Incorporating social determinants of health data into PCORnet®

Keith Marsolo, PhD

Co-PI CDM Core
PCORnet Coordinating Center

Associate Professor
Department of Population Health Sciences
Duke Clinical Research Institute
Duke University School of Medicine

Disclosures

- Investigator on research contracts to Duke University from Pfizer, Boehringer Ingelheim, Novartis, Bristol-Myers Squibb
- Co-inventor – Hive Networks, Inc.

- Duke University is part of the Coordinating Center for PCORnet®, the National Patient-Centered Research Network. PCORnet® has been developed with funding from the Patient-Centered Outcomes Research Institute® (PCORI®). Duke University's participation in PCORnet® is funded through PCORI® Awards (CC2-Duke-2016 and RI-DCRI-01-PS2).
- The statements presented in this work are solely the responsibility of the author and do not necessarily represent the views of other organizations participating in, collaborating with, or funding PCORnet® or of the Patient-Centered Outcomes Research Institute® (PCORI®).

pcornet®

Objectives

- Provide a brief definition of social determinants of health and a background on PCORnet
- Highlight recommendations from a recent Data Convening on expanding social determinants of health data across PCORnet
- Describe in-progress efforts to incorporate social determinants of health data into PCORnet

What are social determinants of health?

- World Health Organization definition:
  - Non-medical factors that influence health outcomes
  - Conditions in which people are born, grow, work, live, and age, and the wider set of forces and systems shaping the conditions of daily life
- Examples of common concepts are shown in the table

Social & Economic Risk Screening Tool	Recommended Social and Behavioral Domains and measures for Electronic Health Records* National Academy of Medicine (formerly Institute of Medicine)[61]	PRAPARE: Protocol for Responding to and Assessing Patient Assets, Risks, and Experiences National Association of Community Health Centers (NACHC)[62]	Accountable Health Communities Screening Tool Centers for Medicare & Medicaid Services[63]
Total # of Questions	24	21	10
Residential address	X	X	
Race/ethnicity	X	X	
Alcohol use	X		
Tobacco use & exposure	X		
Depression	X		
Education	X	X	
Financial resource strain – overall	X		
Household income		X	
Household size		X	
Housing		X	X
Food		X	X
Clothing		X	
Utilities (phone, gas, electric)		X	X
Medicine / health care		X	
Childcare		X	
Transportation		X	X
Neighborhood safety		X	
Interpersonal violence / safety	X	X	X
Physical Activity	X		
Social connections / isolation	X	X	
Stress	X	X	
Migrant / seasonal farmworker		X	
Veteran status		X	
Primary language		X	
Incarceration history		X	
Refugee status		X	
Insurance status		X	

https://www.who.int/health-topics/social-determinants-of-health

Image source: https://liberty.norc.org/content/file?id=62

[61] Committee on the Recommended Social and Behavioral Domains and Measures for Electronic Health Records; Board on Population Health and Public Health Practice; Institute of Medicine. Capturing Social and Behavioral Domains and Measures in Electronic Health Records: Phase 2. Washington (DC): National Academies Press (US); 2015 Jan 8. https://www.ncbi.nlm.nih.gov/books/NBK268995/ doi: 10.17226/18951

[62] NACHC. (2016). PRAPARE®: Protocol for Responding to and Assessing Patient Assets, Risks, and Experiences Paper Version of PRAPARE® for Implementation as of September 2, 2016 http://www.nachc.org/wp-content/uploads/2016/09/PRAPARE_One_Pager_Sept_2016.pdf

[63] Billioux, A., K. Verlander, S. Anthony, and D. Alley. 2017. Standardized Screening for Health-Related Social Needs in Clinical Settings: The Accountable Health Communities Screening Tool. NAM Perspectives. Discussion Paper, National Academy of Medicine, Washington, DC. https://doi.org/10.31478/201705b

PCORnet®: A Network of Networks

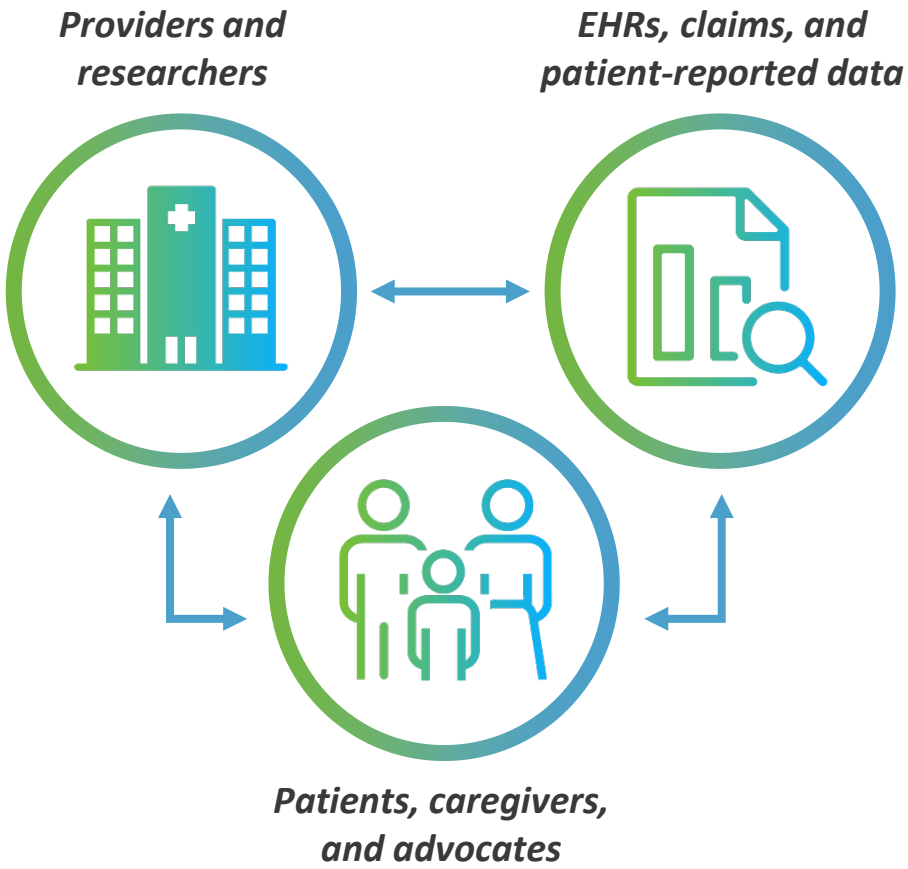
One PCORnet®, Many Possibilities

PCORnet is a national resource, funded by PCORI, where high quality health data, patient partnership, and research expertise deliver fast, trustworthy answers that advance health outcomes.

