NORC at the University of Chicago

WHITE PAPER

OCTOBER 2021

Expanding Social Determinants of Health Data across PCORnet®

Executive Summary

Presented by:
NORC at the University
of Chicago

Presented to:
Patient-Centered Outcomes
Research Institute (PCORI)

Executive Summary

The Patient-Centered Outcomes Research Institute (PCORI®) is undertaking a multi-step assessment process as it considers data infrastructure enhancements to PCORnet®, the National Patient-Centered Clinical Research Network. PCORnet ("the Network") is a distributed research network with a national scope, including Clinical Research Networks (CRNs) that work together as Network Partners and numerous participating health systems that share data within the CRN and across PCORnet. These partnerships allow researchers to query millions of clinical, claims, and registry records to support efficient, high-impact comparative clinical effectiveness research (CER) within a secure, privacy preserving structure.

One of the resources PCORnet uses to enhance the usability and "research readiness" of Network data is the PCORnet common data model (CDM), which standardizes data from the numerous Network Partners into a consistent format. PCORI is considering enhancements to the CDM, as well as other data infrastructure improvements that would increase data availability and accessibility across PCORnet in the short-term, mid-term, and long-term.

PCORI has identified three areas of interest for data infrastructure enhancements:1) social determinants of health (SDOH); 2) patient-report outcomes (PROs) and other patient-generated health data (PGHD); and 3) Centers for Medicare & Medicaid (CMS) claims data.

Under a contract from PCORI, NORC at the University of Chicago is conducting an assessment of opportunities in these three areas, beginning with SDOH. This paper articulates NORC's findings on the current state of SDOH data collection within PCORnet, challenges for expanding PCORnet infrastructure to increase the availability and accessibility of SDOH data within PCORnet, and opportunities for data infrastructure expansion. These findings are based on a survey of health systems participating in PCORnet; a literature review; and key informant interviews (KIIs) with stakeholders associated with PCORI, PCORnet, and SDOH research, and those with expertise in data infrastructure.

Current State. The survey provided insights into the current landscape of SDOH data collection and use among 55 of the 75 health systems participating in PCORnet.

- Nearly all health systems that completed the survey reported that they are collecting and using individual-level SDOH data. However, significant variation exists among the health systems regarding the SDOH domains being collected, the use of standardized terminology to capture SDOH in electronic health records (EHRs), and what data is available to the Network for querying.
- Most health systems report having detailed address data, consisting of street address and 5-digit zip code, but few have the 9-digit zip code available.
- Over half of the health systems who completed the survey report having plans for future SDOH data collection and use, including collecting more SDOH data elements, improving existing screening tools, and implementing new screening tools.
- Several CRNs are using the available data within their networks for robust health disparities research. Their efforts highlight the potential for SDOH research across

PCORnet and potential data infrastructure improvements that would increase Network-wide capacity.

Opportunities. Based on the current state of the PCORnet data infrastructure and existing SDOH data within the Network, challenges and opportunities for expansion and investment emerged in three areas:

- *Opportunities for Leveraging Existing SDOH Data across PCORnet*
  - Incorporating SDOH data elements that can be coded electronically using approved, interoperable standards into the PCORnet Common Data Model: Currently, partners can load demographic information and a limited set of other SDOH-related variables (e.g., language preference) into CDM tables. Data elements could be added to the CDM for SDOH-specific domains that would expand the network's capacity for SDOH research. In addition to prioritizing data elements with existing standards (e.g., food insecurity, housing instability, and transportation), the changes should reflect SDOH data elements that are already being collected by health systems, such as financial hardship, employment status, and intimate partner violence. (Short-term)
  - Leveraging natural language processing to expand SDOH availability: Historically, social needs data are often recorded in unstructured fields rather than discretely captured in structured formats using standards terminologies. Techniques like natural language processing could allow researchers to access SDOH data present in unstructured clinical notes without placing additional data collection burden on the health systems and their staff. (Mid-term)
- *Optimizing Individual-Level and Population-Level Analysis*
  - Supporting address verification and geocoding: The survey found that 91 percent of the health systems are collecting patient street address with 5-digit zip code, but many of the addresses are unverified. Multiple solutions exist to verify addresses, generate zip codes up to 9 digits, and geocode them. Improving the quality of existing address data across the Network and geocoding these data would expand the information available to researchers for individual and population-level analysis of health outcomes and disparities. (Short-term)
  - Enhancing access to research-ready data through linkages to other data sources: The use of aggregate data alone assumes that the structural and neighborhood-level experiences apply across all individuals in a given neighborhood. Geocoding and then linking to community-level data (i.e., data at the county level and smaller) can overcome these biases. Geocoding enables linkages to publicly available surveys (e.g., the American Community Survey [ACS]) and other data sources rich in SDOH and social risk data. (Short-term)
- *Incorporating Other Sources of Social Sector Data across PCORnet*
  - Establishing relationships is critical to engaging with the social sector: Research data access is one piece of a puzzle that must also include benefits to patient health and care and referrals to social services. Through relationship building, the health and social

sectors will identify shared goals for the use of SDOH data and the right mechanisms for accomplishing them. (Short-term)

- Seeking technical solutions to connect to social sector data: At present, there is limited but rapidly developing infrastructure to connect health systems to community-based organizations (CBOs) and other sources of social sector data (e.g., community resource referral platforms). As health systems consider strategies to understand and mitigate the negative effects of social risk on health outcomes, there are emergent efforts to develop the standards and data infrastructure to connect social sector and health data in an interoperable, low-burden, and mutually beneficially way. (Long-term)

Next Steps. These opportunities will be discussed as part of a public-facing webinar on November 4th, 2021, presented in collaboration by NORC, PCORI, and three dyads of technical and industry experts. These experts will provide context and perspective on the three opportunity areas above and engage the audience in a discussion to draw out further suggestions for PCORnet data infrastructure improvements.