

Demystifying Data Sharing

Moderator: Keith Marsolo
Electronic Health Records Core



**NIH PRAGMATIC TRIALS
COLLABORATORY**

Rethinking Clinical Trials®

Topics



- The process NIH Collaboratory Trials used to select a data repository
- Challenges encountered when using repositories
- Tips to streamline depositing data

We will hear from...



Diana Burgess
RAMP



Andrea Cheville
NOHARM



Andrea Cook
BackInAction



Reaching Rural Veterans: Applying Mind-Body Skills for Pain Using a Whole Health Telehealth Intervention (RAMP)

Diana J. Burgess, PhD

Core Investigator, Center for Care Delivery and Outcomes Research, Minneapolis VAHCS

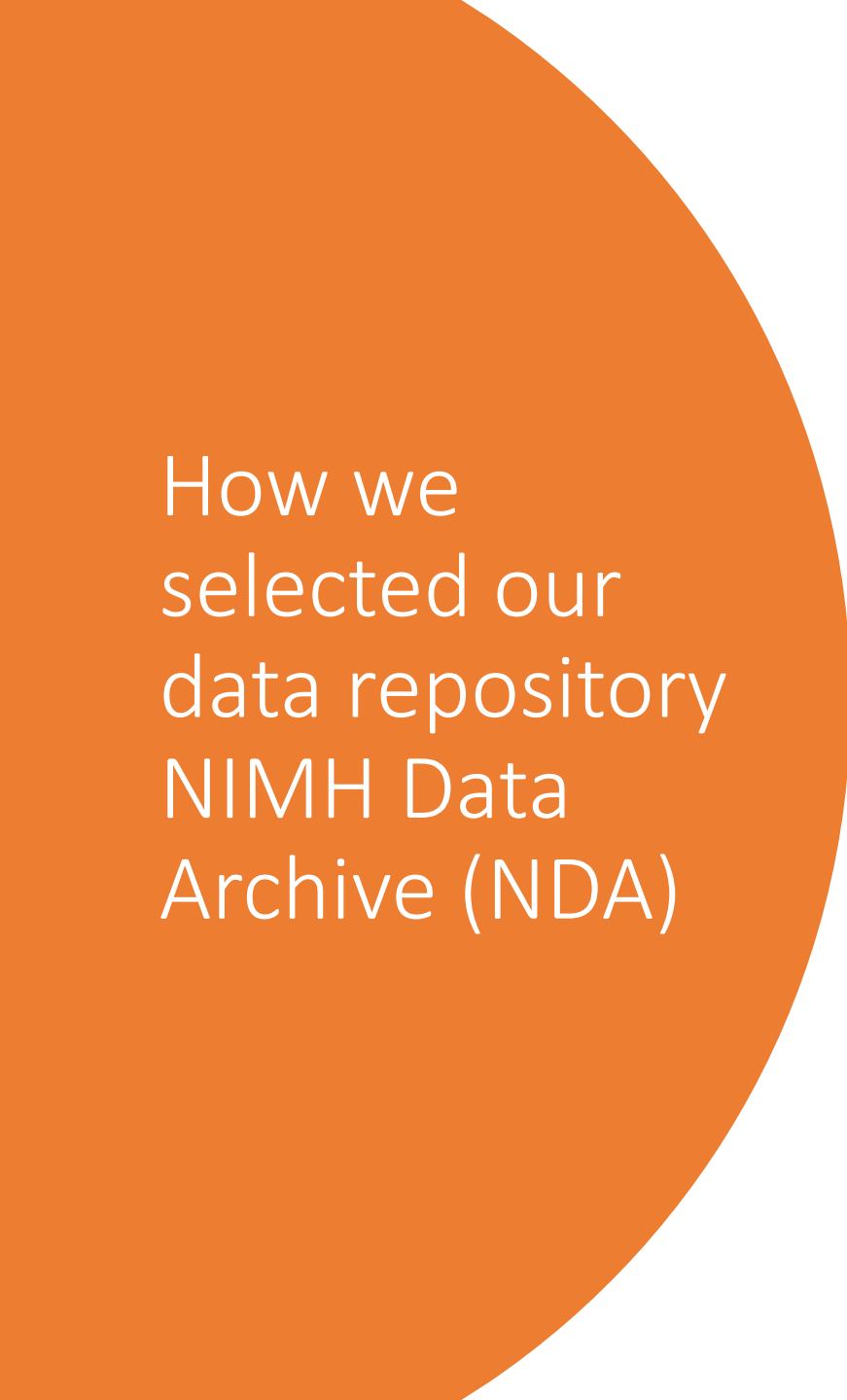
Professor, University of Minnesota Medical School

Director, VA QUERI Complementary and Integrative Health Evaluation Center (CIHEC)


Disclosures

- No relevant financial relationships to disclose.
- Funding from the National Institute of Nursing Research (NINR), as part of the NIH HEAL Initiative (1UG3NR020929-01), with resources from the Minneapolis VA Health Care System and Iowa City VA Health Care System.
- The views expressed in this talk are my own and are not necessarily endorsed by the U.S. Department of Veterans Affairs, or the United States Government.

