evaluation (ASPE), under delegation of authority by the Secretary of HHS, coordinates across relevant federal health programs to build data capacity for patient-centered outcomes research (PCOR), including administering the Office of the Secretary's PCOR Trust Fund (OS-PCORTF). This coordination involves partnerships with agency leaders, scientists, research programs, and data stewards to develop and implement an extensive array of projects that expand data capacity to support PCOR studies. Patient-centered outcomes research aims to generate high-quality evidence about the effectiveness of treatments, services, and other health care interventions on the full range of outcomes that patients, caregivers, clinicians, policymakers, and other stakeholders have identified as important to inform decision making. Data capacity—the availability and suitability of data as well as analytic resources—is foundational to patient-centered outcomes research. Conducting PCOR studies requires timely access to relevant, high-quality data that can be analyzed with rigorous and appropriate research methods to produce scientific evidence. As such, robust data capacity and infrastructure are necessary to improve evidence generation, decision making, and health outcomes for all Americans.

This document represents HHS's Strategic Plan for the OS-PCORTF (2020–2029) to advance its vision of building data capacity for patient-centered outcomes research through coordinated, systematic efforts across federal agencies. The Strategic Plan was developed through a comprehensive review of literature; interviews with agency leaders, program officials, and data stewards; a series of public meetings; and in collaboration with a strategic planning advisory committee of HHS agency representatives. Building on a portfolio of PCOR projects that began in 2011, the Strategic Plan charts a course for continuously improving the capacity for collecting, linking, and analyzing data for PCOR studies. In addition, the Strategic Plan provides a framework for prioritizing and evaluating efforts to build data capacity with respect to four goals that target areas of critical importance.

In addition to providing direction and accountability, the Strategic Plan is designed to be responsive to evolving data infrastructure needs, priorities, and relevant developments, including legislative or policy changes and advances in health care, data science, and the needs of decision makers who use the findings from PCOR studies. The potential strategies outlined for each objective allow for adaptation based on evaluations from both internal and external efforts related to data capacity for patient-centered outcomes research. The Strategic Plan will guide HHS in carrying out its statutory authority to expand data capacity for patient-centered outcomes research through coordination with federal programs.

2   Background

HHS's Strategic Plan for the OS-PCORTF is informed by federal legislation as well as the broader PCOR ecosystem.[a] Patient-centered outcomes research, underway in HHS programs for many years, was catalyzed by the enactment of the Patient Protection and Affordable Care Act (ACA) in 2010, which included authorization for funding to build data capacity for comparative clinical effectiveness research (CER) and PCOR studies.[2] The establishment of the PCORTF was driven in part by a growing recognition of the need for better scientific evidence to inform decisions of patients, clinicians, and policymakers about the balance between the benefits and risks of health care interventions. This section provides a brief background of the PCORTF's authorizing legislation as well as an overview of challenges and opportunities in strengthening data capacity.

2.1   Legislative Background

In its landmark 2001 report, *Crossing the Quality Chasm: A New Health System for the 21st Century*, the Institute of Medicine (IOM), now the National Academy of Medicine, stated that "Americans should be able to count on receiving care that meets their needs and is based on the best scientific knowledge. Yet there is strong evidence that this is frequently not the case."[3] The report noted that, among other things, health care should be patient-centered, responsive to individual patient preferences and needs, and guided by patient values in all clinical decisions.[3]

The IOM was subsequently asked to address critical challenges in generating the evidence needed to inform decision making and improve health outcomes. In 2009, the IOM was charged with recommending national priorities for CER studies, which it defined as the "generation and synthesis of evidence that compares the benefits and harms of alternative methods to prevent, diagnose, treat, and monitor a clinical condition or to improve the delivery of care."[3] The IOM also issued recommendations for the development of methods and infrastructure to support CER studies.[4]

In 2010, the ACA bolstered the emerging paradigm of patient-centered outcomes research, which evolved from comparative clinical effectiveness research (with the terms sometimes used interchangeably), by establishing the PCORTF. Amounts transferred from the PCORTF support the work of the Patient-Centered Outcomes Research Institute (PCORI), Agency for Healthcare Research and Quality (AHRQ), and the Office of the Secretary of HHS to conduct, disseminate, and expand capacity for PCOR and CER studies.[2] Congress reauthorized the PCORTF through 2029 in December 2019.[5]

Each of the organizations supported by the PCORTF has a unique, yet related, set of responsibilities. PCORI, which receives the largest proportion of the PCORTF funding (80 percent), is charged with producing high-integrity, evidence-based information from research. Operating as an independent nonprofit, nongovernmental organization, PCORI funds studies that compare the benefits and harms of health care options to learn which work best given patients' circumstances and preferences, as well as

---

[a] The term "PCOR ecosystem" refers to the participants, processes, resources, technology, and systems involved in patient-centered outcomes research. These elements include, but are not limited to, patients, caregivers, communities, clinicians, researchers, and payers (participants); stakeholder engagement and data governance (processes); data, devices, information technology systems, and methods for design and analysis (technology); and the funding and regulatory drivers that govern interactions and contribute to accountability (resources and systems).

studies to improve methods for PCOR.[6] AHRQ receives 16 percent of the PCORTF funding to disseminate findings from PCOR studies, incorporate evidence into clinical practice, and train researchers in patient-centered outcomes research. AHRQ carries out this work through synthesizing research results (evidence synthesis), developing tools and training to support translation and dissemination of evidence, and working with health care providers and systems to incorporate updated evidence into their practices.[7]

Lastly, as specified in Section 937(f) of the Public Health Service Act, the HHS Office of the Secretary receives 4 percent of the PCORTF to:[2]

> *…provide for the coordination of relevant Federal health programs to build data capacity for comparative clinical effectiveness research, including the development and use of clinical registries and health outcomes research networks, in order to develop and maintain a comprehensive, interoperable data network to collect, link, and analyze data on outcomes and effectiveness from multiple sources including electronic health records.*

2.2 Data Capacity for Patient-Centered Outcomes Research

In the context of patient-centered outcomes research, data capacity refers to the availability and suitability of data, as well as the analytic resources for studying questions that are important to patients, caregivers, clinicians, and policymakers. Although randomized controlled trials (RCTs) are generally seen as the "gold standard" of evidence for decision making, they have important limitations (as with any method of study design and analysis). Historically, these limitations have included relatively small sample sizes, lack of diversity among participants, interventions delivered in highly controlled settings, and a narrow range of outcomes (most often focusing on efficacy).[8] Consequently, results from such RCTs have generally not provided a comprehensive source of evidence about the outcomes and comparative effectiveness of interventions provided in routine health care (i.e., real-world settings).

Strengthening data capacity for PCOR studies through the development, evaluation, and implementation of additional methods for generating, linking, and analyzing data can help to address these limitations and enable core infrastructure functionalities to improve the quality and efficiency of PCOR studies. Such functionalities (see Figure 1) are critical to ensuring that PCOR studies provide valid and reliable evidence for decision making about the relative merits of interventions across diverse patient groups and settings of care.

OS-PCORTF Data Infrastructure Functionalities				
Standardized Collection of Standardized Clinical Data	**Collection of Participant-Provided Information**	**Linking of Clinical and Other Data for Research**	**Use of Clinical Data for Research**	**Use of Enhanced Publicly Funded Data Systems for Research**
Researchers will be able to use standardized clinical data based on common data element standards across research projects and networks, thereby facilitating linkage and aggregation of data across data sources.	Participants, including those in safety net health care organizations, will be able to participate more fully in clinical research by directly providing information (i.e., data points provided by the participant such as patient-reported outcomes).	Researchers will be able to follow patients across the care continuum over time, including those enrolled in clinical trials. Researchers will be able to capture the range of variables influencing health outcomes and link clinical and other types of data (i.e., other clinical data, claims data, participant-provided information, and environmental data) required for research, regardless of where the participant goes.	Researchers will be able to utilize and analyze routinely collected clinical data for implementation of clinical studies (observational and interventional), including data relevant to assessing safety, efficacy, and adherence, as well as genetic data and patient-reported outcomes.	Researchers will be able to readily use, retrieve, link, and aggregate publicly funded data for research due to enhancements in publicly funded data systems.

Figure 1. OS-PCORTF Data Infrastructure Functionalities

Data Sources

Data sources relevant to patient-centered outcomes research include:

- New data collected prospectively for research, such as a clinical trials or observational cohorts
- Clinical data, which are often stored in electronic health records (EHRs)
- Administrative and payer data, including claims and billing data
- National survey data
- Social services data
- Datasets on factors that are relevant to studying health outcomes, such as socioeconomic, environmental, and genomic data

Participant-provided information (PPI), which can be found across multiple data sources (e.g., research studies and EHRs) is also a data asset for patient-centered outcomes research. PPI includes person-generated health data (PGHD), often generated via consumer medical and mobile devices, as well as patient-reported outcomes.

However, the mere availability of such data is not enough; the data must also be of sufficient quality to support research conclusions. Data quality is a multifaceted concept that includes completeness, accuracy, and consistency, and a number of frameworks and metrics have been proposed for assessing data quality for PCOR studies.[9, 10] Researchers must also have access to the relevant data and generally the data must be interoperable to support efficient exchange and use of information across the data lifecycle, databases, and analytic platforms.

Issues of data availability, quality, accessibility, and interoperability continue to present significant limitations for PCOR data capacity. Efforts to develop, harmonize, and implement data standards governing the collection, storage, representation, and exchange of data are therefore critical to addressing these limitations. When supported by the requisite resources and policies, such standards can improve the accuracy, security, and efficiency of data exchange and usage in patient-centered outcomes research.

Data Linkage

Linking datasets is often required to ensure that sufficient data are available to address specific PCOR questions of interest. Use of multiple records and sources can help ensure accuracy in linkage, but this process also increases the risks of privacy losses for the individuals whose data are involved. Improving the accuracy of privacy-preserving record linkage methods (e.g., pseudonymization techniques via hashing algorithms, tokenization[11]) for different data sources and ensuring that these methods remain robust to evolving security threats will increase the usefulness of data available for patient-centered outcomes research.

Data Analysis

In addition to expanding data capacity for PCOR studies via improved linkage methods, researchers are seeking to generate real-world evidence (RWE) on a full range of outcomes through new study designs and analytic approaches.[12] RWE can be generated through both randomized (e.g., pragmatic clinical trials) and non-randomized study designs, with the latter benefiting from advances in analytical methods for causal inference.[13, 14] Similarly, artificial intelligence (AI), which encompasses a range of fields and approaches, has the potential to improve analyses of large volumes of data, including the use of machine learning (ML) methods as well as natural language processing (NLP) methods for unstructured data (e.g., clinical notes). Continued efforts are needed to better understand the strengths and limitations of these approaches and ensure they are used appropriately in PCOR studies.

Finally, given challenges with ensuring timely access to relevant, high-quality data while protecting data privacy and security, researchers and organizations have begun to explore the use of synthetic datasets that realistically simulate variables, values, and relationships. The value of synthetic data as a complement to the use of actual health data depends on the suitability of the data generated, and important challenges remain regarding generation and appropriate use.[15, 16]

3   The OS-PCORTF Strategic Plan

3.1   Mission

The mission statement is rooted in the legislative authority of the OS-PCORTF.

MISSION: Build and strengthen data capacity for patient-centered outcomes research through coordination across agencies and federal programs.

3.2   Vision

The vision statement is the strategic direction, or "north star," of the OS-PCORTF over the next decade.

VISION: Better data for patient-centered outcomes research to improve evidence generation, decision making, and health outcomes for all Americans.

