ASPE | OFFICE OF HEALTH POLICY

OS PCOR TF | OFFICE OF THE SECRETARY PATIENT-CENTERED OUTCOMES RESEARCH TRUST FUND

PROJECT REPORT

Research Data Networks and Patient-Centered Outcomes Research Trends and Opportunities: Scan and Interviews with Key Informants

An Input to the OS-PCORTF Strategic Plan

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

July 2021

Department of Health and Human Services (HHS)

Office of the Assistant Secretary for Planning and Evaluation

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

The Office of Health Policy

The Office of Health Policy (HP) provides a cross-cutting policy perspective that bridges Departmental programs, public and private sector activities, and the research community 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 about this subject, you can visit the HP home page at https://aspe.hhs.gov/office-health-policy-hp or contact the ASPE Project Officers, at:

HHS/ASPE/HP
Room 4447D, H.H. Humphrey Building
200 Independence Avenue, S.W.
Washington, D.C., 20201
Scott.Smith@hhs.gov
Marcos.Trevino@hhs.gov

Authors

Lisa Lang, MPP
Scott R. Smith, Ph.D.
Marcos Trevino, MPH
Sara Wei, MHA
Division of Health Care Quality and Outcomes, Office of Health Policy, Assistant Secretary for Planning and Evaluation, U.S. Department of Health and Human Services; Washington, D.C.

Daniel Chudnov, MLIS, MSBA
Caroline Potteiger
Mariella While, RN, MEd
The MITRE Corporation; McLean, VA

Acknowledgements

We extend our sincere appreciation to all the participants who volunteered their time to participate in the research network interviews that inform this report.

This project was funded under contract number 75FCMC18D0047/75P00120F80242 from the Assistant Secretary for Planning and Evaluation (ASPE), U.S. Department of Health and Human Services. The authors are solely responsible for this document's contents, findings, and conclusions, which do not necessarily represent the views of ASPE. Readers should not interpret any statement in this product as an official position of ASPE or of the U.S. Department of Health and Human Services.

Executive Summary

In December 2019, Congress reauthorized the Patient-Centered Outcomes Research Trust Fund (PCOR Trust Fund) through 2029.[1] Among its other aspects, the reauthorization renews the charge to the HHS Secretary to build data capacity for PCOR, using funds specifically identified for the Office of the Secretary for this purpose (OS-PCORTF). The Secretary delegated authority to the Office of the Assistant Secretary for Planning and Evaluation (ASPE) to accomplish this work, which is carried out in partnership with all HHS agencies and leaders under the auspices of the OS-PCOR Trust Fund portfolio. To guide the OS-PCOR Trust Fund portfolio's efforts to build data capacity over the next 10 years, ASPE is developing a new strategic plan with input from HHS agency experts, leaders, and external stakeholders.

This report focuses on research networks that engage in or support research on patient-centered outcomes and are thus important end users and stakeholders for the OS-PCORTF's work. This report describes an environmental scan of 15 research networks that conduct or are capable of supporting PCOR; the report also reflects semi-structured discussions with principal investigators or leads from eight of these networks. It was conducted in support of development of the strategic plan for the OS-PCORTF. Given the broad landscape of research networks, this report does not endeavor to present a comprehensive picture of all research networks engaged in PCOR, but rather to provide informed insights of potential value to the overall strategic planning process.

Based on our structured discussions, the research networks identified common challenges in accessing and using data for PCOR:

- Lack of high-quality real-world data
- Limited tools and resources for linking data between different sources (e.g., EHRs, claims, PROs)
- Lack of tools to improve data quality and curation
- Constant effort to maintain data quality
- Difficulty of accessing medical claims data, especially Medicaid data

Even the most mature research networks expressed these common challenges. To overcome these challenges, networks have developed a variety of network-specific solutions such as using new tools to collate patient data, applying common data models (CDMs), and monitoring data quality on a constant basis. The networks examined also use HHS data sources, standards, and management requirements in varying degrees in their PCOR work.

The research networks also identified future priorities that pose opportunities for OS-PCORTF, including interest and involvement in enhancing research access to federal health data, expanded to include device and patient-provided information, among other novel sources; strengthening methods and tools to promote and sustain authoritative health data linkage; developing and implementing standard approaches for data quality, consistency, and patient identification; addressing source data workflow strategies for data capture to improve data quality; and addressing 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.

Potential Opportunities for OS-PCORTF to Support PCOR Research as Expressed by the Research Networks

- Enhance research access to federal health data, expanded to include device and patient-provided information, among other novel sources.
- Strengthen methods and tools to promote and sustain authoritative health data linkage.
- Develop and implement standard approaches for data quality, consistency, and patient identification.
- Address source data workflow strategies for data capture to improve data quality.
- 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.

Appendix A includes profiles of all 15 networks in the study.

Table of Contents

List of Figures

List of Tables

1 Background

The 2010 Patient Protection and Affordable Care Act (ACA) is landmark legislation that emphasized building a robust national research program to empower patients and their providers to make more informed healthcare decisions.[1] Comparative clinical effectiveness research—now more routinely referred to as patient-centered outcomes research (PCOR)—has the following features:[2]

- "Assesses the benefits and harms of preventive, diagnostic, therapeutic, palliative, or health delivery system interventions to inform decision making, highlighting comparisons and outcomes that matter to people;
- Is inclusive of an individual's preferences, autonomy, and needs, focusing on outcomes that people notice and care about such as survival, function, symptoms, and health-related quality of life;
- Incorporates a wide variety of settings and diversity of participants to address individual differences and barriers to implementation and dissemination; and
- Investigates (or may investigate) optimizing outcomes while addressing burden to individuals, availability of services, technology, and personnel, and other stakeholder perspectives."

The ACA established the Patient-Centered Outcomes Research Trust Fund (PCORTF) and allocated funding through 2019 for three components: conducting research, disseminating its results, and developing a data research infrastructure for PCOR. These components are the responsibility of, respectively, the Patient-Centered Outcomes Research Institute (PCORI), the Agency for Healthcare Research and Quality (AHRQ), and the HHS Office of the Secretary (OS-HHS).[1] The PCORTF was recently reauthorized for an additional 10 years.[3]

The work described in this report focuses on the third component—developing a research data infrastructure for PCOR. The ACA charges the OS-HHS with coordinating federal programs to "develop and maintain a comprehensive, interoperable data network to collect, link, and analyze data on outcomes and effectiveness."[1] The Assistant Secretary for Planning and Evaluation (ASPE) is responsible for coordinating the development of this research data infrastructure. ASPE administers the Office of the Secretary's share of the Trust Fund (OS-PCORTF) to fund a portfolio of multi-year projects that are usually developed and executed by one or more HHS agencies or offices.[1] Broadly speaking, the OS-PCORTF supports the goal of building HHS data capacity for conducting PCOR. The portfolio's projects have focused on creating data sources and tools, and developing components essential to interoperable data use for research (e.g., standards, services, policies, and governance structures).[4]

The *HHS Strategic Plan for Building Data Capacity for Patient-Centered Research 2017-2021* has helped guide selection of projects funded by the OS-PCORTF.[5] Over the first 10 years of the OS-PCORTF, multiple reports have summarized and evaluated the portfolio, describing its progress and identifying future needs.[202] Newly reauthorized for the next decade, ASPE now seeks to update its OS-PCORTF strategic plan to guide decisions about priorities for advancing data infrastructure in support of PCOR while adapting and responding effectively to the changing landscape and emerging health outcomes challenges. To support strategic planning and portfolio development needs, research, stakeholder engagement, and analysis are underway.

Toward this end, this report is one element of a broader effort to provide ASPE with an up-to-date and comprehensive understanding of the state of data capacity for research on patient-centered outcomes, as well as the needs and strategic priorities of OS-PCORTF stakeholders. The Strategic Plan will be supported by a number of external and internal examinations of strengths, challenges, and opportunities to work within HHS and with other partners to promote more robust data infrastructure for important PCOR, including:

- Challenges and Improvements for PCOR Data Infrastructure: Results from a Stakeholder Prioritization Activity — workshop[6]
- Building Data Capacity for Patient-Centered Outcomes Research: An Agenda for 2021 to 2030 — National Academy of Sciences study group and public workshops (currently in progress)[7]
- Review of Internal HHS Data Capabilities and Priorities for PCOR (in progress)
- Key Informant Interviews with PCOR Research Networks (this document)

For the purposes of this paper, "research network" is a broad term that includes distributed research networks, clinical research data networks, and clinical research registries.[8] The PCORI Methodology Standards Report, the foundation for PCORI's work, established requirements for the design and features of the data and data networks for the PCOR it funds; the standards address the importance of data quality, data integration, use of standard terminologies and common data models, and appropriate privacy, confidentiality, governance, and intellectual property (IP) protections—concerns that continue today.[9]

Many important collaborative research data networks for PCOR began well prior to enactment of the ACA, such as the Healthcare Systems Research Network (HCSRN).[10] The past decade has seen expansion in the number, scope, lives covered, and data sources incorporated in such research networks. Previous reticence to view claims or routine clinical data as highly relevant to clinical trials has evolved in favor of expanded use of real-world data for pragmatic and even randomized trials.[11]

Data collected or affected by HHS and other federal programs and policies have central importance to the work of research networks, particularly for PCOR. Each such data source is provided under the rules defined by its particular legislative or regulatory authorities, and the research and policy communities experience the lack of integration of these data sources.[12] The renewed charge to the OS-PCORTF offers a unique collaborative opportunity to identify and address data problems for the PCOR community that arise at the intersection of federal datasets with each other and with other key data resources for PCOR.

1.1 Purpose

This report focuses on research networks that are key end users of, and sometime participants in, the projects and products of the OS-PCORTF. The researchers in these networks conduct, or are capable of conducting, PCOR among other research, and can help identify pressing data needs, challenges, and priorities for the near and longer term. This study, in conjunction with input from other stakeholders as described above, will inform development of an evidence-based strategy to guide the long-term direction of the OS-PCORTF portfolio. A well-informed and innovative strategy will position HHS to achieve a robust national data infrastructure for PCOR that ultimately supports the vision of empowering patients and providers making informed healthcare decisions together based on evidence from rigorously conducted research.

2 Approach

To identify key findings and opportunities for the OS-PCORTF portfolio from research networks, this report followed four main steps: (1) identify PCOR networks to scan and engage, (2) conduct an environmental scan, (3) hold key informant discussions with a subset of the research networks, and (4) develop research network profiles.

Because research networks are key end users of the OS-PCORTF portfolio's efforts to build data capacity for PCOR, this report focuses on research networks that lead and participate in research on PCOR. As described in detail in this section, profiles of several prominent research networks were developed (Appendix A) based on publicly available materials and discussions with principal investigators from several research networks (and synthesis of their diverse perspectives).

2.1 Methods

Identification of Research Networks

To assess and analyze the PCOR research network landscape, the environmental scan identified a convenience sample of 15 research networks, including five networks with past or current participation in OS-PCORTF projects. The list was reviewed with representatives of various HHS agencies involved with the OS-PCORTF strategic planning efforts.

Semi-structured discussions with principal investigators for eight networks, including three with past or current participation in OS-PCORTF projects, were held. The following factors guided the selection of networks for the key informant discussions: networks whose goal is to generate evidence to inform patient or clinician decision making, groups actively engaged in or supporting PCOR, a variety of sponsoring organizations, and diversity in observational versus clinical trial study design. If a rigorous study were to be performed, an analysis of data approach and scope would ensure a diverse sampling.

Table 1 lists the research networks in the environmental scan and shows their prior or current involvement in OS-PCORTF projects and whether they were part of the key informant subset.

Table 1. PCOR Research Networks in Study

Research Network
1. Accelerating Data Value Across a National Community Health Center Network (ADVANCE)*
2. AHRQ Practice-Based Research Networks (PBRN)*
3. American Society of Clinical Oncology (ASCO) CancerLinQ
4. Electronic Medical Record Support for Public Health (ESPHealth)
5. Food and Drug Administration (FDA) Centers of Excellence in Regulatory Science (CERSI)*
6. FDA Medical Device Epidemiology Network (MDEpiNet)*+
7. FDA Sentinel*+
8. Healthcare System Research Network (HCSRN)
9. National Evaluation System for health Technology Coordinating Center (NESTcc)
10. National Institutes of Health (NIH) All of US*+
11. NIH Clinical and Translational Science Awards (CTSA) Program
12. NIH Collaboratory*
13. Observational Health Data Sciences and Informatics (OHDSI)+
14. Patient-Centered Outcomes Research Network (PCORnet)*+
15. PEDSnet
*Participated in Key Informant Discussion
+Participated in prior or current OS-PCORTF Project

Environmental Scan

An initial review of publicly available information concerning the 15 research networks or registry networks was undertaken using a common protocol. The common protocol was developed to ensure that a minimum set of common information could be extracted across all the research networks of interest. This protocol was divided into six main categories:

- Purpose (functionality and/or area of PCOR focus)
- Composition (e.g., who/what comprises the network, example research projects)
- Governance structure (overall for the network)
- Data sources/elements used
- Network outputs (e.g., products and services)
- Practical impact (e.g., on clinical and/or regulatory decision making or guidelines, evidence-based treatments, patients, and health outcomes/metrics)

Based on this protocol, a template (Appendix B) was designed and used to capture information about the networks during the environmental scan. Materials included in this scan included publicly available information about the networks, mostly from online material.

Discussions with Research Networks

Semi-structured discussions with principal investigators from eight research networks involved 14 people in individual or group discussions. A facilitation guide included questions on the networks'

biggest data-related challenges and plans to address them, current trends and gaps the networks are tracking, current engagement with HHS, and improvements HHS could make to more effectively support PCOR. To ensure accurate reporting, recordings of the discussions supplemented the live notes (a policy is in place for secure storage and destruction of recordings). Appendix C presents the discussion questions. Responses were reviewed, evaluated, and synthesized by the interviewing team. Although all discussion participants knew who was sponsoring the scan and many of the discussion participants were representatives from networks with prior or current involvement in OS-PCORTF projects, the participants did not note any unique perspectives because of this involvement and did not mention any OS-PCORTF products or projects.

Creation of Research Network Profiles

Appendix A. provides the research network profiles developed during the environmental scan and the network discussions.

2.2 Limitations

The principal limitations of the work described in this report:

- The PCOR research network environment is large. Only a convenience sample of PCOR-relevant networks was scanned; the subset of those selected for discussions was even smaller. Therefore, the findings in this report are not intended to be exhaustive.
- The findings and challenges highlighted in this report are limited to what could be discerned in publicly available sources, which may not have been fully up to date.
- The comments attributed to the networks reflect the observations made during the interviews. While wide-ranging, the discussions were constrained by format and time and so may not have provided a complete representation of the sources' assessment of challenges and lessons learned on all matters of research within their networks.

Notwithstanding these limitations, the networks that participated in our discussions are recognized leaders in the field, and we appreciate their participation. Their knowledgeable and expert comments and insights identified common themes and opportunities relevant to both present and future data challenges for PCOR.

3 Findings

The findings in the ensuing subsections illuminate key observations from the research network scan. The findings are organized into seven categories: composition of networks, governance structure, sources of data, primary functionalities and objectives, network outputs, impact, and challenges. Common themes and individual findings are discussed and analyzed for each category.

3.1 Composition of Networks

The environmental scan of the research networks' composition focused on the operators or sponsors of the networks, the populations, and domains they cover, their approaches to data access (centralized or distributed), and the collaboration models they use. The scan information describes key characteristics of the research network landscape and provides context for subsequent findings.

Sponsorship or operation of the research networks is diverse. Some networks are operated or funded by federal entities, such as the FDA, while others are operated by private organizations or professional societies, such as ASCO. For example, NIH and FDA respectively run three of the networks examined, while PCORI supports three of the networks. Due to the wide-ranging purposes and breadth of each of these networks, many different data types and population groups can be studied. Some networks have a broad scope, such as PCORnet and NIH All of Us, while others, such as ADVANCE and PEDSnet, focus on a specific subset of the population, such as low-income or pediatric populations, respectively. This also reflects their mix of research needs and goals in the different levels of granularity of specific types of data that are collected for their research studies. For example, ADVANCE and PEDSnet collect data using expansions of the PCORnet and OMOP common data models (CDMs), respectively, to include data elements specific to their use cases.[13,14] This affords their researchers access to a wider variety of relevant data. Section 3.4.2 discusses research network use of CDMs in more detail.

The research networks have grown to cover large populations. For example, NIH Collaboratory projects span more than 1,000 clinical sites across 90 percent of the United States, and PCORnet data cover more than 70 million people nationwide.[15,16] Similarly, OHDSI, with hundreds of researchers from 30 countries, has access to health records for about 600 million unique patients from across the world.[17] The networks' successes and impact are not limited to the United States but extend to the rest of the world. FDA MDEpiNet, for example, consists of more than 120 national and regional registries from 45 countries, and is continuing to expand globally.[18] By increasing sample size through this expansion, FDA MDEpiNet can study a more diverse population.

The research networks address many domains, ranging from general observational research to specific topics, such as cancer and medical device use. The following examples of research network focus areas demonstrate the breadth of populations of interest, scopes of focus, and research approaches; most networks address more than one of these domains:

- Infrastructure for collaborative research (e.g., NIH Collaboratory and OHDSI)
- Comparative effectiveness research (e.g., PCORnet and OHDSI)
- Safety assessment and monitoring (e.g., FDA CERSI and FDA Sentinel)
- Precision medicine (e.g., NIH All of Us)
- Observational research (e.g., OHDSI)
- Public Health Department monitoring (e.g., ESPHealth)
- Pediatric research (e.g., PEDSnet)
- Translational research (e.g., CTSA)
- Medical device development and evaluation (e.g., FDA MDEpiNet and NESTcc)
- Population-based research (e.g., HCSRN)
- Social determinants of health (e.g., ADVANCE)
- Oncology (e.g., CancerLinQ)
- Primary care (e.g., PBRN)

In addition to the variety of domains, the networks offer different means of access to data. The two main approaches are centralized and distributed data access. In a *centralized* approach, the data

collected from partners are sent to a centralized data warehouse operated by the network. Research network users can then query this centralized database for information. In our convenience sample, NIH All of Us, CancerLinQ, and ESPHealth are examples of a centralized approach where patient records are collected from providers, then stored and queried centrally.

A *distributed* approach means the data stay within network partner firewalls, and summary or individual-level data are only sent to others when queried, with more stringent requirements for sharing individual-level data. PCORnet uses a distributed data approach. The data from PCORnet partners are kept within the "walls" of the institution and typically shared with other organizations as de-identified, aggregate statistics when responding to queries made via a secure Distributed Research Network Query Portal. As a result, PCORnet only shares minimal information to answer any research question.[19] FDA Sentinel and OHDSI are also examples of distributed research networks.

Both methods have benefits and limitations. With the centralized approach, more detailed data can be studied and accessed; however, this approach increases the risk of breach of data privacy and security. In contrast, although the distributed approach is more secure, the types of data available to researchers can be limited. As a result, when network operators determine their data access approach, they must consider the different needs of the sources and users of their featured data and the scope of the research they can support.

Many of the networks examined in this scan collaborate with one another. For example, as shown in Figure 1, researchers in the NIH Collaboratory Distributed Research Network (DRN) use FDA Sentinel's System to query the Sentinel Database. This collaboration is useful in multiple types of research, including observational studies, prospective data collection, and randomized clinical trials.[20]

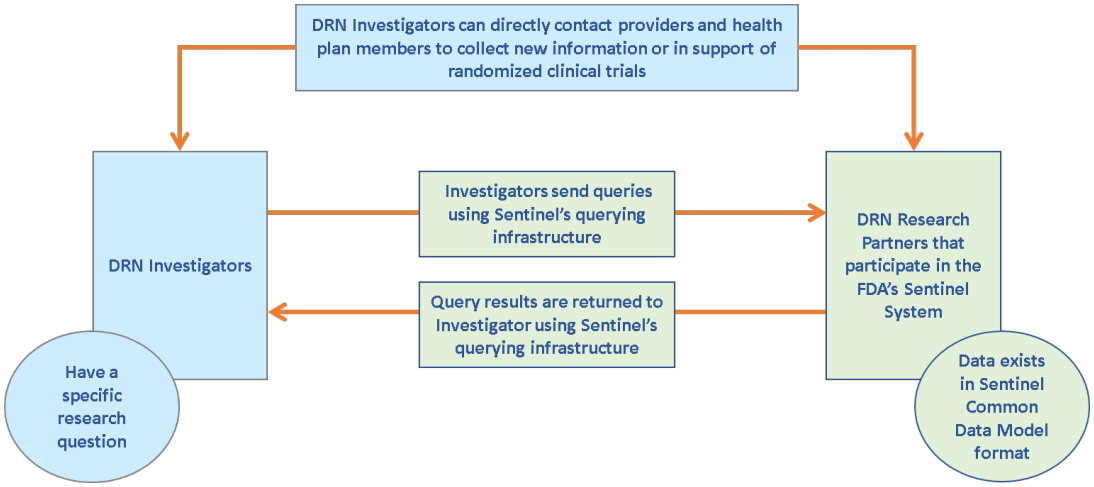

Figure 1. Collaboration Between NIH Collaboratory and FDA Sentinel

Through this partnership, NIH Collaboratory investigators can contact health system members within FDA Sentinel to gather new information for randomized clinical trials.[20] Similarly, FDA Sentinel has many "Collaborating Institutions," such as the Healthcare System Research Network and PCORnet. These Collaborating Institutions help the FDA answer questions and provide healthcare data and scientific, technical, and organizational expertise.[21] Figure 1 illustrates the network collaboration between NIH Collaboratory and its use of the FDA Sentinel System for research.

