UNIVERSITY OF MICHIGAN

# Research Data Stewardship Community of Practice

*The ABCs of CDEs: How Common Data Elements Support Interoperability in Research*

October 24, 2024

RESEARCH DATA STEWARDSHIP INITIATIVE

# Agenda

I. Announcements

II. Presentations

   a. Brief introduction to CDEs - Megan Chenoweth

   b. Researchers' experiences with CDEs

      i. Dr. David Williams
      ii. Dr. James McNally

III. Q & A

# Announcements

- RDSI MCommunity email lists:
  - RDSI Updates [existing] - occasional emails from RDSI working group
  - RDSI Community of Practice [newer!] - preferred communication method from the survey. A more active email group for sharing data-related questions, news, and events.
    - Want to join the new CoP group? Sign up on MCommunity.
    - To email the new CoP group: ResearchDataCoP@umich.edu
- Other announcements?

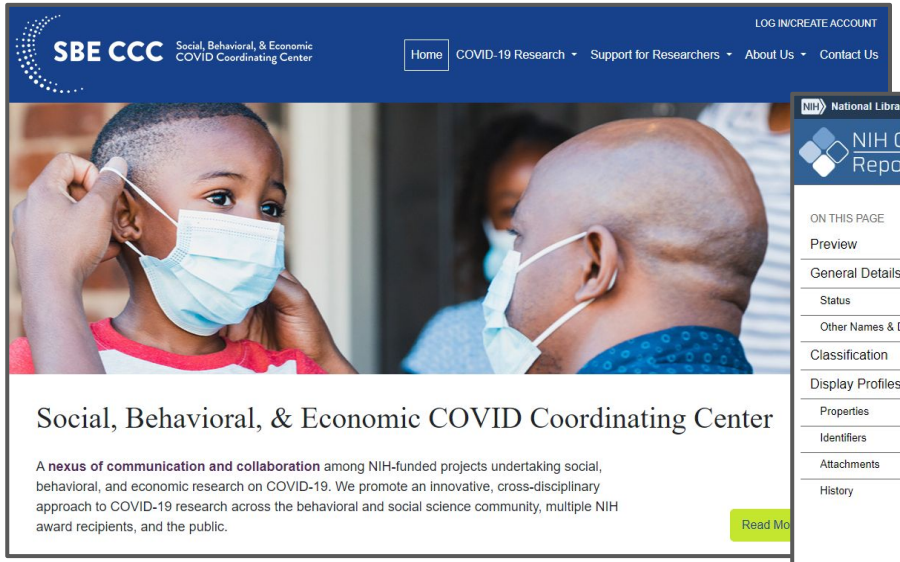
UNIVERSITY OF MICHIGAN

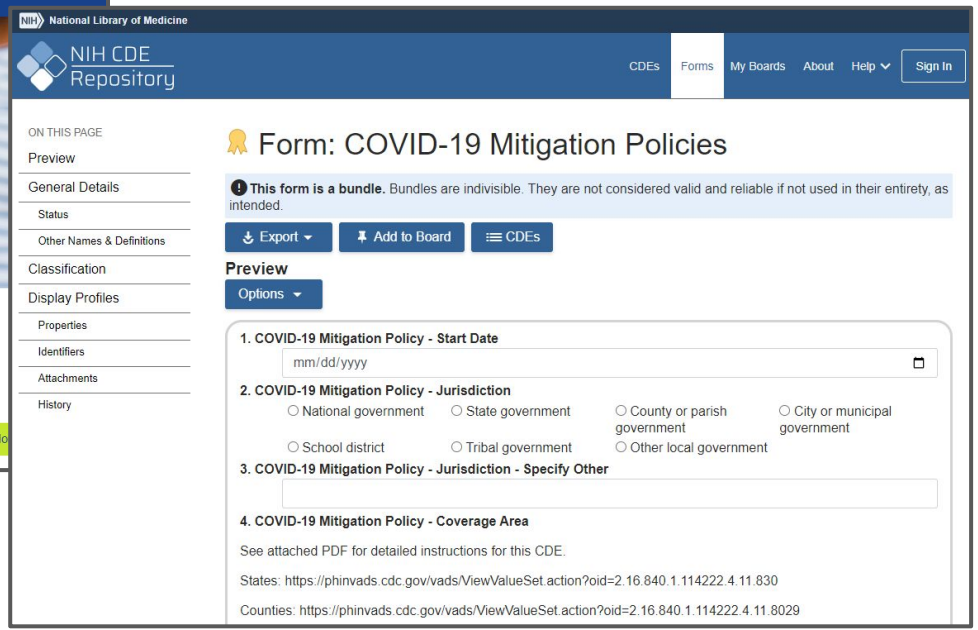
# Introduction to Common Data Elements (CDEs)