3.3   Guiding Principles

These guiding principles are the enduring and distinctive core tenets that ASPE strives to adhere to throughout its work in managing the OS-PCORTF:

- **Person-centeredness:** Ensure that data capacity reflects an inclusive, equitable, and comprehensive approach to meeting evidence needs.
- **Collaboration:** Build effective collaboration with researchers and across HHS agencies to address PCOR priorities.
- **Innovation:** Incubate, evaluate, and build innovative solutions for advancing data capacity.
- **Impact:** Support impactful projects that address critical gaps in data capacity for patient-centered outcomes research.

3.4   Strategic Goals, Outcomes, and Objectives

Robust data capacity and infrastructure are necessary to improve evidence generation, decision making, and health outcomes for all Americans. To achieve this vision, HHS has prioritized four goals that reflect high-priority opportunities to address critical data challenges for patient-centered outcomes research. The four goals, along with their corresponding outcomes, set forth what HHS expects to accomplish over the next decade through the OS-PCORTF portfolio of data projects. The sets of objectives, interim outcomes, and potential strategies outline the pathways for execution and are structured to be responsive to evolving HHS priorities, data infrastructure needs, and scientific advances. As written, the Strategic Plan provides both high-level vision and action-guiding directives for HHS's efforts to ensure the availability and suitability of data and analytic resources for addressing important PCOR questions.

Goal 1: Data Capacity for National Health Priorities

 Build data capacity for patient-centered outcomes research that informs the needs of federal health programs, providers, and the people served by these programs.

The ability to answer PCOR questions is often limited by issues of data availability, quality, accessibility, and interoperability as well as methodological challenges (linkage and analysis). Successfully addressing these challenges requires approaches that are systematic, targeted, and aligned with the specific expertise, capacity, and responsibilities of a particular agency or organization. Under this Strategic Plan, HHS is prioritizing efforts to build and strengthen data capacity related to national health priorities, covering a range of health conditions, populations, and communities, which may evolve over time. Although different priority areas may present somewhat distinct challenges for data capacity (e.g., validation of data quality from emerging sources), they make the issues more tractable and are likely to result in cross-cutting improvements to PCOR data infrastructure and expanded HHS expertise.

Efforts to build and strengthen data capacity for these priorities, and by extension, patient-centered outcomes research more broadly, will also require sustained, successful collaborations and robust dissemination efforts. Collective learning and problem-solving help to ensure that the data, tools, and services developed will advance patient-centered outcomes research by meeting the current and emerging needs of a variety of users, including providers, and increasing knowledge and awareness of resources.

Outcome

Data, tools, and services to improve patient-centered outcomes research relevant to HHS priorities.

Objective 1.1

Expand data capacity to enable patient-centered outcomes research for HHS national health priorities.

Interim Outcome

OS-PCORTF funded awards align with HHS national health priorities and data needed to study PCOR questions.

Potential Strategies[b]

1.1.1 Identify and prioritize the PCOR questions to be studied and the specific data and data infrastructure needed (Ongoing).

1.1.2 Support projects to produce the specific data and data infrastructure needed (Ongoing).

Objective 1.2

Address data infrastructure gaps to improve patient-centered outcomes research for the two research priorities—maternal mortality and intellectual and developmental disabilities—specified in the reauthorization of the PCORTF.

Interim Outcome

Projects addressing data capacity for maternal health and people with intellectual and developmental disabilities are initially prioritized.

Potential Strategies

1.2.1 Identify and prioritize PCOR questions for maternal health and the specific data and data infrastructure needed to address them (Short-term).

1.2.2 Improve the linkage of clinical data and other health data to study maternal health outcomes (Short-term).

1.2.3 Support projects to produce the specific data and data infrastructure needed for maternal health PCOR studies (Long-term).

1.2.4 Identify and prioritize PCOR questions for intellectual and development disabilities and the specific data and data infrastructure needed to address them (Short-term).

1.2.5 Support projects to produce the specific data and data infrastructure needed for intellectual and developmental disabilities PCOR studies (Long-term).

---

[b] Potential strategies are intended to illustrate the types of activities that would achieve the objectives (and are not a definitive or exhaustive list). These strategies span varying time horizons: short-term (expected completion within two to four years); long-term (expected completion in five or more years); or ongoing (continuous activity).

Objective 1.3

Develop collaborations with federal programs, researchers, and data stewards to address high-priority topics relevant to data for patient-centered outcomes research and increase workforce capacity in data science.

Interim Outcome

Expanded and diversified expertise and collaboration models with federal programs, researchers, and data stewards.

Potential Strategies

1.3.1 Develop consortia of agencies working collaboratively to build data capacity for emerging priorities (Short-term).

1.3.2 Diversify and enhance OS-PCORTF expertise through partnership with internship/fellowship programs across HHS (Long-term).

Objective 1.4

Engage end users, including research data networks, PCOR researchers, and other entities within the PCOR ecosystem throughout the OS-PCORTF product development lifecycle.

Interim Outcome

OS-PCORTF data assets and products are developed and tested to be fit for use by end users across the PCOR ecosystem.

Potential Strategies

1.4.1 Require OS-PCORTF applicants and awardees to identify the end users of their products and set forth (and execute) plans to engage them throughout the product development lifecycle and adjust the product so that it is fit for use (Ongoing).

1.4.2 Enhance visibility and accessibility of HHS data assets to the broader PCOR ecosystem (Long-term).

Goal 2: Data Standards and Linkages for Longitudinal Research

 Expand longitudinal data resources that enable patient-centered outcomes research to advance evidence generation.

Given the fragmented nature of data collection, linking datasets is often required to ensure that sufficient data are available to address specific PCOR questions of interest. Data linkage may involve matching patient records across multiple sources as well as resources and time to build an analytic dataset that is comprehensive and longitudinal—namely, following individuals over time and across settings. The ability to match patient records generally depends on the availability and completeness of demographic data fields across the datasets. Different systems may collect different demographic information (or have differently structured fields), and the information may be insufficient, incomplete, or erroneous, resulting in a linked dataset that does not provide an accurate health record. Use of

multiple records and sources can help ensure accuracy in linkage, but such aggregation increases the risks of privacy losses for the individuals whose data are involved.

Improving the availability, accessibility, and suitability of linked data for patient-centered outcomes research may require validated, privacy-preserving methods for data linkage as well as common data standards and harmonized common data models to ensure that the data being linked are interoperable and of sufficient quality (completeness, accuracy, and consistency) to support valid conclusions about the effects of interventions on outcomes. In addition to these more technical challenges, addressing issues of consent for use of linked data, and developing guidance about the reuse of these data for research, are needed to ensure transparency, credibility, and trust in PCOR studies. Policies and governance approaches are needed to support the development and use of longitudinal data sources, including adherence to best practices for data management and stewardship (e.g., Findability, Accessibility, Interoperability, and Reusability [FAIR] principles).[17]

Addressing these challenges with a well-aligned data linkage strategy, including the development of centralized resources for persistent data linkages across data assets, will also support more efficient research by reducing redundant efforts (e.g., researchers would not need to redo data linkages).

Outcome

Accessible, timely, interoperable, linkable, and longitudinal data.

Objective 2.1

Support the development and maintenance of data linkages for patient-centered outcomes research.

Interim Outcome

Access and use of linked federal datasets by research networks/researchers.

Potential Strategies

2.1.1 Coordinate with HHS agencies to continually assess and prioritize data linkage opportunities (Ongoing).

2.1.2 Develop a governance structure to maintain and support data linkages, select pilots, and evaluate impact (Long-term).

2.1.3 Engage HHS partners to build a consensus-based data linkage methodology for patient-centered outcomes research (Long-term).

Objective 2.2

Support and promote ongoing development and adoption of common data standards for patient-centered outcomes research.

Interim Outcome

Recommendations leading to the demonstrated adoption of common data standards in priority areas for patient-centered outcomes research.

> 2.2.1 Regularly scan and develop a crosswalk of data standards to identify issues and opportunities (Ongoing).
>
> 2.2.2 Solicit input from experts to develop a governance structure to promote the use and adoption of data standards (Short-term).
>
> 2.2.3 Evaluate the adoption and use of data exchange standards (e.g., Health Level 7® Fast Healthcare Interoperability Resources®) and their capabilities to advance patient-centered outcomes research priorities (Long-term).

Objective 2.3

Advance the accessibility and improve the quality of longitudinal and complex data for PCOR users.

Interim Outcome

High-quality linked datasets are accessible and used across research projects.

Potential Strategies

> 2.3.1 Solicit input from experts to identify ways to improve data quality and accessibility (Ongoing).
>
> 2.3.2 Identify, develop, and evaluate datasets to improve quality and accessibility (Long-term).

Objective 2.4

Assess the impact of policies related to privacy, security, and consent specific to patient-centered outcomes research.

Interim Outcome

PCOR projects incorporate relevant guidance on privacy, security, and consent policies.

Potential Strategies

> 2.4.1 Critically review privacy-preserving methodologies and report on how these methodologies can be applied to federal data sets for PCOR studies (Long-term).
>
> 2.4.2 Convene an expert group to assess the impact of policies related to privacy, security, and consent practices (Long-term).
>
> 2.4.3 Develop partnerships to identify mechanisms to address privacy, security, and consent concerns (Long-term).

Goal 3: Technology Solutions to Advance Research

Leverage leading technology solutions to improve data capacity for patient-centered outcomes and comparative clinical effectiveness research.

Advances in analytic methods and technology (e.g., computing capacity) can improve the quality and relevance of research findings from PCOR studies, including supporting increased integration and use of the real-world data (RWD) routinely collected on patients' health status and health care delivery (e.g., outside the context of a clinical trial).[12] AI solutions, such as ML and NLP methods, have the potential to address challenges with big data. Demonstrating and evaluating the use of leading technology solutions is imperative; insufficient understanding of the strengths and limitations of these approaches (e.g., accuracy and bias) can worsen health outcomes and disparities.[18] However, when used responsibly, these methods and technology solutions have the potential to improve data capacity in terms of timeliness and patient-centeredness, including expanding the types of data that can be used in PCOR studies. Advances in tools that facilitate the sharing and use of these data for research, such as standards-based application programming interfaces (APIs), further enhance the capacity for PCOR studies to generate stronger RWE and support clinical decision making.

Outcome

Robust RWD across platforms and systems used to generate RWE and expand data usage that informs patient, clinical, and policy decision making.

Objective 3.1

Support implementation of new technology solutions to advance timely access to data for patient-centered outcomes research.

Interim Outcome

ML techniques tested as a privacy-preserving method for accessing Health Information Exchanges (HIEs) in a few sample states.

Potential Strategies

3.1.1   Disseminate through presentations and publications the lessons learned from the OS-PCORTF maternal health project exploring the implementation of new technology tooling to facilitate faster adoption of data standards (Short-term).

3.1.2   Expand the use of ML split learning technique to more HIEs as case studies to facilitate data exchange for patient-centered outcomes research (Long-term).

3.1.3   Develop and implement short-term research projects as use cases and disseminate the findings through presentations and publications (Long-term).

3.1.4   Implement technology solutions for preserving privacy using different data sources for patient-centered outcomes research, such as linked data (EHRs and claims) and surveys (Long-term).

Objective 3.2

Develop and implement a coordinated strategy across OS-PCORTF projects to showcase use of technology solutions expanding PCOR and comparative effectiveness research methodologies.

Interim Outcome

ML and other advanced technology solutions are implemented and evaluated in OS-PCORTF projects.