3.2 Governance Structure

The governance structure of a research network has significant potential to enable or inhibit the flow of information and affect the network's capability to conduct PCOR. In this context, governance can refer to running the individual networks as well as the governance structures that apply to their partners and the data sources they use. Governance involving data includes legal or contractual requirements as well as local policies that control who can contribute, access, or use the data. Access to some types of data may require an individual Institutional Review Board (IRB) approval for each separate use. Several discussion participants noted the potential for Streamlined, Multisite, Accelerated Resources for Trials IRB (SMART IRB) as a step forward.[22] Some individual-level data in the networks, however, cannot be linked because of restrictive terms of use, especially when data are purchased from an outside entity. Similarly, legal or policy restrictions may limit the networks' capability to link social service information with clinical records, making it difficult to track follow-up actions across services.

Governance of large-scale research projects also requires a common and consistent approach to privacy and security, which pose various challenges. Individuals who consent to participate in a research study may be protected by limitations on their consent, which could restrict access to their data that could be of value for many other studies. These concerns can persist even with a move to e-consent approaches. Discussion participants noted that having higher thresholds for data security, limiting data access, and performing analyses to assess the potential for re-identification can significantly lower the risk associated with large datasets while protecting the privacy of individuals.

Research Program Highlight

The NIH All of Us Research Program gathers a large amount of data about individuals centrally, which makes data security and access issues much more significant. To mitigate these issues, the centralized All of Us data approach allows for a consistent approach to access control, where individual researchers can access patient-level data only if they are sponsored by a participating institution that has joined the program and signed a Data Use and Registration Agreement. Other researchers can only access summary-level data from the program, with the option to work with their local institution to sign the agreement.

In distributed research networks using a common data model and query model, only authorized individuals at each data provider organization have access to patient-level information, which typically resides at each institution. Researchers can query the data to answer a specific research question, and each institution returns de-identified results, often at a summary level. This works well for answering targeted questions but limits the types of questions that can be asked. One participant expressed optimism for innovations in technical/statistical methods—known as distributed analysis algorithms—that suggest better privacy protection while enabling patient-level distributed queries with a potential for lowering consent burdens. In all cases, the process of gaining consensus for a data use agreement for new network participants can be time consuming, but this remains a critical step in sharing data for research. One respondent noted that, "often, governance and data use discussions settle around the most conservative common denominator because there's no contractual ability to push them farther than where consensus takes them." At the same time, experience in building these relationships has led to improved language in data use agreements in later studies.

3.3 Sources of Data

Overall, network discussion participants noted success in gaining access to clinical and other sources of data for use in network research studies. Many respondents shared a concern, however, that using data from EHR systems, claims, and most other health-related sources for secondary purposes requires significant continuing attention to data quality. A curation process is necessary to identify disparate practices for recording information across many data entry fields within a single health record (i.e., describing similar encounters in many ways). Even with concerted efforts to organize data using shared data structures such as common data models, network participants must understand whether data linked across sources can be used to represent a true finding and to apply such findings to improve care. For example, different providers within a practice (or different clinical practices) may use common terms in different ways or may record common encounters using different terms. This presents problems when scaling queries to larger datasets collected from many providers, requiring a substantial process to map terms, verify meaning, and check validity. Data that pass validation may later be found erroneous in unexpected ways, even in environments with long-established data workflows. Discussion participants shared that they have substantial experience working to sustain data quality and noted that this work never ends.

Several networks reported success working with claims data obtained through the Centers for Medicare & Medicaid Services (CMS) Virtual Research Data Center (VRDC) and noted its value in obtaining Medicare claims data specifically. Network representatives also reported success working with commercial data products offering de-identified EHR or claims data, although they recognize that different data products are assembled in different ways and may differ in dimensions such as completeness or representativeness.

MDEpiNet representatives described work to expand the scope of Coordinated Registry Networks (CRNs) to include international partners. In addition to many potential benefits of international cooperation, the broader potential for participation in studies of any specific medical device may help to expand sample sizes, potentially increasing the statistical power of results. This can be especially important for studies of high-risk interventions that might otherwise suffer from small sample sizes.

Respondents noted the potential for building large datasets efficiently and inexpensively through participant-provided information (PPI) gathered through consumer-grade products such as smart watches and activity monitors. Large consumer technology companies have started to deploy tools for medical research, including Apple ResearchKit, a software framework for building mobile applications supporting consent, surveys, and data flows, and the recently announced Google Health Studies, a dedicated application that also supports gathering survey answers.[23,24] The broad accessibility of tools like these for the general population seems to offer infrastructure for efficient enrollment and standards-based data sharing. Early studies using these tools involve partnerships with individual care providers and research institutions.

3.4 Research Networks' Purpose and the OS-PCORTF Functionalities

The objectives of the networks highlighted in this scan were both broad and patient centric. The networks aim to improve health outcomes, care delivery, and treatments; to enable device monitoring

to measure and improve effectiveness and safety; and to support collaboration on evidence-based research in all these areas.

In considering the capabilities of a data infrastructure sufficiently robust to support such PCOR objectives, it is useful to consider the core functionalities identified by the OS-PCORTF program (see, for example, *Building Data Capacity for Patient-Centered Outcomes Research in HHS: A Formative Evaluation of 2012-2016 Projects report* .[203]) A robust data infrastructure for PCOR supports:

- Use of Clinical Data for Research
- Standardized Collection of Standardized Clinical Data
- Linking Clinical and Other Data for Research
- Collection of Participant-Provided Information
- Use of Enhanced Publicly Funded Data

These functionalities provided a strong paradigm for reviewing the observations and challenges that arose from the research network discussions and scan.[203] Several themes, discussed in the following subsections, emerged from the environmental scan and discussions with research network representatives along and across these functionalities.

3.4.1 Standardized Collection of Standardized Clinical Data

Each of the research networks in the environmental scan has developed standardized approaches to exchanging de-identified clinical data using standardized common data models, including the FDA Sentinel CDM, the related PCORnet CDM, and the OHDSI Observational Medical Outcomes Partnership (OMOP) CDM.[19,25,26] In many cases, staff at the data-providing organization perform the work of preparing data for exchange through these standards. According to several networks, it is often challenging for a new data provider to begin sharing data for the first time because that organization must develop a team that can manage the stages of data curation necessary to share standardized data for research. After a data provider has acquired this expertise, joining additional networks can be much easier. For instance, due to the wide variety of EHR systems, system access models, and licensing terms for PBRNs, many practices without this in-house capacity work with a third-party service. This third party's expertise with many data standards and clinical systems enables them to manage data curation on behalf of the practice, sometimes through additional agreements with vendors.

Regarding the common data models themselves, although the prominent CDM specifications are similar in scope, they do vary and are not directly compatible. Translators have been shown to work well, however, and several discussion participants noted that most organizations already using one of these standards should find it easy to support another. Additionally, work on harmonization strategies such as the Health Level 7® (HL7®) Common Data Models Harmonization Fast Healthcare Interoperability Resources® (FHIR®) Implementation Guide—developed under the leadership of the FDA, NIH, ONC, and ASPE—provides standards-based mappings and implementation guidance for CDM translation through a common specification.[27] Additionally, FDA Sentinel included an initiative in its 2019–2023 Strategic Plan to enhance interoperability by harmonizing its CDM with others, such as OMOP and the PCORnet CDM.[27,28]

Certain data standards gaps were noted by multiple networks. For example, standardized unique device identifiers (UDIs), which are overseen by the FDA and included in the United States Core Data for Interoperability (USCDI), show great promise for device monitoring, safety, and effectiveness research. Discussion participants reported that nascent clinical systems support for UDI capture at the point of care is developing. Although the trend is growing, it "can't get here fast enough." Observers recognized underlying policy as well as technical challenges. The decision by the FDA in July 2020 to suspend enforcement of UDI labeling requirement until September 2022, however, may suggest that the potential for integrating UDI in support of PCOR may not be realized without additional policy measures to require capture of UDIs on claims forms.[29]

3.4.2 Use of Clinical Data for Research

Using clinical data for research is a primary and continuous activity of all networks. Expanding networks bring new data providers into shared efforts and more experienced participants to support an array of data workflows, including de-identification, quality review, remediation, transformations to common data models and, in several cases, sharing of data sources through application programming interfaces (APIs) based on open standards.

Several discussion participants described productive use of claims data acquired from CMS, commercial data providers, or state or local municipalities, whether as a standalone data source or linked with data extracted from EHR systems. When working with claims data, cohorts and treatments are often selected for study based on patterns in standardized vocabulary codes for diagnoses, procedures, labs, and prescriptions. A similar strategy is often used with EHR data, especially when clinical and claims data are integrated to improve the completeness of patient histories and expand research potential. One discussion participant expressed optimism that clinical data research may grow to incorporate use of additional symptoms often captured outside of diagnosis and procedure codes, including assessments of Activities of Daily Living (ADL), to obtain a more holistic view of health. Respondents also shared an interest in capturing patient social determinants of health (SDOH) data for research use, as described in Section 4.2.

3.4.3 Linking Clinical and Other Data for Research

Many of the networks conduct substantial research using a combination of clinical data extracted from EHR systems and claims data acquired from federal, commercial, and regional sources. Linkages between multiple EHR sources and between EHR and claims data are common goals, although discussion participants expressed wariness about related pitfalls. For example, linked sources present what may seem to be a more complete picture of the longitudinal record for individuals. The possibility of missing information, however, can limit confidence in having a "complete picture." Information could be missing for various reasons, such as a lack of availability of data from some care providers or payers, non-alignment of sources (e.g., covering overlapping time periods), or an unquantified bias in data at hand, such as inadequate representation of complete patient populations.

Some discussion participants reported that they often use and reuse the same datasets or products of EHR or claims data because of their institutional knowledge, familiarity, and access to these sources. This approach provides value and efficiency as compared to the potential costs of acquiring and integrating new data products or resources. It was noted that a tendency to reuse the same data

sources repeatedly rather than integrating new ones could lead to bias and representation issues in some cases.

Several networks shared an interest in working with vital statistics such as death data. Death data represent severe and objective patient outcomes and can help the networks understand the trajectory of patient care. Death data also indicate when networks should stop attempting to look for and link additional records per patient.

3.4.4 Collection of Participant-Provided Information

The networks are engaged in several initiatives to capture and use patient-provided information (PPI). Within the broader community, for example, both the PCORnet and OMOP CDMs include standardized mechanisms for capturing patient-reported outcome (PRO) information from questionnaires and surveys.[19,26] There also is interest in capturing information on individual patient-level SDOH data. Although some discussion participants report progress on capturing components of surveys such as the National Association of Community Health Centers (NACHC) PRAPARE tool within their organizations, they have found enough survey use variation among their own clinics to suggest that network research using this data from multiple organizations may not yet be feasible (as described in more detail in Section 4.2).[30]

One discussion participant reported a twist on this functionality that may be considered an example of participant-collected (i.e., patient-reported) information. Using the innovative Hugo platform, as described below in Section 4.3, new studies are being conducted using patient data collected through a patient-centric workflow, specifically through a process in which the patients themselves can compile and view all their health data and share this data with providers or researchers.[60]

Emerging Patient Data Sources

There is interest in the potential for collecting information from consumer-grade products such as activity and fitness monitors (i.e., wearables), with an eye toward the growing segment of the population using these devices and recognition that it may be possible to collect meaningful data for new studies inexpensively. One discussion participant highlighted the potential for PRO in direct-to-participant trials as a way to gather data more efficiently and with lighter infrastructure requirements. Reliable and consistent workflows for extracting and transmitting data from these devices are nascent, however, and these data are not yet widely available for use in research networks.

There were also some issues that cut across the *Collection of Participant-Provided Information* and other functionalities. For example, the use of information gathered from consumer products such as personal activity monitors is limited by a lack of adoption of standards and standard collection workflows. These standards require development and incentives for adoption at the intersection of the mass-produced consumer-centric product market and the commercial environment for healthcare domain-specific systems and interfaces.

3.4.5 Use of Enhanced Publicly Funded Data Systems for Research

Several respondents noted the use of publicly funded data systems in their network research. Medicare claims data, made available through the CMS Virtual Research Data Center, are commonly used by organizations that invest time into understanding the system and building up authoritative datasets to use in research. Respondents also reported interest in both Medicaid data and state-level claims, which can be more challenging to obtain. Census data from the American Community Survey (ACS) are used in several settings to provide similarly authoritative population-level indicators of social determinants of health, as discussed further in Section 4.2. One respondent highlighted the potential value of data sources brought together in the openFDA service.[31] An additional respondent indicated a plan for integration of data from the National Death Index in the coming year, while another respondent noted use of state-level birth and death vitals, which presented a challenge in obtaining and gaining state Institutional Review Board (IRB) approval for each research protocol.[32] Even with state IRB approval, these datasets may be required to be stored separately, rather than integrated within a CDM-based data warehouse.

3.5 Network Outputs and Dissemination

A key part of the success of research networks is the dissemination of their research findings. It is important for networks to highlight how the data and services they provide are used by the research community and how the networks and their research make an impact. As a result, most of the networks clearly advertise and provide weblinks to publications, project reports, blog posts, videos, and the like. Some research networks are also proactive in sharing their research. For example, PBRN has a listserv anyone can join. Listserv members receive bi-weekly digests and other announcements, highlighting newly published research, webinars, funding opportunities, and other key information.[33] Other common artifacts include conference presentations on research.

Research Network Outputs

- Publications
- Project reports
- Listserv digests and announcements
- Webinars
- Conference presentations
- Training and educational material
- Data quality and analysis tools
- Governance policies
- Documentation on CDMs
- Capability maturity assessment tools
- Brochures for patients and providers
- Support services for researchers

Many research networks have policies that require researchers to state how they will disseminate their study results. For example, when requesting access to CancerLinQ data, researchers must submit an intention to publish the results of their research.[34] Similarly, many networks post guidelines for research publications and require investigators to submit draft publications to a reviewing committee before they are submitted for publication. This is to ensure the investigators properly used the networks' data and

appropriately acknowledged the network. NIH Collaboratory provides detailed guidance for dissemination of research, including distribution approaches for different stakeholders and different frameworks for dissemination.[35] It is critical to circulate research results to reduce the amount of duplication and allow researchers to build on each other's work. As a result, many networks require dissemination of results as part of the research process.

Other common artifacts produced by research networks include training and educational material, data quality and analysis tools, governance policies, and documentation on CDMs. For example, similar to providing guidance on the dissemination of research, NIH Collaboratory provides lessons and tools on many topics relevant to pragmatic clinical trials, including acquiring real-world data, participant recruitment, and patient-reported outcomes.[36] For each of these topics, NIH Collaboratory provides background information, examples of methods, tools to use for specific use cases, characteristics to consider, and more. Within the patient-reported outcome topic, NIH Collaboratory lists out nine different outcome sets and in which domain they should be used.[37] NIH Collaboratory investigators discuss and reach consensus on standard approaches and best practices in the design, conduct, and reporting of pragmatic clinical trials and share it with a broad spectrum of users.[38] Additionally, providing data to the research networks and accessing data the networks possess often requires the use of tools and capabilities that can be unfamiliar to many researchers. For example, the Observational Health Data Sciences and Informatics (OHDSI) project offers software tools that support several steps of the research life cycle from data curation through analysis.[39] Each tool works with the OMOP CDM, guiding researchers through extracting their data, mapping codes, and assessing overall data quality, as well as providing support for various research methods and analysis tools.[30]

In addition, because most of the networks partner with health sites and use data from patients treated at these health sites, the networks provide documentation and brochures for sharing with patients and providers. These brochures communicate why the data are collected and are important. For example, CancerLinQ has patient brochures that describe what CancerLinQ is, why CancerLinQ collects the patient data, provide answers to FAQs, and supply contact information if patients have more questions.[40] This eases the burden for health sites, providers, and researchers to participate in research networks, because the networks already have communications prepared for the patients. As a result, the researchers do not have to create the material from scratch. 4.3presents additional discussion of the importance of patient engagement.

In addition to providing artifacts for the research network users, many research networks provide services to help investigators conduct their research. Some of these services include onboarding/setup, quality management, certification programs, surveillance, project and data management expertise, and training and career development programs. Due to the complexity of data management tools, onboarding services are often necessary to orient teams. Many networks also have training and career development programs to connect and support researchers. For example, PEDSnet has the PEDSnet Scholars Program, with the goal to support the training of clinicians and research scientists to conduct PCOR within learning health systems.[41] While many of these research networks provide access to data, it is important for scientists to understand best practices when conducting research, which is what these programs aim to do. For example, NIH Collaboratory conducts weekly public web seminars at which studies are presented and discussed as a means to disseminate research methods and tools to advance the use of pragmatic trials that rely on real-world data; the Collaboratory routinely has upward of 100–

200 people on a call, an effective approach for getting information about tools out to a community that might benefit from them.[42]

3.6 Research Networks' Impact in the PCOR Community

The networks offer many tools and services that aid investigators in conducting PCOR. It is important to understand how this research impacts the PCOR community. Several networks have developed metrics to monitor how their research programs and projects have made an impact in the PCOR community. For example, FDA CERSI uses a Research Impact Metrics model to assess CERSI research project impact. These assessments consider advancement of regulatory science, dissemination of scientific knowledge, catalyzing action, and informing regulatory decision making as factors of increasing scope/impact in advancing public health.[43]

Types of Research Network Impact

- Advancement of science
- Dissemination of scientific knowledge
- Catalyzing action
- Informing decision making
- Implementation of network
- Reduced clinician burden
- Faster notifications to patients/providers
- Expansion of network systems/tools/programs

Due to the variety of purposes and foci of the networks examined, their impact has been broad, ranging from public health surveillance to federal initiatives. For example, ESPHealth has been fully implemented in Massachusetts. Several clinical partners within Massachusetts use ESP for automated reporting of notifiable disease to MAVEN (the Massachusetts Department of Public Health's surveillance and case management system).[44] This implementation eases clinician burden of reporting disease to the public health department and enables faster notification of disease, allowing a quicker response by the department to help the community. Another example is FDA Sentinel. The Sentinel System is a key feature of the FDA's post-market surveillance system and has informed the FDA in regulatory decision making and helped ensure that products are safe and effective.[45] Based on Sentinel's success, the FDA is expanding the system and creating FDA-Catalyst. FDA-Catalyst supplements the Sentinel System by providing data from interactions between patients and providers.[46] This new program expands the use of real-world evidence in decision making and provides a deeper set of knowledge that can be used to ensure FDA products are safe and effective.

3.7 Challenges

Challenges identified by discussion participants included several categories of barriers to research network participation as well as barriers to data use. These barriers can constrain the potential scope and reach of PCOR, both during the initial phases of creating new networks and for established networks seeking growth and sustainability. Despite these challenges, several respondents shared an appreciation of the benefits of collaboration for access to complex datasets. Table 2 summarizes these barriers to participation and data use.

