



Building a Harmonized Data Market for Longitudinal Data with MIDUS and DDI (3.2)



Overview of Presentation

- Background on MIDUS
 - ▣ Importance of DDI for
 - Harmonization
 - Facilitating complex analysis
- Current Project's Goals
- Implementation of Project Goals
 - ▣ Creating MIDUS DDI 3.2 Instances
 - ▣ Upgrading MIDUS-Colectica Repository/Portal

MIDUS



*Advancing Knowledge
of Factors That
Promote Positive
Health and Resilience*

Baseline: 1995-96

- Harvard
- MacArthur Found.
- N=7,108
- Ages 25-74
- Twins/Siblings

MID-LIFE IN THE UNITED STATES A National Study of Health and Well-Being

Unique Strengths of the MIDUS Study

**In-depth
multidisciplinary
content** achieved
via 5 separate data
collection projects

**Wide age range
(25-74)** facilitates
focus on life course
transitions

MIDUS (Midlife in the U.S.) is a national longitudinal study of how many factors (behavioral, social, psychological, biological, neurological) come together to influence health and well-being as people age from early adulthood into midlife and old age. It was conceived by a multidisciplinary team of scholars interested in understanding aging as an integrative process.

MIDUS Samples

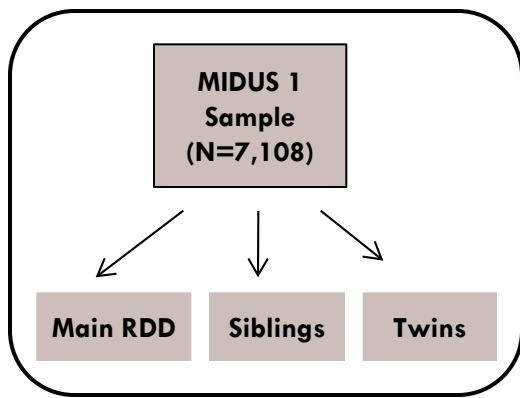
In 1995, MIDUS survey data were collected from a total of 7,108 participants. The baseline sample was comprised of individuals from four subsamples: (1) a national RDD (random digit dialing) sample (n=3 487); (2) oversamples from

In addition, the twin subsample was administered a short screener to assess zygosity and other twin-specific information.