Potential Strategies

3.2.1 Identify opportunities and support projects to pilot, validate, and implement predictive modeling using ML and other AI solutions to inform clinical decision making (e.g., review clinical guidelines to be inclusive for all) (Short-term).

3.2.2 Identify opportunities and support projects that address AI-related algorithmic discrimination and systematic bias in health and health care (Short-term).

3.2.3 Use NLP and ML methods to enhance the identification of stimulant and opioid-involved health outcomes in federal data (Short-term).

3.2.4 Expand the use of ML techniques and other AI solutions to other data relevant to patient-centered outcomes research (e.g., Centers for Medicare & Medicaid Services [CMS] claims, All-Payers' Claims Datasets [APCDs]) to further improve patient-centered outcomes and comparative effectiveness research capacity (Long-term).

Objective 3.3

Use AI solutions to enhance accessibility and interoperability of unstructured data to advance patient-centered outcomes and comparative effectiveness research.

Interim Outcome

AI solutions for integrating unstructured data are tested, validated, and implemented in OS-PCORTF projects.

Potential Strategies

3.3.1 Develop a PCOR learning network to facilitate coordinated communication and exchange of knowledge, including collaboration on exploring the integration of clinical notes and self-reported outcomes into EHRs (Short-term).

3.3.2 Validate and implement strategies for integration of clinical notes and self-reported outcomes into longitudinal RWD through data linkages and adoption of technology solutions (Long-term).

3.3.3 Identify projects to constantly test and validate new technology solutions that improve access to and use of structured and unstructured data to generate RWE, including standards-based APIs (Short-term).

Objective 3.4

Explore the use of new technologies to enhance patient-centeredness in patient-centered outcomes research.

Interim Outcome

ML and other advanced technology solutions explored as options for improving the use of other relevant data, such as self-reported information, to enhance patient-centered outcomes research.

Potential Strategies

3.4.1   In collaboration with experts, develop strategies for seeking and integrating patients' input through EHR tools (e.g., sampling methodologies and survey modes) (Short-term).

3.4.2   Identify projects to test and validate the strategies and methodologies identified in 3.4.1 (Short-term).

3.4.3   Apply ML techniques and other AI solutions to expand the use of relevant data (e.g., EHRs, claims such as CMS and APCDs, linked longitudinal data, and surveys) and improve capacity for patient-centered outcomes and comparative effectiveness research (Long-term).

Goal 4: Person-Centeredness, Inclusion, and Equity

Expand the collection and analysis of socioeconomic, environmental, and other data so all people making health care decisions have the evidence they value about the outcomes and effectiveness of health care.

As the scope of patient-centered outcomes research broadens, additional data are needed to support a more comprehensive perspective on health outcomes, to identify and address disparities, and to examine economic outcomes. Yet existing data do not capture the richness of people's characteristics and experiences. For example, data related to the total cost-of-care and a better understanding of cost considerations are often insufficient. Data for specific populations and communities are sometimes unavailable or are not representative. In other cases, the data might not be timely or might have other gaps that make it difficult to understand the impact of changes over long periods of time. Such data—on both outcomes and potential causal factors—include socioeconomic, environmental, and other data not generally collected within (or for use by) health systems. The availability, quality, and suitability of these data are often insufficient to support PCOR studies.

Addressing these data capacity limitations is therefore necessary for understanding the health and well-being of an individual as a whole (and what is important to that individual) and not just understanding the individual as a patient (person-centeredness); addressing high-priority questions for people who are medically underserved, underrepresented in biomedical research, and disproportionately affected (inclusion and equity);[c] and supporting economic impact analyses, consistent with the expanded scope of outcomes in the reauthorization of the OS-PCORTF.

---

[c] Populations that are medically underserved, underrepresented in biomedical research, and disproportionately affected are those that have been systematically denied full and equal access to critical opportunities and services, which often results in significant disparities in health outcomes.[19]

Outcome

Accurate, relevant, and representative evidence is accessible to individuals; communities; and state, federal, and tribal programs when making health care decisions.

Objective 4.1

Continuously improve data capacity and infrastructure to support people who are medically underserved, underrepresented in biomedical research, and disproportionately affected.

Interim Outcome

High-quality datasets are accessible to conduct PCOR studies on questions important to people who are medically underserved, underrepresented in biomedical research, and disproportionately affected.

Potential Strategies

4.1.1 Prioritize funding projects that focus on harmonizing, standardizing, and increasing uptake of data elements related to race, ethnicity, language, disability, sexual orientation, and gender identity (Short-term).

4.1.2 Identify and strengthen measures, including causes of mortality and morbidity, in consultation with community representatives (Short-term).

4.1.3 Promote adoption of common data models and infrastructure for participatory research with communities that are medically underserved, underrepresented in biomedical research, and disproportionately affected (Long-term).

4.1.4 Collaborate with public (federal, state, and tribal), nonprofit, and private sector entities to develop and sustain data linkages for disparate health and social services data sources to enable a more comprehensive, person-centered understanding of health (Long-term).

4.1.5 Collaborate with researchers and communities to identify questions relevant to health decisions of people who are medically underserved, underrepresented in biomedical research, and disproportionately affected (Long-term).

Objective 4.2

Strengthen the collection and use of environmental data to support populations that are medically underserved, underrepresented in biomedical research, and disproportionately affected.

Interim Outcome

Data for research demonstrate the integration of social, natural, and built environment data.[d]

Potential Strategies

4.2.1 Support efforts to identify, harmonize, and standardize relevant data elements and sources for the social, natural, and built environments (Short-term).

---

[d] Social determinants of health (SDOH) are the social and physical conditions in the environments where people are born, live, learn, work, play, and age. Affecting a wide range of health risks and outcomes, SDOH include conditions in the social environment (e.g., relationships and communities), natural environment (e.g., air and water), and built environment (e.g., buildings, transportation, open spaces, and infrastructure).[20]

4.2.2   Establish collaboration with other government agencies responsible for environmental data to identify use cases and infrastructure requirements to support people who are medically underserved, underrepresented in biomedical research, and disproportionately affected (Long-term).

Objective 4.3

Strengthen the capacity to conduct economic impact analysis within OS-PCORTF funded projects.

Interim Outcome

Data infrastructure is expanded to study the economic impacts of health care for informing decision making.

Potential Strategies

4.3.1   Solicit input from experts on economic variables to expand collection on cost-of-care data (Short-term).

4.3.2   Prioritize funding projects that focus on the collection of cost-of-care data and other economic variables and assessing their relationship to health outcomes (Short-term).

4.3.3   Promote new approaches to PCOR analyses encompassing diverse data, including environmental, economic, and health data, to address socioeconomic outcomes relevant to populations that are medically underserved, underrepresented in biomedical research, and disproportionately affected (Long-term).

4.3.4   Develop measures for the impact of long-term socioeconomic factors on health outcomes for populations that are medically underserved, underrepresented in biomedical research, and disproportionately affected populations (Long- term).

3.5 Strategic Framework

As presented below in Figure 2, the Strategic Framework captures the core elements in the architecture of the Strategic Plan. The foundational elements—mission, guiding principles, and communities—inform the aspirational elements—goals, outcomes, objectives, and vision—and the work that must be accomplished over the next decade to achieve them.
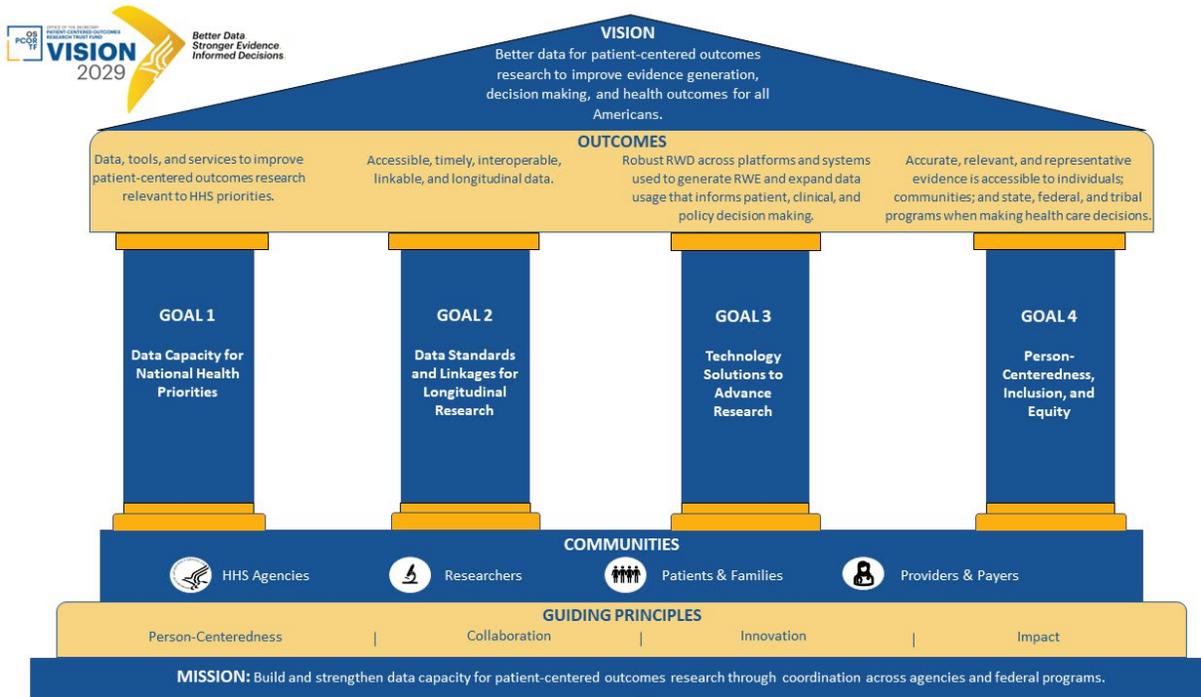

Figure 2. Strategic Framework

The Strategic Plan, including the Strategic Framework, reflects the results of ASPE's robust engagement of and collaboration with stakeholders and federal partners throughout the process (see Appendix A for further details on the approach and methodology). The emphasis on expanding the types of data that are collected, linked, and analyzed for PCOR studies is consistent with the national shift from a narrower *patient-centered* focus to a more inclusive *person-centered* focus that enables a more comprehensive understanding of health and the whole person. This approach is facilitated by advances in technology that have increased the data that can be collected and shared for use in PCOR studies, including data directly generated and shared by individuals (e.g., PGHD). Even with the availability of new types of data, it is critical that *all* data sources are examined to ensure their suitability for research (and analyzed with appropriate methods). The Strategic Plan therefore sets a course for addressing a set of high-priority challenges related to data availability, quality, accessibility, interoperability, linkage, and analysis. Successfully addressing these challenges requires the active participation and coordination of relevant stakeholders, leveraging of necessary resources, and the establishment of trust, credibility, and equity in processes and relationships. As such, ASPE will continue the robust engagement and collaboration throughout the implementation of the Strategic Plan.

4   Implementation

HHS's Strategic Plan for the OS-PCORTF (2020–2029) charts a course for advancing data capacity, including data linkages and analytical methods for PCOR studies, in four targeted goal areas. Successful execution of this Plan requires a systematic, coordinated approach to implementation—most importantly, an integrated set of strategies to achieve the objectives and a framework for ongoing collaborations, monitoring, evaluation, and dissemination of products.