Table 2. Barriers to Network Participation and Data Use

Participation Barriers	Data Use Barriers
Barriers to Entry	**Barriers to Data Sharing**
• Staffing, skills, and tools to curate, transform, and sustain data workflows • Consensus process for data use agreements • Consent process and scope • Ensuring appropriate representation for diverse population	• EHR access limitations • Lack of pre-curated, de-identified, longitudinal, "complete" datasets • Limited access to several categories of data, such as claims, death, UDI, and PPI data • Learning curve for key HHS sources like CMS VRDC
Barriers to Growth	**Barriers to Data Integration**
• "Usual suspects" problem; easier for already-active organizations to join more networks than for new organizations to join their first network • Effort necessary to monitor and sustain data quality does not diminish over time • Bridge support needed to sustain and maintain after initial or project-based funding (such as from PCORTF)	• EHR systems typically do not directly support data curation, CDM transformation, and sharing for PCOR • Common data structures do not guarantee common meaning because clinical practice data capture varies among systems and providers • Researchers' data curation processes rarely inform EHR data quality improvement at the point of care • Linkages must be sustained, requiring additional resources for maintenance

The following examples illustrate these barriers in action. Regarding participation in PCOR, several respondents noted the challenges when new organizations join research networks. A primary barrier is the need to build up the staff, skills, and experience to support the data curation process. In addition, best practices need to be sustained to manage those workflows. These barriers may limit participation of smaller, less well-resourced organizations, which could also constrain innovation and growth. An additional set of challenges surrounds governance, including limits on reusing data collected for a specific use. Network participants must work with their partners to reach consensus on data sharing. New data use agreements are often developed within each network, instead of basing this work on common agreement language and consent practices.

For well-established networks, growth can be limited by the same entry barriers as for newer networks, making it difficult to find and support new data sharing partners and expand research capabilities. Networks need a funding model for sustained operations beyond the initial startup phase. One respondent highlighted the need for additional support beyond early network development phases to facilitate absorption into the larger ecosystem.

A nearly universal sentiment shared by discussion participants is each network's need to focus on data quality in a sustained manner. Even representatives from the most mature networks expressed how keeping a critical eye on data quality (consistency, validity, etc.) is both fundamental to their research and a never-ending task. This work does not end at validation of transformation to a common data model, for example, or even after a tested data flow from an experienced data provider has been established and sustained for years. Because clinical data provided by network partners is shared from many systems that have evolved independently, there is always a possibility that quality concerns may arise due to software upgrades, intermittent downtime, or human error even when a robust program of automated checks is in place. There is also the potential within a single practice, and within stable EHR

systems, that providers can record similar encounters differently in EHRs, further compounding the data quality challenge and requiring review of the meaning of captured data.

Several discussion participants noted that, despite a wealth of experience and the continuing investment of many staff-hours dedicated to this work, there is always something new that can go wrong with data quality. Even when a data curation process results in high-quality data available for network research, one participant reported that successful curation efforts are often managed separately from EHR systems and can lack a feedback loop that would increase the quality of the source data within EHR systems. With more integration, curated data could be used to improve clinical decision support and population management tools that are often developed separately, by different staff, and on schedules not aligned to research needs. Data sufficient for clinical or billing purposes may not be sufficient for all research purposes, but several participants proposed approaches to improve billing or completion of certain records that could be taken to raise the data quality for both use cases. Their comments spoke to potential for tools that could help address more common data quality issues in known and well-used resources, while recognizing that every use will require attention to "local" data quality concerns. A respondent also noted that even in some cases where data is widely available, such as the CMS VRDC, each investigator may still have to start from scratch to prepare their data of interest for their research due to restrictions on sharing.

The research networks also reported challenges in accessing data controlled at the state level, such as Medicaid and vital statistics data. Accessing this data can add unique overhead costs, dictated by state requirements for per-study IRB approvals and, in at least one case noted by a discussion participant, increased systems maintenance costs due to requirements that the data had to be stored on a separate system. Participants also mentioned challenges with getting Medicaid information as well as the potential gap in representation in claims data of the under-insured and uninsured. Where new and developing standards exist, like the growing use of UDI, additional means such as regulatory levers may be required to promulgate use throughout the full information life cycle, including capture within EHRs and reporting.

One discussion participant noted the lack of a universally available, de-identified, "complete" dataset, such as the example of the UK Clinical Practice Research Datalink (CPRD).[47] CPRD is a real-world research service supporting retrospective and prospective public health and clinical studies that is jointly sponsored by the Medicines and Healthcare products Regulatory Agency and the National Institute for Health Research (NIHR), as part of the Department of Health and Social Care. CPRD collects anonymized patient data from a network of GP practices across the UK. Primary care data are linked to a range of other health related data to provide a longitudinal, representative UK population health dataset. The data encompass 60 million patients, including 16 million currently registered patients.[48] Despite the growth of the NIH All of Us database, use of its patient-level data is restricted by design, and a de-identified resource like the CPRD could serve as a publicly available complement. Although there are important differences between the data infrastructures of the United States and United Kingdom that make apples-to-apples comparison difficult, the lack of a commonly available resource of similar scope may be a missed opportunity.

4 Opportunities and Synergy with OS Patient-Centered Outcomes Research Trust Fund

The findings and challenges in Section 3 provide a picture of the research network landscape for PCOR. The data infrastructure innovations, challenges, and opportunities identified through the research network scan and discussions could serve as inputs to inform the new 10-year OS-PCORTF strategic plan. This section addresses unique findings from the research network discussions and organizes them based on common themes identified throughout the scan and interviews. The section concludes with a synthesis of cross-cutting research network findings in two key areas for the OS-PCORTF—treatment of SDOH data and "patient centeredness".

4.1 Network Findings Relevant to HHS Data Capacity for PCOR

During the PCOR research network landscape scan and structured discussions, common challenges in accessing and using data for PCOR were identified. Five common themes were recognized—lack of high-quality real-world data, limited tools and resources for linking data between different sources (e.g., EHRs, claims, PROs), lack of tools to improve data quality and curation, constant effort to maintain data quality, and the difficulty of accessing medical claims data, especially Medicaid data. Many findings in the research network scan that address these categories are listed below.

Lack of High-Quality Real-World Data

- **Access to high-quality, current, and representative real-world data sources remains critical to PCOR.** One discussion participant noted that the current data available are too high in volume and too broad in scope to be useful. For researchers to use real-world data, a high-quality dataset is needed with a representative population, a variety of diseases, and detailed information on medications. Additionally, another discussion participant noted that real-world data sharing would be improved if patient-controlled medical records were expanded and involved consent processes for data sharing. This would also improve the ability for patients to participate in research. A common trend noted by many respondents is the transition from using claims data in research to using EHR data because EHR data have greater detail and granularity.
- **Improved access to real-world data from HHS agencies would save time and increase focus on research.** To build studies off each other, clinical researchers need access to real-world data from HHS agencies. Several discussion participants suggested greater inter-agency collaboration and increased inter-agency data sharing as ways HHS could better enable research networks to study patient outcomes. For example, CMS and the FDA have an inter-agency agreement that provides the FDA more freedom to use CMS data. As a result, FDA Sentinel is able to run analyses with the entire CMS VRDC, which is normally expensive. Limitations on data reuse prevent sharing of VRDC data transformed for research, however, and these limitations do not allow sharing VRDC data that has been transformed to use a CDM, for example. Because of this, all investigators must create their own files within VRDC and transform them independently for their own research.
- **Tools for collecting and linking PPI would provide a more holistic view of the patient experience.** Patient outcomes data need to be collected, even when patients are not part of traditional clinical trials. The use of PPI as a data source in research is becoming more common.

As a result, tools to collect and link PPI to EHRs and claims data should become more robust. With such tools and a more holistic view of the patient experience, researchers would be better able to draw conclusions. Similarly, one discussion participant noted that clinical data research may grow to incorporate use of additional symptoms often captured outside of diagnosis and procedure codes and patient SDOH data for research use.

- **Data captured from devices support evaluation throughout product life cycles.** One potential solution to improve data collection of medical devices is the use of the UDI. This would enable researchers to better understand how medical devices perform in practice. One discussion participant noted that a UDI is especially needed in Medicare claims data and should be made available in the same file where a procedure or surgery was billed.

Limited Tools and Resources for Linking Data Between Different Sources (e.g., EHRs, Claims, PROs)

- **Tools and infrastructure to address governance barriers improve data linkage and interoperability.** For example, leadership within institutions is often reluctant to share data and there is a lack of buy-in and standard agreements within the community to use structured and standardized data. Additionally, some types of data can require an individual IRB approval for each separate use. If terms of use are restricted, especially when data are purchased, governance challenges can also make it difficult to link data associated with a single person. Several discussion participants noted the potential for Streamlined, Multisite, Accelerated Resources for Trials IRB (SMART IRB) as a step forward to ease IRB restrictions.[22] Similarly, legal or policy restrictions may limit the ability of networks to link social service information with clinical records. If there were a unifying infrastructure between healthcare entities and community-level services, data linkages would be improved, and patients would not face the burden of re-entering information multiple times.
- **Innovations in distributed analysis algorithms enable new research.** Several research networks, such as FDA Sentinel, use distributed query with a common data model, allowing patient-level data to remain at each institution while enabling researchers to answer specific research questions using de-identified and often aggregate data. This works well for answering targeted questions but limits the types of questions that can be asked. Improved techniques for distributed analysis would improve privacy protection while enabling patient-level distributed queries with a potential for lowering consent burdens.
- **Resources are needed to link EHR, claims, SDOH, and PRO data outside of the treating institution.** This can be a major problem for research networks that are EHR/institution based because EHR data are incomplete, as EHR systems cannot capture information about care completed outside the institution hosting the EHR. This has implications for longitudinal studies if patients are treated at multiple institutions because researchers cannot link and track patients due to privacy requirements, among other concerns. For example, one discussion participant noted that although it is possible to link patients through social security numbers or health insurance plan numbers, this is problematic because not all patients have a social security number or health insurance. To improve the collection of longitudinal data, a unique patient identifier can be used. However, care obtained at clinics that do not use the unique patient identifier will not be included in research, and handling of duplicates can be an issue.

- **A taxonomy for data model requirements could ease interoperability between EHRs.** EHRs are locally customized and require a "translation" process to work between each other. Although certain initiatives, such as the Care Everywhere project from Epic, have been helpful to integrate care by pulling information from multiple Epic systems into one patient chart, legal challenges inhibit patient data exchange. One respondent noted that while using a common data model helps ease the interoperability between EHRs, it is not a perfect solution because data models may vary significantly by the type of research being conducted (e.g., epidemiology vs. clinical trial research).
- **A standardized set of pipelines to share data would help link data between sources.** Currently, there is a lack of well-established pipelines for rapid exchange of data sources. The process for data sharing is essentially reinvented whenever there is a need. This is inefficient and cumbersome. The data exist, but the pipelines do not.

Lack of Tools to Improve Data Quality and Curation

- **A standardized approach for data curation is a potential opportunity to improve data quality.** Organizations vary in their data curation processes or do not undergo the curation process at all because of financial constraints or lack of understanding of its importance. This can be problematic because different curation processes lead to different levels of data quality. Additionally, one respondent noted that EHR data curation requires more attention to data quality than claims data, as EHR data structures vary more and present more variables than claims data.
- **Tools and processes available to assess, feed-back the assessment to data collectors, and improve data quality could streamline data curation for research.** Publishing guidance for completing data checks or establishing standards and common language for communicating data checking methods could improve data quality curation. One discussion participant noted a recent trend in enhancing the ability to feed-back data quality learnings to the source of collection, so it is clear and transparent. The hope is to improve the quality of clinical data at the point of collection.
- **FHIR will drive more attention to data quality because data will be shared more frequently.** Additionally, one discussion participant noted that although FHIR is helpful for creating sustainable data infrastructure, the process of mapping data to FHIR resources is highly inefficient. However, another discussion participant noted that the use of FHIR is crucial in expanding data linkages, and tools like RedCAP should be utilized.

Constant Effort to Maintain Data Quality

- **PCOR data require effort to restructure and organize**. Many respondents noted that data are gathered for a primary use—usually clinical care or billing—and are fit for that purpose. Research is nearly always a secondary use. Unless there is a concerted effort to structure and organize data, information can be spread over many different combinations of fields within a single health record and described in many different ways. Data quality issues can make it harder to link data across sources, understand whether there is a true finding, or apply those findings to improve care.

- **Improving data quality at the point of collection streamlines data curation for research.** It is a challenge to identify important changes or trends in data that may be valid on their own, but implausible when put in the context of additional data. According to a discussion participant, this typically happens when there are changes in underlying datasets. Some clinicians or providers who collect data and are the primary users of data may not see the importance in data quality. As a simplistic example, a physician reading a record will understand that a height entered in for an adult male as 6 inches was probably entered incorrectly, and the patient is actually 6 feet. However, research networks taking in this data generally experience greater fallout from these issues if they are not caught right away. As a result, researchers should constantly monitor data quality and effectively communicate to data collectors to improve the quality of clinical data at the point of collection.
- **Clinical data is often provided to research networks from a variety of sources.** These sources evolve independently, increasing the risk for changes in the way data are recorded and reported. Additionally, providers have different workflows and can record the same information differently, again increasing the risk for differences in data quality and requiring review of data.

Difficulty of Accessing Medical Claims Data, Especially Medicaid Data

- **Improving access to Medicaid and death and vital statistics data would give researchers more data on some of the highest risk patients, critical to addressing health equity.** Because Medicaid is operated on the state level, each state is using its own claims, it is difficult to obtain comparable claims data for patients. Although there is a national Medicaid repository, it has a significant lag and is incomplete in part due to the more limited capture of Medicaid Managed Care data. Additionally, a discussion participant noted that some state-level data cannot be integrated within their usual data warehouse to be queried by network users and must be housed separately. One discussion participant noted the opportunity to unify Medicaid data across all the states and expand its availability. Additionally, accessing the national death index is time consuming and expensive because some states require IRBs and fees for each use of the data that they contribute to the national death index, limiting access to this data for many researchers. If the national death index were more accessible, it would improve the networks' ability to understand the trajectory of patient care. A respondent also noted that it would be helpful if Medicare Advantage claims data could be shared outside the local environment into a true learning health system. Although some workarounds exist to share data without breaching CMS contracts, many organizations are not willing to take the risk.
- **Improving access to claims data through a centralized repository would allow for claims to be used as a real-world data source in prospective studies.** One discussion participant noted that although claims data exist and are available to the public, claims data can be quite hard to access. As a result, if a centralized repository of tools were developed, researchers would have improved access to curated claims data when conducting prospective studies. This repository could capture previous use and performance of computable phenotypes to avoid misalignment of phenotype definitions in future studies.

These examples offer potential opportunities for HHS, with the OS-PCORTF, and others, to improve data capacity for PCOR over the next 10 years. The full set of opportunities and challenges identified through

this research network scan and discussions will be combined with other inputs to inform the new 10-year strategic plan for the OS-PCORTF.

4.2 Network Use of Social Determinants of Health Data

Several networks report their enhanced focus on incorporating information about social determinants of health. One respondent reported working on a linked database of geo-coded SDOH data from the Census and American Community Survey (ACS) and PCORnet data to support their research studies. The NIH All of Us Research Program is also looking to bring in linkages to SDOH information using a similar strategy, perhaps as soon as this year.

Another recent trend is the use of computable phenotypes, particularly for addressing social determinants of health. Computable phenotypes are clinical conditions or characteristics that can be identified or defined via a query using a defined set of data elements and logical expressions.[49] One discussion participant noted how computable phenotypes could be used with claims data as a real-world data source for use in prospective studies. However, there is a need for a centralized repository that captures previous use and performance of computable phenotypes to avoid misalignment of phenotype definitions. Computable phenotypes are already being used in social determinants of health research. For example, the PhenX (consensus measures for Phenotypes and eXposures) Toolkit provides recommended standard data collection protocols for conducting research and includes a collection specific to social determinants of health.[50]

With its focus on safety net populations, ADVANCE participants report that SDOH is a primary area of research. In addition to having incorporated ACS data into projects, a member organization has developed innovative tools to capture patient-level SDOH information for surveys such as PRAPARE. This support has evolved based on use patterns, enabling different clinics to choose what to answer, and providing flexibility to use different screeners supporting different social determinant domains at the point of care and data collection. This flexibility has sometimes resulted in differences among collected data elements. To increase consistency for potential research use, the implementation has moved toward a roll-up question design, ensuring that responses can be compared at the system level. This regional progress, however, is not yet widely adopted among other organizations within the ADVANCE network, which is looking to gain progress on a national standard and consensus on essential questions before incorporating SDOH into network studies.

The increased attention to SDOH and equity has reinforced the need for data addressing the alignment of healthcare and community-level services (e.g., when a survey response suggests referral to a local agency or service provider). Participants noted, however, that there are many barriers to communications and data sharing between a clinic and an agency. Networks noted that privacy concerns are a main barrier because separate business agreements are often needed between each social service entity, which is burdensome. There are a few data sharing programs aimed at unifying infrastructure between healthcare entities and community-level services, such as Unite Us, but privacy requirements remain an issue.[51] As a result, patients face the burden of re-entering information multiple times, and creating research-quality data for PCOR is a challenge.

4.3 Network Definitions of Patient-Centeredness

A common trend seen throughout many of the research networks examined is an increased push toward patient engagement. In fact, the following research networks all include the patient voice in their governance structure or through special interest committees:

- ADVANCE[36]
- HCSRN[52]
- MDEpiNet[53]
- NESTcc[54]
- NIH All of Us[55]
- PCORnet[56]
- PEDSnet[57]

Including patients throughout the whole research process, including formulating research questions, defining characteristics of study participants, recruitment, outcomes, and dissemination of research, has been a common characteristic seen throughout the networks. The patient voice can be heard via many forms, including committees or panels within the network (such as ADVANCE) or through survey instruments (such as NACHC PRAPARE).[52,58] Network respondents expressed that it is important to understand and incorporate the overall experience of patient care in research studies. For this reason, ADVANCE is working on developing qualitative assessments to understand the overall patient experience. Potential uses of this dataset include creating targeted patient resources, improved PRO instruments, and guidance in generating future research questions.

In addition to using patient feedback in study design, many research projects now use patient-provided data (PPD) as a data source. For example, the current FDA CERSI Yale University-Mayo Clinic study, "Real-world data to assess variation in opioid prescribing and use for acute pain in diverse populations," is designed to help the FDA develop evidence-based recommendations for opioid analgesic-prescribing for specific conditions or procedures. To accomplish this, the study includes a diverse group of patients and examines their use of opioid analgesics to manage acute pain, the patients' trajectories of pain experienced and response to opioids, and how patients dispose of these medications. FDA CERSI has recruited 1,550 patients who have been prescribed short-acting opioid analgesics and asked them to provide information for 180 days on pain control and opioid use through survey questionnaires sent via a platform called Hugo.[60] The patient's electronic medical records and pharmacy data are also connected to the Hugo platform, which links the patient-generated data to the EHR data. This allows researchers to have a more holistic view of the patient experience by including data that might not normally be collected regularly, such as day-to-day pain levels, to be better informed when writing these guidelines.[59] As collection of PPD becomes more common, tools to link PPD to EHRs and claims data (such as Hugo) must become more robust so researchers can have a complete view of the patient experience.

Data Application Highlight

Representatives from FDA CERSI described a new study using an application called Hugo Health. Hugo offers participant/patient integration and selective disclosure of their data sources, presenting a more complete longitudinal view of patient histories than what is normally available, so much that participating providers are able to benefit from the integrated view as well.[60] This approach builds on rights afforded to individuals through the Health Insurance Portability and Accountability Act (HIPAA).

Although the service is still in an early stage, it shows promise. A research network discussion participant noted the potential for the Hugo-based dataflows, which use a consumer-pull model, to offer higher-quality, more complete longitudinal data for use in research studies. This example also stands out for its potential positive feedback loop, in which the value added to data through curation of data from multiple sources for research use is shared back to providers. Providers then obtain access to a more complete view of patient data than what would otherwise be available in their own systems alone.

Although using patient-reported data is becoming more common in research, more studies are needed to determine how the recording of patient-reported information could impact study findings. How a question is asked or how information is recorded may affect the response by the patient and could bias study results. There are many different types of patient-reported outcomes and many ways to report them in a scientific study. As a result, studies are underway to better understand the optimal way to record patient-reported outcomes and to standardize tools and measures for recording them. For example, two of the FDA CERSIs, University of California San Francisco (UCSF)-Stanford and Johns Hopkins University (JHU), conducted a study to develop patient-reported outcomes measures (PROMs) that can be included in FDA submissions for minimally invasive glaucoma surgical (MIGS) devices.[61] It is important that patient-reported outcomes are measured in a standardized, repeatable way to reduce bias and be scientifically accurate.

5 Conclusion

This report presents the successes, challenges, priorities, and future directions of a set of research networks engaged in PCOR. Common data challenges include lack of quality real-world data, a shortage of resources to link data, lack of tools for data quality and curation, the need for constant effort to maintain good data quality, and the difficulty of accessing certain claims data. Even the most mature networks included in this scan face these challenges. To address these challenges, research networks have devised such solutions as using new tools to collate patient data, applying CDMs, providing constant monitoring of data quality, using tiered systems to access data, and expanding access to Medicaid data.

