# DSDR, A Population Health Data Archive: 20 Years of Sharing Research Data and Resources

John E Marcotte, Sarah Rush, & Kelly Ogden-Schuette DSDR, ICPSR, University of Michigan







## The DSDR Team



John E Marcotte
Project Director



Sarah Rush
Data Project Manager



Kelly Ogden-Schuette

Data Project Manager



**Funded by** the Population Dynamics Branch (PDB) of the *Eunice Kennedy Shriver* National Institute of Child Health and Human Development (NICHD), National Institutes of Health (NIH).

Hosted by ICPSR at the University of Michigan.



**DSDR** is a platform for the discovery, access, and sharing of research data for population health studies.

**DSDR** is a locus of information for researchers who collect, analyze, and distribute primary data as well as researchers who analyze secondary data.

**DSDR** is 20 years old in 2024!



## **Discover**

- Data on topics related to:
  - Maternal & Child Health
  - Health Disparities
  - Families & Households
  - Migration & Immigration
  - Cross-national comparisons

### Access

- Download public research data
- Online Analysis
- Apply for access to restricted data

## **Share**

- Deposit data
- Meet NIH Data Sharing requirements
- Provenance
- Preservation

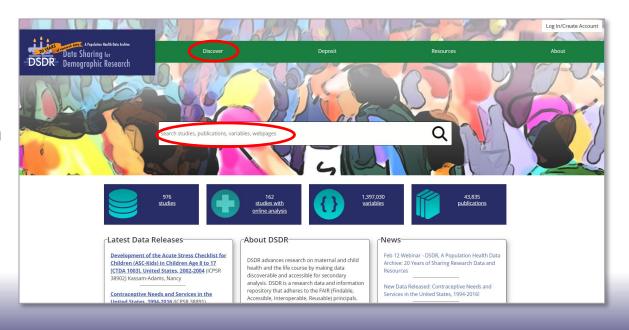


## **Discover Data & Resources**



## Finding Data on the DSDR website

- Searching across studies
  - Study
  - Variable Comparison
  - Publications
- Searching within studies from study home page





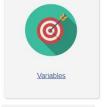
## **Discover Page**

#### Browse by

















#### **Browse by Subject**











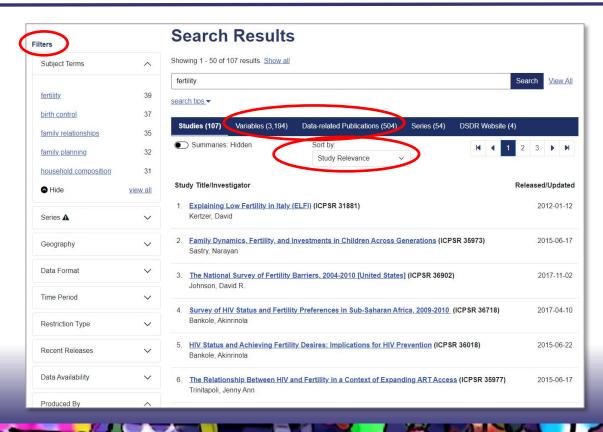






view less ^

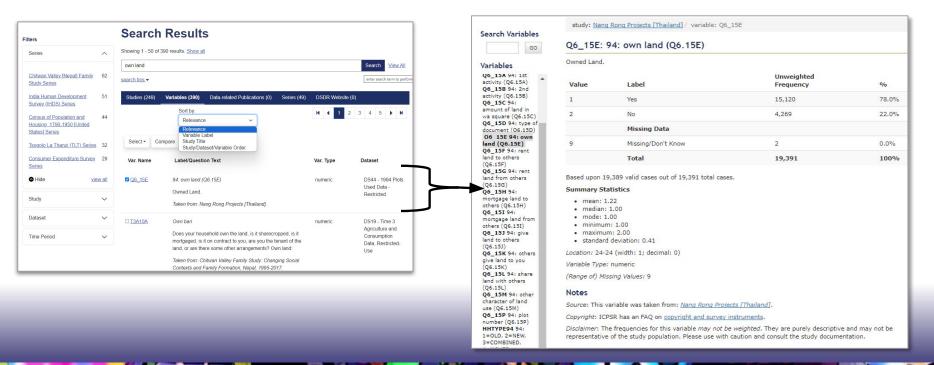




- Searches data holdings
- Switch to Variable or Publication results for search term
- Sort options available
- Narrow results using filters
- Revise search