The strategies include both technical and non-technical activities (e.g., collaboration) of varying time horizons (short-term, long-term, and ongoing). Across the four goals, the focus is on guiding and managing the investments of the OS-PCORTF—namely, the solicitation, funding, monitoring, and evaluation of projects and the dissemination of findings, products, and resources. The selection and implementation of future funding priorities will be informed by ongoing monitoring and evaluation of both the portfolio and advances in health care, data science, and the needs of decision makers. In particular, HHS will focus on critical areas unlikely to be funded or developed by other entities, in order to build a robust and cohesive portfolio of projects that collectively advances the data capacity and infrastructure integral to the success of patient-centered outcomes research.[21]

Moreover, because the Strategic Plan is designed to be responsive to evolving data infrastructure needs, priorities, and other relevant developments, timely evaluation of the various components and activities is necessary to ensure that strategies can be adapted as needed. Evaluation will be guided by appropriate methods and metrics aligned with the distinct nature of data capacity building efforts for patient-centered outcomes research and informed by an understanding of how such efforts are ultimately connected to the health and well-being of all Americans. In this way, the Strategic Plan (and its implementation) can provide an ongoing focus on impact, innovation, collaboration, and person-centeredness for the work conducted under the OS-PCORTF.

5   Conclusion

Generating evidence for informing decisions that improve the health and well-being of all Americans depends on data—data that are high-quality, accessible, and transformable into actionable evidence to strengthen medicine, public health, and social services. Ensuring that these data are available and used to support research, analysis, policymaking, and program delivery for the American people requires robust data capacity and infrastructure. ASPE's work in leading the coordination of collective, systematic efforts to develop, improve, and sustain this data capacity is critical for supporting patient-centered outcomes research.

This Strategic Plan sets forth HHS's approach and vision to fulfilling its distinct statutory mandate to build and strengthen the capacity to collect, link, and analyze data for patient-centered outcomes research over the next decade. Ongoing collaboration and engagement will ensure that the goals and objectives are realized and remain responsive to the needs of the Department and its agencies, federal partners, and the broader PCOR community. This Strategic Plan sets HHS and its partners on a course to achieve its vision for the OS-PCORTF – improving evidence generation, decision making, and health outcomes for all Americans.

Abbreviations and Acronyms

Term	Definition
ACA	Patient Protection and Affordable Care Act
AHRQ	Agency for Healthcare Research and Quality
AI	Artificial Intelligence
APCD	All-Payer Claims Database
API	Application Programming Interface
ASPE	Assistant Secretary for Planning and Evaluation
CER	Comparative Effectiveness Research
CMS	Centers for Medicare & Medicaid Services
EHR	Electronic Health Record
FDA	Food and Drug Administration
FFRDC	Federally Funded Research and Development Center
HHS	U.S. Department of Health and Human Services
HIE	Health Information Exchange
HP	HHS Office of Health Policy
ID/DD	Intellectual and Developmental Disabilities
IOM	Institute of Medicine
ML	Machine Learning
NASEM	National Academies of Sciences, Engineering, and Medicine
NLP	Natural Language Processing
OS	Office of the Secretary
OS-PCORTF	Office of the Secretary – Patient-Centered Outcomes Research Trust Fund
PCOR	Patient-Centered Outcomes Research
PCORI	Patient-Centered Outcomes Research Institute
PCORTF	Patient-Centered Outcomes Research Trust Fund
PGHD	Person-Generated Health Data
PPI	Participant-Provided Information
PRO	Patient-Reported Outcome
RCT	Randomized Controlled Trial
RWD	Real-World Data
RWE	Real-World Evidence
SDOH	Social Determinants of Health
SPEC	Strategic Planning Engagement Council

Appendix A. Approach and Methodology

The Office of the Assistant Secretary for Planning and Evaluation (ASPE) pursued a comprehensive approach to the development of the Strategic Plan using the Kellogg Logic Model.[22] Central to this process was a deliberate iterative approach to increase awareness, facilitate a shared understanding, and solicit feedback from stakeholders[e] to support the development and execution of the Strategic Plan as shown in Figure 3.
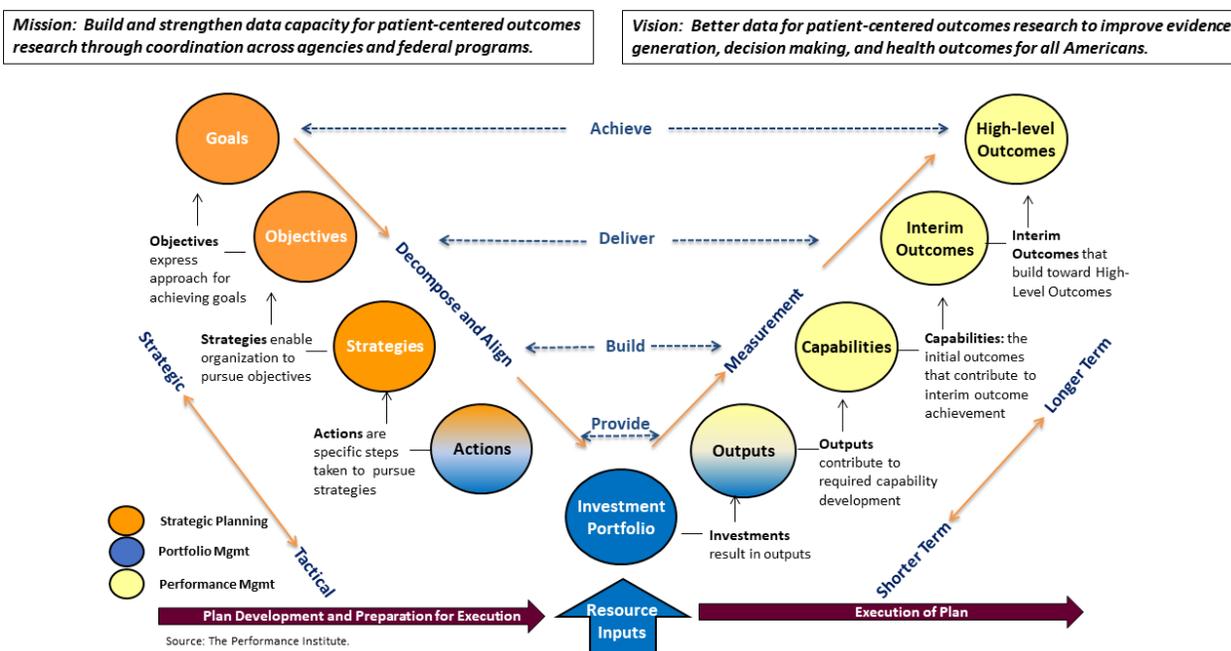
Figure 3. Logic Model Leveraged for OS-PCORTF Strategy Development

A.1   Stakeholder Engagement

ASPE committed to collaboration with robust stakeholder engagement across all phases of the strategic planning cycle through a variety of activities at different levels of intensity. For example, data gathering through an environmental scan and listening sessions was conducted at the beginning of Strategic Plan development. Stakeholders were consulted (e.g., asked for specific input on components of the proposed plan) and invited to review plan components (e.g., read plans and give guidance) throughout its development. Stakeholder engagement activities are summarized in Figure 4 and Table 1 below.

---

[e] Stakeholders are persons or groups who have an interest or concern in a project, activity, or course of action.[23] The term "stakeholder" is used across many disciplines to reflect different levels of input or investment in projects or activities. In this appendix, this term is used when referring to a diverse set of disciplines, communities, or perspectives as a whole.

The findings from each of these efforts can be found in multiple reports, including:

- *Challenges and Improvements for PCOR Data Infrastructure: Results from a Stakeholder Prioritization Activity*[24]
- *Research Data Networks and Patient-Centered Outcomes Research Trends and Opportunities: Scan and Interviews with Key Informants*[25]
- *U.S. Department of Health and Human Services Stakeholder Engagement Report Executive Summary – Analysis and Synthesis of Key Themes*[26]
- *Building Data Capacity for Patient-Centered Outcomes Research: Interim Report One – Looking Ahead at Data Needs (2021)*[27]
- *Building Data Capacity for Patient-Centered Outcomes Research: Interim Report Two – Data Standards, Methods, and Policy (2021)*[28]
- *Building Data Capacity for Patient-Centered Outcomes Research: Interim Report Three – A Comprehensive Ecosystem for PCOR (2022)*[29]
- *Building Capacity for Patient-Centered Outcomes Research (PCOR): Consensus Committee Report – Priorities for the Next Decade (2022)*[21]

Engagement by Data Gathering

MITRE | **Environmental Scan on Current Legislation and Strategies**
Scanned: Legislation, regulations, and strategies
Completion: January 2021

Engagement by Consultation and Review

ASPE | **HHS OS-PCORTF SPEC**
Principal Stakeholder Group: Sixteen representatives from HHS agencies
Completion: February 2022

HHS OS-PCORTF Leadership Council
Principal Stakeholder Group: Advisory entity for ASPE
Ongoing

Engagement by Listening

NORC | **NORC/AcademyHealth Initial Listening Session**
Principal Stakeholder Group: Health services researchers, policy analysts, and other PCOR researchers and users of PCOR products
Completion: August 2020

The National Academies of SCIENCES ENGINEERING MEDICINE | **NASEM Consensus Study Committee Workshops**
Principal Stakeholder Group: Non-governmental stakeholders involved in producing and using data for PCOR and other health, health care, and systems research, policy, and practice purposes
Completion: March 2022

MITRE | **Clinical Registries and Health Outcomes Research Data Network Stakeholder Interviews**
Principal Stakeholder Group: Principal investigators or leads of research data networks
Completion: July 2021

HHS Stakeholder Interviews
Principal Stakeholder Group: HHS Agency stakeholders that produce and use data for PCOR and conduct/participate in PCOR activities
Completion: September 2021

Figure 4. Foundations Informing Strategy Development

Table 1. Strategic Planning Stakeholder Engagement

Strategic Planning Stakeholder Engagement
Engagement by Data Gathering

Environmental Scan on Current Legislation and Strategies

- **Purpose:** Understand the implications of scanned material for building data capacity for patient-centered outcomes research per ASPE's statutory charge, identify opportunities to leverage existing work or partner with federal agencies, and avoid unnecessary duplication with other agencies doing compatible work.
- **Scanned:** Eleven evaluations of 10 pieces of legislation and regulations and 21 strategies.
- **Results:** Twenty-seven opportunities for the Office of the Secretary – Patient-Centered Outcomes Research Trust Fund (OS-PCORTF) in three main categories (Management, Technical, and Cultural) (Appendix B.1).
- **Conducted by:** The Health Federally Funded Research and Development Center (Health FFRDC), operated by The MITRE Corporation.

Engagement by Listening

Initial Listening Session

- **Purpose:** Understand gaps and challenges with the OS-PCORTF infrastructure functionalities and associated goals, as found in multiple reports, such as the *Building Data Capacity for Patient-Centered Outcomes Research in HHS: A Formative Evaluation of 2012–2016 Projects* report.[30]
- **Participants:** Thirty-six health services researchers, policy analysts, and other PCOR researchers and users of PCOR products.
- **Results:** Identification of 87 data infrastructure challenges and 76 data infrastructure improvements, distilled to five common themes (Appendix B.2).
- **Conducted by:** NORC at the University of Chicago and AcademyHealth.

Clinical Registries and Health Outcomes Research Data Network Interviews

- **Purpose:** Understand the common challenges and opportunities for OS-PCORTF to support patient-centered outcomes research identified by the research networks, important end users, and stakeholders.
- **Participants:** Environmental scan of 15 research networks and eight interviews with 14 principal investigators or leads of these networks.
- **Results:** Five common challenges and five areas of opportunity for the OS-PCORTF (Appendix B.3).
- **Conducted by:** The Health FFRDC.