With funding provided by the National Institute on Aging, a longitudinal fol-

MIDUS: Strengths and Complexities

- **Multiple sample waves (longitudinal)**

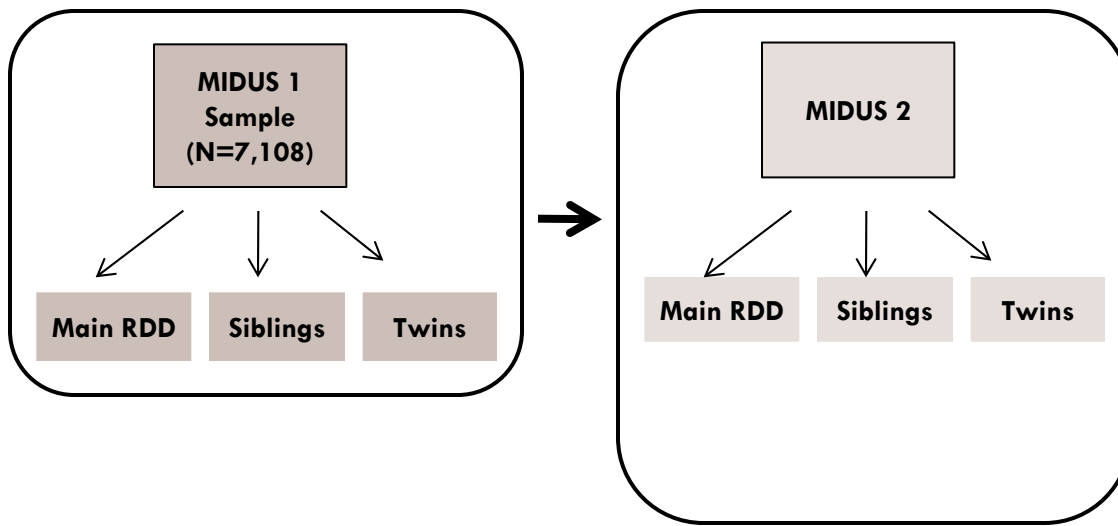


MIDUS Samples and Timelines

1995

2005

2015



MIDUS Samples and Timelines