- **Real-world evidence studies**
- **Comparative effectiveness research**
- **Population health research**
- **Pragmatic research**
- **Health systems research**
- **And more**

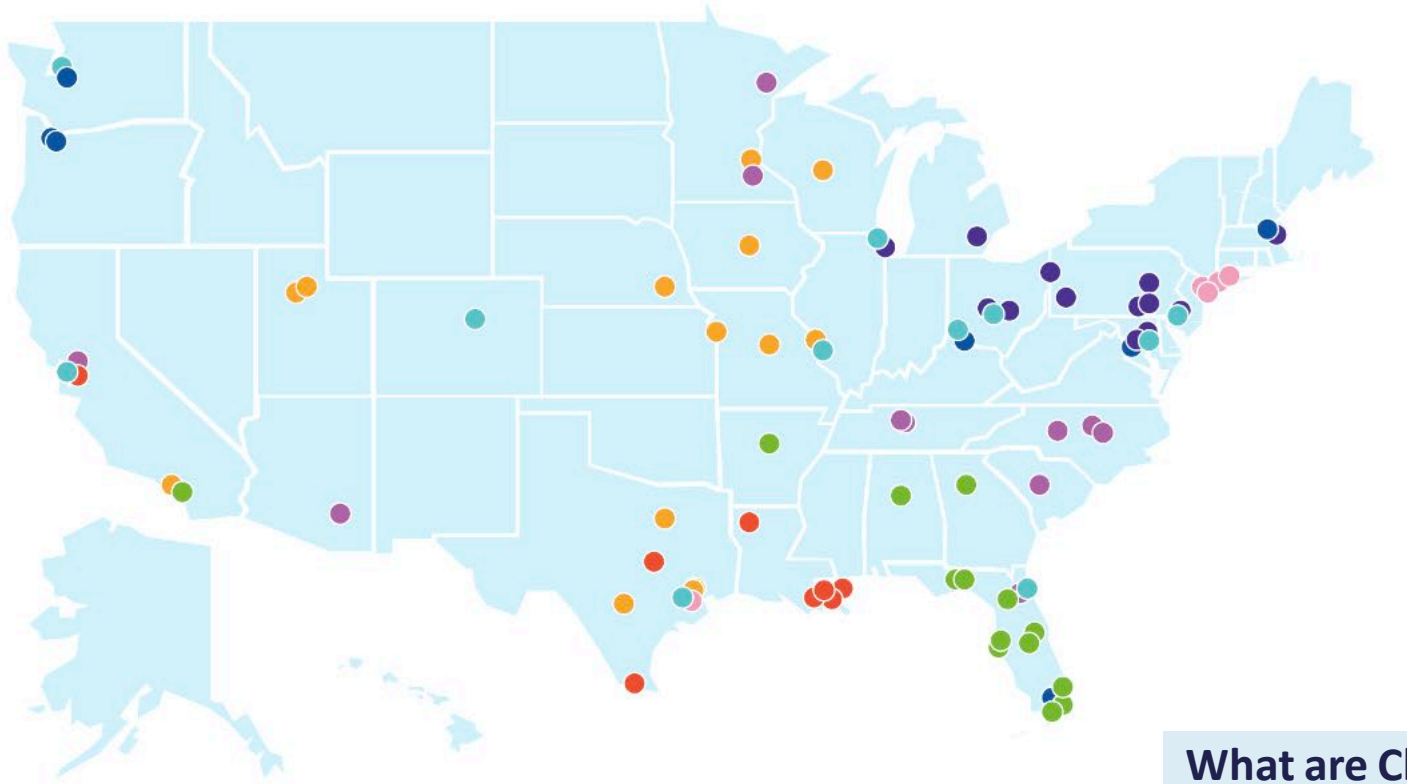
More Than a Data Network

Access to patient partners and **thousands of clinicians** with expert knowledge of PCORnet-enabled data = meaningful research targets and faster answers.



pcornet®

PCORnet® Clinical Research Network locations



- ADVANCE
- GPC
- INSIGHT – NYC
- OneFlorida+
- PaTH
- PEDSnet
- REACHnet
- STAR

What are Clinical Research Networks?

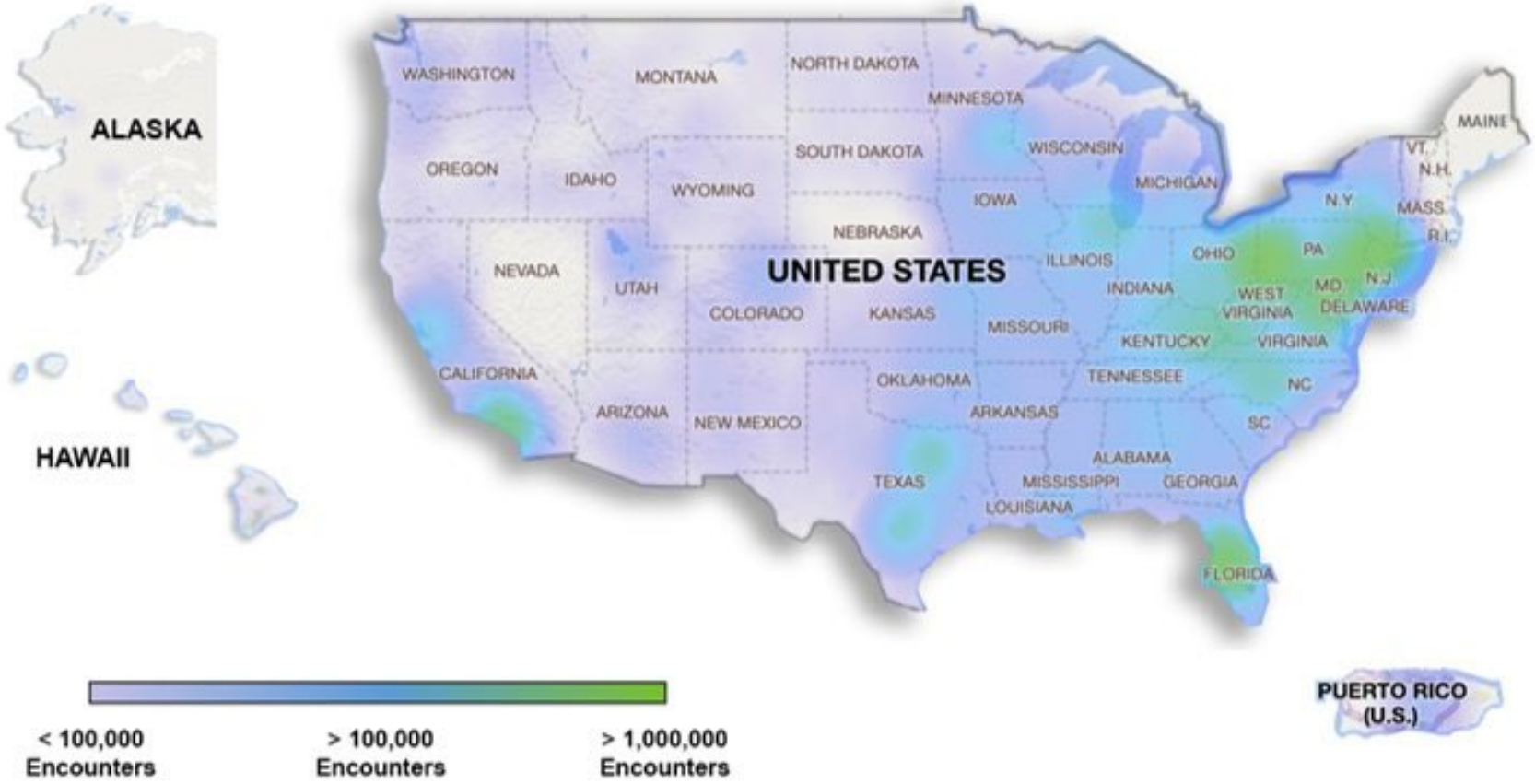
CRNs are groups of diverse healthcare institutions across the U.S., from large academic health centers to local community clinics, united by a commitment to speed patient-centered research via PCORnet.