U.S. Department of Health and Human Services (HHS) Internal Engagements: HHS Stakeholder Interviews

- **Purpose:** Understand key agency priorities related to patient-centered outcomes research; existing work that can be leveraged to improve data capacity; gaps and opportunities; trends and drivers in the environment; and how the current process for creating and funding OS-PCORTF projects could be improved.
- **Participants:** Thirty-two interviews with 62 HHS agency stakeholders who produce and use data for patient-centered outcomes research and conduct/participate in PCOR activities (HHS Agency Leaders, OS-PCORTF Project Leaders, Data Experts).
- **Results:** Twelve common themes across five categories (healthy equity and populations of focus, data infrastructure, collaboration, OS-PCORTF portfolio management, and workforce capacity) (Appendix B.4).
- **Conducted by:** The Health FFRDC.

External Engagements: National Academies of Sciences, Engineering, and Medicine (NASEM) Consensus Study Committee Workshops

- **Purpose:** Identify issues critical to building data capacity for patient-centered outcomes research and generating new evidence to inform health care decisions for the next decade.
- **Participants:** Three public workshops conducted in May and June 2021 with key non-governmental stakeholders involved in producing and using data for patient-centered outcomes research and other health, health care, and systems research, policy, and practice purposes.

Strategic Planning Stakeholder Engagement
• **Results:** NASEM Interim Report One – Looking Ahead at Data Needs identified 12 conclusions across five categories (fundamental data challenges, data fragmentation, health disparities, patient data needs and engagement, and focus on the person as a whole). NASEM Interim Report Two – Data Standards, Methods, and Policy identified nine conclusions across three categories (data standards, methods, and data policy and other data infrastructure considerations). NASEM Interim Report Three - A Comprehensive Ecosystem for PCOR identified eight conclusions across five categories (federal partnerships, state level data and collaborations, clinical trial networks and collaborations, public-private partnerships, and collaborations with patient groups). NASEM Final Report – Priorities for the Next Decade identified 27 conclusions across 10 categories (focus on the person as a whole, data types, fragmentation, data not generated for research, data governance, data access, research practices and analytic methods, project selection, dissemination of results and use, updates to the data infrastructure) (Appendix C).
• **Conducted by:** NASEM-appointed consensus study committee.

Engagement by Consultation and Review
Strategic Planning Engagement Council (SPEC)
• **Purpose:** A group of principal designees across HHS agencies to offer guidance and consultation that inform the Strategic Plan.
• **Participants:** Sixteen principal or designee members from HHS agencies.
• **Results:** Monthly discussions and consultations on strategic plan development to refine and improve the Plan's goals, outcomes, objectives, and strategies.
• **Conducted by:** ASPE Division of Health Care Quality and Outcomes.
HHS OS-PCORTF Leadership Council
• **Purpose:** Ongoing.
• **Participants:** Advisory entity for ASPE Division of Health Care Quality and Outcomes.
• **Results:** Quarterly discussions and consultations on OS-PCORTF themes and direction.
• **Conducted by:** ASPE Division of Health Care Quality and Outcomes.

A.2 Analysis and Synthesis

Following the data gathering and listening activities (see Table 1), inputs and findings were analyzed and synthesized to inform the development of the Strategic Plan.

A.2.1 Analysis of Inputs

ASPE undertook a rigorous and iterative cross-input analysis to craft this Strategic Plan, involving consideration of feedback from the Initial Listening Sessions, an in-depth qualitative analysis of HHS Internal Engagements, conclusions from the NASEM workshops, and feedback from the SPEC and HHS OS-PCORTF Leadership Council.

Specific to the HHS Internal Engagements, the Environmental Scan on Current Legislation and Strategies, and the Clinical Registries and Health Outcomes Research Data Network Stakeholder Interviews, the following steps were conducted to complete the qualitative analysis:

- Entered source documents (interview notes, transcripts, and scan findings) into a qualitative analysis software tool.
- Developed a coding taxonomy based on keywords from interview questions and PCOR concepts discussed with stakeholders.

- Summarized source language to arrive at 1,758 observations. An observation is a brief objective summary statement of a portion of the direct transcription, notes, or findings from the source document.
- Coded, sorted, and grouped like observations and labeled each with a conclusion statement to arrive at 134 conclusions. A conclusion is an objective summary statement of like observations.
- Clustered like conclusions into similar groups to develop 12 common themes in five broad categories focused on ways to enhance data infrastructure for the OS-PCORTF. A theme is a summary statement of like conclusions, written in the form of a problem statement.
- Using the themes and conclusions, the ASPE team crafted a set of candidate goals, outcomes, and objectives, and example strategies for the Strategic Plan.

A.2.2 Synthesis

To synthesize the results across all stakeholder activities, ASPE triangulated inputs and validated the resulting themes and conclusions to support the Strategic Plan components with the SPEC. Triangulation is a qualitative research approach in which multiple methods and data sources are used to develop a comprehensive understanding of the subject matter and to test the validity of the convergence of findings from these different sources.[31]

As part of this process, the ASPE team layered considerations and recommendations generated from the Initial Listening Session, the NASEM workshops, and SPEC feedback into the findings from the qualitative analysis described above. As reports from these inputs were made available, the ASPE team reviewed, compared, and adjudicated the considerations and recommendations against the results of the qualitative analysis. The findings from these inputs largely aligned with the findings of the qualitative analysis, adding richness to the existing themes, conclusions, and Strategic Plan components. In some areas, stakeholder inputs led to the addition and refinement of goals, outcomes, objectives, and potential strategies. Issues of scope and authority (legislative or regulatory constraints) and operational considerations (improvements to the selection and funding of OS-PCORTF-funded projects) also informed the framing and development of the Strategic Plan.

The triangulation, validation, and synthesis of findings were products of monthly collaboration with ASPE and the SPEC. ASPE and the SPEC went through several iterative review and feedback cycles on the goals, outcomes, objectives, and potential strategies to identify, prioritize, and address gaps and challenges with PCOR data capacity.

Appendix B. Findings from Stakeholder Engagement Activities

Through multiple engagement activities, the Office of the Secretary – Patient-Centered Outcomes Research Trust Fund (OS-PCORTF) gathered perspectives and evidence to inform all levels of strategic planning with the participation of internal and external stakeholders[f] and the public. This appendix reflects the high-level findings from four stakeholder activities: *Environmental Scan on Current Legislation and Strategies; Listening Session; Research Data Networks and Patient-Centered Outcomes Research Trends and Opportunities: Scan and Interviews with Key Informants;* and the *HHS Engagement Executive Summary.*[25, 26]

B.1   Environmental Scan on Current Legislation and Strategies

The Office of the Assistant Secretary for Planning and Evaluation (ASPE), with the support of the Health Federally Funded Research and Development Center (Health FFRDC) operated by the MITRE Corporation, conducted an environmental scan of 21 current data strategies, 10 pieces of legislation, and 11 evaluations. Scanned materials were reviewed to understand the implications for building data capacity for patient-centered outcomes research (PCOR) as defined by the OS-PCORTF's statutory charge, to identify opportunities to leverage or partner with federal agencies on their priorities, and to avoid unnecessary duplication with other agencies doing similar work. Opportunities for the OS-PCORTF portfolio were identified and divided into three main categories, as shown in Table 2.

Table 2. OS-PCORTF Portfolio Opportunities Identified from the Environmental Scan

OS-PCORTF Portfolio Opportunities Identified from the Environmental Scan
Management
• More clearly define OS-PCORTF portfolio objectives and specify milestones and time frames.
• Identify metrics to facilitate portfolio-wide assessment, award-specific assessment, and dissemination and translation to ensure that award goals are achieved.
• Consider refinements to the OS-PCORTF proposal process (e.g., weighting of impact, sustainability, novelty, cross-agency collaboration, cross-agency peer review).
• Fund projects that address emerging policy issues enumerated in the reauthorization—intellectual and developmental disabilities (ID/DD), maternal health, and economic burden as a key patient outcome.
• Collaborate regularly with the Patient-Centered Outcomes Research Institute (PCORI) and the Agency for Healthcare Research and Quality (AHRQ) to maintain alignments on prioritization, data needs, and dissemination and training needs that could become components of the OS-PCORTF proposal documentation.
• Engage patients in data collection, procurement, structure, quality, use, and donation, as well as in federally funded PCOR data infrastructure projects.
• Improve OS-PCORTF dissemination efforts to promote awareness of initiatives and products (e.g., cross-agency collaboration, centralized hub for updates/tools/data/standards, external engagements).
• Support other HHS efforts to build data capacity.

---

[f] Stakeholders are persons or groups who have an interest or concern in a project, activity, or course of action.[23] The term "stakeholder" is used across many disciplines to reflect different levels of input or investment in projects or activities. In this appendix, this term is used when referring to a diverse set of disciplines, communities, or perspectives as a whole.

OS-PCORTF Portfolio Opportunities Identified from the Environmental Scan

- Act as a cross-agency convener to foster a shared direction and vision and avoid duplicative work (e.g., via a coordinated, centralized leadership body).
- Expand and enhance engagements beyond portfolio "super users."
- Improve sustainability of projects to enhance uptake, longevity, and impact (e.g., through sustainability planning, early identification of needs and priorities, improved engagement with end users).
- Reconsider what functionalities should be prioritized in the OS-PCORTF strategic framework (e.g., consider adding data quality, governance).
- Establish working groups for each of the five functionalities to advise on relevant federal initiatives that could be leveraged, metrics, and evolving needs.
- Monitor potential disruptors to data management and data science (e.g., digital tools, artificial intelligence [AI]/machine learning [ML] advances, citizen science, social media, cloud), including international and other non-Department of Health and Human Services trends and initiatives (e.g., Defense Advanced Research Projects Agency, U.S. Department of Veterans Affairs, Cybersecurity and Infrastructure Security Agency), and ensure that data ecosystems are designed with flexibility to accommodate new insights and data types.

Technical

- Identify and facilitate access to a broader range of data beyond the data found in electronic health records (EHRs) (i.e., data that are not collected during conventional clinical care—patient-reported outcomes [PROs]/PRO measures data, social determinants of health [SDOH] data, social services data, clinician data, data collected from wearable technology, social media, and through telehealth arrangements)—and integrate these data as part of EHR systems.
- Enable linkages and interoperability of all types of data.
- Develop and implement core outcome sets and key data/metadata (e.g., SDOH, participant-provided information [PPI], device-generated health data from personal devices) that improve relevance of clinical data for research.
- Enhance data quality for research.
- Continue to support data standardization efforts, including for SDOH and other high-priority data types, develop practical guidance on how to harmonize common data models across systems (e.g., promoting more widespread use of Fast Healthcare Interoperability Resources), and apply consistent standards for data collection that enable integration with EHRs (including standards for clinicians to enact within the clinical workflow).
- Support research networks in enabling AI and other analytics that typically require large amounts of centralized, accessible data to train algorithms (e.g., apply federated learning ML technique and other "use-don't-move" analysis models for data).