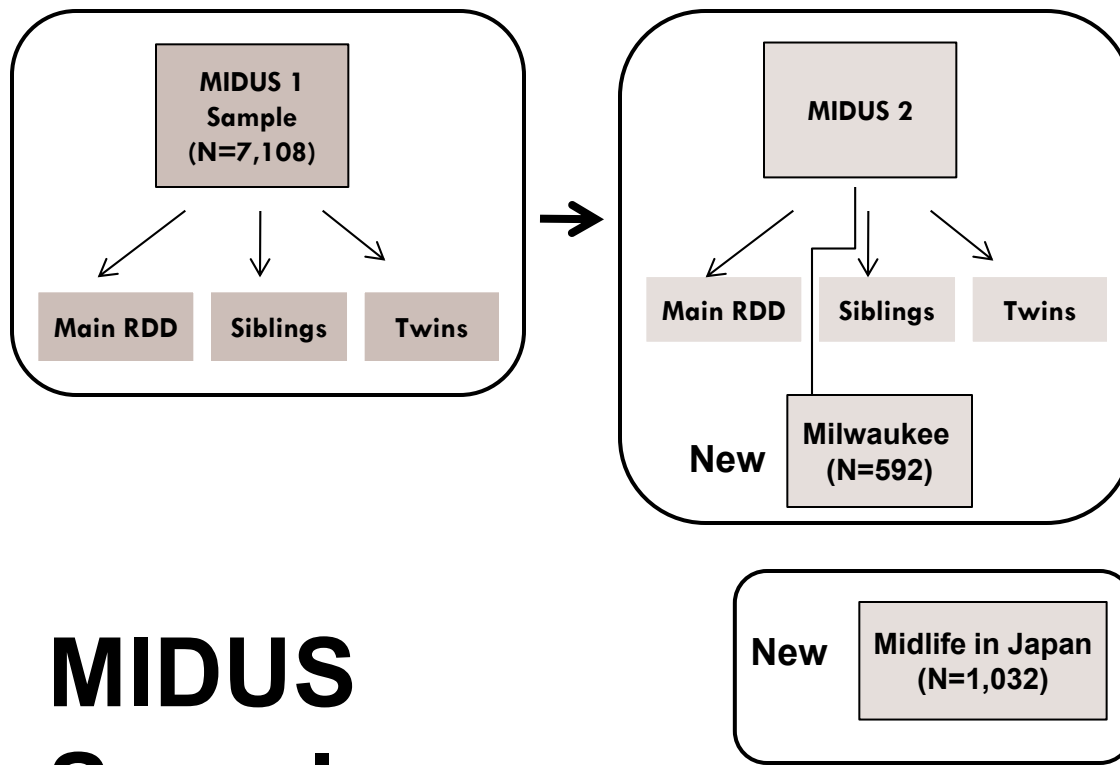
1995

2005

2015

MIDUS: Strengths and Complexities

- ❑ **Multiple sample waves (longitudinal)**
- ❑ **Multiple cohorts**

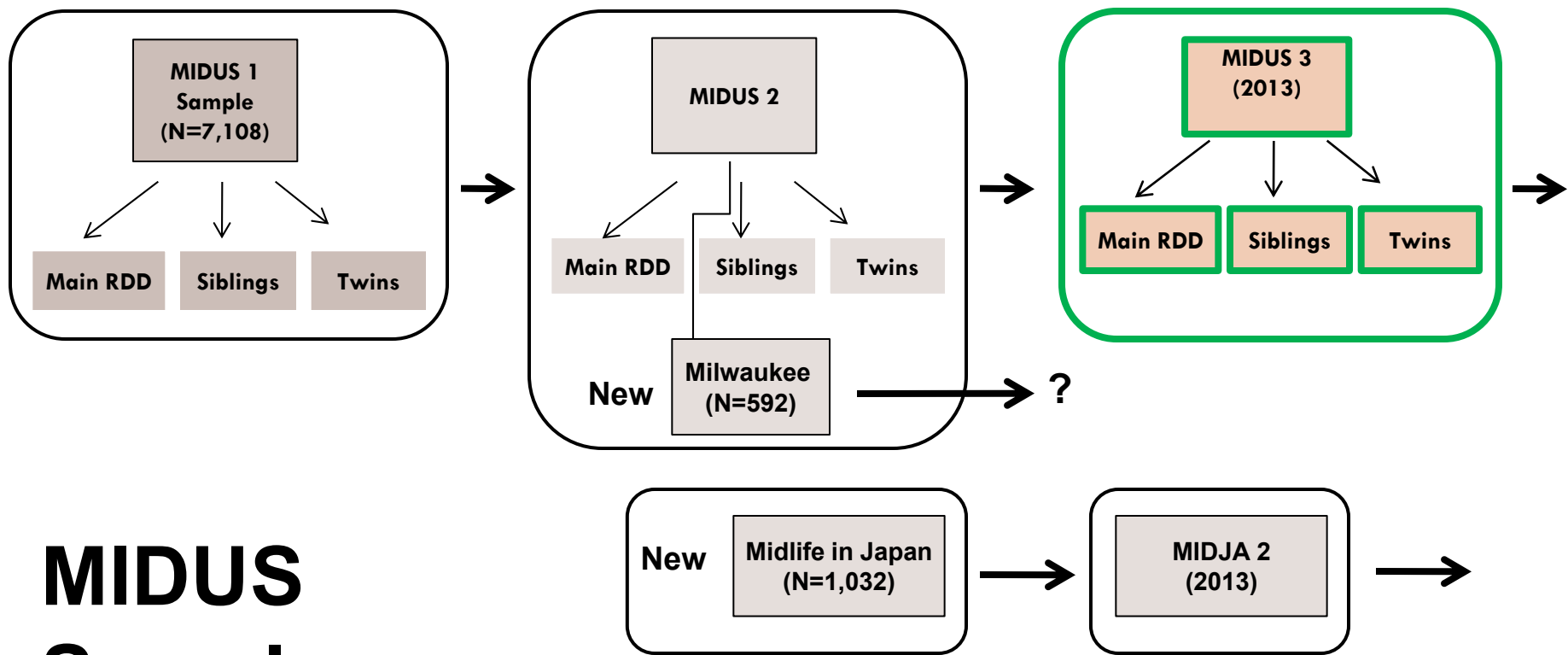


MIDUS Samples and Timelines

1995

2005

2015