Megan Chenoweth

RESEARCH DATA STEWARDSHIP INITIATIVE

# Introduction and Background



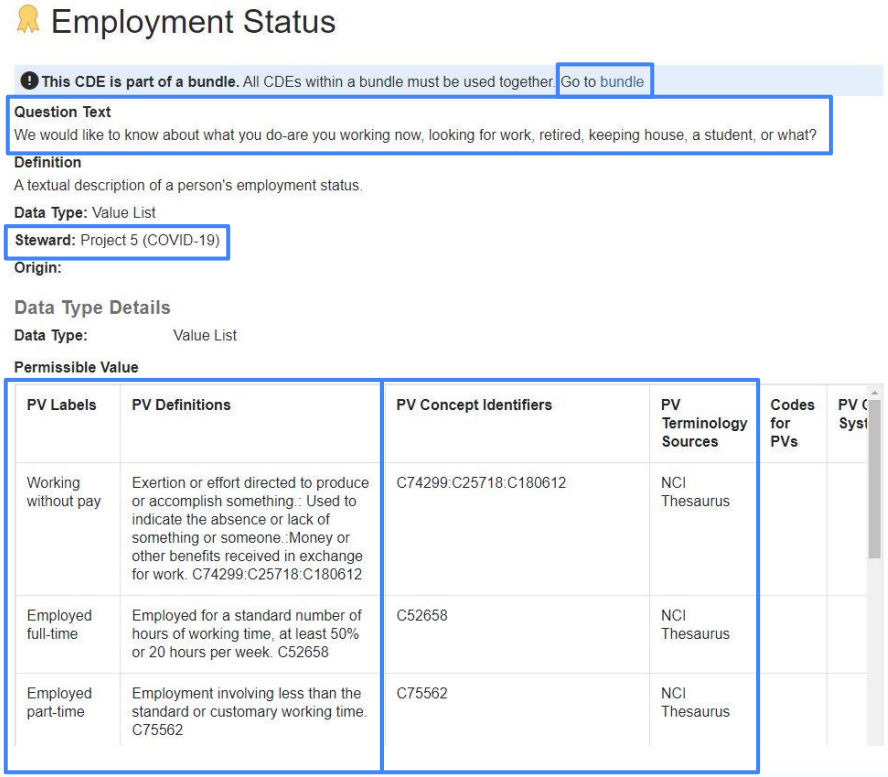



# Defining Common Data Elements

- Standardized questions, variables, or measures
- Specific sets of allowable responses
- Used across multiple studies to ensure consistent data collection

NIH CDE Repository: https://cde.nlm.nih.gov

# Common Data Elements: An Example



## Employment Status

**This CDE is part of a bundle.** All CDEs within a bundle must be used together. Go to bundle

**Question Text**
We would like to know about what you do-are you working now, looking for work, retired, keeping house, a student, or what?

**Definition**
A textual description of a person's employment status.

**Data Type:** Value List

**Steward:** Project 5 (COVID-19)

**Origin:**

### Data Type Details

**Data Type:** Value List

**Permissible Value**

| PV Labels | PV Definitions | PV Concept Identifiers | PV Terminology Sources | Codes for PVs | PV Syst |
|---|---|---|---|---|---|
| Working without pay | Exertion or effort directed to produce or accomplish something.: Used to indicate the absence or lack of something or someone.:Money or other benefits received in exchange for work. C74299:C25718:C180612 | C74299:C25718:C180612 | NCI Thesaurus | | |
| Employed full-time | Employed for a standard number of hours of working time, at least 50% or 20 hours per week. C52658 | C52658 | NCI Thesaurus | | |
| Employed part-time | Employment involving less than the standard or customary working time. C75562 | C75562 | NCI Thesaurus | | |

# Use of Common Data Elements

- Incorporate CDEs into research design
- Often due to funding requirements or other mandates
- Most common in health sciences research, but growing support in the social sciences
- Exact prevalence is unknown, but growing

CDEs Benefits:

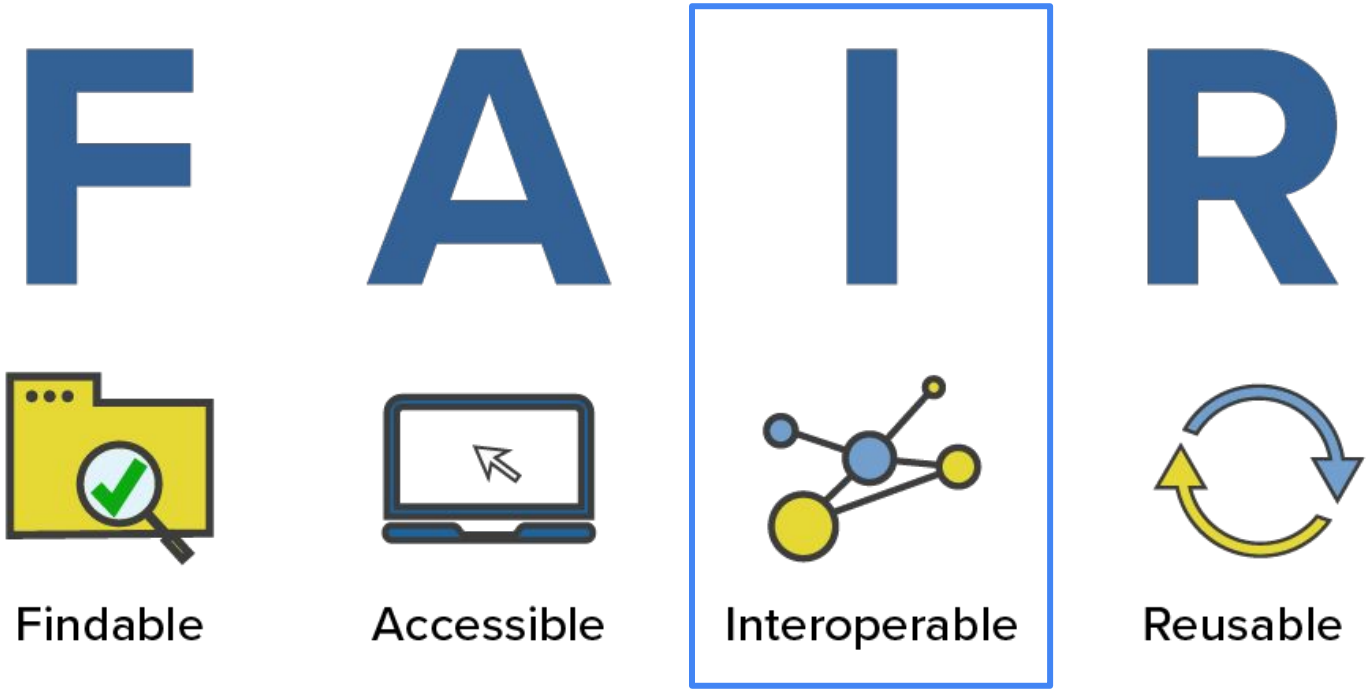
- Facilitate data aggregation and meta-analyses
- Simplify training and operations
- Improve efficiency
- Enhance quality of data collection