Access to Data on a National Scale

13K+
Clinical Sites

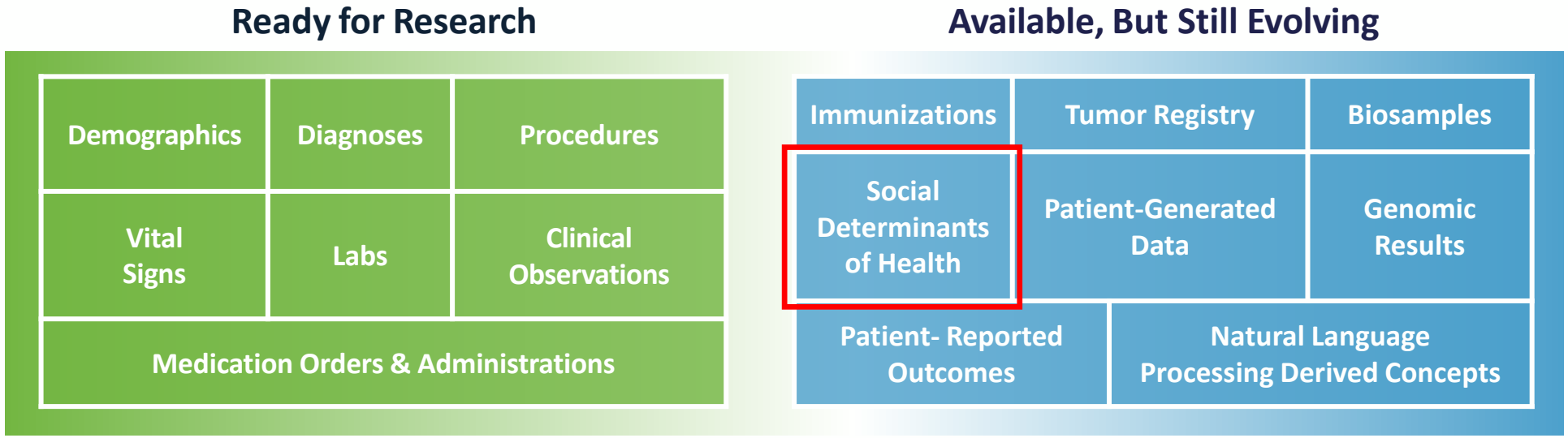
30M
Patient Encounters

30+
PCORnet® Studies



** Map not inclusive of encounters for sites added as part of 2023 PCORnet site expansion*

The PCORnet® Common Data Model (CDM)



Data available from 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

pcornet®

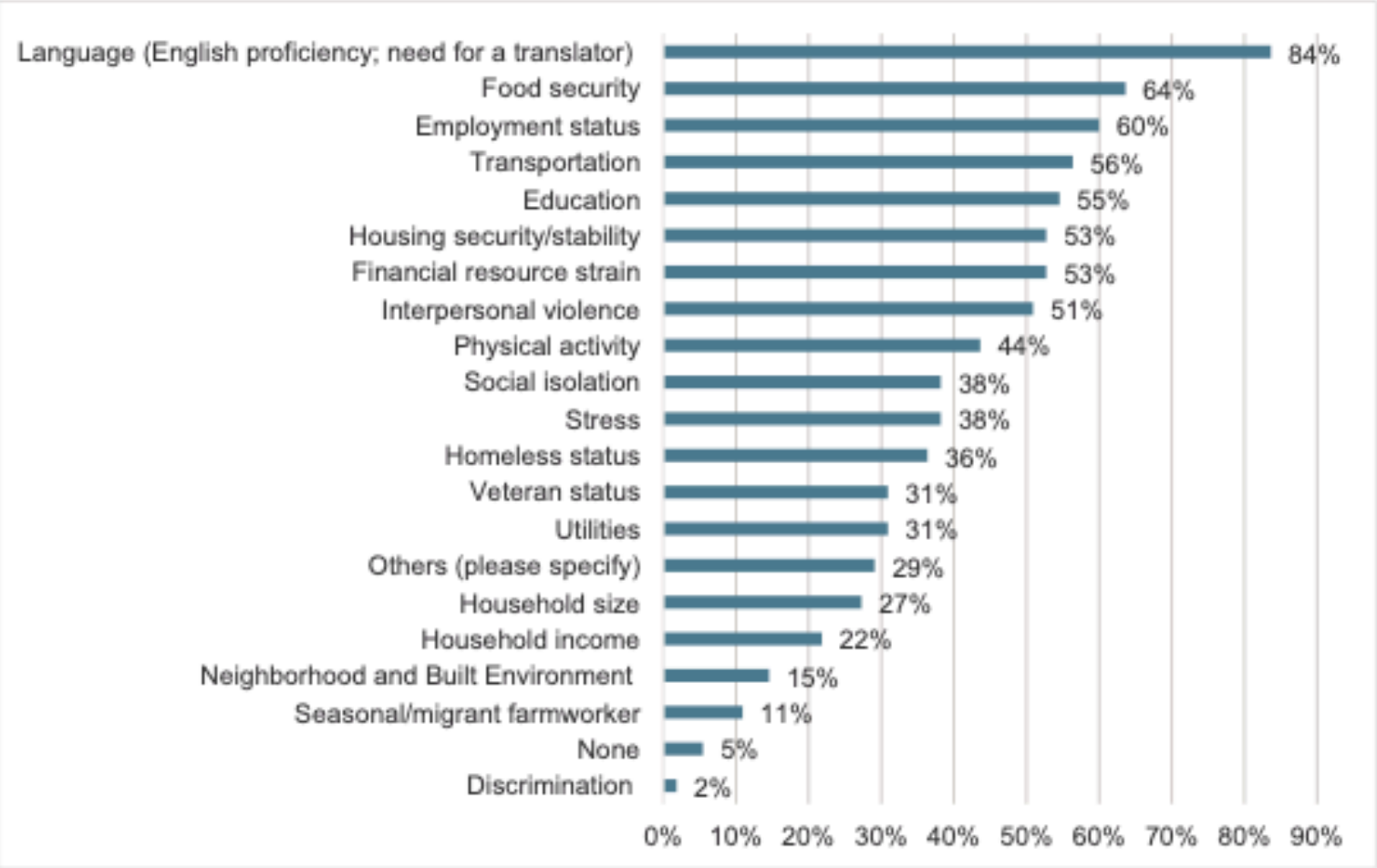
PCORI® Data Convenings

- In 2021 & 2022, PCORI contracted with NORC at the University of Chicago to undertake a series of convenings to consider data infrastructure enhancements to PCORnet®
  - Social determinants of health (SDOH)
  - Patient-reported outcomes & other patient-generated health data
  - Centers for Medicare and Medicaid Services (CMS) claims data
- Social determinants convening built upon efforts of prior PCORnet SDOH workgroup and included survey development, key informant interviews and public webinars

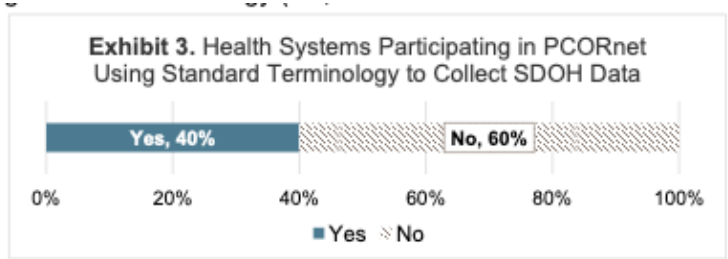
pcornet®

https://www.pcori.org/research-results/2021/pcornet-data-convenings-opportunities-enhance-pcornet-infrastructure-pcor-and-cer

Example findings – availability of patient-level social determinants



* Multiple 'other' responses included patient need for social supports, health literacy, alcohol use, and depression.



Number of respondents: 55

pcornet®

Image source: https://liberty.norc.org/content/file?id=62

Example findings – availability of address information

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

Image source: https://www.pcori.org/sites/default/files/PCORnet-SDOH-Presentation-Slides-508-110421.pdf

Recommendations

Short-Term	• 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	• Use natural language processing to mine unstructured individual-level SDOH data
Long-Term	• Incorporate social sector data, informed by data standards and platforms already in use

pcornet®

Image source: https://www.pcori.org/sites/default/files/PCORnet-SDOH-Presentation-Slides-508-110421.pdf

Recommendations

Short-Term	• 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	• Use natural language processing to mine unstructured individual-level SDOH data
Long-Term	• Incorporate social sector data, informed by data standards and platforms already in use

Focus for today

pcornet

Image source: https://www.pcori.org/sites/default/files/PCORnet-SDOH-Presentation-Slides-508-110421.pdf

Incorporating patient-level SDOH measures

- The PCORnet CDM includes tables that can store patient-level SDOH data (e.g., PRO_CM)
- Adding these data to the CDM generally involves several steps
  - Identifying whether there are codes to represent these measures in standard terminologies (e.g., LOINC, SNOMED)
  - Partners must find the relevant measures within their EHRs and harmonize them to the appropriate code
  - In many EHRs, data may be captured using various workflows over time (e.g., flowsheets, questionnaires, etc.), which can also affect the overall data completeness

pcornet®

Social & Economic Risk Screening Tool	Recommended Social and Behavioral Domains and measures for Electronic Health Records* National Academy of Medicine (formerly Institute of Medicine)[61]	PRAPARE: Protocol for Responding to and Assessing Patient Assets, Risks, and Experiences National Association of Community Health Centers (NACHC)[62]	Accountable Health Communities Screening Tool Centers for Medicare & Medicaid Services[63]
Total # of Questions	24	21	10
Residential address	X	X	
Race/ethnicity	X	X	
Alcohol use	X		
Tobacco use & exposure	X		
Depression	X		
Education	X	X	
Financial resource strain – overall	X		
Household income		X	
Household size		X	
Housing		X	X
Food		X	X
Clothing		X	
Utilities (phone, gas, electric)		X	X
Medicine / health care		X	
Childcare		X	
Transportation		X	X
Neighborhood safety		X	
Interpersonal violence / safety	X	X	X
Physical Activity	X		
Social connections / isolation	X	X	
Stress	X	X	
Migrant / seasonal farmworker		X	
Veteran status		X	
Primary language		X	
Incarceration history		X	
Refugee status		X	
Insurance status		X	

Z-codes

Example – Food security (Hunger Vital Sign™)

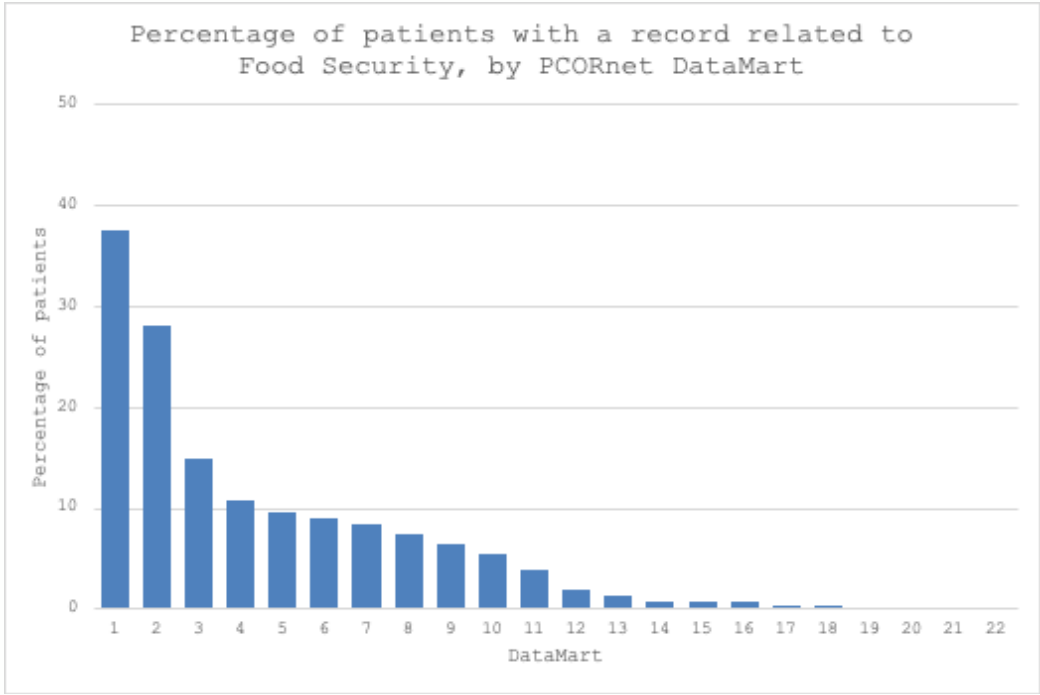
> "Within the past 12 months we worried whether our food would run out before we got money to buy more."

> "Within the past 12 months the food we bought just didn't last and we didn't have money to get more."