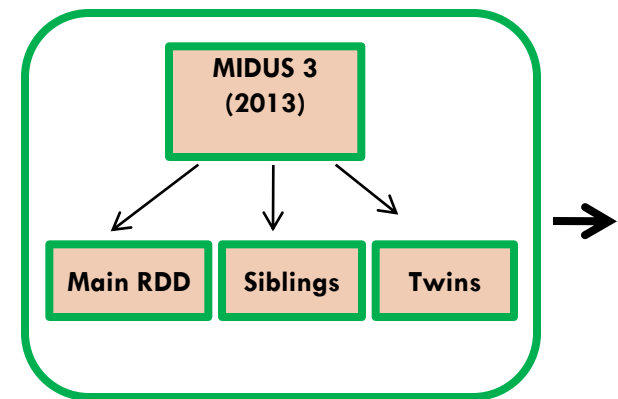
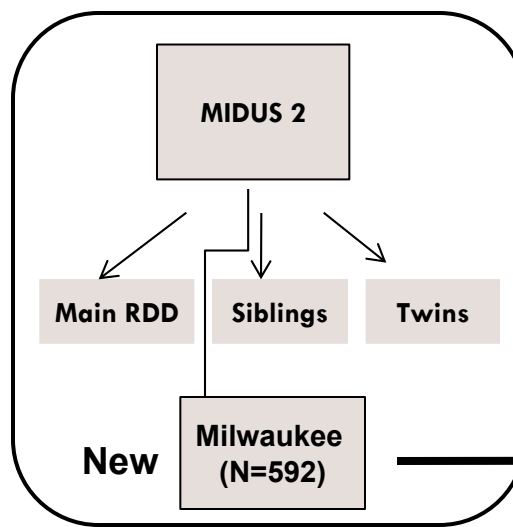
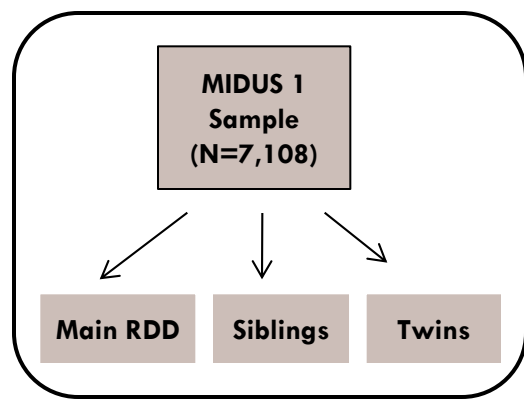


MIDUS Samples and Timelines

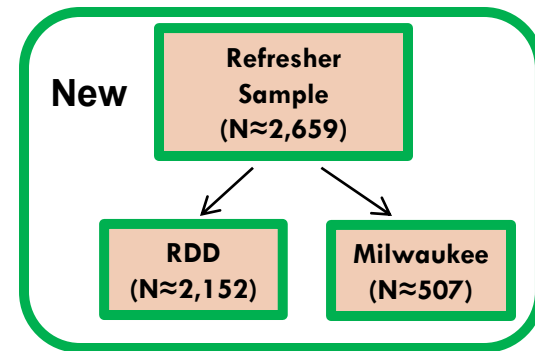
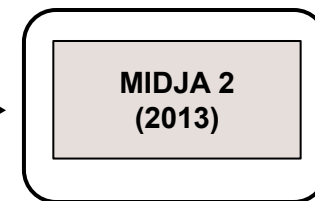
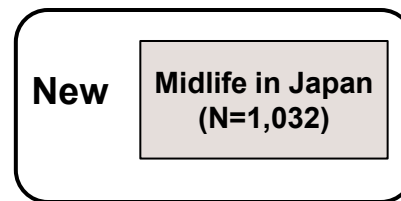
1995