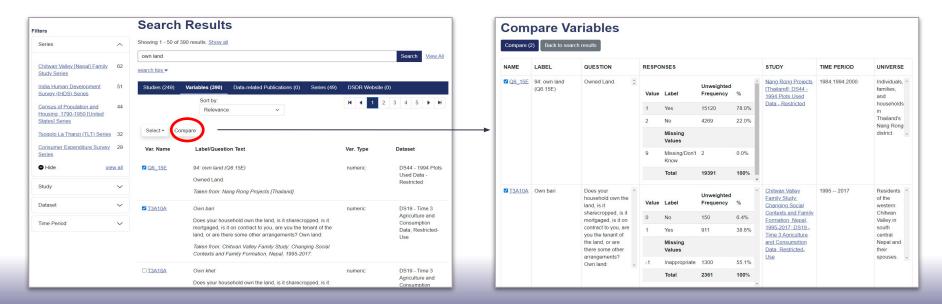


## Variable Search



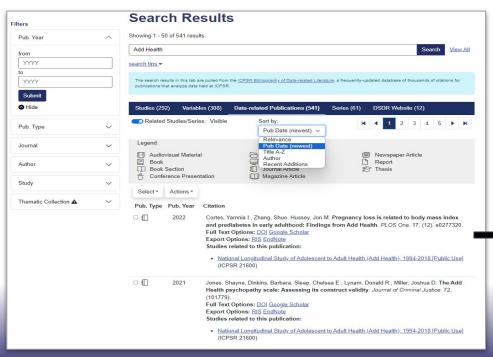


## Variable Comparison





## **Publication Search**



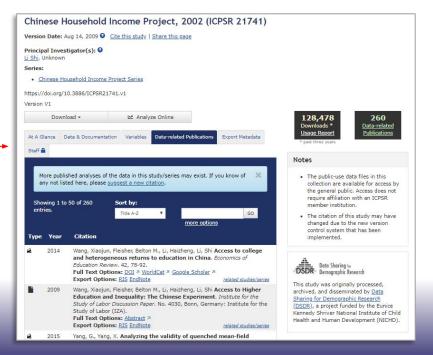
Cortes, Yamnia I., Zhang, Shuo, Hussey, Jon M. Pregnancy loss is related to body mass index and prediabetes in early adulthood: Findings from Add Health. PLOS One. 17, (12), e0277320. Full Text Options: DOI Google Scholar Export Options: RIS EndNote Studies related to this publication:

 National Longitudinal Study of Adolescent to Adult Health (Add Health), 1994-2018 [Public Use] (ICPSR 21600)



## Within Study Searching from the Study Home Page





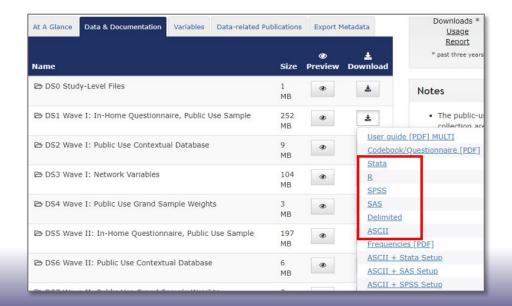


## **Access Data & Resources**



### Download public-use microdata files in multiple formats:

- Stata
- R
- SPSS
- SAS
- Delimited
- ASCII





## Options for Online Data Analysis<sup>†</sup>

#### Full Analysis Capabilities (powered by SDA)

- Wide range of analysis options
- Allows for recoding and computing new variables



#### Freqs/Crosstabs (Legacy)

Streamlined interface for building tables

India Human Development Survey-II (IHDS-II), 2011-12, Household		
Row	select row variable (required)	recode
Column	select column variable	recode
Control	select control variable	recode
Selection Filter	(9	Example: age(18-50)
Weight	No weight	•
Percentaging	☑ Column ☐ Row ☐ Total	
N of cases to display	☑ Unweighted ☐ Weighted	
Summary statistics	☐ Show summary statistics	
	Build the Table Reset	

<sup>†</sup> Available from Study Home page for select studies.



## Top Downloaded Studies<sup>†</sup>

- National Longitudinal Study of Adolescent to Adult Health (Add Health), 1994-2018 [Public Use]
- India Human Development Survey-I (2005), IHDS-II (2011-12), and the IHDS Panel (IHDS 2005, IHDS-II 2011-2012)
- Matlab [Bangladesh] Health and Socioeconomic Survey (MHSS), 1996
- The Mekong Island Population Laboratory (MIPopLab) (2000-06)
- Generations: A Study of the Life and Health of LGB People in a Changing Society, United States, 2016-2019
- TransPop, United States, 2016-2018



## **New Study Additions with LGBTQ Data**<sup>†</sup>

- Generations: A Study of the Life and Health of LGB People in a Changing Society, United States, 2016-2019
- National Couples' Health and Time Study (NCHAT), United States,
   2020-2021
- TransPop, United States, 2016-2018



## Other New Additions and Updates<sup>†</sup>

- Baby's First Years (BFY), New York City, New Orleans, Omaha, and Twin Cities, 2018-2021
- Los Angeles Family and Neighborhood Survey (L.A. FANS), Wave 2, Biomarker Data, 2006-2008
- Project Migrante: Health Status and Access to Health Care Among Migrants on Mexico's Northern Border, 2020-2021
- The New Immigrant Survey Series, United States, 2007-2009: Round 1 (NIS-2003-1) and Round 2 (NIS-2003-2) restricted-use data
- Tsogolo La Thanzi 2 (TLT-2), Malawi, 2015 [Healthy Futures]



## **Highlighted Health Data in DSDR**

- India Human Development Survey (IHDS) (2 studies and a Panel combining the two)
- Los Angeles Family and Neighborhood Survey (L.A.FANS) (14 studies)
- National Longitudinal Study of Adolescent to Adult Health (Add Health), 1994-2018 [Public Use]
- National Survey of Adolescents Series (4 studies)
- NICHD Study of Early Child Care and Youth Development (SECCYD)
   (4 studies)



#### **Documentation Available**

#### **Data Documentation**

- Detailed study descriptions
- Suggested data citation
- Survey instruments and data collection measures (when provided)
- Codebooks
- Frequencies
- Data guides available: Chitwan Valley Family Study Series (CVFS), India Human Development Series (IHDS and IHDS-II), and Add Health

#### Data-related Bibliography

- Journals, reports, books, book chapters, dissertations
- Searchable by study, author, title, year
- Over 43,800 publications related to studies in DSDR
- Full text options available for many journals and reports



## **Three Options for Accessing Restricted Data**

#1
ICPSR sends data to researchers via encrypted download

Researchers come electronically to data via ICPSR's Virtual Data Enclave (VDE)

#2

#3
Researchers come in person to data via ICPSR's Physical Data Enclave (PDE)

For all access methods, researchers must follow protocols for protecting human subjects.