Sheehan J, Hirschfeld S, Foster E, Ghitza U, Goetz K, Karpinski J, Lang L, Moser RP, Odenkirchen J, Reeves D, Rubinstein Y, Werner E, Huerta M. Improving the value of clinical research through the use of Common Data Elements. Clin Trials. 2016 Dec;13(6):671-676. doi: 10.1177/1740774516653238. Epub 2016 Jun 15. PMID: 27311638; PMCID: PMC5133155.

# Where Do CDEs Come From?

1. Originate in research studies
    - May be validated instruments or scales
    - May originate from other sources
2. Identified as possible CDEs via a consensus building process
    - Working groups
    - Revision and feedback
3. Submitted to and reviewed by the NIH CDE Governance Committee

# CDEs and FAIR Data Principles



# CDEs: Learning and Resources

- NLM CDEs repository: NIH Common Data Elements (CDE) Repository
- Other repositories:
  - PhenX: PhenX Toolkit
  - RADx-UP: NIH RADx-UP Common Data Elements
  - NINDS: NINDS Common Data Elements
- On-demand training from NLM:
  - Standardize Your Research Data with the NIH Common Data Element Repository – On Demand | NNLM
  - Common Data Elements: Standardizing Data Collection | NNLM
- Article: Sheehan et al (2016). Improving the value of clinical research through the use of Common Data Elements. https://doi.org/10.1177/1740774516653238

UNIVERSITY OF MICHIGAN

# The HEAL Common Data Elements

Dr. David Williams

RESEARCH DATA STEWARDSHIP INITIATIVE

# The HEAL Common Data Elements

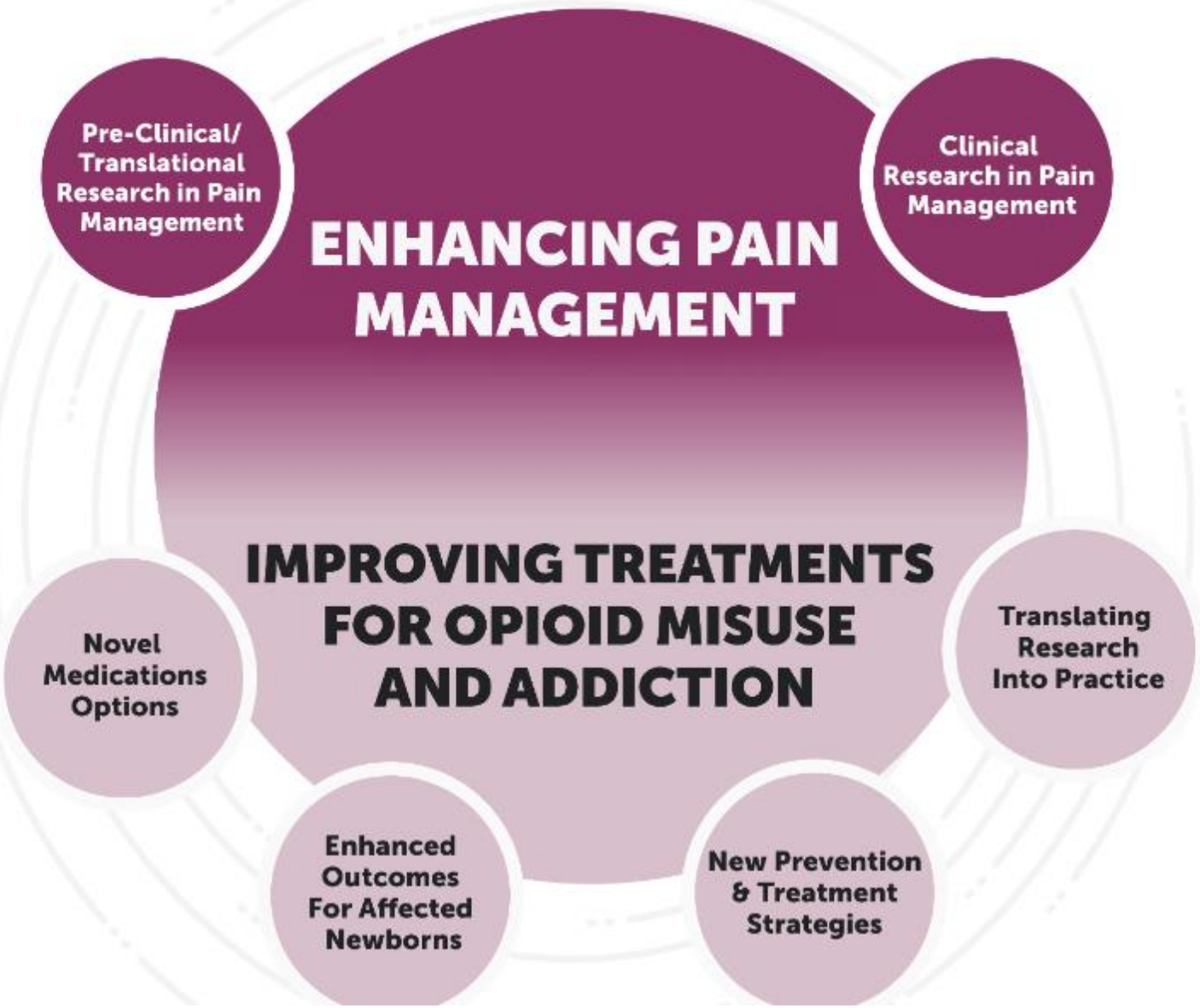
# RDSI Community of Practice Event: The ABCs of CDE's