2005

2015



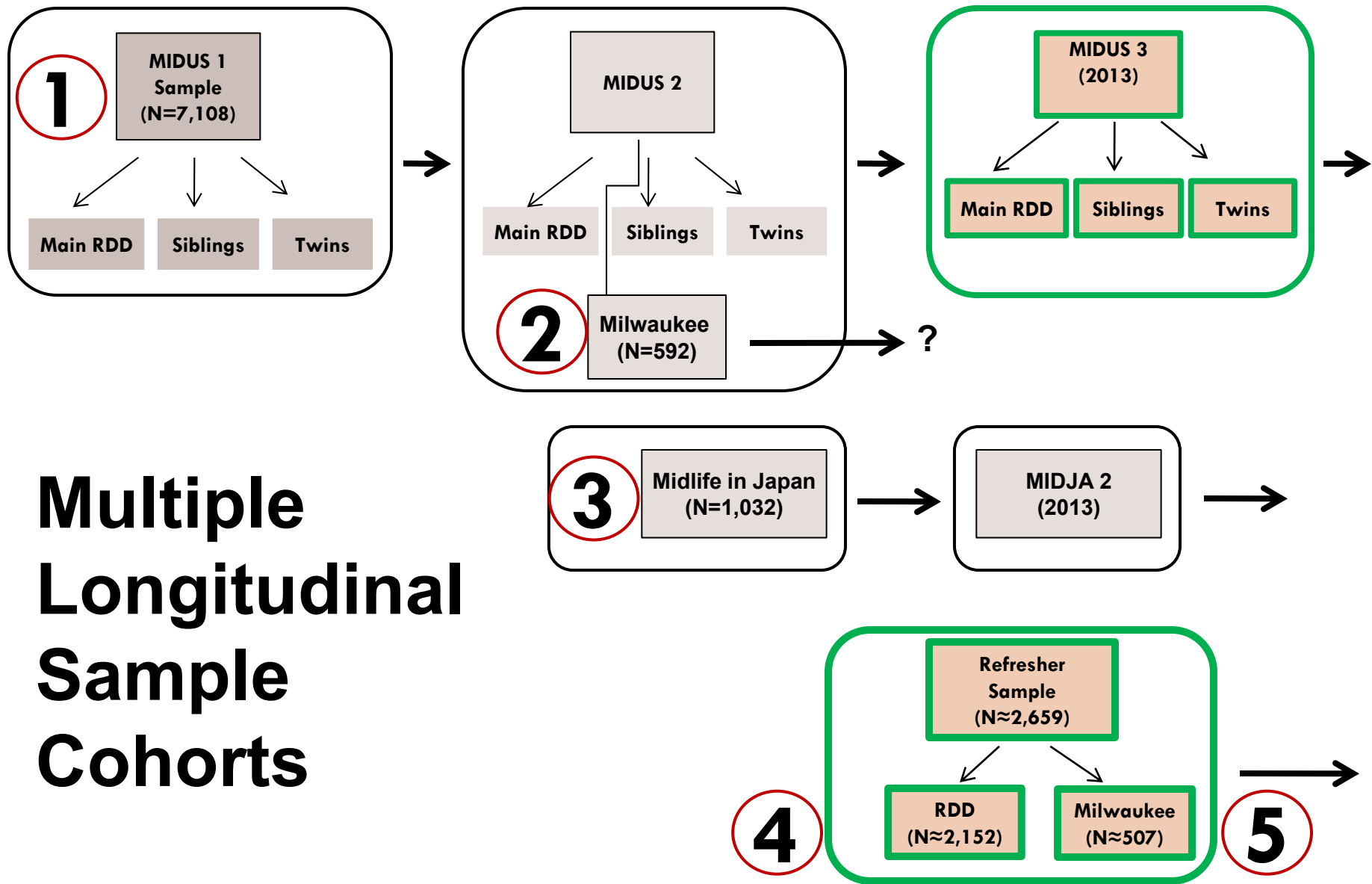
MIDUS Samples and Timelines



1995

2005

2015



Multiple Longitudinal Sample Cohorts

1995

2005

2015

MIDUS: Strengths and Complexities

- ❑ **Multiple sample waves (longitudinal)**
- ❑ **Multiple cohorts**
- ❑ **Multidisciplinary design**
 - ▣ **Aging as integrated bio-psycho-social process**

PROJECT 1

(SURVEY OF A NATIONAL SAMPLE)

Assessed a wide array of psychological constructs (e.g., personality, psychological well-being, positive and negative affect, sense of control, goal orientations) and demographic characteristics (e.g., gender, marital status, socioeconomic standing, employment status), along with extensive health measures (mental and physical).