In addition to challenges, this report identified many potential opportunities for OS-PCORTF to support PCOR research including interest and involvement in enhancing research access to federal health data, expanded to include device and patient-provided information, among other novel sources, strengthening methods and tools to promote and sustain authoritative health data linkage, developing and implementing standard approaches for data quality, consistency and patient identification, addressing source data workflow strategies for data capture to improve data quality, and addressing

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.

Potential Opportunities for OS-PCORTF to Support PCOR Research as Expressed by the Research Networks

- Enhance research access to federal health data, expanded to include device and patient-provided information, among other novel sources.
- Strengthen methods and tools to promote and sustain authoritative health data linkage.
- Develop and implement standard approaches for data quality, consistency, and patient identification.
- Address source data workflow strategies for data capture to improve data quality.
- 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.

Future Engagement and Opportunities to Work with the OS-PCORTF

The continuing push in clinical research toward greater efficiency and achievement of FAIR data standards (data that is Findable-Accessible-Interoperable-Retrievable/Reusable), as well as increasing emphasis on the use of real-world data, all suggest that research networks will continue to be an important stakeholder for HHS data capacity for PCOR in the coming decade. Research networks can help the OS-PCORTF identify challenges that need solutions, based on the PCOR-relevant research gaps and directions they see ahead: their insights on problematic categories of data or potentially beneficial common tools could help inform OS-PCORTF in its deliberations with the rest of HHS. For high-priority PCOR projects that are currently infeasible due to data limitations, the networks could also help inform the OS-PCORTF on the actions needed to make the research possible. Research networks could also support the dissemination of OS-PCORTF-supported tools, datasets, and other products.

Over the course of the next decade, it would likely be beneficial for ASPE and HHS to find ways to sustain engagements such as these discussions as ASPE works to strengthen HHS data capacity for PCOR. Our team would again like to express its appreciation to those who so generously shared their perspectives with us. Engagements such as these will help ensure that the work of the portfolio reflects and addresses the perspectives and needs of these key end users—PCOR researchers—in the years to come.

Abbreviations and Acronyms

Term	Definition
ACA	Patient Protection and Affordable Care Act of 2010
ACS	American Community Survey
ADL	Activities of Daily Living
ADVANCE	Accelerating Data Value Across a National Community Health Center Network
AGING	Advancing Geriatric Infrastructure and Network Growth
AHRQ	Agency for Healthcare Research and Quality
API	Application Programming Interface (e.g., FHIR)
ARIA	Active Risk Identification and Analysis
ASCO	American Society for Clinical Oncology
ASPE	Assistant Secretary for Planning and Evaluation
BMI	Body Mass Index
BUILD	Building UDI into Longitudinal Data for Medical Device Evaluation
CDM	Common Data Model
CER	Comparative Effectiveness Research
CERSI	Centers of Excellence in Regulatory Science
CMS	Centers for Medicare & Medicaid Services
CPRD	UK Clinical Practice Research Datalink
C-PRL	National Center for Pediatric Practice-Based Research and Learning
CRN	Coordinated Registry Network
CTSA	Clinical and Translational Science Awards
DRN	Distributed Research Network
ePRO	Electronic Patient-Reported Outcomes
EAR	Endovascular Aneurysm Repair
ECHOES	Evaluating Control of Hypertension – Effect of Social Determinants
EHR	Electronic Health Record
ESPHealth	Electronic Medical Record Support for Public Health
FDA	U.S. Food and Drug Administration
FDAAA	FDA Amendments Act
FHIR	HL7 Fast Healthcare Interoperability Resources
FORCE-TJR	Function and Outcomes Research for Comparative Effectiveness in Total Joint Replacement
FQHC	Federally Qualified Health Center

Term	Definition
GUDID	Global Unique Device Identification Database
HCN	Health Choice Network
HCSRN	Healthcare System Research Network
HERO	Healthcare Worker Exposure Response & Outcomes
HHS	U.S. Department of Health and Human Services
HIPAA	Health Insurance Portability and Accountability Act of 1996
HIVE	High Performance Integrated Virtual Environment
HP	HHS Office of Health Policy
ICD	Implantable Cardia Defibrillator
IHE	Integrating the Healthcare Enterprise
IRB	Institutional Review Board
JHU	Johns Hopkins University
MARCQI	Michigan Arthroplasty Registry Collaborative Quality Initiative
MCC	Multiple Chronic Conditions
MDIC	Medical Device Innovation Consortium
MDEpiNet	FDA Medical Device Epidemiology Network
MDPH	Massachusetts Department of Public Health
MIGS	Minimally Invasive Glaucoma Surgical
MIPS	Merit-Based Incentive Payment System
ML	Machine Learning
MOSAIC	Meaningful Outcomes and Science to Advance Innovations Center of Excellence
MU	Meaningful Use
NACHC	National Association of Community Health Centers
NCDR	National Cardiovascular Data Registry
NESTcc	National Evaluation System for health Technology Coordinating Center
NIH	National Institutes of Health
OAIC	Older Americans Independence Centers
OCHIN	Oregon Community Health Information Network
OHDSI	Observational Health Data Sciences and Informatics
OHSU	Oregon Health & Science University
OMOP	Observational Medical Outcomes Partnership
OS-HHS	HHS Office of the Secretary
OS-PCORTF	Office of the Secretary - Patient-Centered Outcomes Research Trust Fund

Term	Definition
PBRN	Practice-Based Research Networks
PCOR	Patient-Centered Outcomes Research
PCORI	Patient-Centered Outcomes Research Institute
PCORTF	Patient-Centered Outcomes Research Trust Fund
PGHD	Patient-Generated Health Data
PI	Primary Investigator
PPD	Patient-Provided Data
PPI	Patient-Provided Information
PPP	Public-Private Partnership
PRAPARE	Protocol for Responding to and Assessing Patients' Assets, Risks, and Experiences
PRO	Patient-Reported Outcome
PROM	Patient-Reported Outcomes Measures
QOPI	Quality Oncology Practice Initiative
rAAA	Ruptured Abdominal Aortic Aneurysms
ROI	Return on Investment
RWD	Real-World Data
RWE	Real-World Evidence
SDOH	Social Determinants of Health
SMART	Streamlined, Multisite, Accelerated Resources for Trials
TIDE	Therapeutics Research and Infectious Disease Epidemiology Group
TJRR	Total Joint Replacement Registry
TPLC	Total Product Life Cycle
UCSF	University of California San Francisco
UDI	Unique Device Identifier
UDS	Uniform Data System
UMD	University of Maryland
USCDI	United States Core Data for Interoperability
VRDC	Virtual Research Data Center

Appendix A. Research Network Profiles

The following research networks shown in Table 3 were reviewed and/or participated in discussions to validate and enhance the information gathered to inform strategic planning. All reviews were conducted by accessing publicly available content on the internet (dates of access Nov 2020–Jan 2021). All interviews were conducted by videoconferencing (conducted Dec 2020–Jan 2021).

The following factors guided ASPE's selection of networks for the key informant discussions: include networks whose goal is to generate evidence to inform patient or clinician decision making, include groups actively engaged in or supporting PCOR, achieve variety in sponsoring organizations, and ensure some diversity in observational versus clinical trial study design. The individuals selected to participate in the network interviews included end users (e.g., research directors or investigators) who would likely have knowledge of research data needs and challenges. About half of the key informant group had some current or prior exposure to OS-PCORTF as a participant in a funded project.

- However, there are several noteworthy limitations of the work described in this report. The PCOR research network environment is large. Only a subset of PCOR-relevant networks was scanned; the subset of those selected for discussions was even smaller. Therefore, the findings in this report are not meant to be exhaustive. Additionally, the findings and challenges highlighted in this report are limited to what could be discerned in publicly available sources, which may not have been fully up to date, and observations of individual research network principal investigators during the discussions, constrained by format and time. Despite these limitations, the research network environmental scan and discussions with research network participants help set the context for the OS-PCORTF strategic planning process as conversations with other stakeholders and research networks continue.

Table 3. Research Network Profiles

Scanned	Participated in Discussions	Prior/Current Involvement in PCORTF Projects
ADVANCE	Yes	
ASCO CancerLinQ		
CTSA		
ESPHealth		
FDA CERSI	Yes	
FDA MDEpiNet	Yes	Yes
FDA Sentinel	Yes	Yes
HCSRN		
NESTcc		
NIH All of Us	Yes	Yes
NIH Collaboratory	Yes	
OHDSI		Yes
PBRN	Yes	
PCORnet	Yes	Yes
PEDSnet		

A.1 Accelerating Data Value Across a National Community Health Center Network (ADVANCE)

A.1.1 Overview

According to "ADVANCE Collaborative: What We Do," the Accelerating Data Value Across a National Community Health Center Network (ADVANCE) is led by the Oregon Community Health Information Network (OCHIN) in partnership with Fenway Health, Health Choice Network (HCN), and Oregon Health & Science University.[62] Its goal is to build and maintain a community laboratory of Federally Qualified Health Centers (FQHCs) serving safety net patients, including the uninsured, the under-insured, undocumented immigrants, and other vulnerable populations. Its aims include:

- Integrate data into a single data management system.
- Engage patients and clinicians in comparative effectiveness research.
- Develop electronic systems for participant recruitment and patient-reported data collection.
- Strengthen partnership infrastructure to support Patient-Centered Outcomes Research (PCOR) and support FQHCs.
- Build FQHC network capacity to meet regulatory requirements. [62]

ADVANCE is fully funded by PCORI and is one of the nine large research networks in PCORnet.[63,64]

A.1.2 Composition

As reported at "ADVANCE Data," data in the ADVANCE network comprises more than 6 million patients from 30 states, over 400 cities, and over 1,400 clinic sites.[a] These totals include a substantial number of pediatric and Spanish-speaking patients, and a majority of patients below the federal poverty level, including many using either Medicare or Medicaid, or uninsured.[13]

ADVANCE supports clinician engagement through the OCHIN practice-based research network (PBRN); more information is available using the reference.[65]

ADVANCE participates in the Healthcare Worker Exposure Response & Outcomes (HERO) Registry, investigating healthcare worker safety in support of understanding COVID-19; more information is available using the reference.[66]

A.1.3 Example Research Projects

- **Short- and Long-Term Effects of Antibiotics on Childhood Growth (2016–2018)** – studies the types, timing, and amount of antibiotic use in the first two years of life, BMI and obesity at ages five and ten, and growth trajectories to age five.[67]
- **Evaluating Control of Hypertension – Effect of Social Determinants (ECHOES) (2018–2022)** – uses electronic health records (EHR) data from the Coordinated Registry Networks (CRNs) linked to community-level social determinants information to record changes in hypertension incidence, screening, treatment, and management, and compares states that expanded and did not expand Medicaid.[68]

These projects, along with several others, are documented in more detail under Current ADVANCE Projects.[68]

A.1.4 Innovations

According to ADVANCE, its data warehouse represents "the largest clinical data set on the safety net population in the nation."[69] The explicit inclusion of typically underrepresented patients within ADVANCE stands out from other networks, many of which report challenges in recruiting fully representative populations in their studies.

A.1.5 Key Facts

Table 4 presents the key facts and details for the ADVANCE research network profile.

[a] This count represents clinic sites only within the ADVANCE participant organizations. There are more than 1,400 Federally Qualified Health Centers, each of which may have many clinic sites; the count similarity is coincidental.

Table 4. Key Facts and Details for ADVANCE Research Network

Key Facts	Details
Site	http://advancecollaborative.org/
Began	2015[64]
Sponsors	ADVANCE is led by OCHIN in partnership with Fenway Health, Health Choice Network (HCN), and Oregon Health & Science University[62]
Primary funding	PCORI[64]
PI	Jon Puro, MPA[70]
Data approach	The ADVANCE Research Data Warehouse uses distributed query with an expanded version of the PCORnet CDM with extra fields applicable to the Uniform Data System (UDS) and related purposes. Interested researchers engage through a "Front Door" model common to PCORnet networks. http://advancecollaborative.org/advancedata
Coordinating Center	ADVANCE is part of PCORnet, which has two Coordinating Centers—Duke Clinical Research Institute and Harvard Pilgrim Health Care Institute[63,71]
Publications	http://advancecollaborative.org/?page_id=157

A.2 FDA Centers of Excellence in Regulatory Science (CERSIs)

A.2.1 Overview

The FDA's Centers of Excellence in Regulatory Science and Innovation (CERSIs) are collaborations between the FDA and academic institutions to advance regulatory science through innovative research, training, and scientific exchanges. The goal is to develop tools, standards, and approaches to assess the safety, efficacy, quality, and performance of innovative products. There are currently four FDA CERSIs:[72]

- Johns Hopkins University (JHU)
- University of California San Francisco (UCSF) – Stanford
- University of Maryland (UMD)
- Yale University-Mayo Clinic

FDA has identified three priority areas for CERSI research:[73]

- High-priority topics with needs across product life cycle and relevant demographic groups (including tobacco, individualized therapies, reducing healthcare-associated infection, opioids, oncology, product development)
- Development of methods to improve quality and safety of FDA-regulated products for use (including complex drugs, biological products, medical devices, biocompatibility of medical devices, evaluation of innovative methods)
- Development of methods to improve post-market evaluation of FDA-regulated products (including incorporation of patient input, novel clinical trial designs, leveraging complex and real-world data to inform regulatory decision making, product safety evaluation with special populations, inclusive demographic evaluations)

A.2.2 Composition

The CERSIs are research Centers of Excellence housed at individual or partnered institutions and operate research studies across many domains in collaboration with the FDA. The primary researchers are faculty members, clinicians, scholars, and students at the institutions comprising each CERSI.[72]

All CERSI initiatives relate directly to clinical and regulatory decision making. Stakeholders include all interested parties ranging from device and drug makers to clinical practitioners, educators, researchers, and patient populations.[72]

A.2.3 Example Research Projects

- **Source Data Capture from EHRs (2015–2019)** – uses Standardized Clinical Research Data to develop methods and tools to automate the flow of structured EHR data into external systems and thereby reduce operating costs, save time, and improve data quality for clinical trials (UCSF-Stanford).[74]
- **Real-World Data to Assess Variation in Opioid Prescribing and Use for Acute Pain in Diverse Populations (2019–present)** – studies patients who have been prescribed short-acting opioids and follows them for 180 days to collect information on pain control and opioid use through surveys sent through a novel data-sharing app/platform (Hugo), as well as collection EHR, pharmacy, and wearable data (Yale-Mayo Clinic).[75]

These projects, as well as several others, are documented in more detail under CERSI Research Projects.[76] Each CERSI partner also has a page on CERSI-funded publications.[77,78,79,80]

A.2.4 Innovations

The FDA uses a Research Impact Metrics model to assess CERSI research project impact, considering advancement of regulatory science, dissemination of scientific knowledge, catalyzing action, and informing regulatory decision making as factors of increasing scope/impact in advancing public health.[43] This model offers a framework for evaluating a wide variety of projects and programs.

A.2.5 Key Facts

Table 5 presents the key facts and details for the FDA CERSI research networks.

Table 5. Key Facts and Details for FDA CERSIs

Key Facts	Details
Site	https://www.fda.gov/science-research/advancing-regulatory-science/centers-excellence-regulatory-science-and-innovation-cersis
Began	• UCSF-Stanford – 2014[81] • UMD – 2011[82] • Yale-Mayo Clinic – 2017[83] • JHU – 2014[84]
Sponsors	FDA Office of Regulatory Science and Innovation[72]
Primary funding	FDA[85]
PI	• JHU: G. Caleb Alexander, MD, MS; Tom Colonna JD, PhD; Janet Holbrook, PhD, MPH; Jodi Segal MD, MPH[86] • UCSF-Stanford: Kathy Giacomini, PhD, and Russ Altman, MD, PhD[87] • UMD: William Bentley, ME, PhD, and James Polli, PhD[88] • Yale-Mayo Clinic: Joseph Ross, MD, MPH, and Nilay Shah, PhD[89]
Data approach	There is no overarching, shared data infrastructure. The data standards and approach to data collection/storage/querying vary by project.
Coordinating Center	The CERSIs do not have one overarching Coordinating Center. Instead, each Center is administratively housed at an academic institution.
Publications	https://www.fda.gov/science-research/advancing-regulatory-science/cersi-research-projects

A.3 FDA Medical Device Epidemiology Network (MDEpiNet)

A.3.1 Overview

According to its website hosted by its Coordinating Center at Weill Cornell Medicine, the Medical Device Epidemiology Network (MDEpiNet) is a public-private partnership that combines expertise and resources to advance a national patient-centered device evaluation and surveillance system. Its mission "is to develop and test novel methods, infrastructure, and partnerships for the creation of re-useable real-world data resources and support device evaluation by multiple stakeholders." Its main activities include:

- Conducting studies to better understand how devices perform in the real world.
- Developing methodologies to support the use and creation of real-world evidence.
- Building strategically Coordinated Registry Networks to advance the collection and use of real-world data.
- Collaborating with NESTcc to link CRNs with other data partner networks. [90]

The overall vision of MDEpiNet is to be "a global leader in the development of innovative approaches for robust, relevant, and reliable evidence generation throughout the medical device lifecycle."[91]

A.3.2 Composition

According to MDEpiNet Structure, MDEpiNet consists of more than 160 partnering public and private sector organizations, over 120 registries from 45 countries, and 12 National CRNs and 4 International Registry Consortia, through which access is available to hundreds of millions of patients through such tools as the High Performance Integrated Virtual Environment (HIVE). MDEpiNet comprises more than 780 clinical experts and over 250 methodologists, and it is a collaborator of the National Evaluation System for health Technology Coordinating Center (NESTcc).[18] MDEpiNet is managed by committees on Executive Operations, Scientific Oversight, Sustainability/ROI (Return on Investment), and Patient Engagement.[53]

Each CRN serves to increase medical device knowledge and improve the quality of patient care, focused on a particular clinical area, with working groups that incorporate broad stakeholder participation.[92] The Orthopedic Devices CRN, for example, focuses on patients undergoing hip, knee, shoulder, and spine surgery in the United States, and both links and validates data from multiple sources, such as registries and state and national claims data, in support of comparative effectiveness research for this population. Data providers for Orthopedic Devices CRN include Kaiser Permanente's Total Joint Replacement Registry (TJRR), the Function and Outcomes Research for Comparative Effectiveness in Total Joint Replacement (FORCE-TJR), and the Hospital for Special Surgery registries, as well as the Michigan Arthroplasty Registry Collaborative Quality Initiative (MARCQI).[93] CRNs are led by coordinators, steering committees, and working groups representing participating organizations.

A.3.3 Example Research Projects

- **The Building UDI into Longitudinal Data for Medical Device Evaluation (BUILD) Initiative (2016–2018)** – aimed to extend the use of Unique Device Identifiers (UDIs) for implantable devices supporting cardiac catheterization in multiple hospitals, support research using data about procedures using these devices, and inform future UDI implementation through improved infrastructure and methods for UDI-associated device clinical data capture, surveillance, and registry linkage. This work resulted in a common data model, new UDI-linked databases, and studies based on analyses of data from the new distributed network.[94]
- **MDEpiNet-supported clinical trial (dates active unclear)** – under the auspices of the International Consortium of Vascular Registries, is evaluating the safety and effectiveness of endovascular aneurysm repair (EAR) in the treatment of ruptured abdominal aortic aneurysms (rAAA). This international effort provides real-world evidence of device performance, long-term safety, and outcomes, as well as feedback into the regulatory process.[95]

More information about these projects and additional research is available at the MDEpiNet Programs page and Publications page.[96,97]

A.3.4 Innovations

MDEpiNet has developed tools and standards for broad use, many of which revolve around the use of UDIs, including the Global Unique Device Identification Database (GUDID) administered by the FDA, and its corresponding searchable database, AccessGUDID (hosted at the U.S. National Library of Medicine). Links to these resources, as well as several tools such as vocabularies, Integrating the Healthcare

Enterprise (IHE) profiles, coder data dictionaries, datasets, and both tutorial and planning tools, are available at the MDEpiNet Tools site.[98]

A.3.5 Key Facts

Table 6 presents the key facts and details for MDEpiNet.