Oct 24, 2024


**David A. Williams, Ph.D.**

Professor of Anesthesiology, Medicine, Psychiatry and Psychology
Associate Director, Chronic Pain and Fatigue Research Center
Co-Director, Research Development, Michigan Institute for Clinical Health Research (MICHR)
Director, Network-based Research Unit, MICHR
University of Michigan Medical Center
Ann Arbor, Michigan

# NIH HEAL INITIATIVE RESEARCH OVERVIEW



# HEAL Research to Enhance Pain Management Across the Scientific Spectrum

| Basic Science | Preclinical | Clinical | Implementation/ Dissemination |
|---|---|---|---|
| **Drug Discovery**<br>New targets for non-opioid medications | | | **Preventing Chronic Pain**<br>Identify risk for developing chronic pain |
| | **Pre-Screening Platforms**<br>Novel molecules, biologics, devices to relieve pain | | **Clinical Studies**<br>Determine safety and efficacy of new treatments |
| | **Therapeutics Development**<br>Develop or improve drugs & devices to relieve pain | | **Effectiveness Trials**<br>Define and measure effectiveness for various populations and settings |
| | **Biomarkers & Endpoints**<br>- Accurate pain measurement<br>- Patient-centered outcomes | | **Pragmatic Trials**<br>Integrating treatments into real-world clinic practice |

# HEAL CDEs

USASP
US ASSOCIATION FOR THE STUDY OF PAIN

PUBLISHED BY
ELSEVIER

The Journal of Pain, Vol 23, No 3 (March), 2022: pp 370–378
*Available online at www.jpain.org and www.sciencedirect.com*

Focus Article

## NIH's Helping to End Addiction Long-term℠ Initiative (NIH HEAL Initiative) Clinical Pain Management Common Data Element Program

Laura Dover Wandner,* Anthony F. Domenichiello,† Jennifer Beierlein,* Leah Pogorzala,* Guadalupe Aquino,‡ Andrew Siddons,* Linda Porter,* and Jane Atkinson‡
NIH Pain Consortium Institute and Center Representatives

# Core – Adult CDEs

## Adult Acute Pain*

| Pain Intensity | Pain Interference | Physical Functioning/ QOL | Sleep | Pain Catastrophizing | Depression | Anxiety | Global Satisfaction with Treatment | Substance Use Screener |
|---|---|---|---|---|---|---|---|---|
| BPI Pain Severity | BPI Pain Interference | PROMIS Physical Functioning Short Form 6b | PROMIS Sleep Disturbance 6a + Sleep Duration Question | Pain Catastrophizing Scale – Short Form 6 | PHQ-2 | GAD-2 | PGIC | TAPS 1 |

## Adult Chronic Pain*

| Pain Intensity | Pain Interference | Physical Functioning/ QOL | Sleep | Pain Catastrophizing | Depression | Anxiety | Global Satisfaction with Treatment | Substance Use Screener |
|---|---|---|---|---|---|---|---|---|
| PEG | | PROMIS Physical Functioning Short Form 6b | PROMIS Sleep Disturbance 6a + Sleep Duration Question | Pain Catastrophizing Scale – Short Form 6 | PHQ-2 | GAD-2 | PGIC | TAPS 1 |

*Questions are required to be asked at two time points
*Monitoring opioid use will be required by HEAL, however, the method by which it will be assessed is still pending. PIs will be asked to monitor opioid use (including dosage) by appropriate PRO, EHR, or other measures.

Wandner, L. HEAL Pain Common Data Elements (CDE Program. NIH Heal Initiative. (slide media). April 12, 2022. Laura.Wandner@NIH.gov

# Core – Pediatric CDEs

## Pediatric Acute and Chronic Pain*

| | Pain Intensity | Pain Interference | Physical Functioning / QOL | Sleep | Pain Catastrophizing | Depression | Anxiety | Global Satisfaction with Treatment | Substance Use Screener |
|---|---|---|---|---|---|---|---|---|---|
| Child | BPI Pain Severity | BPI Pain Interference | PedsQL Inventory | AWS + Sleep Duration Items | Pain Catastrophizing Scale for Children | PHQ-2 | GAD-2 | PGIC | NIDA Modified Assist Tool - 2 |
| Parent | | | | | Pain Catastrophizing | PHQ-2 | GAD-2 | | |

Wandner, L. HEAL Pain Common Data Elements (CDE Program. NIH Heal Initiative. (slide media). April 12, 2022.
Laura.Wandner@NIH.gov

# Re-released* The Core Demographics (Adult and Pediatric) - Domains

- Date of Birth
- Age
- Sex at Birth
- Gender Identity
- Ethnicity, Race
- Highest Level of Education
- Employment Status
- Relationship Status
- Annual Household Income
- Applied for Disability Insurance
- Pain Duration

*Demographics are, for the most part, CDISC compliant

Wandner, L. HEAL Pain Common Data Elements (CDE Program. NIH Heal Initiative. (slide media). April 12, 2022. Laura.Wandner@NIH.gov

# Supplemental Questionnaires

## What are supplemental CDEs?

A comprehensive set of screening tools selected by HEAL PIs for use in their studies. They may be unique to the study; e.g., pain condition-specific or study-relevant measures.

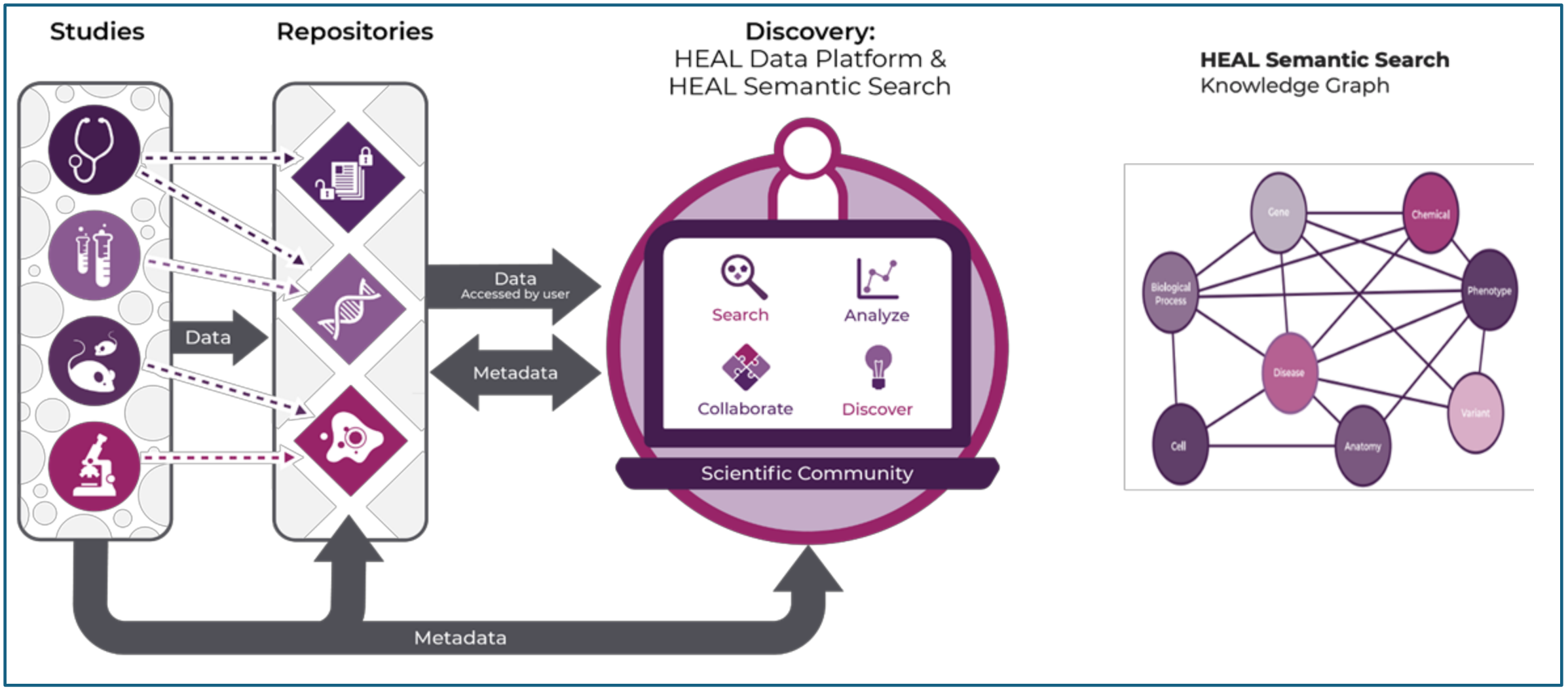
The NIH has collected supplemental questionnaires from all clinical pain programs and/or studies. ~600 distinct supplemental questionnaires have been accepted into the HEAL CDE program for this initiative.

**Supplemental CDEs are <u>not required for the HEAL funded pain studies</u>, but the benefits include:**

- See questionnaires being used within similar studies
- Coding and collecting data (in addition to the core CDEs) for the HEAL Data Ecosystem

| | | | | |
|---|---|---|---|---|
| The Migraine Disability Assessment (MIDAS) questionnaire assesses headache-related disability with the aim of improving migraine care. | • File midas-crf.docx (23.14 KB)<br>• File midas-cde.xlsx (21 KB) | English | • Supplemental | Headache/Migraine |

# HEAL Data Ecosystem Strategy and Tools



UNIVERSITY OF MICHIGAN

# So You Made a CDE: Now What?

Dr. James McNally

RESEARCH DATA STEWARDSHIP INITIATIVE

# So you made a CDE: Now what?

**James. W. McNally**
**NACDA Program on Aging**

**jmcnally@umich.edu**

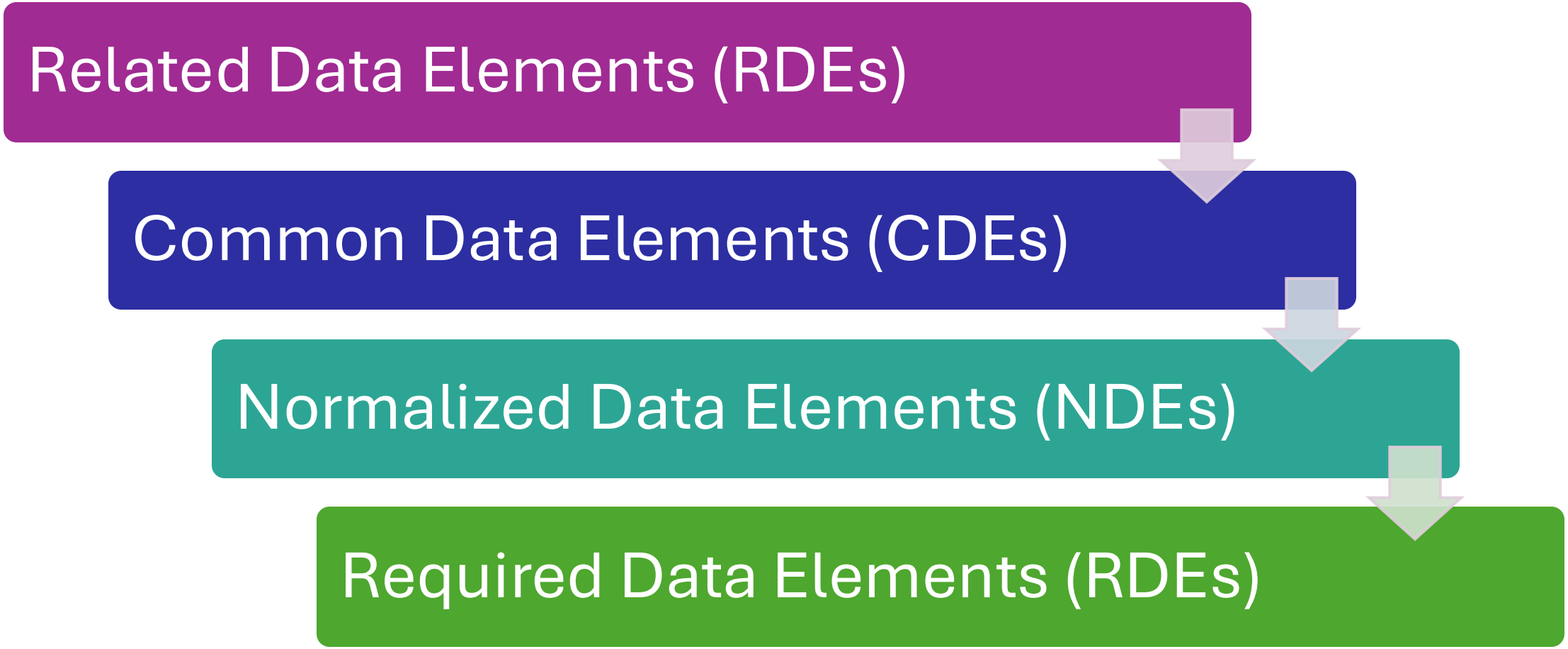
***RDSI Community of Practice Event:***
***The ABCs of CDEs***

***October 24, 2024***

# ABSTRACT


- Common Data Elements (CDEs) is a general term used to capture a wide array of constructs.
- CDE repositories are often a broad catalogue of variable that have been used to measure an outcome with little or no guidance.
- It may only be a search engine that lists related variables.
- The provenance of many CDEs is uncertain.
- Steps are being taken to bring more formality into the way CDEs are organized and presented.
- Much more work needs to be done.

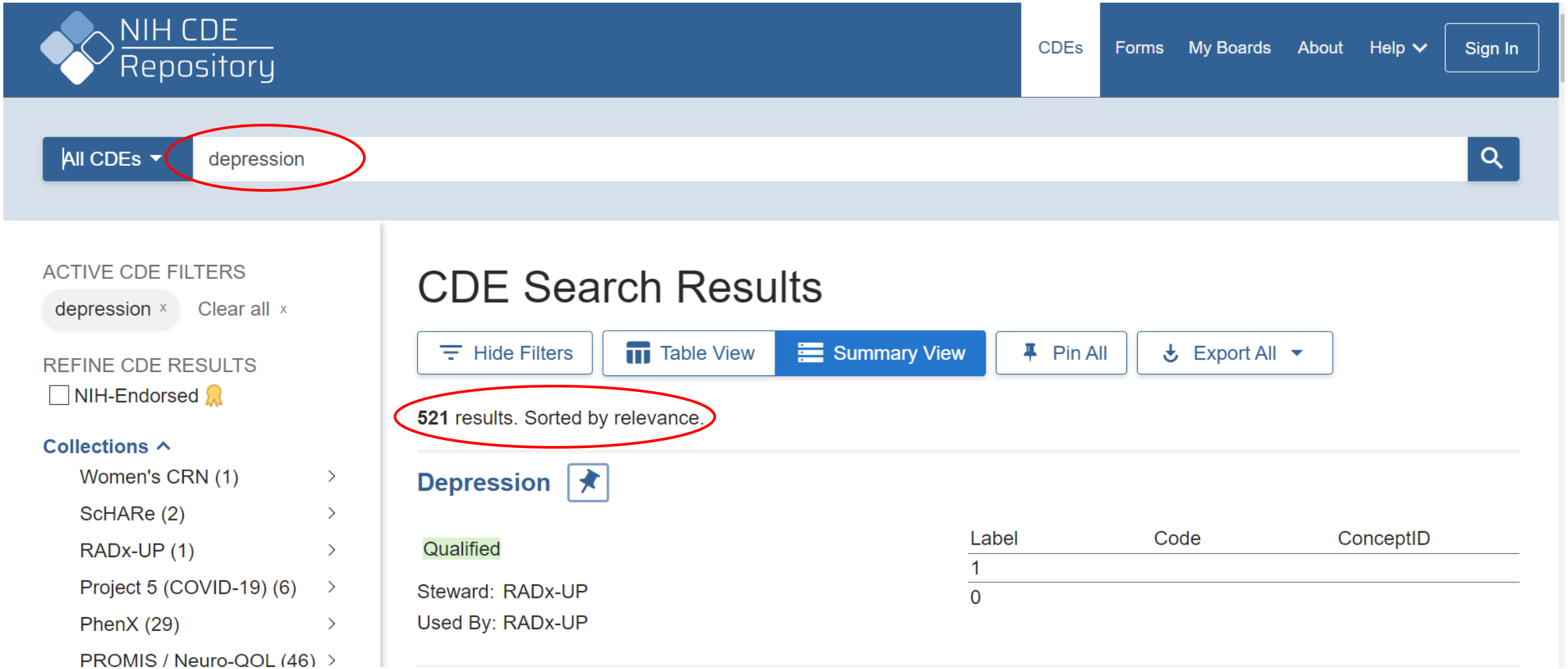
# IDEAL VARIABLE TRANSITION STAGES



# Related Data Elements (RDEs)

- RDEs represent the majority of variables, including those maintained in a “CDE” repository. They are a centralized catalogue of variables.
- RDEs are connected by a general measurement structure that seek to identify specific outcomes or behaviors.
- They rarely share a One-to-One comparability.
- They may employ unrelated scales, collection metrics, definitions and variable ranges.
- Anyone can create a RDE and potentially add it to a CDE repository.
- There is no requirement that any specific measure be used.

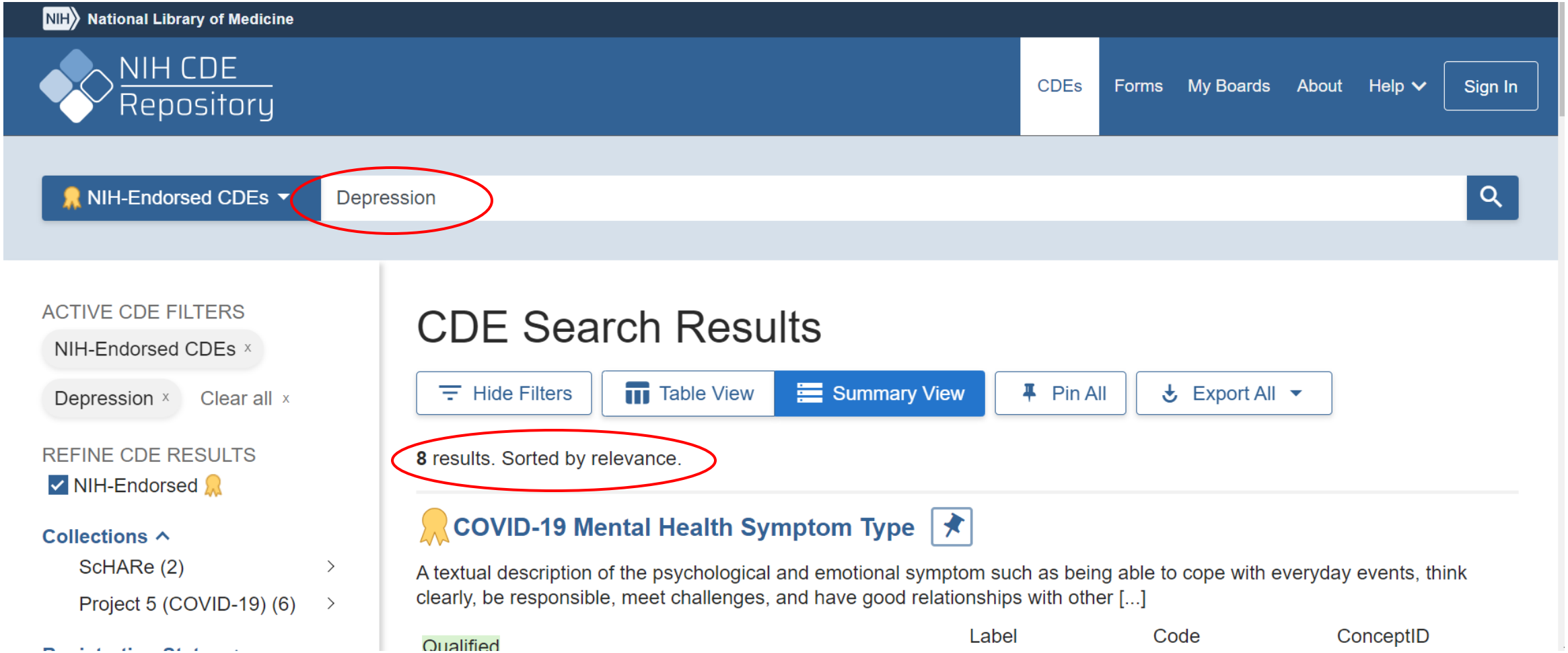
# RDE example in NIH CDE Repository



# Common Data Elements (CDEs)

- CDEs represented a structured refinement of RDEs.
- CDEs share similar, if not identical, variable elements.
- CDEs lend themselves to harmonization across variables and studies.
- Content experts can validate their similarities in a systematic manner and emphasize their value as shared constructs.
- There is no requirement to use specific CDEs and they are often aggregated within larger RDE collections.
- CDEs need to be formally separated from RDEs to encourage use.

# NIH-Endorsed CDEs: An example



# CDE Variable Comparisons – Match Types

**Directly Comparable:**
No manipulation of data needed for comparison.
Question phrasing and values match.

| NSHAP | NHATS |
|---|---|
| CON_WEIGHT | CURRWEIGH |
| Weight (lbs) | How much do you currently weigh? |
| Actual weight in pounds | Actual weight in pounds |

**Need Harmonization:**
Variables must be adjusted to be compared.
Questions more/less match, but values differ.

| CON_GLASSES | GLASSESWR |
|---|---|
| Did respondent wear glasses to read chart? | Do you wear glasses or contacts to help you see things at a distance? |
| 1 no | |
| 2 no, but R usually wears glasses/contacts for distance | 1 yes |
| 3 yes | 2 no |

**One to Many / Many to One:**
Multiple variables in one series combine to match a single variable in another.

| CON_GUPPROB_4 | CANE | WALKER |
|---|---|---|
| Respondent uses a cane or walker. | Respondent used a cane. | Resondent used a walker. |
| 0 no | 1 yes | 1 yes |
| 1 yes | 2 no | 2 no |

**Related Concepts:**
Subject matter is related, but variables do not precisely measure the same concept.

| CON_GASP | PROBBREAT |
|---|---|
| Has anyone ever told you that you stop breathing or gasp for breath during sleep? | In the last month, did you have any breathing problems, including shortness of breath or difficulty breathing? |
| 0 no | 1 yes |
| 1 yes | 2 no |

# Normalized Data Elements (NDEs)

- NDEs represent a further refinement of the CDE construct.
- NDEs are variables whose use has have achieved general consensus within a discipline as valid and generalizable to the broader population.
- Demography has many NDEs, ASDR, Index of Dissimilarity, Life Tables.
- Physiatry also has many NDEs, Kessler 6, PHQ-9, CES-D.
- Gerontology also has NDEs, ADLs IADLs, Ecog, Barthel Scale.
- You use these NDEs they way they are presented.
- The PhenX Toolkit is an example of an emerging NDE system.

# ADL and IADL Checklist

- ADL and IADL represent a generally agreed upon series of questions that capture function and independent living status.

**ADL Checklist**

| ADL Function | Independent | Needs Help |
|---|---|---|
| Bathing | | |
| Dressing | | |
| Transferring, e.g., from bed to chair | | |
| Toileting | | |
| Grooming | | |
| Feeding oneself | | |

**IADL Checklist**

| ADL Function | Independent | Needs Help |
|---|---|---|
| Using the telephone | | |
| Preparing meals | | |
| Managing household finances | | |
| Taking medications | | |
| Doing laundry | | |
| Doing housework | | |
| Shopping | | |
| Managing transportation | | |

# Example of NDEs encouraged by NIH funding

# Required Data Elements (RDEs)

- RDEs represent the most formalized and structured examples of variable classification.
- RDEs are comprehensive catalogs used as a navigational tool for healthcare professionals, researchers, and patients.
- RDEs are routinely reviewed by content experts and updated as science changes and formal definitions change.
- You use RDEs if you want to be published or paid.
- The International Classification of Disease (ICD) standardized codes and the Diagnostic and Statistical Manual of Mental Disorders (DSM) are examples of RDE systems.

**The International Classification of Diseases (ICD) is first published in 1893 by the World Health Organization (WHO) as the ICD-1.**

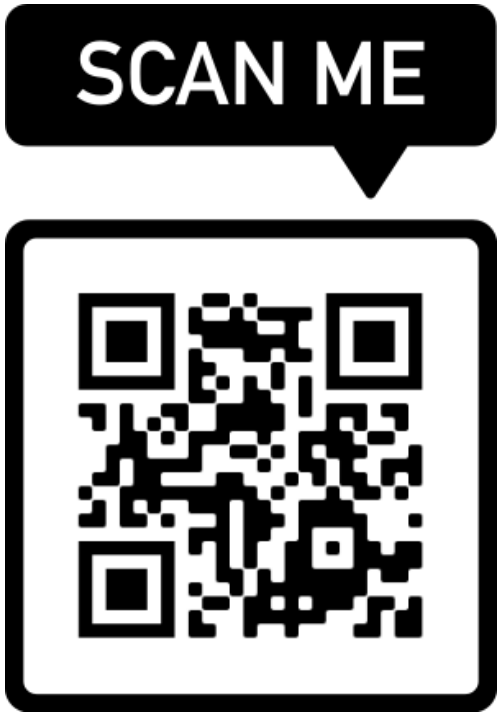
RDEs are not static but change is structured.

| ICD-9 | ICD-10 |
|---|---|
| 3-5 characters in length | 3-7 characters in length |
| Approximately 13,000 codes | Approximately 68,000 available codes |
| First digit may be alpha (E or V) or numeric; digits 2-5 are numeric | Digit 1 is alpha; digits 2 and 3 are numeric; digits 4-7 are alpha or numeric (alpha digits are not case sensitive) |
| Limited space for adding new codes | Flexible for adding new codes |
| Lacks detail | Very specific |
| Lacks laterality | Has laterality (i.e., codes identifying right vs. left side of the body) |
| Use same code for every visit | Has possibility of identifying initial encounter, subsequent encounter; or sequela |
| Only 4 codes were reported on a claim form | Up to 12 codes can be reported on a claim form |

# Conclusions

- Our current system of CDEs is an unstructured mess.
- The use of specific variables is optional in all but the oldest, and most formalized variable systems.
- There is a structured framework we can use to think about “CDEs” in a more structured manner.
- Fitting a variable within this framework makes it possible to rigorously plan and identify metrics to encourage the use of one variable versus another variable.
- Its not a CDE if you are the only one using it, even if it is in a repository.

# Thank you! Questions?




Scan this QR Code to get to our LinkTree (https://linktr.ee/nacdadata):

- NACDA website
- NACDA Colectica Portal
- YouTube playlist (researcher interviews and presentation recordings)
- Newsletter signup
- NIH data sharing resources


- **jmcnally@umich.edu**

UNIVERSITY OF MICHIGAN

# Q & A

RESEARCH DATA
STEWARDSHIP INITIATIVE

# Event evaluation

Please fill out our survey:

https://myumi.ch/W5Z2D

UNIVERSITY OF MICHIGAN

# Thank you for attending!

CoP co-chairs:

- Sara Samuel: henrysm@umich.edu
- Megan Chenoweth: mmchenow@umich.edu

Learn more about RDSI:
https://research.umich.edu/research-data-stewardship/

RESEARCH DATA STEWARDSHIP INITIATIVE