LOINC | food security | Search

RESULTS 39 | DISPLAYING 200 1-39 | FILTER | VIEW List Card | EXPORT

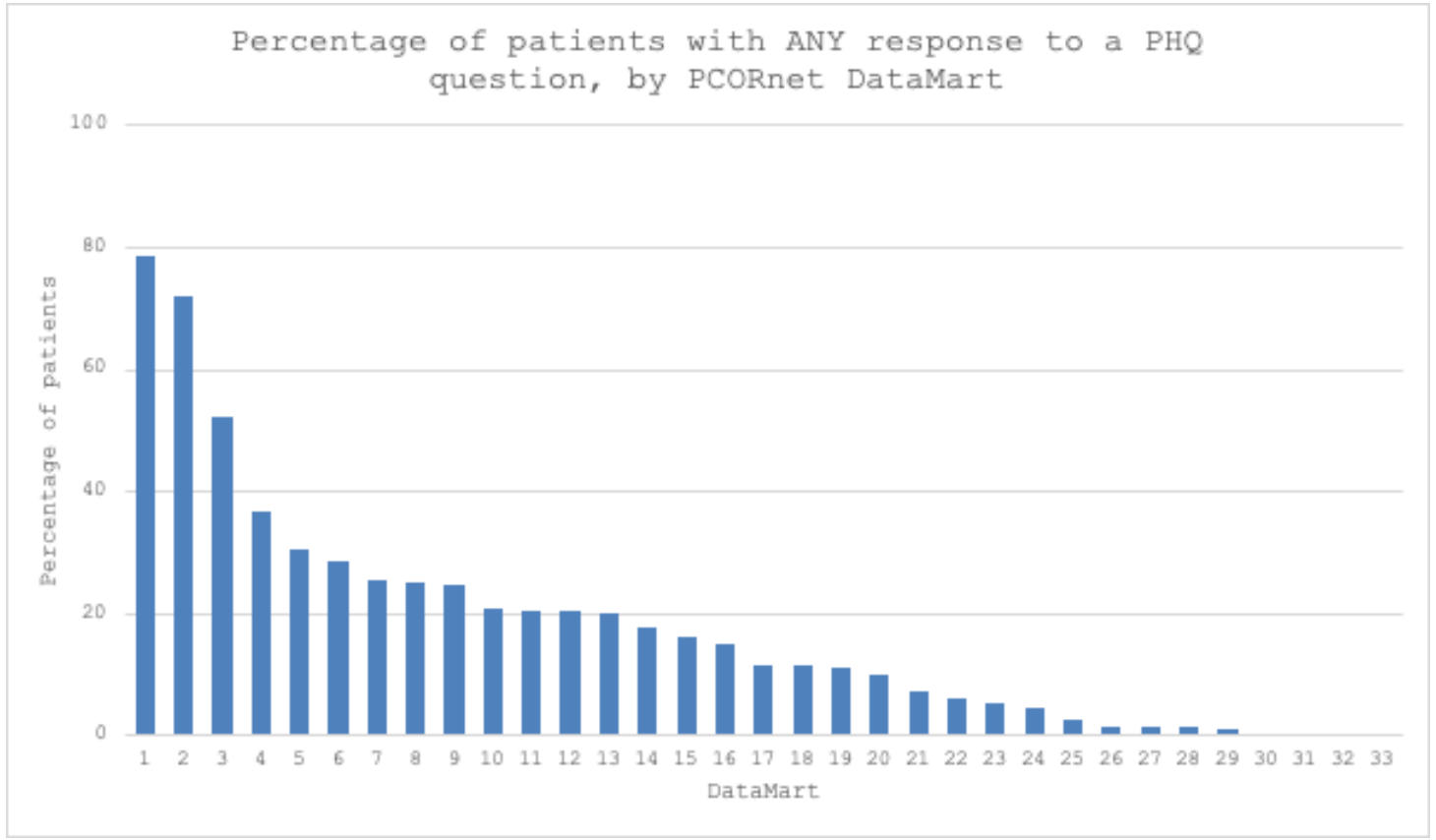
Status	LOINC	Long Common Name	Component	Property	Timing	System	Scale	Method	Class	Type	Example UCUM Units	Order/ Observation	Version First Released
	88122-7	Within the past 12 months we worried whether our food would run out before we got money to buy more [U.S. FSS]	Within the past 12Mo we worried whether our food would run out before we got money to buy more	Find	Pt	^Patient	Ord	U.S. Food Security Survey	SURVEY.SDOH			Observation	2.63
	88123-5	Within the past 12 months the food we bought just didn't last and we didn't have money to get more [U.S. FSS]	Within the past 12Mo the food we bought just didn't last and we didn't have money to get more	Find	Pt	^Patient	Ord	U.S. Food Security Survey	SURVEY.SDOH			Observation	2.63
	95248-1	In the last 12 months, we couldn't afford to eat balanced meals [U.S. FSS]	In the last 12Mo, we couldn't afford to eat balanced meals	Find	12Mo	^Patient	Ord	U.S. Food Security Survey	SURVEY.SDOH			Observation	2.68
	95249-9	In the last 12 months, did you ever cut the size of your meals or skip meals because there wasn't enough money for food [U.S. FSS]	In the last 12Mo, did you ever cut the size of your meals or skip meals because there wasn't enough money for food	Find	12Mo	^Patient	Ord	U.S. Food Security Survey	SURVEY.SDOH			Observation	2.68



Includes all DataMarts who had loaded any records related to food security by July 2023 (n=22). Number of patients with a record in the ENCOUNTER table used to calculate denominator.