MODE: 30-minute Phone Interview and Two 50-page Self-Administered Questionnaires

THE MIDUS II PROJECTS

PROJECT 2

(Daily Diary Study)

8 days of daily experience obtained via phone interviews.

(e.g., time use, physical health symptoms and substance use, work productivity, psychological distress)

4 days of salivary cortisol

PROJECT 3

(Cognitive Functioning)

Phone-based cognitive battery

(e.g., episodic verbal memory, working memory, verbal ability and speed, fluid intelligence/reasoning, speed of processing, episodic verbal memory/forgetting)

Face-to-face assessment of cognitive capacities

PROJECT 4

(Biomarkers)

2-Day Clinic Visit:

Biomarkers—neuroendocrine, cardiovascular, immune, bone

Physical exam

Medical history

Medications

Sleep assessments

Laboratory challenge study—heart-rate

variability, blood pressure, cortisol

PROJECT 5

(Neuroscience)

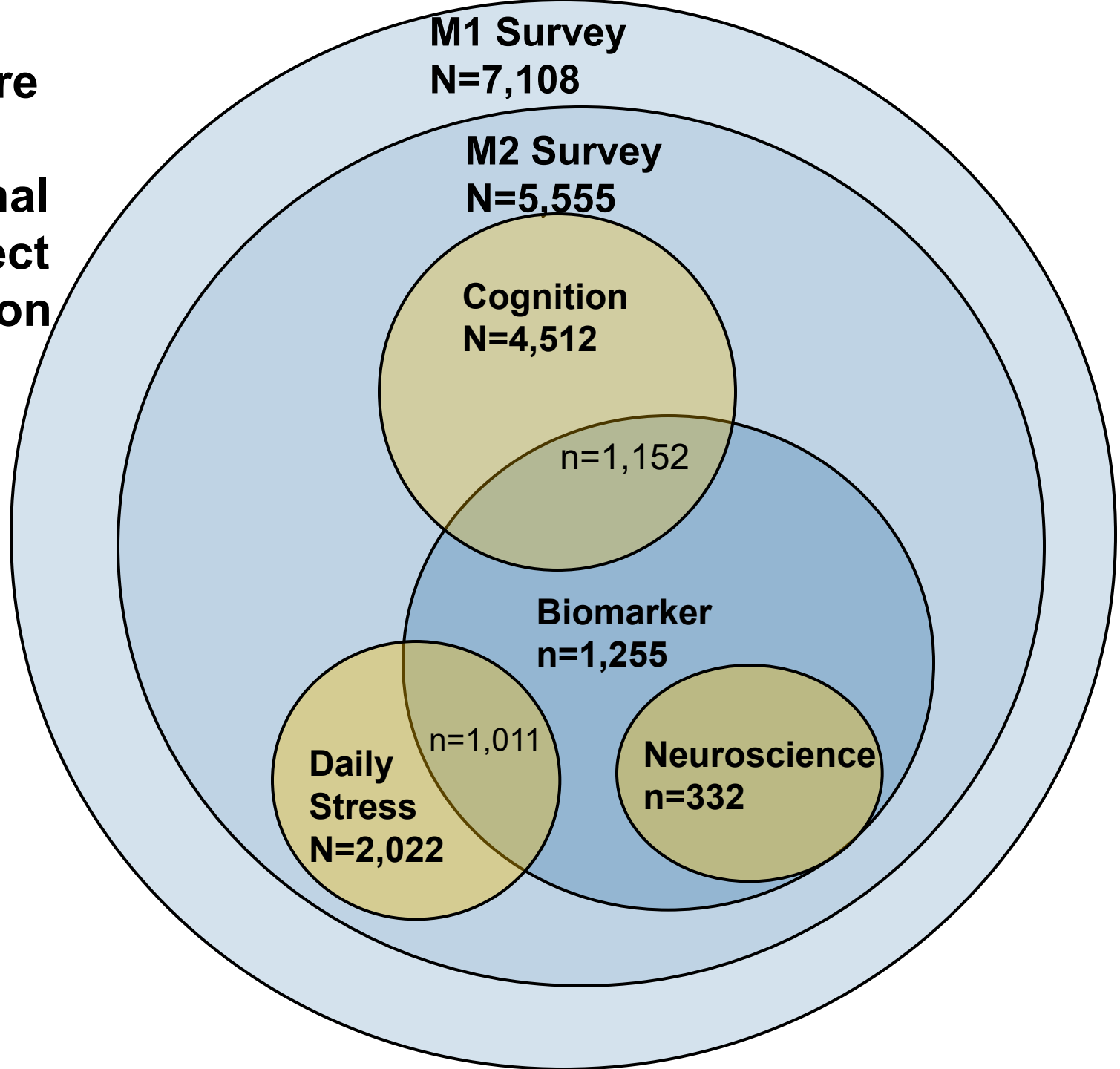
Affective reactivity & recovery:

- baseline electroencephalography (EEG)
- task-related EEG
- task-related electromyography (EMG; eyeblink startle response, post auricular startle reflex, corrugator supercilli activity)
- structural MRI of neuroanatomy
- task event-related fMRI

MIDUS Core

-

**Longitudinal
Multi-project
Participation**



MIDUS: Strengths and Complexities

- ❑ **Multiple sample waves (longitudinal)**
- ❑ **Multiple cohorts**
- ❑ **Multidisciplinary design**
 - ▣ **Aging as integrated bio-psycho-social process**
- ❑ **Wide use of MIDUS – Open Data philosophy**
 - ▣ **#1 data download at NACDA**
 - ▣ **Top 10 data download at ICPSR**
 - ▣ **500 publications**

Current DDI Efforts



MIDUS Metadata Repository/Portal

<http://midus.colectica.org/>

Current Project goals

Under a DDI rubric...

1. Harmonization (internal)

- Clarify related nature of longitudinal and cross-cohort variables (improving search function)
- Provide information/procedures for reconciliation

2. Customized Data Extract (CDE)

- Allow researchers to focus on variables of interest
- Facilitate accurate merges across numerous datasets

Harmonization

WiscMail x midus_p1 x MR_AGGR x Home, co x Study Cat x ddi DDI Mov x New Tab x www.eddi x CaptureIn x

https://docs.google.com/spreadsheet/ccc?key=0AmjuyyBzwvcodHI0VVdnTnRwUE1CaXY3bmX2X0VmNEE#gid=0

Apps Personal Work DDI UW-Madison N... google drive - ... Google Scholar Home : Oxford ... MIDUS Colectic... MIDUS - Midlif...