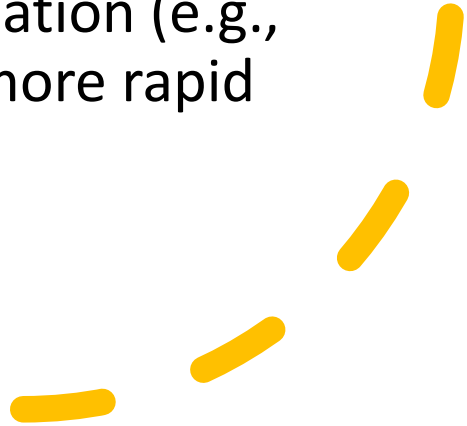


A large orange circle on the left side of the slide, partially cut off by the edge.

How we selected our data repository NIMH Data Archive (NDA)

- Context: UG3 transition report submitted 4/28/2025
 - Choices: Vivli, NDA, and ICPSR
 - Wanted repository that was easy for future analysts to access & would work with VA restrictions
 - Selected NIMH Data Archive (NDA) as our data repository because it:
 - Accepts HEAL-funded data free for 120 studies (Vivli has a fee of \$4,000 starting 2025)
 - Provides support including curation and harmonization services (Vivli has support for free but curation for a fee and ICPSR is free but no support).
- 
- A decorative yellow dashed line in the bottom right corner, consisting of several curved segments.

Challenges

- Time it takes to attend seminars, discuss with team and weigh the different options (pros/cons)
 - Personnel resources are needed on other time-sensitive project tasks
 - Opportunity costs making this decision that could have been better used on other project tasks.
 - Navigating process for securing institutional approvals
 - Would have liked more concise presentation of repository options with summary information (e.g., bottom line), so we could have made a more rapid decision.
- 

NOHARM: Nonpharmacological Options in post-operative Hospital- based And Rehabilitation pain Management pragmatic trial



Andrea Cheville, MD, MSCE
Professor of Physical Medicine and Rehabilitation
Mayo Clinic

Jon Tilburt, MD
Professor of Medicine
Mayo Clinic

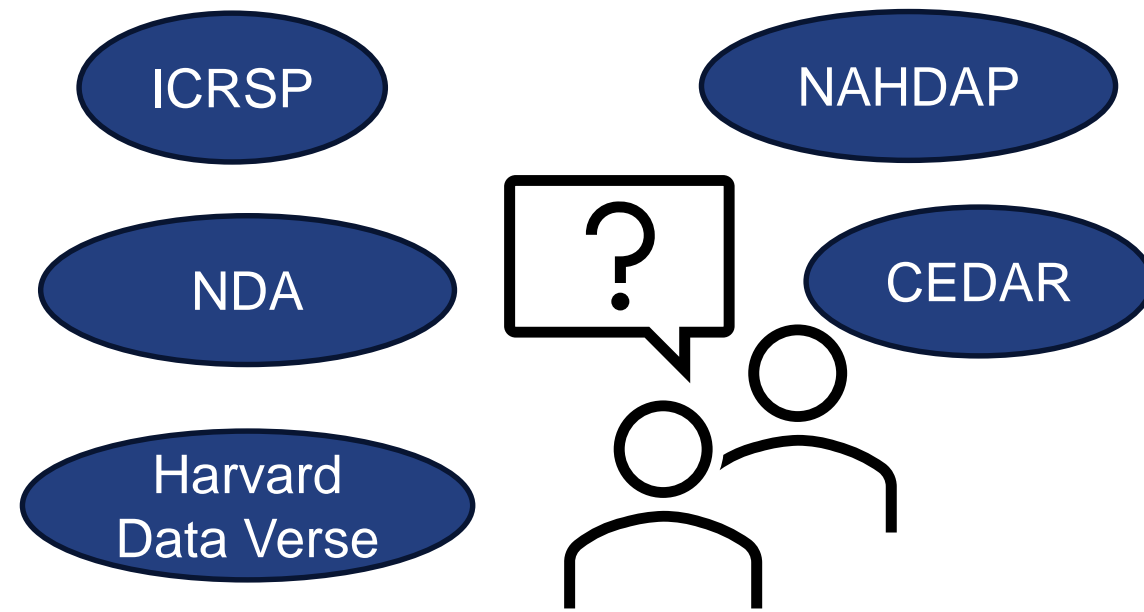


**NIH PRAGMATIC TRIALS
COLLABORATORY**

Rethinking Clinical Trials®

Selecting a repository

- Costs
 - Upload(s)
 - Curation
 - ICRSP \$15K
 - Maintenance
- Institute, award, data type, etc. specific
- Formatting specifications
 - Differences between analytic data set and repository requirements
 - Challenging to ascertain
- Research team experience and understanding
 - Don't know what we don't know
 - Point person may require repository-specific training
- Support
 - May be at repository-, institutional-, mechanism-, and/or consortium-level



HEAL-compliant Repository Selection Guide

<https://www.healdatafair.org/resources/guidance/selection>

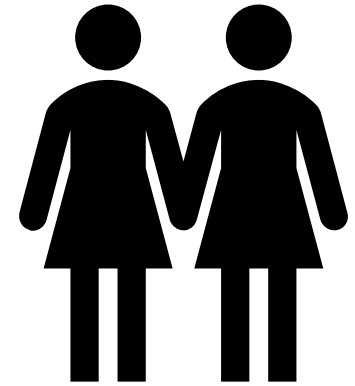
Repository	Descriptive Tags	Organism	IC/Program	Links
NIMH Data Archive	Human Subjects, Imaging	Human	NIMH	Guidance¹ Overview
NICHD DASH	Clinical Trial, Biospecimens	Human	NICHD ²	Guidance Overview
NAHDAP	Human Subjects	Human	Some NIDA studies	Guidance Overview
NIDA Data Share	Clinical Trial	Human	NIDA CTN ²	Guidance Overview
NIDDK Central Repository	Human Subjects, Biospecimens	Human	NIDDK ²	Guidance Overview
SPARC	Brainstem, Spinal Cord, Peripheral Nerv...	Unrestricted	RE-JOIN, PRECISIO...	Guidance Overview
Pennsieve	Generalist, Imaging, Microphysiology, O...	Unrestricted		Guidance Overview
ICPSR	Social and behavioral data	Human		Guidance Overview
Vivli	Clinical Trial, Generalist	Human		Guidance Overview
...	Guidance Overview

1 row selected

Rows per page: 100 1-29 of 29

Leveraging support resources

- Identifying sources of support
 - Email announcements
 - Conferences
 - Annual HEAL meeting
 - HEAL Data Ecosystem
 - Web sites
 - Searchable interfaces, <https://www.healdatafair.org/resources/guidance/selection>
 - Colleagues
 - Collaboratory
 - Program officers
- Modes of support
 - Email exchange
 - HEAL Data Stewards Webinars
 - Training videos
 - Webinars
 - Telephone & video conference
- Differentiating across resources
 - General vs. niche
 - Commodity specific-, e.g. metadata, dictionary
 - “Step” specific, e.g., form completion, curation, transfer



What and when to share

- CT.gov
 - \leq 12 months following
 - Extension possible but challenging to obtain
- Pilot data
 - How much to invest?
 - Harvard Data Verse
- Data dictionary
 - Pilot vs. analytic vs. repository data set
- Meta data
 - HEAL stewards will curate metadata
 - CEDAR, <https://metadatacenter.org/>
 - Study-level Metadata
 - Variable-level Metadata
 - Common Data Elements
- Pre-/post- publication
 - Staggered approach
 - Primary and secondary outcome reporting
 - Exploratory analyses
 - Data not included in repository

BackInAction

- Repository Choice: NIMH Repository
 - Wanted a repository to assure data was being used for research purposes per consent language
 - Wanted repository to restrict to non-profit use for research per consent
- Challenges:
 - One of four of our Health Care System site did not have proper consent language so could not submit their data to a public repository
 - Did not submit acupuncturist IDs since we did not inform community accupunturist that we may share their data (cannot replicate analysis but so little correlation within acupuncturist (ICC<0.001) not large implications)
 - Re-map data to requirements was very time consuming and resource intensive to assure data was accurate and deidentified (e.g. repository requires a lot of dates so to de-identify we moved those dates to be 01/01/Year to conform to deidentification by year)
 - Analytic Code for analysis cannot be used on public datasets given they change names of almost all variables (we could upload by creating aliases, but output is their naming convention) so less useful for future researchers
- Tips:
 - Choose Repository early so that naming convention and formatting align
 - Review Consent forms across all sites to assure data sharing is appropriately outlined (standardized consent forms ideally).

Discussion



**NIH PRAGMATIC TRIALS
COLLABORATORY**

Rethinking Clinical Trials®