Table 6. Key Facts and Details for MDEpiNet Research Network

Key Facts	Details
Site	https://www.mdepinet.net/
Began	2010[91]
Sponsors	Global public-private partnership under cooperative agreement with FDA and FDA Center for Devices and Radiological Health[91]
Primary funding	FDA[99]
PI	Art Sedrakyan, MD, ScD, PhD, Population Health Sciences, Weill Cornell Medical College[90]
Data approach	Varies among individual projects. Core activities and services at the Coordinating Center include construction of medical device information libraries for long-term surveillance, linkage of registry data with claims and EHR data, and advanced text processing of unstructured data. The central High-Performance Integrated Virtual Environment (HIVE) supports biomedical data standardization, harmonization, and analytics along with data deposit, management, and computation. As one example, the Cardiac Devices CRN combines EHR, claims, and registry data to support a variety of device safety, surveillance, and methodology studies engaging international and regional partnerships. https://www.mdepinet.net/hive https://www.mdepinet.net/methodologicalsolutions https://www.mdepinet.net/cardiac
Coordinating Center	Weill Cornell Medicine[100]
Publications	https://www.mdepinet.net/publications

A.4 FDA Sentinel

A.4.1 Overview

In 2007, Congress passed the FDA Amendment Act (FDAAA), which mandated the FDA to create a system to assess the safety of approved medical products. In response to this mandate, the FDA created the Sentinel Initiative. The overall purpose of Sentinel is to work with public, academic, and private entities to monitor the safety of FDA-regulated products, including drugs, vaccines, biologics, and medical devices.[101]

The Sentinel System comprises healthcare organizations, called Data Partners, that collect healthcare data for their patients in the form of medical billing information and EHRs. These data are then transformed locally into the Sentinel Common Data Model format and sent to the Sentinel Operations

Center and integrated into the Sentinel Distributed Database, which is the collection of datasets from many different Data Partners all in the Sentinel Common Data Model. Because the data are all the in the same standardized format, the FDA can analyze the data to study relationships and patterns, which has enabled the FDA to conduct research studies on FDA-regulated products.[102]

A.4.2 Composition

The Sentinel Distributed Database has more than 70 million members, with both medical and drug coverage, that are currently accruing new data. The database also includes billions of records on pharmacy dispensing, medical encounters, and test results.[103]

The FDA leads the Sentinel System, meaning the FDA has decision-making authority and handles the administration of Sentinel System contracts, establishes strategic priorities, issues project requests, and reviews all work products.[21]

The Sentinel System has 16 Data Partners who have data in the Sentinel Common Data Model. In addition, the Sentinel System has 24 Collaborating Institutions that help to answer the FDA's questions and provide healthcare data and scientific, technical, and organizational expertise. The Collaborating Institutions are led by Harvard Pilgrim Health Care Institute and include the Health Care Systems Research Network and PCORnet (including ADVANCE and PEDSnet).[21]

A.4.3 Example Research Projects

- **Active Risk Identification and Analysis (ARIA) (2016–present)** – analysis used on FDA-regulated medical products to assess a known serious risk related to the use of the drug, to assess signals of serious risk related to the use of the drug, or to identify an unexpected serious risk. When the analysis is finished, the FDA posts the results and details in ARIA Assessments & Impact pages.[104]
- **FDA-Catalyst (2017–present)** – supports surveillance and research of marketed medical products by supplementing the Sentinel System to include data from interactions with patients and/or providers. Key FDA Catalyst projects include collection of patient-provided information (PPI) via mobile devices for comparative effectiveness research (CER) and drug safety research, randomized clinical trials for atrial fibrillation treatment, and development of an E-Consent app for drug-based COVID-19 treatment studies.[46]

These projects, as well as others, can be found on the FDA Sentinel Publications & Presentations page.[105] FDA Sentinel also has a page for Meetings, Workshops & Training resources.[106]

A.4.4 Innovations

The Sentinel System is the "largest multi-site distributed database in the world" dedicated to medical product safety.[45] The FDA launched the Mini-Sentinel Pilot in 2009 and by 2011, the distributed dataset reached 100 million patients. Today, the Sentinel Distributed Database has a total of 351.8 million unique patient identifiers. It is important to note, however, that if patients move between health plans, they may have more than one patient identifier. As a result, there may be more patient identifiers than unique patients included in the database.[45,101]

The Sentinel Operations Center leads the development of the Sentinel Common Data Model, which is used by 16 Data Partners. The Sentinel System has developed routine querying tools to allow users to

identify potential medical product safety concerns, characterize populations of interest, characterize the use of medical products, and perform epidemiologic analyses to assess impact on health outcomes.[102]

A.4.5 Key Facts

Table 7 presents the key facts and details for the FDA Sentinel System

Table 7. Key Facts and Details for FDA Sentinel System

Key Facts	Details
Site	https://www.sentinelinitiative.org/
Began	2008[107]
Sponsors	FDA[21]
Primary funding	FDA[101]
PI	Richard Platt, MD, MSc[108]
Data approach	Sentinel uses distributed query based around the Sentinel Common Data Model. Participant use of data from external organizations is limited to Sentinel activities rather than through explicit Data Use Agreements, with the FDA retaining unlimited rights for access and use. https://www.sentinelinitiative.org/methods-data-tools/sentinel-common-data-model https://www.sentinelinitiative.org/about/principles-policies/principles-policies-data
Coordinating Center	Three distinct Coordinating Centers:[109] • Sentinel Operations Center – overall operations and partnerships, led by Harvard Pilgrim Health Care Institute[110] • Sentinel Innovation Center – innovations to advance Sentinel, led by Harvard Pilgrim Health Care Institute[111] • Community Building and Outreach Center – communication, collaboration, and stakeholder involvement, led by Deloitte Consulting[112]
Publications	https://www.sentinelinitiative.org/news-events/publications-presentations

A.5 NIH All of Us

A.5.1 Overview

According to its website, the All of Us Research Program aims to collect and enable research using data from at least one million people living in the United States, with goals including better health, speeding up health research discoveries, reducing health disparities, improving health equity, and enabling novel individualized healthcare. The creation of a centralized database using data representative of the diverse population is intended to inform studies in many clinical areas, focusing on risk factors and treatments appropriate for patients with different backgrounds in the context of the Precision Medicine Initiative.[113,114]

A.5.2 Composition

At present, the All of Us Research Hub reports over 367,000 participants having joined, with more than 272,000 participants having completed initial program steps including consent, EHR data sharing, physical measurements, and donating at least one biosample. This population includes representation of racial, ethnic, sexual, and gender minorities, as well as people with low income or limited education, and further stands out as it is sourced from entirely patient-provided information.[115]

Researchers may come from organizations that have signed a data use agreement with All of Us and apply through a formal data access request, agreement, and training process. A process for signing up a new institution is available.[116] A subset of anonymized aggregate data is available to the public through a defined data browser application, but not for download. Data at the individual patient level is only available to registered researchers, with a third level of controlled access requiring additional approval.[117]

A.5.3 Example Research Projects

The All of Us Publications page lists six published papers at present, all of which relate to the administration of the program itself, reflecting the recency of this nascent program.[118] Several hundred research projects using data from the Registered Tier are listed in the Research Projects Directory, covering a breadth of clinical practice areas, disease groups, subpopulations such as those having a common ethnicity, and technical methods.[119]

A.5.4 Innovations

For direct volunteer participants, the All of Us Operational Protocol specifies sharing of data from providers to the All of Us Data and Research Center via Sync for Science (based on HL7 FHIR) among other products. Additional tools for provider organizations to convert data to the Observational Medical Outcomes Partnership (OMOP) Common Data Model are slated for development.[120]

The participant intake process includes patient surveys that cover demographics and behavioral topics and suggests that additional surveys may be sent to participants. These might include SDOH information, as suggested by the surveys listed at the Survey Explorer, which includes questions covering demographics, behavioral history, and access to care that may overlap with SDOH topics.[121]

A.5.5 Key Facts

Table 8 presents the key facts and details for the NIH All of Us Program.

Table 8. Key Facts and Details for NIH All of Us

Key Facts	Details
Site	https://allofus.nih.gov/ https://www.researchallofus.org/
Began	Initial awards made in 2016; national enrollment began in 2018[114]
Sponsors	NIH[113]
Primary funding	NIH[113]
PI	Joshua Denny, MD, MS, is the CEO of the All of Us Research Program[122]
Data approach	Data are collected into a centralized data warehouse.[120] A tiered access model allows the public to review aggregate summary statistics and allows researchers registered through a partner organization to access longitudinal patient-level data. https://www.researchallofus.org/data-tools/data-access/
Coordinating Center	Vanderbilt University Medical Center, working with Broad Institute and Verily, with several sub-awardees[123]
Publications	https://www.researchallofus.org/publications/

A.6 NIH Collaboratory

A.6.1 Overview

The overall mission of the NIH Health Care Systems Research Collaboratory is to "strengthen the national capacity to implement cost-effective large-scale research studies that engage healthcare delivery organizations as research partners." The Collaboratory aims to create a new infrastructure for collaborative research to ensure providers and patients can make decisions based on the best available clinical evidence.[124]

The Collaboratory engages with healthcare systems to support and design demonstration projects that address major public health questions. These projects help establish best practices and provide proof-of-concept designs for others involved in pragmatic clinical research.[124]

The NIH Collaboratory Distributed Research Network (DRN) allows collaboration between investigators based in health systems that participate in multiple research networks, including the FDA Sentinel System. The DRN uses the Sentinel System's data, methods, tools, and querying infrastructure and allows for investigators to contact health system members to gather new information for randomized clinical trials.[20]

A.6.2 Composition

There are five Core Working Groups within the Collaboratory that support all demonstration projects and initiatives:[124,125]

- Biostatistics and Study Design
- Electronic Health Records

- Health Care Systems Interactions
- Patient-Centered Outcomes
- Ethics and Regulatory

The NIH Collaboratory Demonstration Projects are active across 90 percent of the United States and include more than 1,100 clinical sites.[16]

The Collaboratory partners with several other organizations, including PCORnet, Health Level Seven International, Agency for Healthcare Research and Quality, National Academy of Medicine, Office of the National Coordinator for Health Information Technology, and the FDA (including the Sentinel Initiative).[124]

A.6.3 Example Research Projects

- **Collection of Patient-Provided Information Through a Mobile Device Application for Use in Comparative Effectiveness and Drug Safety Research (2017)** – a DRN member, Kaiser Permanente Washington, partnered with the FDA MyStudies App to collect exposures and healthcare outcomes from pregnant women.[126]
- **Mother-Infant Linkage Table (2018–2019)** – the DRN linked over 4 million mothers and infants, allowing the possibility to study maternal health and prenatal exposure to birth outcomes.[127]

In addition to these projects, the NIH Collaboratory publishes all protocols, consent documents, public use datasets, computable phenotypes, analytic code, outcome papers, and other materials to a Data Resource Sharing page.[128]

A.6.4 Innovations

The NIH Collaboratory has created "a Living Textbook for pragmatic clinical trials" as a living resource to guide researchers interested in pragmatic clinical trials and research that engages healthcare delivery organization partners. It contains information, tools, and lessons learned to provides knowledge on pragmatic clinical trials from the development of the research question to dissemination of results. The Living Textbook is intended for a broad audience: clinical trialists, healthcare professionals and administrators, and academics, among others.[38]

A.6.5 Key Facts

Table 9 presents the key facts and details for NIH Collaboratory.

Table 9. Key Facts and Details for NIH Collaboratory

Key Facts	Details
Site	https://rethinkingclinicaltrials.org/
Began	2012[129]
Sponsors	NIH[124]
Primary funding	NIH Common Fund[124]
PI	Helene Langevin, MD, CM, and Richard J. Hodes, MD[124]
Data approach	The NIH Collaboratory DRN uses the FDA Sentinel System, specifically its "data, methods, tools, and querying infrastructure," a distributed query approach using the Sentinel CDM as described in Section A.4.5. Authorized investigators from research partner organizations can initiate studies, which are opt-in for each partner organization. https://rethinkingclinicaltrials.org/nih-collaboratory-drn/
Coordinating Center	• Provides national leadership and technical expertise in research; works with the NIH to produce, document, and disseminate standards for research; creates infrastructure that facilitates multicenter studies and use of electronic health data; supports synergy among projects and groups; and coordinates communication and dissemination.[124] • Members – Duke Clinical Research Institute, Harvard Pilgrim Health Care Institute, Kaiser Permanente Washington Health Research Institute, Johns Hopkins Berman Institute of Bioethics[124]
Publications	https://rethinkingclinicaltrials.org/demonstration-projects/

A.7 Practice-Based Research Networks (PBRN)

A.7.1 Overview

The Agency for Healthcare Research and Quality (AHRQ) defines Practice-Based Research Networks (PBRNs) as "groups of primary care clinicians and practices working together to answer community-based health care questions and translate research findings into practice. PBRNs engage clinicians in quality improvement activities and an evidence-based culture in primary care practice to improve the health of all Americans."[130] The PBRN Registry further enables PBRN participants to connect, learn, and conduct research together. PBRNs have been funded both through a series of programs exclusively reserved for PBRNs as well as through non-exclusive competitive grants from AHRQ and other federal agencies.[131]

A.7.2 Composition

The AHRQ PBRN Registry lists 185 PBRNs with member practices representing all 50 states and more than 25 countries as of August 2020. The Registry covers nearly 30,000 practices, more than 150,000 clinicians, and 86-plus million patients as of 2015.[132]

In 2012, AHRQ funded infrastructure support grants to establish eight PBRN Centers for Primary Care Practice-Based Research and Learning (P30 Centers) that share these goals:[133]

- Bring together multiple PBRNs to leverage resources and stimulate innovation in improving delivery and organization of primary care
- Nurture partnerships
- Conduct research
- Disseminate knowledge

The P30 Centers include 53 participating PBRNs.

A.7.3 Example Research Projects

- **Collaborative Care to Reduce Depression and Increase Cancer Screening Among Low-Income Urban Women (PCM3) (2013–2019)** – N^2 (a network of safety-net PBRNs) partnered with its member PBRNs to conduct community-based research and learning projects. The study compared two evidence-based interventions examining the effectiveness of improving cancer screening and PROs. This project was funded by PCORI.[134,135]
- **Stage 3 of the Meaningful Use (MU) Incentive Program (2013–2014)** – MOSAIC (Meaningful Outcomes and Science to Advance Innovations Center of Excellence) partnered with the National Commission on Quality Assurance and the New York City Department of Health and Mental Hygiene to examine the Stage 3 MU Program through studies evaluating and proposing strategies for EHR innovations to increase the value of MU objectives.[136,137]

A.7.4 Innovations

The P30 Centers bring together multiple PBRNs to leverage resources and stimulate innovation in improving the delivery and organization of primary care. While each P30 Center shares common features and goals, they each have a unique purpose and focus. For example, the focus of the National Center for Pediatric Practice-based Research and Learning (C-PRL) Center is on child-based research, and the N2 Center is focused on community health centers that serve the underrepresented population.[133,134,138]

A.7.5 Key Facts

Table 10 presents the key facts and details of the PBRNs.

Table 10. Key Facts and Details for PBRN

Key Facts	Details
Site	https://pbrn.ahrq.gov/
Began	History dates back several decades; recent exclusive funding programs date prior to 2000[139]
Sponsors	AHRQ (primary, as of early 2000s)[131]
Primary funding	AHRQ[130]
PI	Rebecca A. Roper, MS, MPH, is Director, PBRN Initiative, AHRQ[140]
Data approach	The many PBRNs follow different approaches to collecting, harmonizing, and querying data. A 2014 publication on PBRN Research Good Practices co-authored by seven participating PBRNs described a wide variety of tools and techniques for gathering data and initiating studies, from sharing spreadsheets to standardizing data models.[141] This sampling reflects the decades-long evolution of PBRNs, which began prior to modern-day EHR systems. Research is typically initiated by investigators from participating organizations, with data access often limited to participants as well, although practices vary among PBRNs.
Coordinating Center	Distributed model, coordinated by each PBRN
Publications	https://pbrn.ahrq.gov/tools-and-resources/pbrn-literature

A.8 Patient-Centered Clinical Research Network (PCORnet)

A.8.1 Overview

In 2013, PCORI established the national Patient-Centered Clinical Research Network (PCORnet), with a goal to "improve the nation's capacity to conduct health research, particularly comparative effectiveness research (CER), efficiently by creating a large, highly representative network for conducting clinical outcomes research." The PCORnet infrastructure is centered around five types of research—real-world evidence studies, pragmatic clinical trials, population health research, health systems research, and studies on how to best engage with patients in research.[15]

PCORnet has created the PCORnet Common Data Model, a data standard used by partners of the network. PCORnet partners map EHR data, patient-reported outcomes, health claims, and other data sources to the PCORnet data model, generating real-world evidence about the comparative clinical effectiveness of therapies, diagnostics, and prevention strategies.[142] The network uses a distributed approach; the data are not collected into a single data pool but rather stay behind each partner's firewall. Researchers can perform queries on the data.[19]

A.8.2 Composition

PCORnet is a partnership of nine Clinical Research Networks (including ADVANCE and PEDSnet), two Health Plan Research Networks, and a Coordinating Center. PCORnet includes data from more than 70 million people across the United States.[143]

PCORnet also has "patient partners," who are patient representatives on the PCORnet Steering Committee. These patients provide input and collaborate with health professionals to accelerate study enrollment rates, improve protocol design, think of research aims, and strengthen dissemination of results.[56]

A.8.3 Example Research Projects

- **Balance of Effectiveness and Minimal Risk of Bleeding in Aspirin Use (2016–2020)** – studied two different doses of aspirin on risk of bleeding. For this study, PCORnet patient partners created a newsletter for enrolled patients, revised study communication to make it more understandable for a patient audience and engaged with clinicians on what aspects of the study were most important to improve participation rates.[144,145]
- **Heart Failure Medication on Symptom Outcomes (2018–2019)** – studied a change of heart failure medication on outcomes from a symptom perspective. Researchers used retrospective EHR data available via PCORnet and patient-reported outcomes (PRO) via electronic patient-reported outcomes (ePRO) forms.[144,146]

A.8.4 Innovations

The PCORnet Common Data Model is a key tool used by many research networks. Other common data models (e.g., OHDSI OMOP) are mapped to and from the PCORnet CDM for exchange and compatibility.[19]

PCORnet has a "Front Door" strategy for partnership. This means anyone interested in using PCORnet's infrastructure and collaborating on patient-centered research can create a request. Types of requests include data requests, connection with other collaborators, help with understanding what type of PCORnet resources are available, and PCORnet study designation. Once a request is submitted, the PCORnet Coordinating Center sets up a consultation to clarify the request and connects the submitter with the correct contact.[147]

A.8.5 Key Facts

Table 11 presents the key facts and details of PCORnet.

Table 11. Key Facts and Details for PCORnet

Key Facts	Details
Site	https://pcornet.org/
Began	2013[148]
Sponsors	Coordinating Center Leads – Duke Clinical Research Institute and Harvard Pilgrim Health Care Institute[71]
Primary funding	PCORI[148]
PI	Thomas Carton, PhD, MS, Steering Committee Chair[71]
Data approach	PCORnet uses distributed query based around the PCORnet Common Data Model (CDM). Interested researchers engage through a "Front Door" model that enables anyone to make requests, although data access is limited to participating researchers. https://pcornet.org/data/ https://pcornet.org/front-door/
Coordinating Center	• Leads data and engagement activities and supports overall network infrastructure[71] • Members – Duke Clinical Research Institute and Harvard Pilgrim Health Care Institute[71]
Publications	https://pcornet.org/research/

A.9 CancerLinQ

A.9.1 Overview

CancerLinQ is a non-profit subsidiary of the American Society of Clinical Oncology (ASCO). It is a research network focused on improving oncology patient care by analyzing real-world data collected from healthcare systems across the country. Data are extracted from the EHR, harmonized and normalized, and then added to a CancerLinQ database for researchers to access.[149]

In addition to sharing data with researchers, CancerLinQ provides many services to users, including onboarding/set-up, quality management, dashboard/report tools, data exploration and reporting tools, QOPI Certification submission, and transmission of Merit-Based Incentive Payment System (MIPS) Quality Measure data.[150]

A.9.2 Composition

Many health practices, including large institutions, community practices, safety net hospitals, and academic medical centers send real-world data to CancerLinQ. A wide range of researchers can access de-identified aggregated datasets, including researchers from medical specialty societies, grant organizations, academic institutions, federal/state/local governments, and ASCO members.[151,152]

A.9.3 Example Research Projects

- **Annual Trends in Opioid Prescribing for Patients with Metastatic Non-Small Cell Lung Cancer (2010–2017)** – studied the decline of Schedule II prescription rates and the increase of Schedule IV and III prescription rates on the impact of management of cancer pain.[153]
- **Development of an Algorithm Using Natural Language Processing to Identify Metastatic Breast Cancer Patients from Clinical Notes (dates unclear, published 2020)** – developed an algorithm using natural language processing to extract metastatic status and site of metastasis from clinical notes. This is helpful for clinical trial matching and outcomes research.[153]

A.9.4 Innovations

SmartLinQ is a tool created by CancerLinQ that creates dashboards, reports, and quality measure tracking tools for practice teams to visualize and proactively measure and track quality of care. Practices can use a set of measures and track scores to improve quality care.[150]

The SmartLinQ tool also contains the SmartLinQ Quality Oncology Practice Initiative (QOPI) Certification Program Pathway, which allows practices to track, monitor, and submit measure scores to the QOPI Certification Program. Managers can access the measures, identify opportunities to improve measure performance, make corrections, and track scores daily.[150]

A.9.5 Key Facts

Table 12 key facts and details of CancerLinQ.

Table 12. Key Facts and Details for CancerLinQ

Key Facts	Details
Site	https://www.cancerlinq.org/
Began	2014[149]
Sponsors	ASCO[149]
Primary funding	ASCO[149]
PI	Sean Khozin, MD, MPH, is CancerLinQ CEO[149]
Data approach	CancerLinQ uses a centralized model, where data are collected from participating practice EHR systems and curated into a standardized, de-identified system for exploration and research. Researchers from participating organizations can explore summary statistics on key datasets prior to requesting full access to data, provided through a commercial cloud service provider. https://www.cancerlinq.org/about/resources https://discovery.cancerlinq.org/request-data
Coordinating Center	No formal Coordinating Center
Publications	https://meetinglibrary.asco.org/

A.10 Clinical and Translational Science Awards (CTSA)

A.10.1 Overview

The NIH Clinical and Translational Science Awards (CTSA) Program is designed to develop innovative solutions that will improve the efficiency, quality, and impact of the process for turning observations in the laboratory, clinic, and community into interventions to improve the health of individuals and the public. CTSA goals include:[154]

- Train and cultivate the translational science workforce
- Engage patients and communities in every phase of the translational process
- Promote integration of special and underserved populations
- Increase quality/efficiency of translational research
- Advance the use of informatics

A.10.2 Composition

CTSA is a community of medical institutions (called hubs) that work together to provide resources, mentoring, and opportunities to perform research. Fifty institutions currently receive CTSA funding.[155]

A.10.3 Example Research Projects

- **Streamlined, Multisite, Accelerated Resources for Trials Institutional Review Board (SMART IRB) (2016–present)** – meant to be a single IRB platform for multi-site clinical studies. NIH policy requires all NIH-funded multi-site clinical studies to use a single IRB.[22,156]
- **Common Metrics Initiative (2015–present)** – an initiative meant to assess and optimize CTSA's Program impact on national health. The initiative establishes a set of standard evaluation measures across the hubs to focus program activities, streamline data collection, and demonstrate measurable progress.[157,158]

A.10.4 Innovations

CTSA includes the Trial Innovation Network, which is meant to address roadblocks in clinical trials. The Network features a single IRB system, master contracting agreements, quality-by-design approaches, and evidence-based strategies for recruitment and engagement. There are three types of partners involved in this Network: Trial Innovation Centers, Recruitment Innovation Centers, and CTSA Program Hubs.[159]

A.10.5 Key Facts

Table 13 presents the key facts and details of the CTSA Program.

Table 13. Key Facts and Details for CTSA Program

Key Facts	Details
Site	https://ncats.nih.gov/ctsa/about
Began	2006[160]
Sponsors	NIH – National Center for Advancing Translational Sciences[154]
Primary funding	NIH[154]
PI	Michael G. Kurilla, MD, PhD[161]
Data approach	CTSA Program hubs collaborate on initiatives such as the National COVID Cohort Collaborative (N3C). N3C is a centralized resource supporting COVID-19 research, collected using data mapped to an implementation of the OMOP CDM. Access to distinct limited, de-identified, and synthetic datasets derived from this resource is available to researchers from United States–based institutions who register, sign a Data Use Agreement, and complete training and other request-dependent requirements. https://ncats.nih.gov/n3c/about/data-overview
Coordinating Center	Center for Leading Innovation and Collaboration (CLIC) – implemented by The University of Rochester[162] CTSA Program Data to Health Coordination Center (CD2H) – implemented by The Oregon Health & Science University, Northwestern University, University of Washington, Johns Hopkins University School of Medicine, Sage Bionetworks, The Scripps Research Institute, Washington University in St. Louis, The University of Iowa, and The Jackson Laboratory[162]
Publications	https://ncats.nih.gov/ctsa/action

A.11 Electronic Medical Record Support for Public Health (ESPHealth)

A.11.1 Overview

Electronic Medical Record Support for Public Health (ESPHealth) is an open-source software platform that organizes and maps EHR data, analyzes the data for conditions of public health interest, and can transmit either case reports or aggregate summaries to health departments. It serves as the platform for a distributed data network that can be queried by authorized public health officials to assess conditions of interest and can aggregate in a secure and transparent fashion under the oversight and control of the data owner.[163]

A.11.2 Composition

ESPHealth has been fully implemented in Massachusetts. Automated reporting of notifiable diseases occurs from several clinical partners via ESP to MAVEN (MDPH's surveillance and case management system). There is also longitudinal reporting for chronic diseases. Massachusetts has also implemented

aggregate-level querying and reporting capabilities at three clinical partners, which provide care for more than 20 percent of the state population.[44]

In addition to Massachusetts, Ohio, Texas, Washington, and North Dakota have sites that have implemented ESPHealth.[44]

A.11.3 Example Research Projects

- **Targeted condition alerts (dates unclear, published 2019)** – helps ensure public health safety by identifying patients with chlamydia and gonorrhea who have not received recommended treatments or follow-up testing per Centers for Disease Control and Prevention guidelines.[44]
- **RiskScape (ongoing)** – web-based visualization tool enables users to explore summaries of conditions of interest filtered or stratified by demographic and health characteristics, using a centralized database of deidentified individual-level data receiving monthly updates.[44]

A.11.4 Innovations

ESPHealth helps practices fulfill their obligation to report selected communicable diseases to public health agencies. ESP automatically detects most high-frequency notifiable diseases and can submit case reports to participating health departments. The Massachusetts Department of Public Health (MDPH) automatically integrates ESP case reports into its electronic case management system. Reporting includes data on symptoms, comorbidities, treatments, and pregnancy status. Reports are sent to health department surveillance systems in HL7 format in an encrypted fashion. The records can be longitudinal for chronic infections. Practices using this in Massachusetts are eligible for Meaningful Use Credits for reporting to a Specialized Registry.[164]

ESPHealth is also used for vaccine adverse event reporting. ESP prospectively monitors vaccinated patients for anything that may indicate a possible adverse reaction to the vaccine. If this is detected, the system invites the clinician to comment and/or send an automated, pre-populated electronic case report to VAERS as an HL7 message.[164]

A.11.5 Key Facts

Table 14 presents the key facts and details of ESPHealth.

Table 14. Key Facts and Details for ESPHealth

Key Facts	Details
Site	https://www.esphealth.org/
Began	2005[165]
Sponsors	Developed by the Therapeutics Research and Infectious Disease Epidemiology Group (TIDE) in the Department of Population Medicine at Harvard Medical School and Harvard Pilgrim Health Care Institute working with the Massachusetts Department of Public Health[166]
Primary funding	Funding mostly comes from the MA Department of Public Health and the CDC. Other funding has come from the Office of the National Coordinator for Health Information Technology and the NIH.[166]
PI	Richard Platt, MD, MSc, is the Department Chair of TIDE[167]
Data approach	The ESP software package is installed by providers who map their local data for import using ESP-specific models. Surveillance queries are then run against distributed ESP instances, and case reports are automatically sent to the public health department via HL7. Providers can also participate in a distributed query system to support additional, non-automated research. https://www.esphealth.org/how-esp-works https://www.esphealth.org/resources/faq
Coordinating Center	Coordinating Center is part of TIDE in the Department of Population Medicine at Harvard Medical School and Harvard Pilgrim Health Care Institute[166]
Publications	https://www.esphealth.org/resources/publications-and-presentations

A.12 Healthcare Systems Research Network (HCSRN)

A.12.1 Overview

The Healthcare Systems Research Network (HCSRN) conducts population-based research based on EHR, claims, and administrative healthcare data. HCSRN can support pragmatic and traditional clinical trials across multiple health sites. Health site member organizations maintain control of their own data via a federated model. As a result, member organizations agree to make certain data available for research, but there is no centralized database to store data.[201]

A.12.2 Composition

HCSRN represents more than 2,000 scientists and research staff from an array of disciplines.[168] Members are non-profit healthcare delivery systems with embedded research units whose scientists are dedicated to public domain research. Members typically have an integrated delivery system, the ability to define patient populations, and access to electronic medical and administrative records.[169]

Research is conducted within the Network and with external collaborators. Typical collaborators include other HCSRN members, academic medical centers, Clinical and Translational Science awardees, regional and disease-specific research networks and public health departments.[170]

A.12.3 Example Research Projects

- **AGING Initiative (2014–present)** – The goal of the AGING (Advancing Geriatric Infrastructure and Network Growth) Initiative is to bridge the HCSRN with the Claude D. Pepper Older Americans Independence Centers (OAIC) to create a national resource to nurture and advance an interdisciplinary research agenda focused on older adults with multiple chronic conditions (MCC).[171]

- **SUPREME DM (2010–present)** – is a comprehensive, longitudinal clinical registry of a population of approximately 1.1 million insured patients with diabetes mellitus. The project has developed a similar database of all members without diabetes from 11 integrated healthcare delivery systems. These databases can be used for surveillance and research. The registry covers the period 2005–2012, draws from demographic and clinical data elements in EHR and other system databases, captures patient-reported data where it is already being routinely collected, and adds calculated data on medication adherence. The SUPREME-DM Network consists of a multi-disciplinary network of nearly 30 diabetes researchers from both the HCSRN and academic centers.[172]

A.12.4 Innovations

HCSRN uses the CTSA SMART IRB Platform to manage and adhere to the oversight provisions for multi-site research.[173] HCSRN has special interest groups (e.g., addiction, aging, genomics, and patient engagement) for researchers to connect and collaborate.[170]

A.12.5 Key Facts

Table 15 presents the key facts and details for HCRSN.

Table 15. Key Facts and Details for HCRSN

Key Facts	Details
Site	http://www.hcsrn.org/en/
Began	1994[10]
Sponsors	HCSRN is led by members of multiple organizations[174]
Primary funding	Member dues[175]
PI	Jeanette May, PhD, MPH, Executive Director[174]
Data approach	The HCSRN uses a distributed query approach based around its Virtual Data Warehouse (VDW) CDM. Researchers from HCSRN member organizations may conduct studies after agreed-upon requirements are met. http://www.hcsrn.org/en/About/Data/
Coordinating Center	No formal Coordinating Center
Publications	http://www.hcsrn.org/en/Collaboration/Consortia/

A.13 National Evaluation System for health Technology Coordinating Center (NESTcc)

A.13.1 Overview

The National Evaluation System for health Technology Coordinating Center (NESTcc) is an independent, neutral, Coordinating Center driving quality and efficiency in the use of real-world data (RWD) to inform medical device development and evaluation. The main goal is to improve real-world evidence (RWE) studies to enhance regulatory and clinical decision making. NESTcc curates data sources to meet objectives and ensure transparent, traceable RWD provenance. NESTcc also informs and streamlines research for the generation of high-quality evidence by establishing core datasets, using common definitions, and outlining data quality and methods that expand the possibilities of RWD to answer questions of interest.[176,177]

A.13.2 Composition

NESTcc holds a dual role in the medical device ecosystem as a Coordinating Center and as a collaborative community. The data NESTcc handles is from a patient population of 157 million in the United States, 21,973 practices/clinics, and 256 hospitals/medical centers.[177]

In 2018 NESTcc established a research data network activity.[176] NESTcc partners with Network Collaborators, such as MDEpiNet, to collect, curate, and analyze data to make sure it is fit for purpose.[177]

A.13.3 Example Research Projects

- **Post-Market Medical Device Surveillance with a Novel mHealth Platform (2017–2019)** – tests the feasibility of using an mHealth app for post-market surveillance in patients (1) after sleeve gastrectomy and (2) after catheter-based atrial fibrillation ablation. Outcomes collected included enrollment times, patient participation, dropout, completion of patient-reported outcome measure queries, and user satisfaction and burden.[178]
- **ICD Registry DELTA Active Surveillance Pilot Study (2013–2018)** – A surveillance system was developed to monitor ongoing clinical datasets to detect emerging differences in safety or efficacy of medical devices. This project is an observational study that assesses the validity of DELTA surveillance in monitoring high-energy implantable cardiac defibrillator (ICD) leads and utilizes the propensity-matched survival method applied dynamically to the National Cardiovascular Data Registry (NCDR) ICD Registry.[179,180]

A.13.4 Innovations

NESTcc produces Test-Cases that reflect the diversity of types of medical devices available and the different uses of data in pre-market and post-market settings. This includes projects along the pre-market approval regulatory pathways, across nine disease areas, and throughout the medical device Total Product Life Cycle (TPLC). There are two main objectives: (1) explore feasibility for medical device ecosystem stakeholders to work RWD Sources and NESTcc's initial set of Network Collaborators and (2) identify areas where NESTcc could play a role in reducing transaction costs (contracting, IRB, data sharing agreements, publication policies, etc.).[181]

A.13.5 Key Facts

Table 16 presents the key facts and details of the NESTcc.

Table 16. Key Facts and Details for NESTcc

Key Facts	Details
Site	https://nestcc.org/
Began	2016[176]
Sponsors	MDIC[176]
Primary funding	FDA[176]
PI	Robert Griffin, PhD – Director of Research[182]
Data approach	NESTcc supports a research network that comprises large-scale providers and research networks including MDEpiNet and PEDSnet, and can coordinate research using query systems supported by network members using multiple standards and common data models, including Sentinel, PCORnet, and OMOP. A publicly available form allows anyone to submit their project interests, and NESTcc will work to review submissions and manage engagements from assembling a research team through evaluating final results. https://nestcc.org/research-network/ https://nestcc.org/real-world-evidence/
Coordinating Center	NESTcc is a Coordinating Center for MDIC[176]
Publications	https://nestcc.org/demonstration-projects/ https://nestcc.org/test-cases/

A.14 Observational Health Data Sciences and Informatics (OHDSI)

A.14.1 Overview

Observational Health Data Sciences and Informatics (OHDSI) is a multi-stakeholder, international, interdisciplinary collaborative dedicated to eliciting the value of health data through large-scale analysis.[183] The mission of OHDSI is to improve health by empowering a community to collaboratively generate the evidence that promotes better health decisions and better care.[184] All solutions are open source.[183]

A.14.2 Composition

The network comprises researchers/participants from a wide range of disciplines (including clinicians, data management and statistics/data science, researchers, epidemiologists, programmers/engineers, and students/post-docs). There are roughly 180 researchers, more than 100 databases, and one billion patient records.[185,186]

A.14.3 Example Research Projects

- **Hydroxychloroquine and azithromycin treatment risk study (analyzed data from 2000–2020)** – published in *Lancet Rheumatology*, showing increased risk of complications in both short and long term, depending on application.[187,188]
- **Side effects from chlorthalidone (analyzed data from 2001–2018)** – showed side effects from chlorthalidone are more serious than hydrochlorothiazide in treatment for lowering blood pressure, published in *JAMA Internal Medicine*.[187,189]

A.14.4 Innovations

OHDSI developed the OMOP CDM, which is widely used among data research networks.[190] The Book of OHDSI details the OHDSI community, standards, and helpful tools.[191]

A.14.5 Key Facts

Table 17 presents the key facts and details of OHDSI.

Table 17. Key Facts and Details for OHDSI

Key Facts	Details
Site	https://ohdsi.org/
Began	2014[17]
Sponsors	Multi-stakeholder led[192]
Primary funding	Projects are funded by a variety of sources[192]
PI	• Christian Reich, MD, PhD (IQVIA), is a Principal Investigator of OHDSI and was PI and Program Manager of OMOP[185] • George Hripcsak, MD (Columbia), is PI of the OHDSI Coordinating Center[185]
Data approach	OHDSI studies are conducted using a distributed query approach based on OHDSI's OMOP CDM. The volunteer-based community welcomes research participation from the general public, who are encouraged to connect with other researchers through OHDSI forums and to follow published guidance for conducting studies. https://ohdsi.org/data-standardization/ https://ohdsi.org/who-we-serve/ https://ohdsi.github.io/TheBookOfOhdsi/StudySteps.html#StudySteps
Coordinating Center	Coordinating Center is at Columbia University's Department of Biomedical Informatics, and has received support from Takeda Pharmaceuticals, Janssen R&D, and AstraZeneca[193]
Publications	https://data.ohdsi.org/OhdsiStudies/

A.15 PEDSnet

A.15.1 Overview

The purpose of PEDSnet is to conduct research to improve the health and lives of children. Research types include observational research and clinical trials among health systems.[194] There are four strategic goals until 2021:[195]

- Grow the data network to include 10 percent of U.S. children
- Reduce study start-up time to less than 60 days
- Attain financial sustainability
- Provide training opportunities in the science of pediatric learning health systems

A.15.2 Composition

PEDSnet conducts pediatric research across all specialties. The community comprises hospitals and healthcare organizations, researchers, clinicians, and patients. PEDSnet makes data available in two common data models—PEDSnet CDM and PCORnet CDM. The PEDSnet CDM subsumes all PCORnet elements and addresses data elements not yet addressed in PCORnet CDM.[14]

A.15.3 Example Research Projects

- **Effects of Age, Sex, Race/Ethnicity, and Allergy Status in Obesity-Related Pediatric Asthma (2009–2015)** – Obesity in children increases the risk for new asthma. How age, sex, race/ethnicity, and allergy status affect the relationship between obesity and asthma is unclear. This study describes the relationship between high body mass index (BMI) and incident asthma.[196,197]
- **PKIDS (2020–present)** – an observational, prospective cohort study to compare stone clearance and patient experiences for surgical treatment to remove kidney stones. The study also compares re-treatment, and unplanned healthcare encounters across treatments within three months after surgery.[198,199]

A.15.4 Innovations

The PEDSnet Scholars Program is a grant rewarded under the AHRQ-PCORI Institutional Mentored Career Development Program. The goal is to support the training of clinicians and research scientists to conduct PCOR within learning health systems.[41]

Patients are also included in the Executive Management Team and Engagement Workgroup team. The engagement committee makes up parents and healthcare providers and they review/give input on research requests. Parents/patients are engaged in formulating research questions; defining characteristics of study participants, comparators, and outcomes; and conducting/monitoring the research.[57]

A.15.5 Key Facts

Table 18 presents the key facts and details for PEDSnet.

Table 18. Key Facts and Details for PEDSnet

Key Facts	Details
Site	https://pedsnet.org/
Began	2015[200]
Sponsors	PCORI[200]
Primary funding	PCORI[200]
PI	Christopher B. Forrest, MD, PhD[200]
Data approach	PEDSnet uses a distributed query approach using the PEDSnet CDM, which is based on the OMOP CDM and supports all PCORnet CDM elements and additional components needed by pediatric investigators. PEDSnet supports studies using the PCORnet CDM as well. Investigators can seek data consultations and research collaborations through a "Front Door" model common to PCORnet networks. https://pedsnet.org/data/common-data-model/ https://pedsnet.org/research/access-pedsnet/
Coordinating Center	Executive Management Team oversees Coordinating Center activities
Publications	https://pedsnet.org/research/publications/

Appendix B. Research Network Data Collection

The following outline describes the framework used to capture information about each of the networks highlighted in this report and summarized in the profiles in Appendix A. The objectives for this data collection process included preparing for discussions with network representatives and helping to identify themes, trends, and challenges to inform strategy development for the OS-PCORTF. Although the environmental network scan was not intended to create an exhaustive index or catalog of research networks and relevant information about them, the data collection framework contributed to capturing a level of detail that could serve additional purposes.

Purpose (functionality and/or area of PCOR focus)

1. Gap(s) addressed
2. Research objectives

Composition

3. Who/what comprises the network (e.g., researchers, populations and subpopulations, data providers, clinical domains)?
4. Example research projects (brief description of 2-3 illustrative projects)

Data Sources/Elements Used

5. Patient-provided information
6. Data standards used
7. Social determinants of health

Network Outputs

8. How are research findings shared (e.g., reports, publications, briefs, webinars, conferences)?
9. What other artifacts are produced (e.g., tools, standards, datasets)? How are they shared?
10. What services are provided (e.g., training)? How are they shared?

Practical Impact (e.g., on clinical and/or regulatory decision making or guidelines, evidence-based treatments, patients, health outcomes/metrics)

11. Stakeholders impacted (e.g., subpopulations)

Governance Structure

12. Network end users (primary and secondary)
13. Formal (funded) Coordinating Center (Y/N)?
14. Approach to data collection/storage/querying: centralized vs. federated/distributed
15. Funding approach for research projects (e.g., decision/award process, timeline, frequency, performance monitoring)

16. Coordination mechanisms (e.g., data use agreements, membership rules)
17. Is data privacy/security addressed? If so, what issues are covered or discussed?
18. Are patients engaged in the network? If so, how?

Other

19. Hyperlink(s)
20. Data approach (prospective, existing, both)
21. Federal or external owner
22. Primary Investigator (PI)
23. Dates active
24. Additional notes/innovations of note

Appendix C. Semi-Structured Discussion Questions

Introduction

1. Tell me about your role within [network]. *If they have had multiple roles, ask about their history with the network.*

Questions About Their Operating Environment

2. What are [your network's] biggest data-related challenges? What plans, if any, does your network have to address them?
    - *Probe*: For example, in data tools, services, support, governance, participation, training, incentives?
    - *Probe*: Do you have a sense of the timeframe or resources needed to overcome these challenges?
3. What trends is your research program tracking and responding in order to strengthen the ability to conduct research that informs decisions by patients and clinicians? What gaps do you see in the present environment related to data network functionality?
    - *Probe*: For example, forces driving research, network development, and data infrastructure, whether on the technical, policy/regulatory, consumer, or health system demand side?
    - *Probe:* What innovative health information technologies (health IT) and methodologies, such as artificial intelligence tools, machine learning, privacy preserving linkage approaches, and collaborative scientific platforms, might present opportunities for researchers studying patient outcomes?
4. How does your network currently engage with HHS on data that supports PCOR studies?
5. What could HHS do to better enable research networks like yours to study patient outcomes?
6. Is there anything else you think that ASPE should know about building data capacity for networks like yours to improve PCOR research in the future?

Supplemental Questions

7. What major functionalities do data networks need to support research on the effectiveness of interventions on health outcomes?
    - *Probe*: For example, a new functionality could be the ability to follow patient outcomes longitudinally over time and across different databases.
8. What improvements are needed in data infrastructure functionality over the next 10 years?
    - *Probe:* For example, policies, governance, standards, services, technology.

Supplemental Questions for Networks that Conduct Research

9. Are there research questions your network would like to be able to explore, but the available data or methods limit your ability to answer them?

List of References

1. 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
2. Patient-Centered Outcomes Research Institute. (2012, May 08). *Patient-centered outcomes research*. https://www.pcori.org/research-results/about-our-research/patient-centered-outcomes-research
3. Further Consolidated Appropriations Act, 2020, Publ. L. No. 116-94, 133 Stat. 2534 (2019). https://www.congress.gov/116/plaws/publ94/PLAW-116publ94.pdf
4. Assistant Secretary for Planning and Evaluation. *Patient-Centered Outcomes Research Trust Fund.* https://aspe.hhs.gov/patient-centered-outcomes-research-trust-fund. Accessed February 2021.
5. Department of Health and Human Services. (2017). *HHS strategic plan for building data capacity for patient-centered research 2017-2021*.
6. NORC at the University of Chicago, AcademyHealth. (2020). *Challenges and improvements for PCOR data infrastructure: Results from a stakeholder prioritization activity*. Office of the Assistant Secretary for Planning and Evaluation, Office of Health Policy. https://aspe.hhs.gov/system/files/pdf/259016/PCOR-Data-Infrastructure.pdf
7. The National Academies of Sciences, Engineering, and Medicine. *Building data capacity for patient-centered outcomes research: An agenda for 2021 to 2030.* https://www.nationalacademies.org/our-work/building-data-capacity-for-patient-centered-outcomes-research-an-agenda-for-2021-to-2030. Accessed February 2021.
8. Kim, K.K., Browe, D.K., Logan, H.C., Holm, R., Hack, L. & Ohno-Machado, L. (2014). Data governance requirements for distributed clinical research networks: triangulating perspectives of diverse stakeholders. *Journal of the American Medical Informatics Association*, *21*(4), pp.714-719.
9. Patient-Centered Outcomes Research Institute (PCORI) Methodology Committee. (2019). *The PCORI Methodology Report.* https://www.pcori.org/sites/default/files/PCORI-Methodology-Report.pdf
10. Rahm, A.K., Ladd, I., Burnett-Hartman, A.N., Epstein, M.M., Lowery, J.T., Lu, C.Y., Pawloski, P.A., Sharaf, R.N., Liang, S.Y. & Hunter, J.E. (2019). The Healthcare Systems Research Network (HCSRN) as an Environment for Dissemination and Implementation Research: A Case Study of Developing a Multi-Site Research Study in Precision Medicine. *eGEMs*, 7(1).
11. Marsolo, K.A., Brown, J.S., Hernandez, A.F., Hammill, B.G., Raman, S.R., Syat, B., Platt, R. & Curtis, L.H. (2020). Considerations for using distributed research networks to conduct aspects of randomized trials. *Contemporary clinical trials communications*, *17*, p.100515.
12. Ozanich, G. & Ramos, C. (2020). *Maximizing data interoperability and integration to support value-based care.* Office of the Assistant Secretary for Planning and Evaluation. https://aspe.hhs.gov/system/files/pdf/264456/Data-Integration-Brief.pdf
13. The ADVANCE Collaborative. *ADVANCE data.* http://advancecollaborative.org/advancedata. Accessed February 2021.
14. PEDSnet. *Common data model*. https://pedsnet.org/data/common-data-model/. Accessed February 2021.

15. The National Patient-Centered Clinical Research Network. *About.* https://pcornet.org/about/. Accessed February 2021.

16. National Institutes of Health. *Demonstration projects.* NIH Collaboratory Living Textbook of Pragmatic Clinical Trials. https://rethinkingclinicaltrials.org/demonstration-projects/. Accessed February 2021.

17. Observational Health Data Sciences and Informatics. *Who we are.* https://ohdsi.org/who-we-are/. Accessed February 2021.

18. MDEpiNet. *MDEpiNet structure.* https://www.mdepinet.net/structure. Accessed February 2021.

19. The National Patient-Centered Clinical Research Network. *Data.* PCORnet. https://pcornet.org/data/. Accessed February 2021.

20. Platt, Richard & Curtis, Lesley. *NIH Collaboratory Distributed Research Network (DRN)*. NIH Collaboratory Living Textbook of Pragmatic Clinical Trials. https://rethinkingclinicaltrials.org/nih-collaboratory-drn/. Accessed February 2021.

21. U.S. Food and Drug Administration. *Who is involved.* Sentinel. https://www.sentinelinitiative.org/about/who-involved. Accessed February 2021.

22. National Center for Advancing Translational Sciences, National Institutes of Health. (2018, October 26). *NCATS SMART IRB platform*. https://ncats.nih.gov/ctsa/projects/smartirb

23. Apple Inc. (2021). *ResearchKit*. https://developer.apple.com/researchkit/

24. Google. (December 9, 2020). *Advancing health research with Google Health Studies*. https://blog.google/technology/health/google-health-studies-app/

25. U.S. Food and Drug Administration. *Sentinel Common Data Model.* Sentinel. https://www.sentinelinitiative.org/methods-data-tools/sentinel-common-data-model. Accessed February 2021.

26. The Observational Health Data Sciences and Informatics. *OMOP Common Data Model.* https://ohdsi.github.io/CommonDataModel/. Accessed February 2021.

27. HL7 International – Biomedical Research and Regulation Work Group. (2019). *Common Data Models Harmonization FHIR Implementation Guide*. https://build.fhir.org/ig/HL7/cdmh/

28. U.S. Food and Drug Administration. (2019). *Sentinel System five-year strategy 2019-2023*. https://www.fda.gov/media/120333/download

29. U.S. Food and Drug Administration. (2020, June 30). *Unique device identification system (UDI system).* https://www.fda.gov/medical-devices/device-advice-comprehensive-regulatory-assistance/unique-device-identification-system-udi-system

30. The National Association of Community Health Centers. *Protocol for Responding to Assessing Patients' Assets, Risks, and Experiences (PRAPARE).* https://www.nachc.org/research-and-data/prapare/. Accessed February 2021.

31. U.S. Food and Drug Administration. *openFDA*. https://open.fda.gov/. Accessed February 2021.

32. Centers for Disease Control and Prevention. (2021, February 8). *National Death Index.* National Center for Health Statistics. https://www.cdc.gov/nchs/ndi/index.htm

33. Agency for Healthcare Research and Quality. *PBRN listserv.* Practice-Based Research Networks. https://pbrn.ahrq.gov/tools-and-resources/pbrn-listserv. Accessed February 2021.

34. ASCO CancerLinQ. (2020). *CancerLinQ Discovery data access policy*. https://s3.amazonaws.com/files.cancerlinq.org/prod/s3fs-public/2020-07/Data%20Access%20Policy%20%28CLQ%20Discovery%29%285.27.2020%29.pdf

35. National Institutes of Health. *Dissemination approaches for different stakeholders*. NIH Collaboratory Living Textbook of Pragmatic Clinical Trials. https://rethinkingclinicaltrials.org/chapters/dissemination/dissemination-different-stakeholders/dissemination-to-different-stakeholder-introduction/. Accessed February 2021.

36. National Institutes of Health. *Living Textbook table of contents.* NIH Collaboratory Living Textbook of Pragmatic Clinical Trials. https://rethinkingclinicaltrials.org/table-of-contents/. Accessed February 2021.

37. Zigler, C., O'Brien, E., & Weinfurt, K. *Section 3: NIH HEAL, FDA, and other core outcome sets.* NIH Collaboratory Living Textbook of Pragmatic Clinical Trials. https://rethinkingclinicaltrials.org/chapters/conduct/real-world-evidence-patient-reported-outcomes-pros/nih-heal-fda-and-other-core-outcome-sets/. Accessed February 2021.

38. National Institutes of Health. *Welcome to Rethinking Clinical Trials: A Living Textbook of Pragmatic Clinical Trials.* NIH Collaboratory Living Textbook of Pragmatic Clinical Trials. https://rethinkingclinicaltrials.org/welcome/. Accessed February 2021.

39. Observational Health Data Sciences and Informatics. *Software Tools*. https://www.ohdsi.org/software-tools/. Accessed February 2021.

40. ASCO CancerLinQ. *Resources*. https://www.cancerlinq.org/index.php/about/resources. Accessed February 2021.

41. PEDSnet. *Scholars*. https://pedsnet.org/pedsnet-scholars/. Accessed February 2021.

42. National Institutes of Health. *Grand rounds*. NIH Collaboratory Living Textbook of Pragmatic Clinical Trials. https://rethinkingclinicaltrials.org/grand-rounds-hub/. Accessed February 2021.

43. U.S. Food and Drug Administration. (2018, July 13). *CERSI Research Impact Metrics.* https://www.fda.gov/science-research/advancing-regulatory-science/cersi-research-impact-metrics

44. Electronic Medical Record Support for Public Health. *Who's using ESP.* https://www.esphealth.org/installations. Accessed February 2021.

45. U.S. Food and Drug Administration. (2019, October 18). *FDA's Sentinel Initiative.* https://www.fda.gov/safety/fdas-sentinel-initiative

46. U.S. Food and Drug Administration. *FDA-Catalyst projects.* https://www.sentinelinitiative.org/methods-data-tools/fda-catalyst-projects. Accessed February 2021.

47. (UK) Medicines & Healthcare products Regulatory Agency. (2021). *Clinical Practice Research Datalink*. https://www.cprd.com

48. National Institute for Health Research. *Clinical Practice Research Datalink.* https://www.cprd.com/. Accessed February 2021.

49. Richesson, R.L., Wiley, L.K., Gold, S., & Rasmussen, L. *Electronic health records-based phenotyping.* NIH Collaboratory Living Textbook of Pragmatic Clinical Trials. http://rethinkingclinicaltrials.org/chapters/conduct/electronic-health-records-based-phenotyping/electronic-health-records-based-phenotyping-introduction/#:~:text=In%20the%20context%20of%20electronic,data%20elements%20and%20logical%20expressions. Accessed February 2021.

50. PhenX Toolkit. *Social determinants of health collections.* https://www.phenxtoolkit.org/collections/view/6. Accessed February 2021.

51. Unite Us. *Unite Us platform*. https://uniteus.com/platform/. Accessed February 2021.

52. The ADVANCE Collaborative. *Patient engagement.* http://advancecollaborative.org/?page_id=161. Accessed February 2021.

53. MDEpiNet. *Committees*. https://www.mdepinet.net/committees. Accessed February 2021.

54. NEST Coordinating Center. *NEST Coordinating Center (NESTcc) charter.* https://nestcc.org/wp-content/uploads/2020/05/NESTcc-Charter_v4.pdf. Accessed February 2021.

55. All of Us Research Program, National Institutes of Health. *Core values.* https://allofus.nih.gov/about/core-values. Accessed February 2021.

56. The National Patient-Centered Clinical Research Network. *Engagement.* https://pcornet.org/engagement/. Accessed February 2021.

57. PEDSnet. *Policies*. https://pedsnet.org/resources/governance/. Accessed February 2021.

58. National Association of Community Health Centers, Association of Asian Pacific Community Health Organizations, & Oregon Primary Care Association. (2019). *PRAPARE implementation and action toolkit.* https://www.nachc.org/wp-content/uploads/2020/08/NACHC_PRAPARE_ALL-Updated-8.24.20.pdf

59. Center for Outcomes Research & Evaluation. *Yale University-Mayo Clinic CERSI*. Yale School of Medicine. https://medicine.yale.edu/core/current_projects/cersi/research/. Accessed February 2021.

60. Hugo Health. *Hugo*. https://hugo.health/. Accessed February 2021.

61. U.S. Food and Drug Administration. (2018, March 29). *Patient and provider views on clinical endpoints: A qualitative preference study of minimally invasive glaucoma surgical (MIGS) devices*. https://www.fda.gov/science-research/advancing-regulatory-science/patient-and-provider-views-clinical-endpoints-qualitative-preference-study-minimally-invasive

62. The ADVANCE Collaborative. *What we do.* http://advancecollaborative.org/?page_id=149. Accessed February 2021.

63. The ADVANCE Collaborative. *Who we are.* http://advancecollaborative.org/?page_id=147. Accessed February 2021.

64. Patient-Centered Outcomes Research Institute. *Accelerating Data Value Across a National Community Health Center Network (ADVANCE).* https://www.pcori.org/research-results/2015/accelerating-data-value-across-national-community-health-center-network. Accessed February 2021.

65. The ADVANCE Collaborative. *Clinician engagement.* http://advancecollaborative.org/?page_id=163. Accessed February 2021.

66. The ADVANCE Collaborative. *Healthcare Worker Exposure Response & Outcomes (HERO).* http://advancecollaborative.org/?page_id=1717. Accessed February 2021.

67. OCHIN. (2018). *PCORnet obesity observational study | Short and Long-Term Effects of Antibiotics on Childhood Growth*. http://advancecollaborative.org/wp-content/uploads/2017/09/PCORnet-Obesity-Observational-Study-Research-Summary-161116.pdf

68. The ADVANCE Collaborative. *Current ADVANCE projects.* http://advancecollaborative.org/?page_id=155. Accessed February 2021.

69. The ADVANCE Collaborative. *Frequently asked questions.* http://advancecollaborative.org/?page_id=151. Accessed February 2021.

70. The ADVANCE Collaborative. *The ADVANCE update: September 2019.* http://advancecollaborative.org/?p=1571. Accessed February 2021.

71. The National Patient-Centered Clinical Research Network. *Governance.* PCORnet. https://pcornet.org/governance/. Accessed February 2021.

72. U.S. Food and Drug Administration. (2018, September 06). *Centers of Excellence in Regulatory Science and Innovation (CERSIs).* https://www.fda.gov/science-research/advancing-regulatory-science/centers-excellence-regulatory-science-and-innovation-cersis

73. U.S. Food and Drug Administration. (2018, June 05). *FDA Center/Office regulatory science research priority areas for CERSI Program.* https://www.fda.gov/science-research/advancing-regulatory-science/fda-centeroffice-regulatory-science-research-priority-areas-cersi-program

74. U.S. Food and Drug Administration. (2021, January 26). *Source data capture from electronic health records (EHRs): Using standardized clinical research data* (OneSource phase 1). https://www.fda.gov/science-research/advancing-regulatory-science/source-data-capture-electronic-health-records-ehrs-using-standardized-clinical-research-data

75. U.S. Food and Drug Administration. (2019, July 30). *Real-world data to assess variation in opioid prescribing and use for acute pain in diverse populations*. https://www.fda.gov/science-research/advancing-regulatory-science/real-world-data-assess-variation-opioid-prescribing-and-use-acute-pain-diverse-populations

76. U.S. Food and Drug Administration. (2021, January 26). *CERSI research projects.* https://www.fda.gov/science-research/advancing-regulatory-science/cersi-research-projects

77. Johns Hopkins Bloomberg School of Public Health. *Publications.* https://www.jhsph.edu/research/centers-and-institutes/center-of-excellence-in-regulatory-science-and-innovation/publications/. Accessed February 2021.

78. University of California San Francisco. *UCSF-Stanford Center of Excellence in Regulatory Science and Innovation (CERSI).* https://pharm.ucsf.edu/cersi/publications. Accessed February 2021.

79. CERSI, University of Maryland. *Publications*. https://cersi.umd.edu/publications. Accessed February 2021.

80. Center for Outcomes Research & Evaluation. *Yale University-Mayo Clinic CERSI*. Yale School of Medicine. https://medicine.yale.edu/core/current_projects/cersi/press_and_publications_list/. Accessed February 2021.

81. Gadye, Levi. (2020, January 23). *FDA, academia, and industry join forces to shepherd new therapies to patients.* UCSF School of Pharmacy. https://pharmacy.ucsf.edu/news/2020/01/fda-academia-industry-join-forces-shepherd-new-therapies-patients

82. Tune, Lee & Ceraul, Becky. (2012, March 16). *UM-CERSI awards Inaugural Innovation grants to UMD, UMB and FDA collaborators*. CERSI, University of Maryland. https://cersi.umd.edu/news/story/umcersi-awards-inaugural-innovation-grants-to-umd-umb-and-fda-collaborators

83. Kashef, Ziba. (2017, January 31). *FDA grant establishes Yale-Mayo Clinic center to advance regulatory science.* YaleNews. https://news.yale.edu/2017/01/31/fda-grant-establishes-yale-mayo-clinic-center-advance-regulatory-science

84. Johns Hopkins Bloomberg School of Public Health. (2014, May 5). *Johns Hopkins joins the Food and Drug Administration's Centers of Excellence in Regulatory Science and Innovation (CERSI).* https://www.jhsph.edu/news/news-releases/2014/johns-hopkins-joins-the-food-and-drug-administration-centers-of-excellence-in-regulatory-science-and-innovation-cersi.html

85. U.S. Food and Drug Administration. (2014). *Centers of Excellence in Regulatory Science and Innovation (CERSI) established in 2014. About FDA.* https://www.accessdata.fda.gov/scripts/FDATrack/view/track_project.cfm?program=oc-chief-scientist&id=OCS-ORSI-New-CERSI

86. Johns Hopkins Bloomberg School of Public Health. *People.* https://www.jhsph.edu/research/centers-and-institutes/center-of-excellence-in-regulatory-science-and-innovation/people/. Accessed February 2021.

87. University of California San Francisco. *About*. UCSF-Stanford Center of Excellence in Regulatory Science and Innovation (CERSI). https://pharm.ucsf.edu/cersi/about. Accessed February 2021.

88. CERSI, University of Maryland. *About M-CERSI*. https://cersi.umd.edu/about. Accessed February 2021.

89. Center for Outcomes Research & Evaluation. *Yale University-Mayo Clinic CERSI*. Yale School of Medicine. https://medicine.yale.edu/core/current_projects/cersi/leadership_contact/. Accessed February 2021.

90. Weill Cornell Medicine, Population Health Sciences. *Medical Device Epidemiology Network (MDEpiNet).* https://phs.weill.cornell.edu/research-collaboration/research-programs/mdepinet. Accessed February 2021.

91. MDEpiNet. *The Medical Device Epidemiology Network (MDEpiNet).* https://www.mdepinet.net/. Accessed February 2021.

92. MDEpiNet. *Coordinated registry networks.* https://www.mdepinet.net/coordinated-registry-networks. Accessed February 2021.

93. MDEpiNet. *Orthopedic devices.* https://www.mdepinet.net/ortho. Accessed February 2021.

94. MDEpiNet. *UDI adoption.* https://www.mdepinet.net/build. Accessed February 2021.

95. MDEpiNet. *Registry based trials*. https://www.mdepinet.net/uab. Accessed February 2021.

96. MDEpiNet. *Programs*. https://www.mdepinet.net/programs. Accessed February 2021.

97. MDEpiNet. *Publications*. https://www.mdepinet.net/publications. Accessed February 2021.

98. MDEpiNet. *Tools*. https://www.mdepinet.net/tools. Accessed February 2021.

99. Star Metrics. (2013). *Advancing medical device postmarket surveillance infrastructure and epidemiologic.* Federal RePORTER. https://federalreporter.nih.gov/Projects/Details/?projectId=544275&ItemNum=3346&totalItems=3846&searchId=b850241613a74a58962c0bd1a1edd5d4&searchMode=Smart&page=34&pageSize=100&sortField=ProjectNum&sortOrder=desc&filters=$Agency;FDA$ProjectType;p&navigation=True

100. MDEpiNet. *Coordinating center*. https://www.mdepinet.net/coordinating-center. Accessed February 2021.

101. U.S. Food and Drug Administration. *About the Food and Drug Administration (FDA) Sentinel Initiative*. Sentinel. https://www.sentinelinitiative.org/about. Accessed February 2021.

102. U.S. Food and Drug Administration. *How Sentinel gets its data*. Sentinel. https://www.sentinelinitiative.org/about/how-sentinel-gets-its-data. Accessed February 2021.

103. U.S. Food and Drug Administration. *Key database statistics*. Sentinel. https://www.sentinelinitiative.org/about/key-database-statistics. Accessed February 2021.

104. U.S. Food and Drug Administration. *Ongoing ARIA assessments*. Sentinel. https://www.sentinelinitiative.org/assessments/aria-overview/ongoing-aria-assessments. Accessed February 2021.

105. U.S. Food and Drug Administration. *Publications & presentations*. Sentinel. https://www.sentinelinitiative.org/news-events/publications-presentations. Accessed February 2021.

106. U.S. Food and Drug Administration. *Meetings, workshops, & trainings*. Sentinel. https://www.sentinelinitiative.org/news-events/meetings-workshops-trainings. Accessed February 2021.

107. U.S. Food and Drug Administration. (2020, March 09). *FDA's Sentinel Initiative – Background.* https://www.fda.gov/safety/fdas-sentinel-initiative/fdas-sentinel-initiative-background

108. Health Care Systems Research Network. *Sentinel initiative.* http://www.hcsrn.org/en/Collaboration/Consortia/msn.html. Accessed February 2021.

109. U.S. Food and Drug Administration. *Sentinel structure* Sentinel. https://www.sentinelinitiative.org/about/sentinel-structure. Accessed February 2021.

110. U.S. Food and Drug Administration. *Sentinel Operations Center (SOC).* Sentinel. https://www.sentinelinitiative.org/about/sentinel-structure/sentinel-operations-center-soc. Accessed February 2021.

111. U.S. Food and Drug Administration. *Sentinel Innovation Center (IC).* Sentinel. https://www.sentinelinitiative.org/about/sentinel-structure/sentinel-innovation-center-ic. Accessed February 2021.

112. U.S. Food and Drug Administration. *Community Building and Outreach Center (CBOC)* Sentinel. https://www.sentinelinitiative.org/about/sentinel-structure/community-building-and-outreach-center-cboc. Accessed February 2021.

113. All of Us Research Program, National Institutes of Health. *All of Us Research Program overview*. https://allofus.nih.gov/about/all-us-research-program-overview. Accessed February 2021.

114. All of Us Research Program, National Institutes of Health. *About.* https://allofus.nih.gov/about. Accessed February 2021.

115. All of Us Research Program, National Institutes of Health. *Data snapshots.* All of Us Research Hub. https://www.researchallofus.org/data-tools/data-snapshots/. Accessed February 2021.

116. All of Us Research Program, National Institutes of Health. *Institutional agreements.* All of Us Research Hub. https://www.researchallofus.org/institutional-agreements/. Accessed February 2021.

117. All of Us Research Program, National Institutes of Health. *Data access and use.* All of Us Research Hub. https://www.researchallofus.org/data-tools/data-access/. Accessed February 2021.

118. All of Us Research Program, National Institutes of Health. *All of Us publications.* All of Us Research Hub. https://www.researchallofus.org/publications/. Accessed February 2021.

119. All of Us Research Program, National Institutes of Health. *Research project directory.* All of Us Research Hub. https://www.researchallofus.org/research-projects-directory/?section=abc. Accessed February 2021.

120. All of Us Research Program, National Institutes of Health. (2018). *All of Us Research Program operational protocol*. https://allofus.nih.gov/sites/default/files/aou_operational_protocol_v1.7_mar_2018.pdf

121. All of Us Research Program, National Institutes of Health. *Survey explorer.* All of Us Research Hub. https://www.researchallofus.org/data-tools/survey-explorer/. Accessed February 2021.

122. All of Us Research Program, National Institutes of Health. *NIH All of Us research program staff.* https://allofus.nih.gov/about/who-we-are/nih-all-us-research-program-staff. Accessed February 2021.

123. All of Us Research Program, National Institutes of Health. *Data and research center.* https://allofus.nih.gov/funding-and-program-partners/data-and-research-center. Accessed February 2021.

124. National Institutes of Health. *About NIH Collaboratory*. NIH Collaboratory Living Textbook of Pragmatic Clinical Trials. https://rethinkingclinicaltrials.org/about-nih-collaboratory/. Accessed February 2021.

125. National Institutes of Health. *Core working groups*. NIH Collaboratory Living Textbook of Pragmatic Clinical Trials. https://rethinkingclinicaltrials.org/cores-and-working-groups/. Accessed February 2021.

126. U.S. Food and Drug Administration. (2017, January 03). *Collection of patient-provided information through a mobile device application for use in comparative effectiveness and drug safety research*. Sentinel. https://www.sentinelinitiative.org/methods-data-tools/fda-catalyst-projects/collection-patient-provided-information-through-mobile

127. U.S. Food and Drug Administration. (2019, January 13). *Mother-infant linkage table*. Sentinel. https://www.sentinelinitiative.org/methods-data-tools/sentinel-common-data-model/mother-infant-linkage-table

128. National Institutes of Health. *Data and resource sharing.* NIH Collaboratory Living Textbook of Pragmatic Clinical Trials. https://rethinkingclinicaltrials.org/data-and-resource-sharing/. Accessed February 2021.

129. Meyers, Catherine. (2017, November 03). *The NIH Collaboratory launches a new resource on methods and best practices for pragmatic clinical trials*. NCCIH Research Blog. https://www.nccih.nih.gov/research/blog/the-nih-collaboratory-launches-a-new-resource-on-methods-and-best-practices-for-pragmatic-clinical-trials

130. Agency for Healthcare Research and Quality. *About PBRNs.* Practice-Based Research Networks. https://pbrn.ahrq.gov/about. Accessed February 2021.

131. Agency for Healthcare Research and Quality. *PBRN funding history.* Practice-Based Research Networks. https://pbrn.ahrq.gov/about/pbrn-funding-history. Accessed February 2021.

132. Agency for Healthcare Research and Quality. *Practice-Based Research Networks.* https://pbrn.ahrq.gov/. Accessed February 2021.

133. Agency for Healthcare Research and Quality. *Centers for primary care practice-based research and learning: Nurture partnerships, method development, and trans-network collaborations.* Practice-Based Research Networks. https://pbrn.ahrq.gov/pbrn-profiles/p30-centers. Accessed February 2021.

134. Agency for Healthcare Research and Quality. *N^2: Building a network of safety net PBRNs.* Practice-Based Research Networks. https://pbrn.ahrq.gov/sites/default/files/docs/page/N2.pdf. Accessed February 2021.

135. Patient-Centered Outcomes Research Institute. (2020). *Comparing two programs to increase cancer screening for women with depression.* https://www.pcori.org/research-results/2013/comparing-two-programs-increase-cancer-screening-women-depression

136. Agency for Healthcare Research and Quality. *MOSAIC: Meaningful Outcomes and Science to Advance Innovations Center of Excellence.* Practice-Based Research Networks. https://pbrn.ahrq.gov/sites/default/files/docs/page/MOSAIC.pdf. Accessed February 2021.

137. Agency for Healthcare Research and Quality. *Care coordination enabled by health information technology: What will it take? (District of Columbia).* Digital Healthcare Research. https://digital.ahrq.gov/ahrq-funded-projects/care-coordination-enabled-health-information-technology-what-will-it-take. Accessed February 2021.

138. Agency for Healthcare Research and Quality. *National Center for Pediatric Practice-Based Research and Learning (C-PRL).* Practice-Based Research Networks. https://pbrn.ahrq.gov/sites/default/files/docs/page/C-PRL.PDF. Accessed February 2021.

139. Agency for Healthcare Research and Quality. *History of PBRNs.* Practice-Based Research Networks. https://pbrn.ahrq.gov/about/history-pbrns. Accessed February 2021.

140. Agency for Healthcare Research and Quality. (2013). *Primary care practice-based research networks (PBRN).* https://www.ahrq.gov/professionals/prevention-chronic-care/improve/pbrn/index.html

141. North American Primary Care Research Group (NAPCRG). (2014). *Practice-based Research Network Research Good Practices*. https://www.napcrg.org/media/1196/prgp-2014-09-29.pdf

142. Patient-Centered Outcomes Research Institute. (2020). *PCORnet, The National Patient-Centered Clinical Research Network.* https://www.pcori.org/sites/default/files/PCORI-PCORnet-Fact-Sheet.pdf

143. The National Patient-Centered Clinical Research Network. *Network.* PCORnet. https://pcornet.org/network/. Accessed February 2021.

144. The National Patient-Centered Clinical Research Network. *Research.* PCORnet. https://pcornet.org/research/. Accessed February 2021.

145. Duke University. (2020). *Aspirin dosing: A patient-centric trial assessing benefits and long-term (ADAPTABLE).* ClinicalTrials.gov. https://clinicaltrials.gov/ct2/show/NCT02697916

146. Novartis Pharmaceuticals. (2019). *Patient reported outcomes investigation following initiation of drug therapy with Entresto (Sacubitril/Valsartan) in heart failure (PROVIDE-HF).* ClinicalTrials.gov. https://clinicaltrials.gov/ct2/show/NCT03387163

147. The National Patient-Centered Clinical Research Network. *Front door.* PCORnet. https://pcornet.org/front-door/. Accessed February 2021.

148. Fleurence, R. L., Curtis, L. H., Califf, R. M., Platt, R., Selby, J. V., & Brown, J. S. (2014). Launching PCORnet, a national patient-centered clinical research network. *Journal of the American Medical Informatics Association*. https://academic.oup.com/jamia/article/21/4/578/2909226

149. ASCO CancerLinQ. *About us*. https://www.cancerlinq.org/about. Accessed February 2021.

150. ASCO CancerLinQ. *CancerLinQ service menu*. https://s3.amazonaws.com/files.cancerlinq.org/prod/s3fs-public/2019-08/CANCERLINQ%20-%20Services%20Menu_08222019.pdf. Accessed February 2021.

151. ASCO CancerLinQ. *Oncology practices*. https://www.cancerlinq.org/solutions/oncology-practices. Accessed February 2021.

152. ASCO CancerLinQ. *Researchers*. https://www.cancerlinq.org/solutions/researchers. Accessed February 2021.

153. ASCO CancerLinQ. (2020). *CancerLinQ Discovery scientific evidence.* https://s3.amazonaws.com/files.cancerlinq.org/prod/s3fs-public/2020-05/CLQ_ScientificEvidenceDeck_Discovery.pdf

154. National Center for Advancing Translational Sciences, National Institutes of Health. (2019, September 18). *About the CTSA program*. https://ncats.nih.gov/ctsa/about

155. National Center for Advancing Translational Sciences, National Institutes of Health. (2021, January 12). *CTSA program hubs*. https://ncats.nih.gov/ctsa/about/hubs

156. SMART IRB. *About us*. https://smartirb.org/about-us/. Accessed February 2021.

157. National Center for Advancing Translational Sciences, National Institutes of Health. (2019, May 05). *Common Metrics Initiative*. https://ncats.nih.gov/ctsa/projects/common-metrics

158. Daudelin, D.H., Peterson, L.E., Welch, L.C., Chandler, R., Pandley, M., Noubary, F., Lee, P.L., & Selker, H.P. (2020). Implementing common metrics across the NIH Clinical and Translational Science Awards (CTSA) consortium. *Journal of clinical and translational science*, 4(1), pp.16-21.

159. National Center for Advancing Translational Sciences, National Institutes of Health. (2021, January 12). *Trial Innovation Network*. https://ncats.nih.gov/ctsa/projects/network

160. Liverman, C.T., Schultz, A.M., Terry, S.F., & Leshner, A.I. (2013). The CTSA program at NIH: Opportunities for advancing clinical and translational research. *National Center for Biotechnology Information.* https://www.ncbi.nlm.nih.gov/books/NBK169198/

161. Clinical & Translational Science Awards Program. *About*. National Center for Advancing Translational Sciences. https://ctsa.ncats.nih.gov/. Accessed February 2021.

162. Clinical & Translational Science Awards Program. *Initiatives*. National Center for Advancing Translational Sciences. https://ctsa.ncats.nih.gov/initiatives/. Accessed February 2021.

163. Electronic Medical Record Support for Public Health. *What is it?* https://www.esphealth.org/. Accessed February 2021.

164. Electronic Medical Record Support for Public Health. *Features.* https://www.esphealth.org/features. Accessed February 2021.

165. International Aging Research Portfolio. *Enhancing public health through electronic medical and personal health rec*. http://agingportfolio.org/projects/project/3P01HK000016-03S1. Accessed February 2021.

166. Electronic Medical Record Support for Public Health. *About us.* https://www.esphealth.org/contact. Accessed February 2021.

167. Department of Population Medicine. *Therapeutics research and infectious disease faculty/staff.* Harvard Pilgrim Health Care Institute. https://www.populationmedicine.org/research/therapeutics-research-infectious-disease/staff. Accessed February 2021.

168. Health Care Systems Research Network. *Mission and vision.* http://www.hcsrn.org/en/About/. Accessed February 2021.

169. Health Care Systems Research Network. *Membership.* http://www.hcsrn.org/en/About/Membership/. Accessed February 2021.

170. Health Care Systems Research Network. *Collaboration with HCSRN.* http://www.hcsrn.org/en/Collaboration/. Accessed February 2021.

171. Health Care Systems Research Network. *AGING initiative.* http://www.hcsrn.org/en/Collaboration/Consortia/AGING_Initiative.html. Accessed February 2021.

172. Health Care Systems Research Network. *SUPREME DM.* http://www.hcsrn.org/en/Collaboration/Consortia/dm.html. Accessed February 2021.

173. Health Care Systems Research Network. *IRB review.* http://www.hcsrn.org/en/Resources/IRB/. Accessed February 2021.

174. Health Care Systems Research Network. *Leadership & organization.* http://www.hcsrn.org/en/About/Leadership/. Accessed February 2021.

175. Greene, S. (2019, February). *The Health Care Systems Research Network: A closer look at the capabilities, resources, and benefits that comprise our research infrastructure* [PowerPoint slides]. The Health Care Systems Research Network. http://www.hcsrn.org///asset/1e97f45d-5996-4e3a-9f8d-846d532eba2e/Closer%20Look%20at%20the%20HCSRN%20Capabilties%20-%20SDRF%20February%202019.pdf

176. NESTcc. *About us.* https://nestcc.org/about/about-us/. Accessed February 2021.

177. NESTcc. *Research network*. https://nestcc.org/research-network/. Accessed February 2021.

178. NESTcc. *Post-market medical device surveillance with a novel mHealth platform*. https://nestcc.org/portfolio-item/post-market-medical-device-surveillance-with-a-novel-mhealth-platform/. Accessed February 2021.

179. NESTcc. *ICD registry DELTA active surveillance pilot study.* https://nestcc.org/portfolio-item/icd-registry-delta-active-surveillance-pilot-study/. Accessed February 2021.

180. Grantome. *Active surveillance of cardiovascular devices: The multi-registry DELTA network.* https://grantome.com/grant/NIH/U01-FD004963-02. Accessed February 2021.

181. NESTcc. *Test-Cases.* https://nestcc.org/test-cases/#toggle-id-4-closed. Accessed February 2021.

182. NESTcc. *Team.* https://nestcc.org/about/nestcc-team/. Accessed February 2021.

183. Observational Health Data Sciences and Informatics. *OHDSI.* https://ohdsi.org/. Accessed February 2021.

184. Observational Health Data Sciences and Informatics. *Mission, vision & values.* https://ohdsi.org/who-we-are/mission-vision-values/. Accessed February 2021.

185. Observational Health Data Sciences and Informatics. *Collaborators.* https://www.ohdsi.org/who-we-are/collaborators/. Accessed February 2021.

186. Observational Health Data Sciences and Informatics. *The Book of OHDSI: Chapter 20: OHDSI as a research network.* https://ohdsi.github.io/TheBookOfOhdsi/NetworkResearch.html#ohdsi-as-a-research-network. Accessed February 2021.

187. Observational Health Data Sciences and Informatics. *OHDSI news and updates.* https://www.ohdsi.org/ohdsi-news-updates/. Accessed February 2021.

188. Lane, J.C., Weaver, J., Kostka, K., Duarte-Salles, T., Abrahao, M.T.F., Alghoul, H., Alser, O., Alshammari, T.M., Biedermann, P., Banda, J.M. & Burn, E. (2020). Risk of hydroxychloroquine alone and in combination with azithromycin in the treatment of rheumatoid arthritis: a multinational, retrospective study. *The Lancet Rheumatology*, *2*(11), pp.e698-e711.

189. Observational Health Data Sciences and Informatics. *Recommended diuretic causes more side effects than similar hypertension drug, per recent LEGEND study.* https://www.ohdsi.org/ohdsi-news-updates/chlorthalidone-side-effects-study/. Accessed February 2021.

190. Observational Health Data Sciences and Informatics. *Data standardization.* https://www.ohdsi.org/data-standardization/. Accessed February 2021.

191. Observational Health Data Sciences and Informatics. *The Book of OHDSI.* https://ohdsi.github.io/TheBookOfOhdsi/. Accessed February 2021.

192. Observational Health Data Sciences and Informatics. *Frequently asked questions.* https://ohdsi.org/who-we-are/frequently-asked-questions/. Accessed February 2021.

193. Observational Health Data Sciences and Informatics. *Support for OHDSI* https://ohdsi.org/who-we-are/support-for-ohdsi/. Accessed February 2021.

194. PEDSnet. *About*. https://pedsnet.org/about/. Accessed February 2021.

195. PEDSnet. *Our vision*. https://pedsnet.org/about/our-vision/. Accessed February 2021.

196. PEDSnet. *Publications*. https://pedsnet.org/research/publications/. Accessed February 2021.

197. Lang, J.E., Bunnell, H.T., Lima, J.J., Hossain, M.J., Wysocki, T., Bacharier, L., Dempsey, A., Ulrich, L., Test, M.R. & Forrest, C.B. (2019). Effects of age, sex, race/ethnicity, and allergy status in obesity-related pediatric asthma. *Pediatric pulmonology*, *54*(11), pp.1684-1693.

198. PEDSnet. *About the PKIDS study*. https://pedsnet.org/pkids/what-pkids/. Accessed February 2021.

199. Children's Hospital of Philadelphia. (2021). *Pediatric kidney stone (PKIDS) care improvement network (PKIDS).* ClinicalTrials.gov. https://clinicaltrials.gov/ct2/show/NCT04285658?term=NCT04285658&draw=2&rank=1

200. Patient-Centered Outcomes Research Institute. (2019, November). *PEDSnet.* https://www.pcori.org/research-results/2015/pedsnet

201. Health Care Systems Research Network. *Data resources.* http://www.hcsrn.org/en/About/Data/. Accessed February 2021.

202. Assistant Secretary for Planning and Evaluation. *Patient-Centered Outcomes Research Trust Fund reports.* https://aspe.hhs.gov/patient-centered-outcomes-research-trust-fund-reports. Accessed February 2021.

203. 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