midus_p1_variable crosswalk_20120927 ☆

File Edit View Insert Format Data Tools Help All changes saved in Drive

No other viewers Comments Share

fx RAACB3WK

	A	B	C	D	E	F	G	H
1	M1 Variable Name	M2 Variable Name	MKE Variable Name	MR Variable Name	MKER Variable Name	M3 Variable Name	J1 Variable Name	Comparability Notes (between MIDJA and MIDUS)
398				RA1PB2DN	RAACB2DN	C1PB2DN		Question not included in MIDJA instrument.
399				RA1PB2DU	RAACB2DU	C1PB2DU		Question not included in MIDJA instrument.
400	A1PB3A	B1PB3A	BACB3A	RA1PB3A	RAACB3A	C1PB3A		Question not included in MIDJA instrument.
401	A1PB3B	B1PB3B	BACB3B	RA1PB3B	RAACB3B	C1PB3B		Question not included in MIDJA instrument.
402	A1PB3C	B1PB3C	BACB3C	RA1PB3C	RAACB3C	C1PB3C		Question not included in MIDJA instrument.
403	A1PB3D	B1PB3D	BACB3D	RA1PB3D	RAACB3D	C1PB3D		Question not included in MIDJA instrument.
404	A1PB3E	B1PB3E	BACB3E	RA1PB3E	RAACB3E	C1PB3E		Question not included in MIDJA instrument.
405	A1PB3F	B1PB3F	BACB3F	RA1PB3F	RAACB3F	C1PB3F		Question not included in MIDJA instrument.
406	A1PB3G	B1PB3G	BACB3G	RA1PB3G	RAACB3G	C1PB3G		Question not included in MIDJA instrument.
407	A1PB3H	B1PB3H	BACB3H	RA1PB3H	RAACB3H	C1PB3H		Question not included in MIDJA instrument.
408	A1PB3I	B1PB3I	BACB3I	RA1PB3I	RAACB3I	C1PB3I		Question not included in MIDJA instrument.
409	A1PB3J	B1PB3J	BACB3J	RA1PB3J	RAACB3J	C1PB3J		Question not included in MIDJA instrument.
410	A1PB3K	B1PB3K	BACB3K	RA1PB3K	RAACB3K	C1PB3K		Question not included in MIDJA instrument.
411		B1PBWORK		RA1PB3WK	RAACB3WK	C1PB3WK		Question not included in MIDJA instrument.
412	A1PB5	B1PB4N	BACB4N	RA1PB4N	RAACB4N	C1PB4N		Question not included in MIDJA instrument.
413	A1PB5	B1PB4M	BACB4M	RA1PB4U	RAACB4M	C1PB4U		Question not included in MIDJA instrument.
414		B1PB5	BACB5	RA1PB5	RAACB5	C1PB5		Question not included in MIDJA instrument.
415	A1PB4A	B1PB5	BACB5	RA1PB5A	RAACB5A	C1PB5A		Question not included in MIDJA instrument.
416	A1PB4B	B1PB5	BACB5	RA1PB5B	RAACB5B	C1PB5B		Question not included in MIDJA instrument.
417	A1PB4C	B1PB5	BACB5	RA1PB5C	RAACB5C	C1PB5C		Question not included in MIDJA instrument.
418	A1PB4D	B1PB5	BACB5	RA1PB5D	RAACB5D	C1PB5D		Question not included in MIDJA instrument.
419	A1PB4E	B1PB5	BACB5	RA1PB5E	RAACB5E	C1PB5E		Question not included in MIDJA instrument.
420	A1PB4F	B1PB5	BACB5	RA1PB5F	RAACB5F	C1PB5F		Question not included in MIDJA instrument.
421	A1PB4G	B1PB5	BACB5	RA1PB5G	RAACB5G	C1PB5G		Question not included in MIDJA instrument.
422				RA1PB5AB	RAACB5AB	C1PB5AB		Question not included in MIDJA instrument.
423	A1PB6	B1PB6	BACB6	RA1PB6	RAACB6	C1PB6	J1SE1	
424	A1PB8_2	B1PB7	BACB7	RA1PB7	RAACB7	C1PB7		Question not included in MIDJA instrument.
425	A1PB8_3	B1PB7A	BACB7A	RA1PB7A	RAACB7A	C1PB7A		Question not included in MIDJA instrument.

+ ☰ Phone + SAQ Comparability Class

29282-0001-D....pdf CHANGELIST_3....xlsx MESAProt0002....doc MESAProt0002....doc -RIVERINVS - 1....pdf Show all downloads...

Harmonization

- Concordance table
 - ▣ Includes “Comparability notes” and “Comparability class”
 - ▣ Example: Variable A1 PA30 “time since last BP test”
 - “M1 is not directly comparable with M2, MKE, MR, MKER, M3: M1 responses were coded as number of months, while other waves broke out number and unit separately.”
 - Offer code/algorithm for reconciliation

Custom Data Extract

- Customized dataset
 - ▣ Search variables, use shopping basket
 - ▣ Allow variables from across MIDUS
 - Merge different datasets
 - Different formats (csv, SPSS, SAS, Stata)
 - