

Identifying Population Health Data Resources Within a Large Data Repository

JAMES W. MCNALLY, SUSAN LEONARD, AND MARGARET LEVENSTEIN
UNIVERSITY OF MICHIGAN, INTER-UNIVERSITY CONSORTIUM
FOR POLITICAL AND SOCIAL RESEARCH (ICPSR)
UNIVERSITY OF MICHIGAN, ANN ARBOR

Population Health

“Population Health” is a relatively new attempt to integrate the growing wealth of data on both the community and the individual level into seamless models that describe:

- ▶ Health Risks: Personal Behaviors, as well as Social, Economic and Environmental Factors
- ▶ Health Treatments: How health problems are treated from ER care to Personalized Medicine to Non Traditional Medicines
- ▶ Health Outcomes: Survival, Quality of Life, Long Term Results

Collectively these provide a metric for the success of the Health Care Provision System

Population Health



Reframing Kindig and Stoddart (2003), “Population Health” can be defined as the collection, organization, and aggregation of health data across multiple information platforms and technology resources. The analysis of these data into a unified, interoperable data files can lead to improvements in social, clinical and economic decisions resulting in the improved health of an entire human population.

Or More Simply...

Population Health is a
“conceptual framework for
thinking about why some
populations are healthier
than others.”

Population Health



From our perspective:

Population Health then is about DATA and how to organize and integrate multiple levels of information to better understand individual health outcomes.

Data Driven answers are what we do...

NACDA Program on Aging



NACDA is part of ICPSR

The NACDA data archive has spent the past 35 years identifying, gathering and organizing data on aging and the aging life course so we can share these resources with the research community.

NACDA currently supports approximately 1,600 individual studies on aging and health and adds 30 to 50 new studies to our collections each year.

Inter-university Consortium for Political and Social Research (ICPSR)

[FIND DATA](#) ▾ [START SHARING DATA](#) ▾ [MEMBERSHIP](#) ▾ [SUMMER PROGRAM](#) ▾ [TEACHING & LEARNING](#) ▾ [DATA MANAGEMENT & CURATION](#) ▾



- DATA & PUBLICATIONS
- EVERYTHING ELSE

[Log In/Create Account](#)



ICPSR

A partner in social science research

ICPSR Collections

- ▶ Currently, ICPSR maintains a data archive of more than 250,000 files of research in the social and behavioral sciences, with growing collections in epidemiology, biomedical research, and clinical studies. ICPSR hosts 21 specialized collections of data in education, aging, criminal justice, substance abuse, terrorism, and other fields. NACDA represents part of this organizational structure that seeks to identify, structure and disseminate data that supports the study of the aging lifecourse.
- ▶ Locating and organizing this information in a systematic way is the starting point of making data discoverable and more effective for research use.

Figure 1: Population Health Data Collection Framework

NACDA identifies and organizes four primary sources of gerontological data associated with Population Health Research:

- 1) Behavioral and Social Research (BSR) Data
- 2) Biomedical Data
- 3) Genomic Data
- 4) Electronic Health Records (EHR)

NACDA currently manages one of nation's largest repositories of longitudinal BSR health and socioeconomic data.

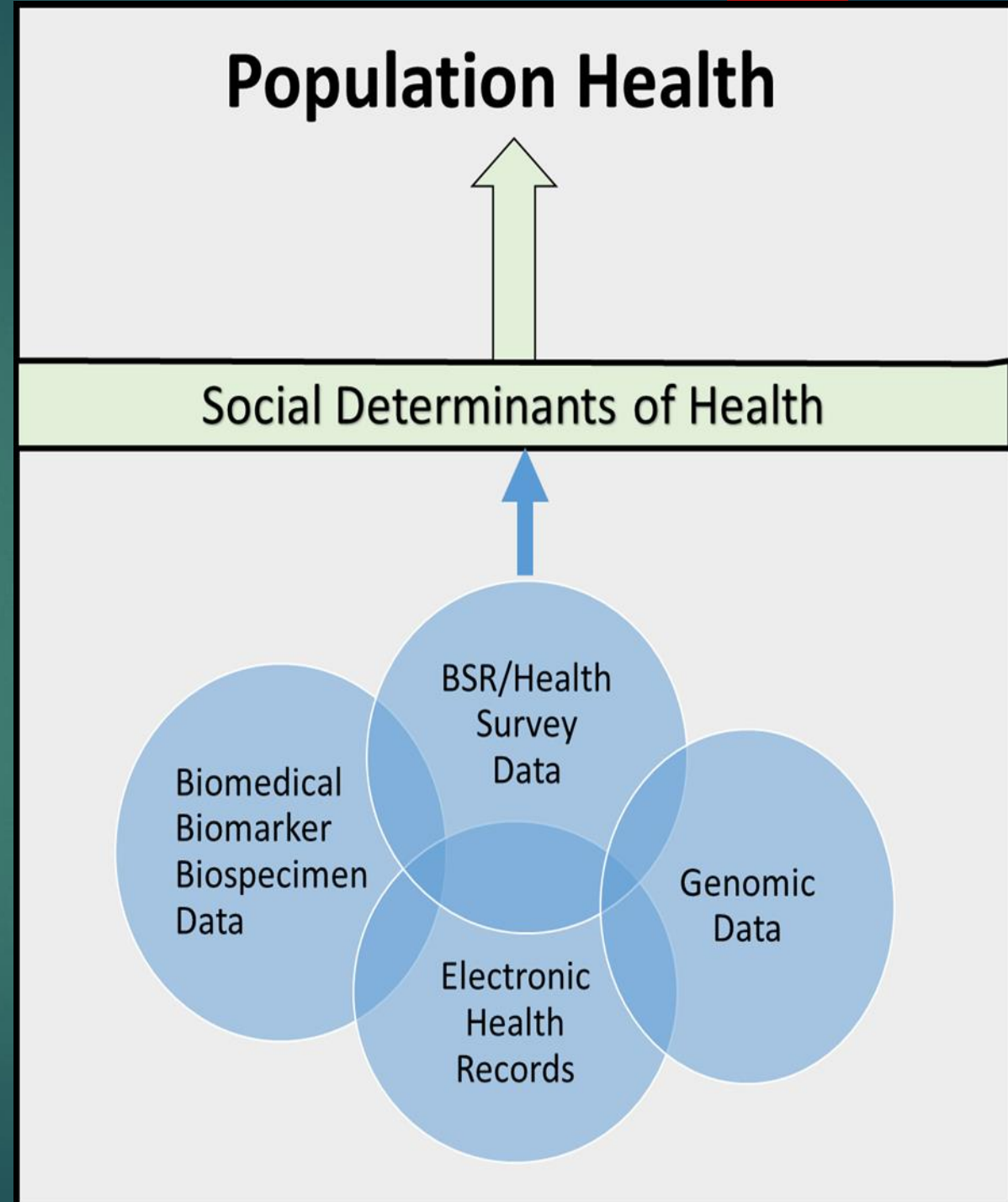
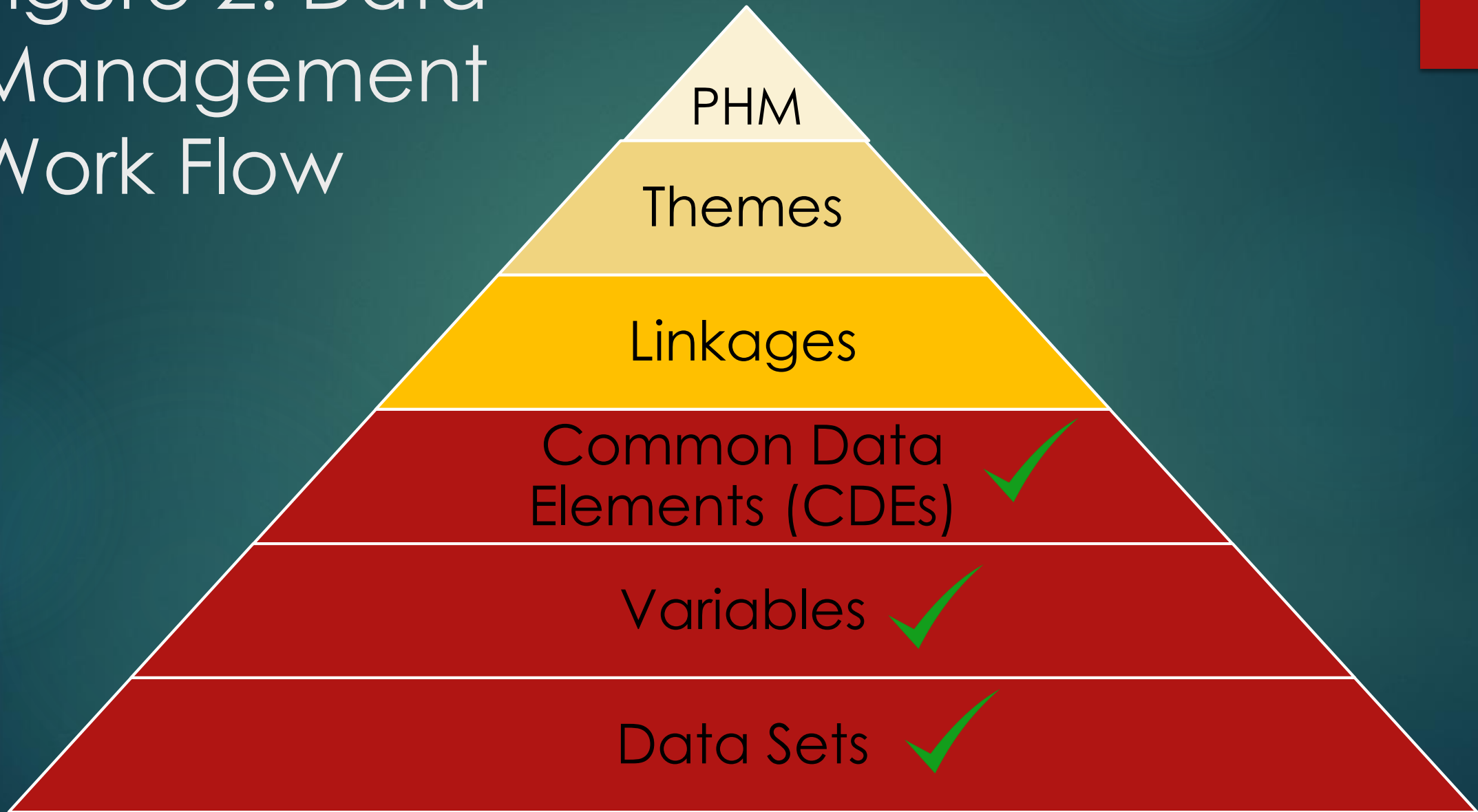


Figure 2: Data Management Work Flow



Common Data Elements (CDEs)

- ▶ CDEs are standardized terms for the collection and exchange of **data**. CDEs are metadata; they describe the type of **data** being collected, not the **data** itself.
- ▶ The broad goal of the ICPSR CDE project is to enhance the variable level metadata for studies across all collections to improve discoverability when researchers are developing specific projects or research frameworks.
- ▶ With funding from OSBRR, we are now completing a pilot project to improve variable discoverability based upon common search ontologies using a CDE approach.

Table 1: Use Case Searches for CDE Ontology

Use case	Search terms	Variables
1) Researcher interested in the relationship between self-reported health and dementia.	self-reported health, dementia, Alzheimer's	
	self-reported health	163
	dementia	2154
	Alzheimer's	1834
1) Alzheimer's researcher generating a survey and looking for best-case examples of question text and order.	dementia, Alzheimer's	
	dementia	2154
	Alzheimer's	1834
1) Undergraduate writing a paper on health challenges in older ages.	health, elderly, dementia	
	health	76622
	elderly	7539
	dementia	2154
1) Government or policy worker looking for factual guidance (e.g., what is the rate of dementia in my town, and what might the load be going forward?).	dementia, age	
	dementia	2154
	age	149123
1) Researcher looking for twin studies to investigate the role of genetics in dementia.	twins, dementia	
	twins	3026
	dementia	2154
1) Undergraduate student looking for facts to be included in a paper on caregiver health.	caregiver health, Alzheimer's	
	caregiver health	14
	Alzheimer's	1834

Moving up the Triangle

CREATING LINKAGES
AND
IDENTIFYING THEMES

Table 2: Social Determinants of Health

Economic Stability	Neighborhood and Physical Environment	Education	Food	Community and Social Context	Health Care System
Employment	Housing	Literacy	Hunger	Social integration	Health coverage
Income	Transportation	Language	Access to healthy options	Support systems	Provider availability
Expenses	Safety	Early childhood education		Community engagement	Provider linguistic and cultural competency
Debt	Parks	Vocational training		Discrimination	Quality of care
Medical bills	Playgrounds	Higher education			
Support	Walkability				

Health Outcomes

Mortality, Morbidity, Life Expectancy, Health Care Expenditures, Health Status, Functional Limitations

Table 3: Mockup of Study Listings Identifying Data Sets with Social Determinants of Health Variables

Study	Economic Stability	Neighborhood Environment	Education	Food	Social Context	Health Care Systems	Total
STUDY 1	100%	100%	100%	100%	100%	100%	100%
STUDY 2	100%	90%	90%	100%	50%	90%	87%
STUDY 3	80%	90%	40%	85%	50%	80%	71%
STUDY 4	10%	25%	80%	30%	10%	20%	29%
STUDY 5	5%	10%	85%	15%	20%	20%	26%

Conclusions and Next Steps

- ▶ This process has come far in the past 20 years, but much work needs to be done.
- ▶ The current universe of CDE repositories provide an excellent framework but fail to capture common SBR measures and variables that are central to health and Population Health.
- ▶ Definitions matter. For work to move forward we need to agree on cross-disciplinary definitions of common measures.
- ▶ We are at a point where we can build forms of Population Health Models. We have the data.
- ▶ What we lack is common agreement upon measures that encourage replication and validation of emerging studies.

NACDA Program on Aging

NACDA operates on the internet 24 hours a day, 365 days a year.

- We welcome you as researchers to use our data.
- Visit us at:
www.icpsr.umich.edu/icpsrweb/NACDA
- Or Google “NACDA Aging”
- ▶ *Twitter* [@NACDA_Aging](https://twitter.com/NACDA_Aging)