

Expanding Social Determinants of Health Data across PCORnet®

Current State and Future Opportunities

November 4, 2021



Opening Remarks

Prashila Dullabh

Agenda

01 Opening Remarks

02 PCORnet Overview

03 Current Landscape of SDOH Data Collection

04 Opportunities for Leveraging Existing
SDOH Data

05 Opportunities for Incorporating Social Sector
Data

06 Opportunities for Optimizing Individual-
and Population-Level Analysis

07 Final Thoughts and Q&A



PCORI is exploring opportunities to enhance PCORnet data infrastructure

- Increasing the availability and accessibility of high value data for research
- Identifying short-term, mid-term, and long-term infrastructure needs
- Topic areas:



**Social
Determinants of
Health**



**Patient Reported
Outcomes &
Patient-Generated
Health Data**



**Data Linkage
Capacity to CMS
Claims Data**

SDOH Workstream

Informed by Previous Work

- PCORI SDOH Work Group recommendations
- 2018 Survey

Current Work



SDOH Webinar Objectives:



Review findings
on the
landscape of
SDOH data
collection across
PCORnet



Discuss **strategies**
to increase
SDOH data
availability and
accessibility
across PCORnet



Discuss and
prioritize
opportunities
for data
infrastructure
expansion
across PCORnet

Webinar Speakers

Opportunities for
Leveraging Existing
SDOH Data



Keith Marsolo
Co-Investigator
PCORnet Coordinating Center



Jon Puro
Principal Investigator
ADVANCE

Opportunities for
Incorporating Social
Sector Data



Jacob Reider
President, Healthy
Alliance IPA
Alliance for Better Health



Thomas Carton
Principal Investigator
REACHnet

Opportunities for
Optimizing Individual- and
Population-Level Analysis



Betsy Shenkman
Principal Investigator
OneFlorida



Erika Cottrell
Investigator
OCHIN

PCORnet®, The National Patient-Centered Clinical Research Network

Nik Koscielniak, PhD, MPH
Program Officer
Research Infrastructure



- **Funded by PCORI in 2013**
- **Faster, more efficient**
 - Facilitate broad conduct of comparative clinical effectiveness research (CER) of national importance
- **National resource**
 - Accessible via central gateway to researchers within and outside of the Network
- **Harness health data and engage patients and communities**
 - Electronic health records, claims data, patient-reported outcomes, etc.
- **Nationally representative patient populations**
 - Support a wide range of research activities

PCORI's Prioritizing Principles for PCORnet Infrastructure



Patient-Centered

- I. Strengthen the central role of patient and caregiver engagement to produce true partnership in the full research process
- II. Attend to and strengthen the diversity of the patient populations served and the inclusion of care sites that serve underserved populations

National Scope

- III. Recognize and strengthen the unique ability of PCORnet to generate definitive evidence through studies of national scope; prioritize investments that strengthen that capability particularly for studies focused on PCORI *Strategic Research Priorities*
- IV. Build on the unique capabilities of the PCORnet data structures, prioritizing investments that will align with the PCORI *Strategic Research Priorities*

Governance and Partnerships

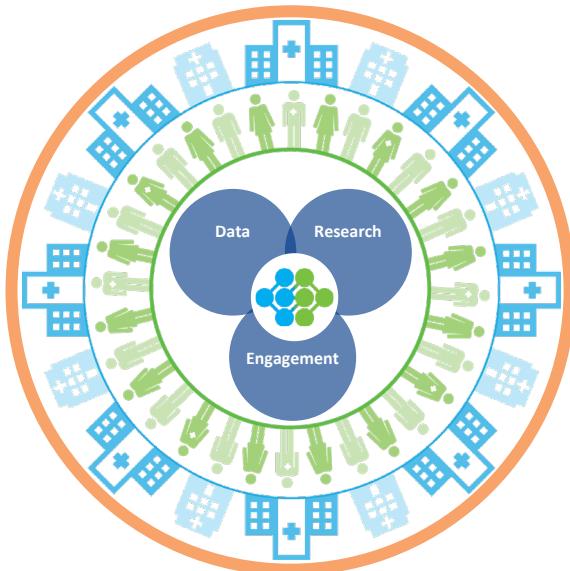
- V. Continue to develop the shared governance model which builds trust among participating networks and is coupled with close milestone-based monitoring of PCORI's infrastructure funds. As appropriate, encourage the development of mechanisms for cost recovery for infrastructure use related to projects supported by other funders
- VI. Recognize the importance of optimum distribution of core coordinating functions, with the dual goals of optimizing effective network function and capitalizing on the distributed expertise and capabilities of this complex network
- VII. Recognize, enable and promote the value of PCORnet to contribute to a learning health care system through effective partnerships with all stakeholders. Strengthen the educational resources needed to achieve this goal
- VIII. Recognize, enable and optimize the value of PCORnet for building partnerships with Federal Health Agencies

PCORnet is a “network of networks” that harnesses the power of partnerships



Trust is built. Answers are accelerated. Health improves.

Connecting
people and data



to make informed
decisions together

**Clinical
Research
Networks
(CRNs)**



**Patient
Partners**



**Coordinating
Center**



**A national
infrastructure for
patient-centered
clinical research**

PCORnet Common Data Model

Data are Research Ready



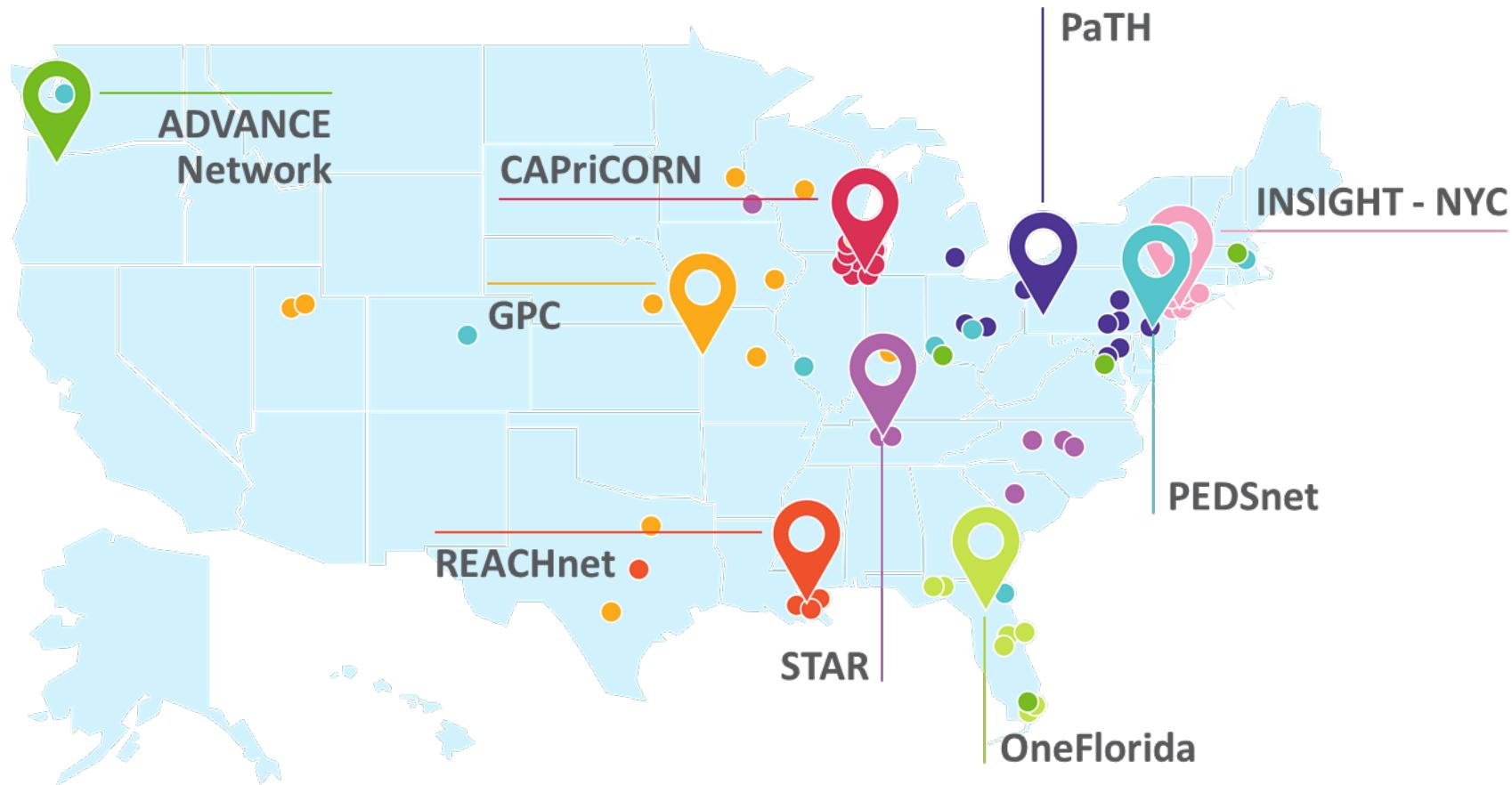
- Lots of data is great, but for it to be useful it has to be standardized across systems. The PCORnet Common Data Model standardizes data into a single language, enabling fast insights, including:

Ready for Research				Available, But Still Evolving			
Death Data	Diagnoses	Medication Orders		Geocodes	Social Determinants of Health	Tumor Registry	Biosamples
Claims	Labs	Demographics	Procedures	Patient-Reported Outcomes	Genomic Results	Natural Language Processing Derived Concepts	Patient-Generated Data

Data available from several Clinical Research Networks, in the PCORnet Common Data Model and ready for use in research.

Data available at some Clinical Research Networks, may or may not be in the PCORnet Common Data Model and require additional work for use in research.

Nine PCORnet CRNs can help researchers conduct research more efficiently. Researchers can access data from everyday medical encounters from more than 60 million people across the United States.



PCORnet Front Door

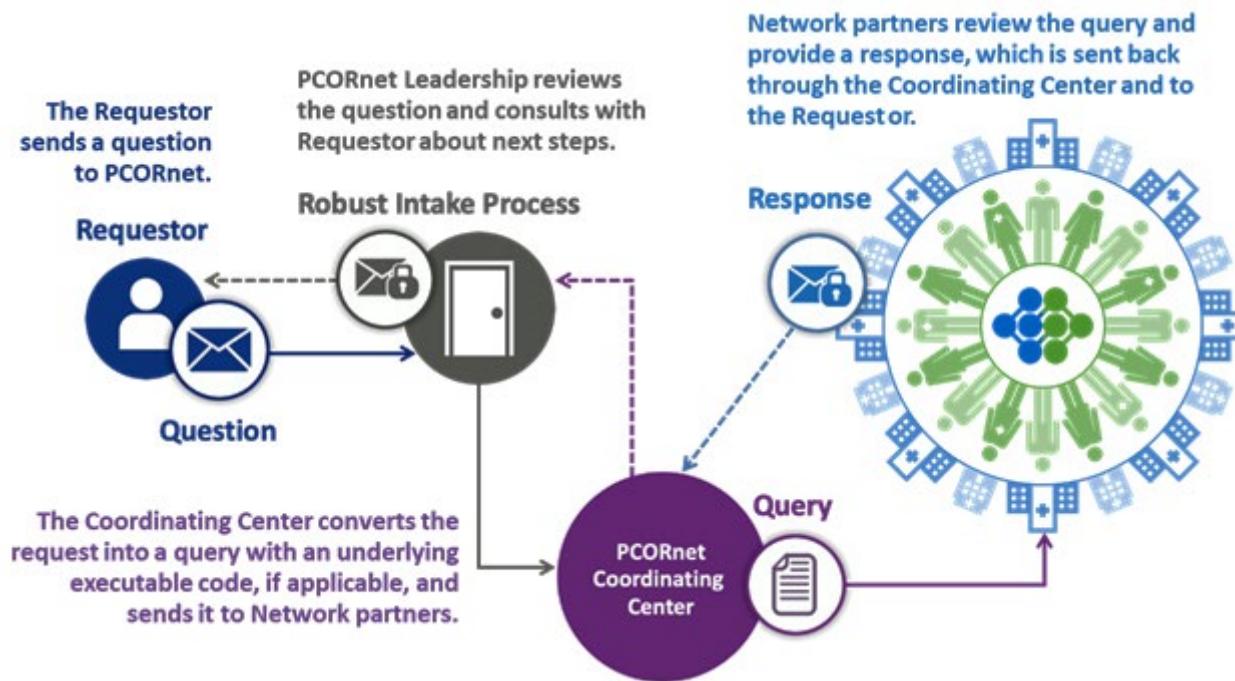


- Access point for investigators, patient groups, healthcare organizations, clinicians, government, industry, sponsors, and all stakeholders seeking to leverage the PCORnet infrastructure and collaborate on patient-centered research

- **Data Network Requests**
- **Network Collaborator Requests**
- **Study Feasibility Review**



- Data curation processes to ensure data are research ready
- Query process is performed locally at the network site
- The data remains at the network site behind institutional firewalls, maintaining security
- Quarterly data refreshes

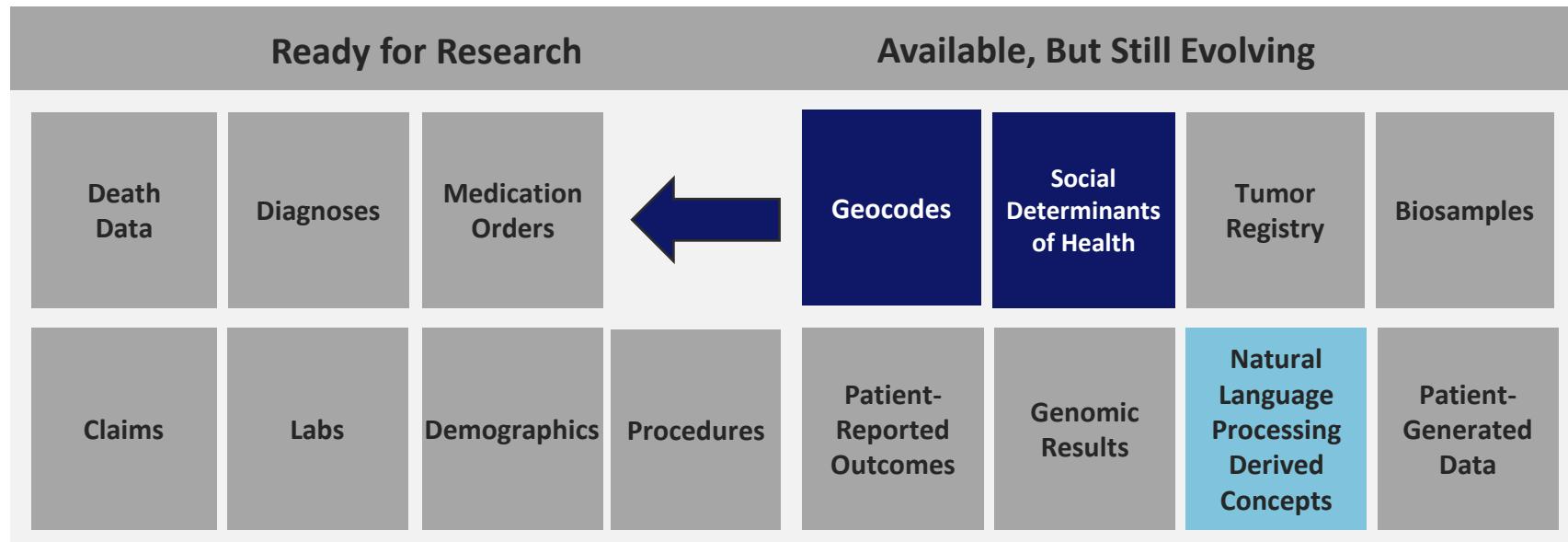


PCORnet Common Data Model

SDOH Data that are Research Ready



- Enhance the PCORnet CDM to make SDOH data and geocoding research ready



Data available from several Clinical Research Networks, in the PCORnet Common Data Model and ready for use in research.

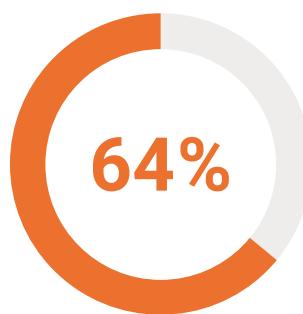
Data available at some Clinical Research Networks, may or may not be in the PCORnet Common Data Model and require additional work for use in research.

Current Landscape of SDOH Data Use and Collection across PCORnet

Lauren Hovey

Health systems are collecting some form of individual-level SDOH data (n=55)

Food Security



35 health systems collect data on food security

Transportation



31 health systems collect data on transportation

Housing security/stability



29 health systems collect data on housing security or stability

Financial resource strain



29 health systems collect data on financial resource strain

Gaps in SDOH data collection exist across participating health systems

40%

**Health Systems
Using Standard
Terminology
to Collect
SDOH Data**

(22 health systems)

53%

**Health Systems
Using an SDOH
Data Collection
Framework**

(29 health systems)

20%

**Health Systems
Have SDOH Data in
the EHR for >80%
of Patients**

(11 health systems)

Address data is broadly collected across PCORnet, with some limitations:

Address-level data



50 health systems collect a combination of street address and/or 5-digit ZIP code



- The data are limited for 9-digit ZIP code

Key opportunities for SDOH data capture:

Short-Term	<ul style="list-style-type: none">• Incorporate SDOH-related data elements into the CDM• Improve address data and increase geocoding• Support linkages between PCORnet data resources and publicly available data• Establish partnerships with community-based organizations (CBOs)
Mid-Term	<ul style="list-style-type: none">• Use natural language processing to mine unstructured individual-level SDOH data
Long-Term	<ul style="list-style-type: none">• Incorporate social sector data, informed by data standards and platforms already in use

Opportunities for Leveraging Existing SDOH Data across PCORnet

Jon Puro & Keith Marsolo

Prioritizing SDOH domains and data elements to be added to the CDM



- Heterogeneous SDOH data collected by health systems across the Network
- Varying SDOH standardization at the health systems level
- Consensus-based SDOH standards under development:
 - Food insecurity, housing instability, transportation access, financial hardship, employment status, veteran status, intimate partner violence, social isolation

Prioritizing SDOH domains and data elements to be added to the CDM



Opportunity

- Incorporate data elements collected by PCORnet data partners into the Common Data Model:
 - Language, education, food security, employment status, transportation, housing security, financial resource strain



Questions for Discussants

1. Where do you see near-term opportunities to expand the CDM to include other SDOH domains?
2. What opportunities are there for the PCORnet Coordinating Center to better facilitate loading of SDOH data into the CDM?
3. What technical considerations should inform PCORI's planning as it looks to expand the use of SDOH data across the Network?



Questions for Audience

1. Where are the most promising opportunities to expand the CDM given existing SDOH data across the Network?
2. How can PCORI strengthen the Network's existing SDOH data so that it is "research-ready"?

Standardizing and extracting additional SDOH data collected through EHRs



Landscape

- Limited use of standard terminologies for SDOH data
- SDOH data captured in free text fields of EHRs
- Some CRNs are using natural language processing to mine unstructured fields



Opportunity

- Use natural language processing to leverage unstructured data



Questions for Discussants

1. What has been your experience with using natural language processing?
2. What is the viability of natural language processing for the Network and what would be needed to scale this method?
3. What other approaches could be used to derive value from unstructured data?



Questions for Audience

1. What challenges and opportunities do you see for the Network in extracting SDOH data using natural language processing?
2. What do you think is needed to scale natural language processing methods across the Network?



Poll Question

**Incorporating SDOH elements into the CDM would increase the availability of SDOH data to researchers using the Network.
Which domains would you prioritize for inclusion?**

- Food insecurity, housing instability, transportation access, financial hardship, employment status, veteran status, intimate partner violence, social isolation

Opportunities for Incorporating Other Sources of Social Sector Data across PCORnet

Thomas Carton & Jacob Reider

Expanding SDOH data collection beyond PCORnet Network Partners



Landscape

- Limited data infrastructure to connect health systems with community-based organization (CBO) data
- Social sector data are an untapped source of SDOH data for PCORnet
- Health systems are interested in using social sector data



Opportunity

- Link CBO data to data from PCORnet by leveraging existing platforms
- Connect to social sector data via other mechanisms



Questions for Discussants

1. What opportunities exist to connect Network Partners with community resource referral platforms?
2. What do you see as the opportunities for accessing and/or linking to CBO data?
3. What kind of infrastructure would be needed?
4. What kind of social sector data would be valuable to augment PCORnet data?



Questions for Audience

1. What are the most promising types of CBO data (e.g., housing services, food pantry referrals) to integrate into the PCORnet CDM?
2. Where do you see the most promising opportunities for PCORnet in linking to social sector data? Why?



Poll Question

Please indicate which social sector data you see as having the greatest potential to enhance PCORnet's SDOH research:

- a. Data sets from social service agencies (e.g., HUD)
- b. Data from community-based organizations within a given geographical region
- c. Data from commercial platforms (e.g., UniteUs, Signify)
- d. Data from Health Information Exchange Organizations (HIOs) or Community Information Exchanges (CIEs)
- e. Other (please specify)

Opportunities for Optimizing Individual-Level and Population-Level Analysis

Betsy Shenkman & Erika Cottrell

Improving address data to support geocoding and linkages



Landscape

- 91 percent of the health systems are collecting patient street address with 5-digit zip code (NORC 2021 survey)
- Geocoding individual-level data supports population-level analysis
- Open source and commercial products are available for geocoding



Opportunity

- Deploy mechanisms to improve address data and geocode addresses to support SDOH research



Questions for Discussants

How do we get to a geocoded address and/or 9-digit ZIP code
Network-wide?

1. What tools to verify/correct and geocode patient address are already in use within the Network?
2. Are these tools scalable for broader use—why or why not?
3. Where would they fit best within the PCORnet architecture?
4. What would be the pros and cons of the open source and commercial solutions when deployed across PCORnet?



Question for Audience

What other strategies could be employed to improve address information and geocode within PCORnet?

Facilitating Data Linkages Using Geocodes



Landscape

- Linking individual-level data within PCORnet to publicly available population-level data can enhance its research utility
- A key consideration for address-based geocoding and linkages is that sources of aggregate SDOH data vary in their geographical specificity
 - Depending on the research use case, different levels of geographical specificity may be desired



Opportunity

- Support linkages between PCORnet data resources and publicly available data sources



Questions for Discussants

1. What infrastructure is needed to facilitate linkages between population-level data and PCORnet data resources?
2. What other high value data sets would enhance PCORnet's individual-level data?



Question for Audience

What high value, population-level data sets would enhance PCORnet's SDOH research capacity? For example:

- CDC/ATSDR's Social Vulnerability Index [SVI]
- American Community Survey data
- National Death Index and state-level vital statistics
- Medicaid and Medicare claims data
- Environmental data



Poll Question

Of the opportunities discussed, which would you prioritize as a more immediate need?

- a. Deploy mechanisms to improve address data
- b. Geocode addresses to support SDOH research
- c. Support linkages between PCORnet data resources and publicly available data sources

Final Thoughts and Q&A

Prashila Dullabh & Lauren Hovey

Thank you.



Research You Can Trust™

Nik Koscielniak
Program Officer
nkoscieldniak@pcori.org

Prashila Dullabh
Vice President & Senior Fellow
Dullabh-Prashila@norc.org