## Share Data & Data-Related Resources



## Reasons to Deposit Data with DSDR

- Advance science by making data available for secondary analysis
- Meet funding requirements
- Long-term, secure preservation
- Dissemination: exposure to a broad range of audiences
- Persistent identifiers to aid in data discoverability

- Data curation to enhance & add value
- Disclosure Risk Review
- Bibliography of data-related citations
- Announcements publicizing data
- Ongoing support for researchers
- Deposit consultations



## **Deposit Options**

#### **Standard**

- Data and documentation deposited exclusively with DSDR
- DSDR becomes sole source of data
- Restricted data must be standard deposit

#### **Mirror**

- The same data and documentation are available from multiple archives; one authoritative source
- Public access only
- Reach international researchers who speak different languages

#### **Variable Metadata**

- Variables become searchable through DSDR while data and documentation remain at study site
- On-going studies with frequent updates

#### **Study Metadata**

- Union Catalog
- Data available from another source becomes discoverable through DSDR
- Researchers can search and compare studies



## **Data Submissions to DSDR**

DSDR acquires research data, study and variable-level metadata, scales and researcher syntax related to **maternal and child health**, **health disparities**, and the **human lifecycle** within the following priority areas set by NICHD:

#### #1

Data resulting from support by the NICHD Population Dynamics Branch (PDB)

#### #2

Data that provide opportunities for researchers working within the scientific scope of the NICHD PDB that were supported by sources other than the NICHD PDB

#### #3

Data resulting from support by the NICHD Division of Extramural Research from sources other than the PDB, with an emphasis on behavioral and social science research data



## Help with NIH Data Management & Sharing Plans

- Letters of Support (send requests to DSDR@icpsr.umich.edu)
- Available from the ICPSR <u>Share NIH Data</u> page:
  - Template NIH DMS Plan: includes information on sharing NIH data with DSDR/ICPSR (copy and paste language)
  - Budgeting: get a curation cost estimate from ICPSR
    - Allow 10 business days for an estimation
  - Webinar recordings



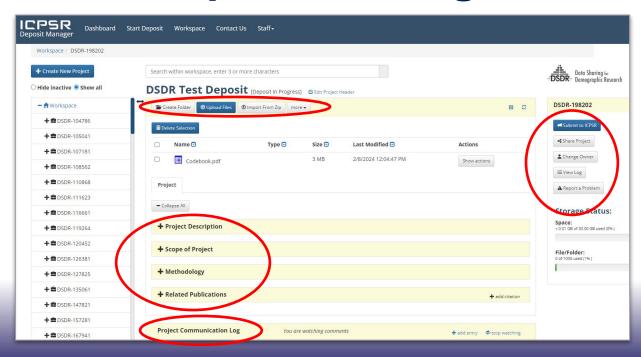


## What to Include in Your Deposit

- Data Files
  - Quantitative: SPSS, Stata, or SAS submissions encouraged with defined variable labels, value labels, and missing data codes.
  - Qualitative, video data, images, and geo-spatial data also accepted
- Documentation
  - Codebooks (data dictionaries), data collection instruments, and guides
  - Respondent consent form, project summaries, and data-related citations
- Study Description (Metadata)
  - Řequired: Study title, principál investigator(s), summary
  - Suggested: study design, sample, methodòlógy



## **Deposit Manager Workspace**



- Begin on DSDR Deposit page
- Customizable workspaces
- Study metadata
- Ability to save work
- Status tracking
- Communication log



## 20 Years of DSDR

+008

studies (data accessible through DSDR)

170+

union
catalogs
(data accessible
outside DSDR)

>1 million

searchable variables

43k+

data-related publications



## Looking to the Future...

- New data:
  - Additional datasets for The Chitwan Valley Family Study series
  - Additional Age 1 and Age 2 datasets for Baby's First Years
  - Additional datasets for The New Immigrant Survey series
- Developing more resources to help researchers with:
  - Human subject protection and disclosure risk mitigation
  - Meeting NIH Data Sharing and Management Policy requirements
- New types of data: Video data, Images, Qualitative
- Additional data guides





## **Connect with DSDR**

Stop by and see us in PAA Annual Meeting
April 17-20

**Researcher Support** 

(734) 763-8937

DSDR@icpsr.umich.edu

Subscribe to DSDR News for the latest updates!