Cultural

- Work with "on the ground" research teams and pilot sites to surface legal, financial, governance, and technical challenges encountered while implementing and testing products.
- Address data governance issues, including data provenance, control, and ownership.
- Address barriers to implementation of common standards (e.g., cost, time, lack of incentives—such as grant requirements, proliferation of unaligned standards) and PROs/PRO measures (e.g., burden, clinical workflow, stakeholder buy-in).
- Develop policy framework for data privacy and security that is balanced with data access and interoperability/linkage needs, including guidance and standards for researchers to help maximize use of PPI while maintaining privacy and security (e.g., patient self-sovereignty over data).
- Integrate financial drivers and other incentives into the strategic framework, especially to promote open science and data sharing, use, and linkage for PCOR studies.
- Increase access to data assets by reducing barriers (e.g., cost, lags in data availability); developing cross-agency and public "sandboxes" where researchers can collaborate and explore cross-agency datasets;

OS-PCORTF Portfolio Opportunities Identified from the Environmental Scan
establishing public-private partnerships; developing case studies and best practices for cross-sector collaboration on data linkage/exchange; and supporting creation, promotion, and dissemination of a single catalog on the process for acquiring federal data assets.
• Evaluate research network investments to avoid wasteful redundancies and identify synergies/gaps.
• Develop strategies to encourage adoption and use of PROs by clinicians and patients (e.g., build into clinical workflow, minimize burden, perform timely calculation of PRO results, help patients/clinicians interpret and apply data).
• Address the "cycle of exclusion" of PRO/PPI collection from vulnerable or underserved populations (e.g., due to lack of technology access, limited incentives for participation, attrition), as well as limitations of EHRs which exclude populations and individuals that do not access health care.
• Contribute to advancing research on evaluating the effectiveness and implementation of social care practices and policy initiatives.

B.2   Listening Session

ASPE sought to gather perspectives on patient-centered outcomes research data infrastructure needs and priorities from a diverse group of stakeholders through an online prioritization activity supported by NORC at the University of Chicago and AcademyHealth. The goal of the activity was to provide feedback on PCOR data infrastructure gaps and priorities for the next decade.[24]

This stakeholder group of nearly 40 participants—with a wide range of occupational backgrounds, including policy, health care delivery, research, and informatics—participated in the online prioritization activity. Participants included:

- PCORI stakeholders
- Members of AcademyHealth interest groups (e.g., Health Information Technology, Learning Health System, Public Health Systems Research, Quality and Value)
- Health system representatives (Department of Veterans Affairs Health Services Research and Development Service, Kaiser Permanente, Geisinger, Sanford Health)
- Technical advising bodies (e.g., Electronic Data Methods Forum, Methods and Data Council, State University Partnership Learning Network, Medicaid Medical Directors Network)
- Industry representatives (e.g., Cerner, Epic)
- Public health communities (e.g., American Medical Informatics Association 2020 Scientific Program Committee Leadership members, National Interoperability Collaborative, Data Across Sectors for Health)
- Members of the OS-PCORTF Technical Expert Panel

Participants were asked to review the five functionalities that define and form the basis for robust PCOR data infrastructure, central to ASPE's strategic framework for building data capacity for patient-centered outcomes research, and provide comments on challenges and improvements in each area, as well as generate and rank a list of potential activities.[24] The prioritization activity occurred in three parts: (1) a virtual listening session, (2) an idea generation activity using the Codigital platform, and (3) an online prioritization activity using the Codigital platform.

The virtual listening session consisted of an overview of the activity, the background on OS-PCORTF, and a draft set of questions that would guide the idea generation activity and the prioritization activity. Both

the idea generation and online prioritization activity generated a total of 87 data infrastructure challenges and 76 data infrastructure improvements. For each functionality, participants identified five top challenges and five top improvements, developed a summary statement, and identified additional key themes that warranted discussion.

Across the five functionalities, participants returned to five common themes that focused on the need to:

1. **Enhance consistency in data standardization.** Participants raised multiple challenges and improvements related to the issue of consistency or transparency in data standardization. Participants focused on the need for consistent processes for collecting, cleaning, and presenting data. They also highlighted the importance of promoting adoption of and adherence to standards across the health system after they are developed.

2. **Improve access to SDOH data that are not routinely collected during care delivery.** Participants sought resources to support the standardized collection of SDOH data and expressed a need for expanded access to federal datasets to support research inquiries related to SDOH, including zip code-level data on neighborhood characteristics.

3. **Improve ability to access, integrate, and use PPI, particularly those data generated from medical devices and wearables.** Stakeholders noted the importance of accessing PPI, including PROs and patient-generated health data, from medical devices to support their research inquiries. Participants focused on the need to develop and disseminate standards to support PPI data collection and analysis (including the collection of PROs) and the aggregation and integration of PPI into electronic health records. Participants also sought mechanisms to promote collection and use of PROs among patients and clinicians.

4. **Increase access to federal datasets, with an emphasis on access to de-identified datasets**. Across the functionalities, access to data sources was a prominent theme. As previously noted, participants focused on access to SDOH data resources, including federal datasets with SDOH data, across multiple functionalities. Participants broadly highlighted the need for increased access to federal data resources. This topic was particularly prominent for the subject of using federal databases for research, where participants focused on the need for easily accessible, de-identified federal datasets that include SDOH data. Participants underscored the need for access to surveillance data, an emerging topic given the current Coronavirus Disease 2019 global pandemic.

5. **Expand collaboration across organizations at the local, state, and federal levels.** Participants highlighted the need for collaboration to leverage and enhance existing data sources and infrastructure. Collaboration was discussed at both the meso level (e.g., collaboration to enable cross-sector data sharing) and the macro level (e.g., regulatory frameworks, enhanced federal data assets, and development of and incentives for standards adoption). Based on the challenges and improvements they submitted, participants foresaw the need for widespread cooperation to make data available and useful for research, while maintaining the privacy and security of patient health information.

The priority challenges and improvements identified in this activity generally aligned with ASPE's existing strategic framework. Participants did not signal a need to expand the definitions of the five functionalities or significantly adjust ASPE's strategic framework. Discussion included alignment with

OS-PCORTF projects and alignment with ASPE milestones for progress. Overall, the activity served to underscore the importance and relevance of the framework and the work of the OS-PCORTF.

The compilation of these findings can be found in this report: *Challenges and Improvements for PCOR Data Infrastructure: Results from a Stakeholder Prioritization Activity.*[24]

B.3 Clinical Registries and Health Outcomes Research Data Network Interviews

In support of ASPE, the Health FFRDC completed an environmental scan of 15 health outcomes research networks that conduct or are capable of supporting PCOR studies. In addition, eight interviews were held with 14 principal investigators or leads of these research networks to understand the common challenges and opportunities for OS-PCORTF to support PCOR studies.

Five common challenges and future priorities that represent opportunities for the OS-PCORTF to support patient-centered outcomes research were identified, as shown in Table 3.

Table 3. Challenges and Potential Opportunities for OS-PCORTF to Support PCOR Identified by the Research Networks

Challenges and Potential Opportunities for OS-PCORTF to Support PCOR as Expressed by the Research Networks	
Challenges	**Potential Opportunities**
Lack of high-quality real-world data (RWD).	Enhance research access to federal health data, expanded to include device and participant-provided information, among other novel sources.
Limited tools and resources for linking data between different sources (e.g., EHRs, claims, PROs).	Strengthen methods and tools to promote and sustain authoritative health data linkage.
Lack of tools to improve data quality and curation.	Develop and implement standard approaches for data quality, consistency, and patient identification.
Constant effort to maintain data quality.	Address source data workflow strategies for data capture to improve data quality.
Difficulty of accessing medical claims data, especially Medicaid data.	Address potential for bias against low-resource providers and their patients due to lags and inconsistencies in federal data available on managed care patients in Medicare and Medicaid and other sources.

The compilation of these findings can be found in this report: *Research Data Networks and Patient-Centered Outcomes Research Trends and Opportunities: Scan and Interviews with Key Informants.*[25]

B.4 HHS Stakeholder Interviews

The Health FFRDC, on behalf of ASPE, completed 32 interviews with 62 HHS agency stakeholders who produce and use data for patient-centered outcomes research and conduct/participate in PCOR activities (HHS Agency Leaders, OS-PCORTF Project Leaders, Data Experts). The goal of these interviews

was to understand key agency priorities related to patient-centered outcomes research, existing work that can be leveraged to improve data capacity, gaps and opportunities, trends and drivers in the environment, and how the current process for creating and funding OS-PCORTF projects could be improved.

From the interviews, 12 common themes were developed in five broad categories focused on ways to enhance data infrastructure for the OS-PCORTF, as shown in Table 4.

Table 4. Themes Across Topic Areas from the HHS Stakeholder Interviews

	Themes Across Topic Areas from the HHS Stakeholder Interviews
	1. There is a need to apply a health disparities and equity lens across platforms, projects, and data sources/core datasets to address research needs for underserved, underrepresented, and at-risk groups and populations. 2. Data for PCOR studies are insufficient in quantity, quality, and linkages for many population segments (e.g., COVID-19, ID/DD, Maternal/Child Health, Opioid Use).
	3. Patient-centered outcomes research requires the OS-PCORTF to be at the leading edge of new technologies and methods (e.g., AI, including ML and natural language processing, and other advanced analytics) to address issues with data access, analysis, and sharing. 4. Strategies are needed for using RWD for practical application for patient-centered outcomes research. 5. There is a need for efficient and sustainable/reusable data infrastructure that addresses standardization, liquidity, and transparency. 6. There is a continued need to invest in sustainable, robust data linkages to demonstrate intervention-associated outcomes. 7. There are limitations in the completeness, timeliness, access, cost, exchange, and granularity of data for effective patient-centered outcomes research. 8. There is a need to develop policies and governance approaches to follow patients over time and across settings that protect privacy and identity.
	9. There is a need to improve ASPE's collaboration infrastructure to build, scale, and sustain data capacity for PCOR studies.
	10. The OS-PCORTF must respond quickly to shifting departmental priorities and policies by funding projects to meet emerging needs. 11. There are challenges with the OS-PCORTF funding application process that affect the quality of applications.
	12. The current data science workforce across HHS limits the ability of agencies to engage in data analysis and sustain project outputs for patient-centered outcomes research.

The compilation of these findings can be found in this report: *U.S. Department of Health and Human Services Stakeholder Engagement Report Executive Summary – Analysis and Synthesis of Key Themes.*[26]

Appendix C. NASEM Consensus Study Committee: Building Data Capacity for Patient-Centered Outcomes Research

The Assistant Secretary for Planning and Evaluation (ASPE) asked the National Academies of Sciences, Engineering, and Medicine (NASEM) to appoint a consensus study committee to identify issues critical to building data capacity and infrastructure for patient-centered outcomes research (PCOR) over the next decade. The study was a collaboration of three units of NASEM: Committee on National Statistics, Board on Health Care Services, and Computer Science and Telecommunications Board. The consensus study committee had a diverse membership, with 15 members whose expertise and experience spanned the broad fields of (1) patient-centered outcomes research; (2) research methods, statistics, and demography; (3) computer science and data infrastructure; and (4) patient engagement and patient perspectives.

The committee organized three workshops to collect input from stakeholders on the PCOR data infrastructure. The workshop series brought together stakeholders to (1) look ahead at data needs, discuss the ability of the PCOR data infrastructure to meet those needs, and identify priorities going forward; (2) discuss data standards, methods, and policies that could make the PCOR data infrastructure more useful; and (3) discuss research and data collaborations that could make the PCOR data infrastructure more useful. A series of three interim reports and one final report summarize the discussion and committee findings and conclusions, each contributing to the development of this Strategic Plan.

Building Data Capacity for Patient-Centered Outcomes Research:

- Interim Report One – Looking Ahead at Data Needs (September 2021)[27]
- Interim Report Two – Data Standards, Methods, and Policy (October 2021)[28]
- Interim Report Three – A Comprehensive Ecosystem for PCOR (January 2022)[29]
- Final Report – Priorities for the Next Decade (May 2022)[21]

Table 5 summarizes the 12 conclusions in five broad categories identified in NASEM Interim Report One.

Table 5. NASEM Interim Report One Conclusions

Conclusions: Looking Ahead at Data Needs
Fundamental Data Challenges
• **Conclusion 2-2:** The data available for patient-centered outcomes research are often collected for reasons other than research, which limits their usefulness. Opportunities exist for increasing the utility of the data infrastructure by carefully considering the multiple uses to which the data might be applied. • **Conclusion 4-3:** Researchers encounter substantial barriers to accessing existing data for patient-centered outcomes research. Facilitating and simplifying data access could further increase the usefulness of data for research.
Data Fragmentation
• **Conclusion 2-3:** Existing data on the social determinants of health (SDOH) are found in a variety of databases. Barriers to linking across these data silos represent a major challenge to understanding how SDOH affect health outcomes.

Conclusions: Looking Ahead at Data Needs
● **Conclusion 4-2:** The data available for patient-centered outcomes research are fragmented across a variety of databases. Expanding data linkages could greatly increase the usefulness of these data for research.
Health Disparities
● **Conclusion 2-1:** Health disparities can occur across a broad range of characteristics and populations. Data limitations affect the ability to identify and understand these disparities in many areas. Data for specific populations are sometimes unavailable or are not representative. In other cases, the data might not be timely or might have other gaps that make it difficult to understand the impact of changes over long periods of time.
● **Conclusion 2-4:** Existing data do not capture the richness of people's characteristics and experiences. While such limitations are to be expected, opportunities exist for capturing data that are better able to characterize these complexities. A robust data infrastructure builds on the strengths of what is available today and has the flexibility to adapt, both as measures and terminologies become obsolete and as new technologies emerge.
● **Conclusion 2-5:** Prioritizing and improving the collection of data can lead to a better understanding of health disparities and to potential solutions for reducing disparities.
Patient Data Needs and Engagement
● **Conclusion 3-1:** The patient-centered outcomes research data infrastructure has not reached its full potential to provide data that can answer questions that matter to patients and enable them to make informed decisions. Information about the cost-of-care was highlighted among the types of data that would be particularly useful.
● **Conclusion 3-2:** Dissemination and translation of the research findings could be greatly enhanced by using forms of communication that are relevant to those outside of the research community.
● **Conclusion 4-4:** Making the data more visible and more widely accessible could enable patients and communities to use the information in ways that reduce health disparities, complementing research efforts in this area.
● **Conclusion 4-5:** Data needs related to the total cost-of-care and a better understanding of cost considerations are areas that deserve more attention.
Focus on the Person as a Whole
● **Conclusion 4-1:** Broadening the focus from the patient to the person more generally would enable a more comprehensive approach to the data infrastructure and a better understanding of the outcomes and impacts that matter to people.

Table 6 summarizes the nine conclusions in three broad categories identified in NASEM Interim Report Two.

Table 6. NASEM Interim Report Two Conclusions

Conclusions: Data Standards, Methods, and Policy
Data Standards
● **Conclusion 2-1:** Standards are most useful when their development is driven by their potential uses and a clear concept of the value they can contribute.
● **Conclusion 2-2:** ASPE could add significant value in the area of standards for patient-centered outcomes research by: – Continuing to promote the development of a data infrastructure and an implementation strategy that facilitates the use of standards and access to the data;

Conclusions: Data Standards, Methods, and Policy

– Convening stakeholder meetings to enhance communication and work towards developing a common language for standards;
    – Facilitating accessibility to the data and collaborations with existing organizations working in this area; and
    – Leading efforts to catalogue and exemplify data standards and analytic standards.
● **Conclusion 2-3:** While data standards are important to conducting patient-centered outcomes research, applying standards to the analytic methods as well is important to facilitate the reliability and reproducibility of study results.
● **Conclusion 2-4:** An international perspective is an important consideration for the patient-centered outcomes research data infrastructure, and the infrastructure focused on standards specifically would benefit from building on work that happens internationally.

Methods

● **Conclusion 3-1:** The ability to adopt a longitudinal, comprehensive perspective of an individual's journey could open new opportunities for patient-centered outcomes research. The shift could be facilitated by focusing on efforts to:
    – Simplify integration of data across the research data ecosystem;
    – Address challenges posed by the limitations associated with health identifiers;
    – Incorporate person-generated data into health data systems; and
    – Leverage real-world data to expand the timeline view of a person's health-related experiences.
● **Conclusion 3-2:** Observing scientific best practices, including those of transparency and ethical use of data, is essential to generate trust in patient-centered outcomes research among all stakeholders, including the public and researchers. This is important both for observational data and for emerging data sources and methods.
● **Conclusion 3-3:** The results of patient-centered outcomes research (and research in general) are only replicable and are most useful when the underlying data and comprehensive research documentation (such as analytic code) are made available for use by others.

Data Policy and Other Data Infrastructure Considerations
● **Conclusion 4-1:** Building and maintaining trust among the people and communities whose data are being sought for research is essential for high-quality data. Including representatives of consumers and patients in the research process to understand how to measure health impacts that matter to individuals is an important component in building trust. ● **Conclusion 4-2:** This is an opportune time to revisit and update the legislation and rules governing data privacy and the sharing of data for research.

Table 7 summarizes the eight conclusions in three broad categories identified in NASEM Interim Report Three.

Table 7. NASEM Interim Report Three Conclusions

Conclusions: A Comprehensive Ecosystem for PCOR
Federal Partnerships
● **Conclusion 2-1:** Collaboration among federal agencies and between federal agencies and other partners (such as states, patient groups, and others) is essential for continuing to build the PCOR data infrastructure. The areas where additional collaboration would be particularly useful include:

Conclusions: A Comprehensive Ecosystem for PCOR

– Increasing consistency in the use of standards for data interoperability and element definitions;
    – Addressing barriers that hinder data linkages, such as the limitations associated with health identifiers and mitigating potential selection biases resulting from linkage error;
    – Balancing the burden of the data collections and disclosure risks with the value of the datasets;
    – Communicating the usefulness of the data collections to those who are asked to provide data about themselves and those who collect the data;
    – Promoting discussion and education about fitness for use of the data; and
    – Working with stakeholders and patients to promote sharing of data.
● **Conclusion 2-2:** There is a need to increase awareness among all stakeholders about new data infrastructure developments funded by the PCOR Trust Fund. Increased awareness will enhance the efficiency and effectiveness of research, which will increase the impact of the investments made in infrastructure development.

State-Level Data and Collaborations
● **Conclusion 3-1:** There are opportunities to learn from what states have accomplished in building data capacity. ● **Conclusion 3-2:** The usefulness of data available for PCOR could be increased by the sharing and adoption of best practices among the states for the data collected, their quality, and ease of access.
Clinical Trial Networks and Collaborations
● **Conclusion 4-1:** Infrastructure investments could enhance the utility of data routinely generated in the course of care for clinical trials.
Public-Private Partnerships
● **Conclusion 5-1:** Successful partnerships across healthcare systems require participant trust, clear evidence of mutual benefit, and ability to control risk.
Collaborations with Patient Groups
● **Conclusion 6-1:** Patient groups can be helpful partners in all aspects of PCOR, including engaging patients in order to improve research participation and the impact of results. ● **Conclusion 6-2:** Patient-directed disease registries can be a source of in-dept**h**, longitudinal, prospective clinical and patient-reported data that are not available from other data sources.

Table 8 summarizes the 27 conclusions in ten categories identified in the Final NASEM Report.

Table 8. NASEM Final Report Conclusions

Conclusions: Priorities for the Next Decade
Focus on the Person as a Whole
● **Conclusion 3-1:** Broadening the focus from the patient to the person more generally and from populations to communities would enable a more comprehensive approach to the data infrastructure and a better understanding of the outcomes that matter to people.
Including High Priority Types of Data in the Data Infrastructure
● **Conclusion 3-2:** A variety of data types were identified that are less likely to be available or easily accessible in the PCOR data infrastructure, including data on mortality, cost-of-care, social determinants of health, and disability status, as well as other characteristics of people associated with disparities in health outcomes.

Conclusions: Priorities for the Next Decade

Increased attention to filling gaps in the availability of these data will enhance the utility of the infrastructure for answering questions that matter to people and will enable research on potential intervention targets.

- **Conclusion 3-3:** An area with opportunities for additional expansion is the collection of patient- and person-generated data and the routine integration of these data into data platforms that can be used both for research and for other purposes, including regulatory decision-making and to inform shared decision making.
- **Conclusion 3-4**: Patient-directed disease registries can be a source of in-depth, longitudinal, prospective clinical and patient-reported data that are not available from other data sources.
- **Conclusion 3-5:** Assembling a comprehensive longitudinal record of individuals' health journeys, which also includes the social context of their lives to the extent possible, would facilitate more far-reaching outcomes research.

Addressing Fragmentation

- **Conclusion 3-6:** The data available for PCOR are fragmented across a variety of databases. Expanding data linkages could greatly increase the usefulness of these data for research.
- **Conclusion 3-7:** Collaboration among federal agencies and between federal agencies and other partners to address barriers that hinder data linkages, such as the limitations associated with the lack of unique health identifiers and patient or person matching, will improve the PCOR data infrastructure. The usefulness of data available for PCOR could further be increased by sharing and adopting best practices among the states concerning the collection of data, data quality, and ease of access.

Data Not Designed to Advance Knowledge

- **Conclusion 3-8:** Standards are most useful when their development is driven by their potential uses and a clear concept of the value they can contribute.
- **Conclusion 3-9:** Taking an international perspective is important for the development of a PCOR data infrastructure; in particular, the infrastructure focused on standards would benefit from building on work that happens internationally.
- **Conclusion 3-10:** ASPE, in collaboration with other partners and stakeholders, could add significant value in the area of standards for PCOR by**:**
  – Continuing to promote the development of a data infrastructure and an implementation strategy that facilitate the use of standards and access to the data;
  – Convening stakeholder meetings to enhance communication and work toward developing a common language for standards;
  – Facilitating access to the data and collaborations with existing organizations working in this area;
  – Leading efforts to catalogue and exemplify data standards and analytic standards for a holistic view of individuals' health; and
  – Increasing consistency in the use of standards for data interoperability and element definitions.
- **Conclusion 3-11:** Prioritizing projects that address fidelity or use of standards may convey greater value for the PCOR infrastructure than developing new standards.

Governing Data Access

- **Conclusion 3-12:** This is an opportune time to revisit and update the legislation and rules governing data privacy and the sharing of data for research.
- **Conclusion 3-13:** Governance challenges that create barriers to developing the PCOR infrastructure can be found at all levels of the system. Data availability could be increased by exploring challenges at the local level, including variable interpretations of federal laws and regulations, and by identifying approaches to address those challenges.

Data Access Options

- **Conclusion 3-14:** Investments in identifying mechanisms for facilitating the ability of researchers, patients, and other people to access data will contribute to increased use of the PCOR infrastructure.

Conclusions: Priorities for the Next Decade
• **Conclusion 3-15:** Building and maintaining trust among the people and communities whose data are being sought for research is essential for producing high-quality data, and patient groups can be helpful partners in these efforts. Including representatives of patients and other people in the research process to understand how to measure health impacts that matter to individuals is an important component in building trust. Providing value back to data donors, such as through the sharing of research results, could help underscore the importance and benefits of the information to stakeholders, including individuals, families, clinicians, and communities, in addition to enabling them to use the information in ways they find relevant. These uses could play a particularly important role in reducing health disparities, complementing research efforts in this area.
• **Conclusion 3-16:** Successful data sharing partnerships across health care systems and government agencies require participant trust, clear evidence of mutual benefit, and the ability to control risk.
Advancing Research Practices and Analytic Methods
• **Conclusion 3-17:** PCOR products would be enhanced by investing in methods that are essential for the conduct of PCOR, such as including persons throughout the research continuum, addressing problems of missing data, improving study designs, ensuring appropriate inference from methods utilizing observational data, and addressing structural bias in data systems and studies.
• **Conclusion 3-18:** Applying best practices to the analytic methods used in PCOR is important to facilitate the reliability and reproducibility of study results.
• **Conclusion 3-19:** The results of PCOR are only replicable and most useful when the underlying data and comprehensive research documentation (such as analytic code) are made available for use by others.
Project Selection to Support the Data Infrastructure Framework
• **Conclusion 3-20:** The development of the data infrastructure might be enhanced and critical gaps could be filled by proactively identifying necessary projects in areas that examine the overall framework for the PCOR data infrastructure, particularly in the context of broader issues such as the balance between privacy and increased data use.
• **Conclusion 3-21:** Investments in areas unlikely to be funded or developed by other entities may have a particularly high value.
• **Conclusion 3-22:** Investments in projects that have potential use and application beyond the condition or disease for which they are proposed will accelerate the use of the infrastructure.
Dissemination of Results and Use of the Data Infrastructure
• **Conclusion 3-23:** There is a need to increase awareness among all stakeholders about new data infrastructure developments funded by the Office of the Secretary PCOR Trust Fund. Increased awareness will enhance the efficiency and effectiveness of research, which in turn will increase the impact of the investments made in infrastructure development.
• **Conclusion 3-24:** Investments in implementing and disseminating infrastructure tools and products will accelerate the achievement of overall PCOR infrastructure goals.
• **Conclusion 3-25**: Dissemination and translation of the research findings could be greatly enhanced by using forms of communication that are relevant to those outside the research community.
Updating the Data Infrastructure
• **Conclusion 3-26:** Explicitly focusing on improved health as the goal of the PCOR infrastructure may be a useful way to prioritize projects and target infrastructure investments.
• **Conclusion 3-27:** A tighter feedback loop with the external end-users and developers of evidence would enhance the value of data infrastructure investments. Examining what evidence was generated due to ASPE interventions and identifying what impact it had on policy and knowledge would also help close the gaps observed in realizing the potential of the PCOR data infrastructure.

Appendix D. OS-PCORTF Strategic Planning Engagement Council

The table below lists the U.S. Department of Health and Human Services (HHS) agencies involved in the Strategic Planning Engagement Council (SPEC) and the specific principal or designee member.

Table 9. OS-PCORTF Federal Partner Organizations and Representatives Involved in the SPEC

Participating HHS Agencies and Offices	SPEC (2020–2022)
Administration for Children and Families	Brett Brown
Administration for Community Living	Jennifer Johnson
Agency for Healthcare Research and Quality	David Meyers Karin Rhodes
Assistant Secretary for Preparedness and Response	Leremy Colf Ian Watson
Centers for Disease Control and Prevention	Tim Carney
Centers for Medicare and Medicaid Services	Andy Shatto
U.S. Food and Drug Administration	Daniel Caños Tina Morrison
Health Resources and Services Administration	Maushami DeSoto Sarah Potter
Indian Health Service	Susy Postal
National Institutes of Health	Wendy Weber
Office of the National Coordinator for Health Information Technology	Steven Posnack Stephanie Garcia
Substance Abuse and Mental Health Services Administration	Christine Cichetti

List of References

[1] U.S. Department of Health & Human Services. *About HHS*. https://www.hhs.gov/about/index.html. Accessed October 2021.

[2] 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

[3] Wolfe, A. (2001). Institute of Medicine report: Crossing the quality chasm: A new health care system for the 21st century. *Policy, Politics, & Nursing Practice*, 2(3). https://pubmed.ncbi.nlm.nih.gov/25057539/

[4] Institute of Medicine. (2011). *Learning what works: Infrastructure required for comparative effectiveness research: Workshop summary.* The National Academies Press. https://doi.org/10.17226/12214

[5] Further Consolidated Appropriates Act, 2020, Publ. L. No. 116-94, 133 Stat. 2534 (2019). https://www.congress.gov/116/plaws/publ94/PLAW-116publ94.pdf

[6] Patient-Centered Outcomes Research Institute. (2021, March 30). *About PCORI*. https://www.pcori.org/about/about-pcori

[7] Agency for Healthcare Research and Quality. (2016, October). *AHRQ's role in the Patient-Centered Outcomes Research (PCOR) cycle*. https://www.ahrq.gov/pcor/dissemination-of-pcor/ahrq-role-in-pcor-cycle.html

[8] Institute of Medicine. (2012). *Ethical and scientific issues in studying the safety of approval drugs.* The National Academies Press. https://doi.org/10.17226/13219

[9] Zozus, M., Hammond, E., Green, B., Kahn, M., Richesson, R., Rusincovitch, S., Simon, G., & Smerek, M. (2014). *Assessing data quality for healthcare systems data used in clinical research.* NIH Collaboratory. https://dcricollab.dcri.duke.edu/sites/NIHKR/KR/Assessing-data-quality_V1%200.pdf

[10] Kahn, M.G., Callahan, T.J., Barnard, J., Bauck, A.E., Brown, J., Davidson, B.N., Estiri, H., Goerg, C., Holve, E., Johnson, S.G., & Liaw, S.T. (2016). A harmonized data quality assessment terminology and framework for the secondary use of electronic health record data. *Egems*, *4*(1). https://doi.org/10.13063/2327-9214.1244

[11] Regenstrief Institute. *PPRL linkage honest broker*. https://www.regenstrief.org/n3c-lhb/. Accessed February 2022.

[12] U.S. Food and Drug Administration. (2018). *Framework for FDA's real-world evidence program.* https://www.fda.gov/media/120060/download

[13] Franklin, J., Platt, R., Dreyer, N., London, A. J., Simon, G. E., Watanabe, J. H., Horberg, M., Hernandez, A., & Califf, R. (2021). When can nonrandomized studies support valid inference regarding effectiveness or safety of new medical treatments? *Clinical Pharmacology & Therapeutics*. https://doi.org/10.1002/cpt.2255

[14] Concato, J., Lawler, E., Lew, R., Gaziano, M., Aslan, M., & Huang, G. (2010). Observational methods in comparative effectiveness research. *The American Journal of Medicine, 123*(12), e16-e23. https://www.sciencedirect.com/science/article/abs/pii/S0002934310008466?via%3Dihub

[15] Tucker, A., Wang, Z., Rotalinti, Y., & Myles, P. (2020). Generating high-fidelity synthetic patient data for assessing machine learning healthcare software. *Npj Digital Medicine, 3*(147). https://doi.org/10.1038/s41746-020-00353-9

[16] Chen, R., Lu, M., Chen, T., Williamson, D., & Mahmood, F. (2021). Synthetic data in machine learning for medicine and healthcare. *Nature Biomedical Engineering, 5*, 493-497. https://doi.org/10.1038/s41551-021-00751-8

[17] Wilkinson, M. D., Dumontier, M., Aalbersberg, I. J., Appleton, G., Axton, M., Baak, A., ... & Mons, B. (2016). The FAIR Guiding Principles for scientific data management and stewardship. *Scientific data*, *3*(1), 1-9. https://doi.org/10.1038/sdata.2016.18

[18] Chen, I., Pierson, E., Rose, S., Joshi, S., Ferryman, K., & Ghassemi, M. (2021). Ethical machine learning in healthcare. *Annual Review of Biomedical Data Science. 4*, 123 – 144. https://doi.org/10.1146/annurev-biodatasci-092820-114757

[19] President Joseph R. Biden Jr. (2021). Executive Order on Advancing Racial Equity and Support for Underserved Communities Through the Federal Government. Executive Office of the President of the United States. https://www.whitehouse.gov/briefing-room/presidential-actions/2021/01/20/executive-order-advancing-racial-equity-and-support-for-underserved-communities-through-the-federal-government/

[20] Healthy People 2030. *Social determinants of health*. Office of Disease Prevention and Health Promotion. https://health.gov/healthypeople/objectives-and-data/social-determinants-health. Accessed February 2022.

[21] National Academies of Sciences, Engineering, and Medicine. (2022). *Building data capacity for patient-centered outcomes research: Consensus Committee Report – Priorities for the Next Decade*. The National Academies Press. https://doi.org/10.17226/26489.

[22] W.K. Kellogg Foundation. (2004). *W.K. Kellogg Foundation: Logic model development guide.* https://www.wkkf.org/resource-directory/resources/2004/01/logic-model-development-guide

[23] Gateway to Health Communication. *Preferred terms for select population groups & communities.* Centers for Disease Control and Prevention. https://www.cdc.gov/healthcommunication/Preferred_Terms.html. Accessed February 2022.

[24] AcademyHealth and NORC at the University of Chicago. (2020). *Challenges and improvements for PCOR data infrastructure: Results from a stakeholder prioritization activity.* Assistant Secretary for Planning and Evaluation. https://aspe.hhs.gov/sites/default/files/migrated_legacy_files//197426/PCOR-Data-Infrastructure.pdf

[25] Lang, L., Smith, S., Trevino, M., Wei, S., Chudnov, D., Potteiger, C., & While, M. (2021). *Research data networks and patient-centered outcomes research trends and opportunities: Scan and interviews with key informants*. Assistant Secretary for Planning and Evaluation. https://aspe.hhs.gov/sites/default/files/2021-07/PCOR%20Research%20Network%20Landscape%20Report_Final_07_14_2021_508.pdf

[26] Banger, A. K., Booth, G. M., Campbell, N. R., Hull, S. C., Potteiger, C. A., Vreeman, D., Lang, L., Smith, S. R., Trevino, M., Wei, S. (2021). U.*S. Department of Health and Human Services stakeholder engagement report executive summary – Analysis and synthesis of key themes*. Assistant Secretary for Planning and Evaluation. https://aspe.hhs.gov/sites/default/files/documents/0015922ac33989b8b655683c003a7256/hhs-stakeholder-engagement-exec-summary.pdf

[27] National Academies of Sciences, Engineering, and Medicine. (2021). *Building data capacity for patient-centered outcomes research: Interim report one – Looking ahead at data needs.* The National Academies Press. https://doi.org/10.17226/26297

[28] National Academies of Sciences, Engineering, and Medicine. (2021). *Building data capacity for patient-centered outcomes research: Interim report two – Data standards, methods, and policy.* The National Academies Press. https://doi.org/10.17226/26298

[29] National Academies of Sciences, Engineering, and Medicine. (2022). *Building data capacity for patient-centered outcomes research: Interim report three – A comprehensive ecosystem for PCOR.* The National Academies Press. https://doi.org/10.17226/26396

[30] Amoozegar, J., Blumenfeld, B., Brown, S., Ortiz, A., Renaud, J., Richardson, J., & West, S. (2017). *Building data capacity for patient-centered outcomes research in HHS: A formative evaluation of 2012-2016 projects*. Division of Healthcare Quality and Outcomes Office of Health Policy/ASPE/HHS. https://aspe.hhs.gov/system/files/pdf/259016/ASPEPCORTFEvaluation.pdf

[31] Carter, N., Bryant-Lukosius, D., DiCenso, A., Blythe, J., & Neville, A. (2014). The use of triangulation in qualitative research. *Oncol Nurs Forum, 41*(5), 545-547. https://doi.org/10.1188/14.ONF.545-547