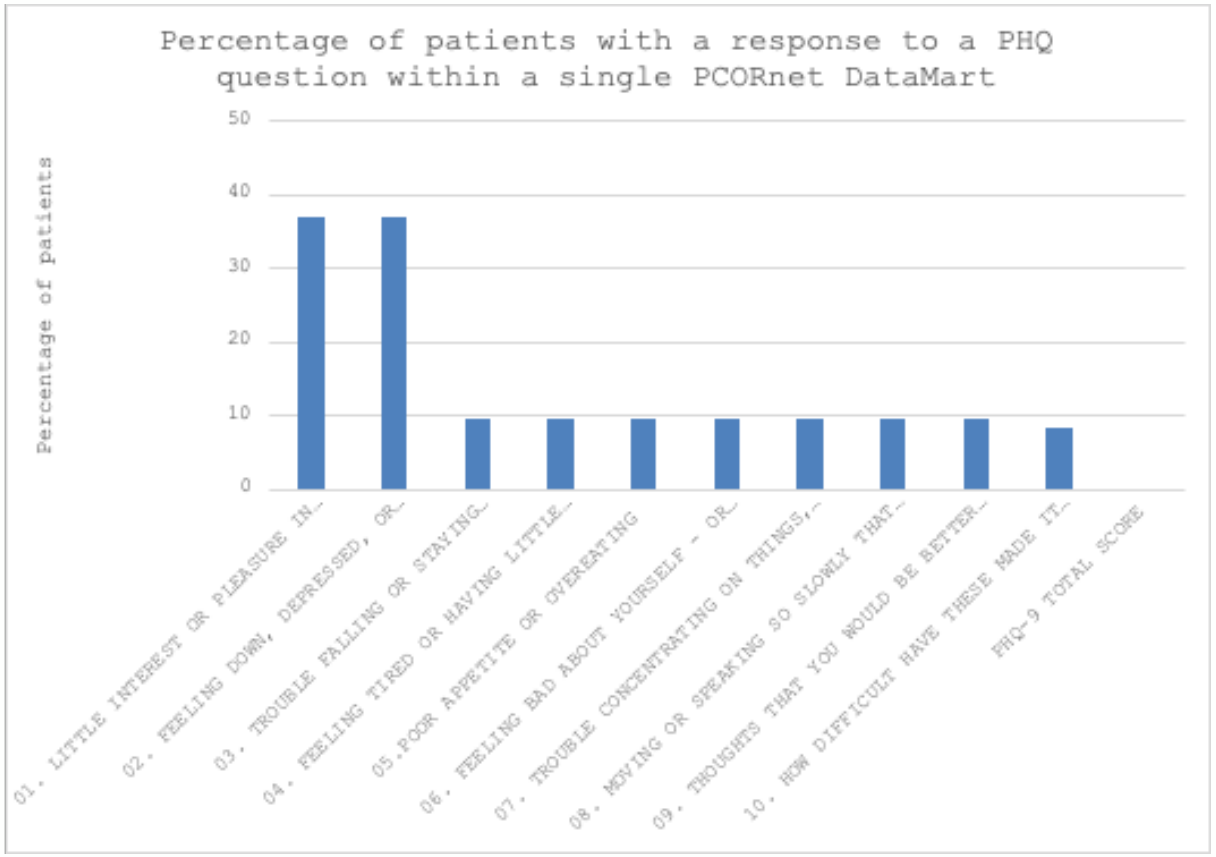
https://childrenshealthwatch.org/public-policy/hunger-vital-sign/

Example – Depression (Patient Health Questionnaire)



Includes all DataMarts who had loaded any PHQ records July 2023 (n=33). Number of patients with a record in the ENCOUNTER table used to calculate denominator.

PHQ availability by question



PHQ-2

PHQ-8

PHQ-9

PHQ-9 Total Score

Additional measure if >0 on any question

Over the last *2 weeks*, how often have you been bothered by any of the following problems?
(use "✓" to indicate your answer)

	Not at all	Several days	More than half the days	Nearly every day
1. Little interest or pleasure in doing things	0	1	2	3
2. Feeling down, depressed, or hopeless	0	1	2	3
3. Trouble falling or staying asleep, or sleeping too much	0	1	2	3
4. Feeling tired or having little energy	0	1	2	3
5. Poor appetite or overeating	0	1	2	3
6. Feeling bad about yourself—or that you are a failure or have let yourself or your family down	0	1	2	3
7. Trouble concentrating on things, such as reading the newspaper or watching television	0	1	2	3
8. Moving or speaking so slowly that other people could have noticed. Or the opposite — being so figety or restless that you have been moving around a lot more than usual	0	1	2	3
9. Thoughts that you would be better off dead, or of hurting yourself	0	1	2	3
add columns		+	+	
(Healthcare professional: For interpretation of TOTAL, please refer to accompanying scoring card).	TOTAL:			

10. If you checked off *any problems*, how *difficult* have these problems made it for you to do your work, take care of things at home, or get along with other people?

- Not difficult at all
- Somewhat difficult
- Very difficult
- Extremely difficult

Image source: https://med.stanford.edu/fastlab/research/imapp/msrs/_jcr_content/main/accordion/accordion_content3/download_256324296/file.res/PHQ9%20id%20date%2008.03.pdf

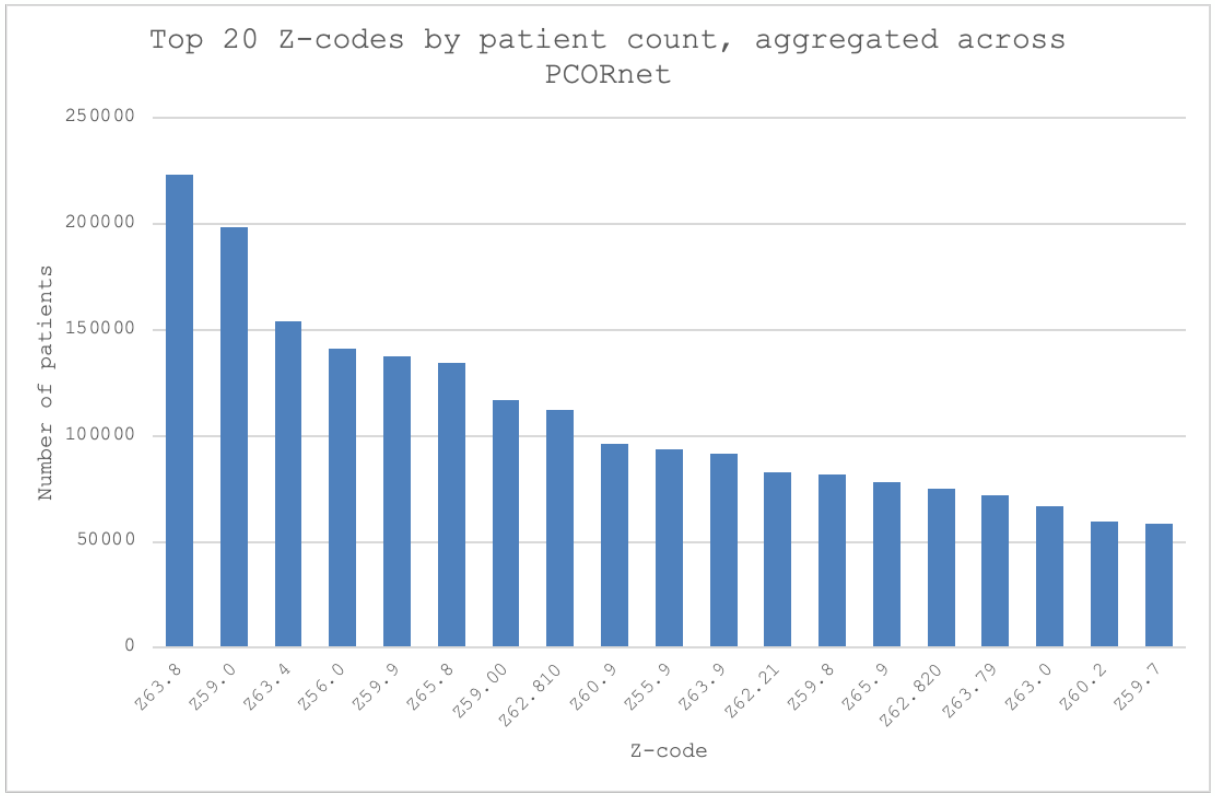
Insurance status

- Insurance status is often considered a surrogate measure of SDOH
- Within the PCORnet CDM (and most electronic health records), insurance is captured at an encounter level, with Payer Name recorded, not necessarily the payer type or payer class.
- These raw values must be harmonized to the CDM value set (based on the Source of Payment Typology)
- Completeness is highly variable across Network Partners – ranging from 0% to 100% missing
- As part of a targeted improvement effort, working with the network to improve the quality of these data

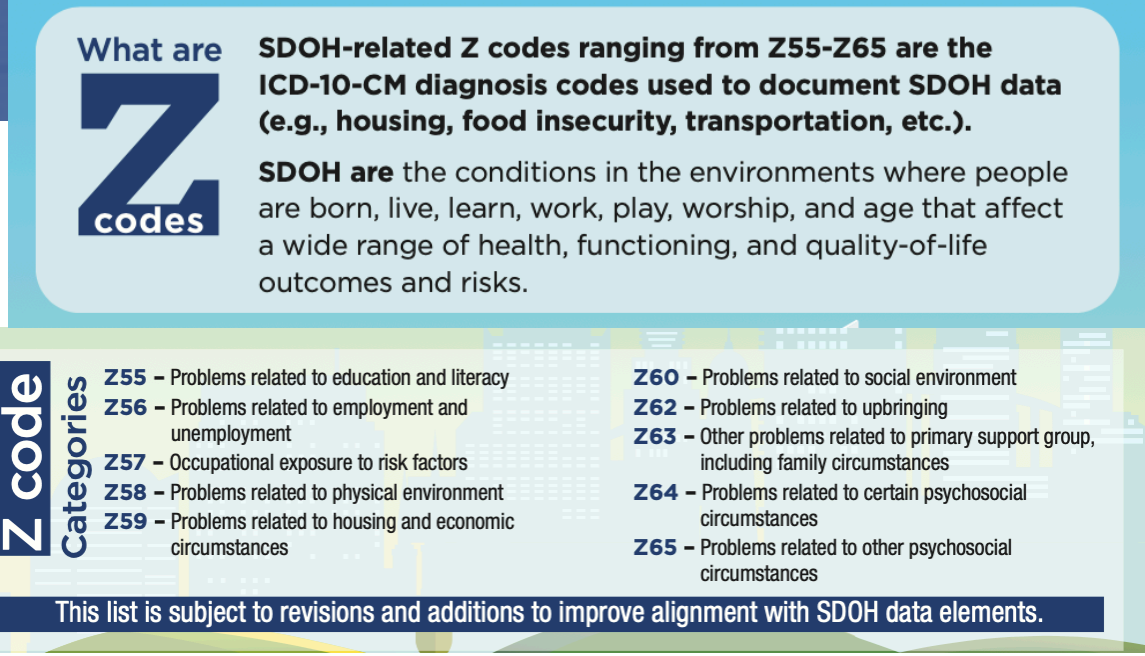
Example Payer Names
ANTHEM ALLIANCE
ANTHEM BLUE ACCESS
ANTHEM BLUE ACCESS CHOICE
ANTHEM BLUE PREFERRED HMO/PLUS
ANTHEM MEDICARE
ANTHEM PATHWAY
ANTHEM PATHWAY X
MOLINA EXCHANGE
MOLINA HEALTH CARE OF OHIO
MOLINA HEALTH CARE TRANSPLANT
MOLINA HNC
MOLINA HNCC TRANSPLANT
MOLINA MYCARE OH DUALS MEDICAID SECONDARY
MOLINA MYCARE OHIO DUAL OPTIONS

pcornet®

What about Z-codes?



Aggregate counts across all of PCORnet. Total denominator of all patients with any diagnosis code: ~114M



Code	Description
Z63.8	Other specified problems related to primary support group
Z59.0	Homelessness
Z63.4	Disappearance and death of family member.
Z56.0	Unemployment, unspecified
Z59.9	Problem related to housing and economic circumstances, unspecified
Z65.8	Other specified problems related to psychosocial circumstances
Z59.00	Homelessness unspecified
Z62.810	Personal history of physical and sexual abuse in childhood
Z60.9	Problem related to social environment, unspecified
Z55.9	Problems related to education and literacy, unspecified
Z63.9	Problem related to primary support group, unspecified
Z62.21	Child in welfare custody
Z59.8	Other problems related to housing and economic circumstances
Z65.9	Problem related to unspecified psychosocial circumstances
Z62.820	Parent-biological child conflict
Z63.79	Other stressful life events affecting family and household
Z63.0	Problems in relationship with spouse or partner
Z60.2	Problems related to living alone
Z59.7	Insufficient social insurance and welfare support

https://www.cms.gov/files/document/zcodes-infographic.pdf

Address data & geocoding; linkage to external sources

- At network partners, the PCORnet CDM is considered a Limited Data Set – it contains dates of service and permitted elements of address including State and Zip Code
  - County was added in CDM v6.1 (in production July 2023)
- Most network partners also have access to full patient addresses, with some capabilities to geocode to census tract, latitude/longitude, etc.
  - These geocodes cannot be directly queried as part of network-wide prep-to-research queries, but can be linked to other datasets and those values can be more readily leveraged

Survey response from January 2023 – Does your organization have a process to generate geocodes?

Response	Total
Yes	37
No	8
Unsure	1
Other	5

Survey response from January 2023 – Level of geocodes available (25 responses in total)

Response	Total
Census tract	24
Census block	19
Census block group	23
9-digit zip code	14
Latitude / Longitude	22

Using zip code to report socioeconomic status in prep-to-research queries

- Incorporated a reference table into the query tools to assign patients to a quartile based on the Area Deprivation Index for their 5-digit zip code
- Note: Area Deprivation Index is not validated for 5-digit zip codes, but still provides some insight into the PCORnet population

This report contains results from 61 network partners that responded by December 5, 2022

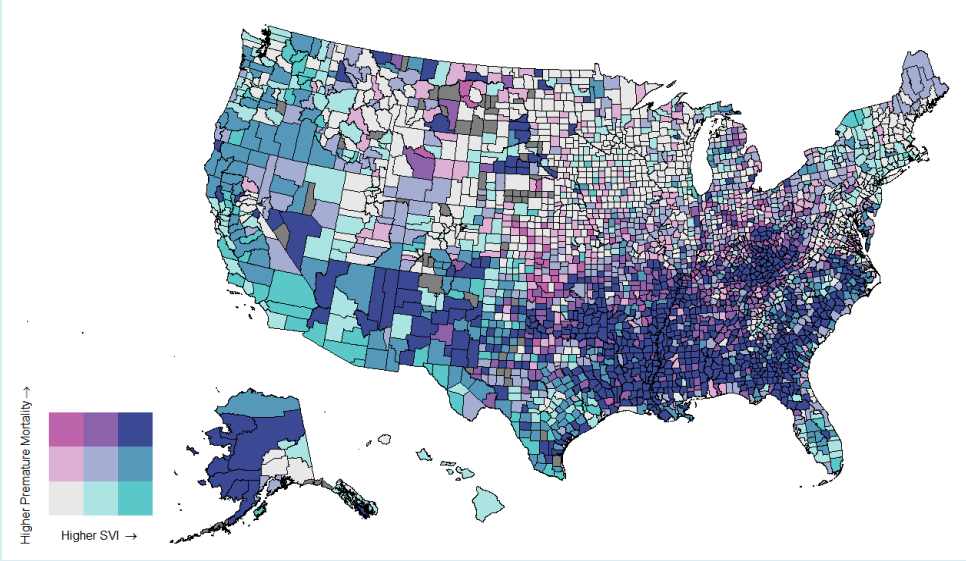
Table 1. Aggregated Baseline Table for All Cohorts

	Anxiety diagnosis		Asthma diagnosis		Breast cancer diagnosis		Heart failure diagnosis	
	N Mean	% StdDev	N Mean	% StdDev	N Mean	% StdDev	N Mean	% StdDev
Characteristics								
Number of unique patients	3,448,502		1,584,973		320,528		800,795	
Demographics								
Age at Health Event of Interest								
Mean Age	43.4	19.7	38.4	21.4	63.5	12.7	68.8	14.8
00-19	545,377	16%	463,059	29%	65	0%	8,156	1%
20-44	1,281,712	37%	452,878	29%	25,757	8%	46,065	6%
45-64	976,182	28%	396,748	25%	135,117	42%	219,759	27%
65-74	392,660	11%	171,296	11%	95,256	30%	214,488	27%
75+	252,571	7%	100,992	6%	64,333	20%	312,327	39%
Sex								
Female	2,318,804	67%	969,373	61%	318,051	99%	377,028	47%
Male	1,128,252	33%	615,271	39%	2,449	1%	423,689	53%
Other/Missing	1446	0%	329	0%	28	0%	78	0%
Race								
American Indian or Alaska Native	19,154	1%	9,480	1%	1,138	0%	3,780	0%
Asian	60,840	2%	33,792	2%	8,635	3%	10,785	1%
Black or African American	378,103	11%	358,653	23%	41,439	13%	160,725	20%
Native Hawaiian or Other Pacific Islander	6,409	0%	4,332	0%	417	0%	1,542	0%
White	2,626,454	76%	954,595	60%	236,875	74%	550,357	69%
Other/Missing	357542	10%	224,121	14%	32,024	10%	73,606	9%
Hispanic								
No	2,667,594	77%	1,193,416	75%	258,161	81%	665,958	83%
Yes	342,017	10%	228,293	14%	22,078	7%	51,281	6%
Other/Missing	438891	13%	163,264	10%	40,289	13%	83,556	10%
Area Deprivation Index (ADI) quartile based on 5-digit zip to ADI mapping (see Notes)								
ADI unavaiable for 5-digit zip	843	0%	313	0%	38	0%	180	0%
ADI quartile 1 (lowest deprivation)	1,125,796	33%	459,067	29%	124,518	39%	209,492	26%
ADI quartile 2	705,352	20%	284,931	18%	63,884	20%	161,331	20%
ADI quartile 3	748,237	22%	328,373	21%	64,441	20%	195,396	24%
ADI quartile 4 (highest deprivation)	454,701	13%	292,446	18%	39,650	12%	157,681	20%
5-digit zip missing	413,573	12%	219,843	14%	27,997	9%	76,715	10%

AHRQ Social Determinants of Health Database

- "One stop" standardized community level SDOH data from multiple public sources
  - Social context, economic context, education, healthcare context, physical infrastructure
- Purpose
  - Make community-level SDOH data easier to use
  - Account for differences across areas
  - Identify effective interventions
  - Inform efforts to improve health, equity
- Linkable by geography
  - County level (2009-2020)
  - Zip code level (2011-2020)
  - Census Tract level (2009-2020)

Premature Mortality and Social Vulnerability Index across Counties in the U.S., 2018



Source: AHRQ SDOH Database, version 1, from 2018 County Health Rankings and 2018 CDC SVI. Darker shading is higher.

https://www.ahrq.gov/sdoh/data-analytics/sdoh-data.html

This work was supported by the Office of the Secretary Patient-Centered Outcomes Research Trust Fund Under Interagency Agreement 750119PE0K0036

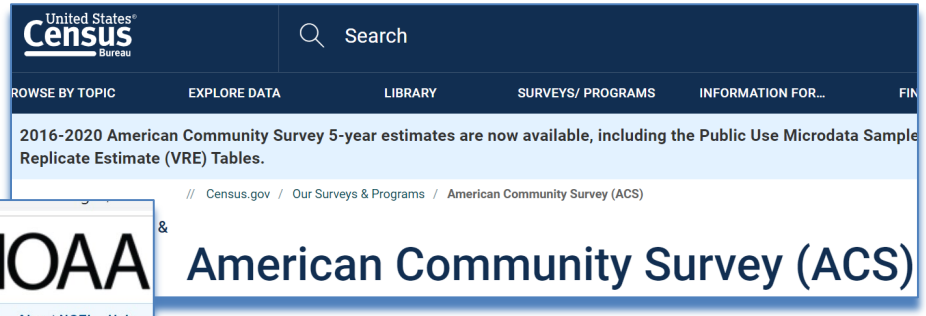
Slides used permission from
Patricia Keenan, PhD
Senior Researcher, AHRQ

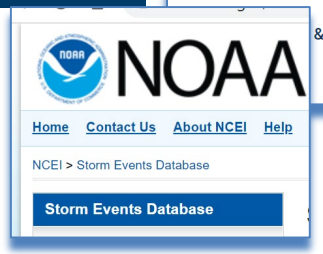
Community-level SDOH Variables Organized by Domains and Topics

Social Context	Economic Context	Education	Physical Infrastructure	Healthcare Context
• Demographics • Disability • Immigration • Living conditions • Segregation • Socioeconomic disadvantage indices	• Employment • Income • Poverty	• Attainment • Education funding • Literacy • Numeracy • School system	• Access to Exercise • Crime • Environment • Food access • Housing • Industry composition • Internet connectivity • Migration • Social services • Transportation	• Characteristics of healthcare – facilities – providers • Distance to providers • Health behaviors • Health care quality • Health insurance status • Utilization and cost • Health outcomes

Source: AHRQ SDOH Database, version 1.

Examples of SDOH Database Data Sources





CDC/ATSDR Social Vulnerability Index

Summary

- PCORnet has demonstrated that patient-level SDOH data can be incorporated to the CDM
  - Data availability is dependent on adoption & utilization by health systems
  - May be suitable for studies on targeted populations, but will depend on collection practices at a given health system
- Area-level measures can provide population-level SDOH insights
  - 5-digit zip and county can be included in Limited Data Sets, and are more easily used in distributed analytics (e.g., prep-to-research queries)
  - Capabilities for geocoding exist at many institutions, but will require involvement of local personnel to generate values based on census tract or latitude/longitude – may be best suited for specific studies

pcornet®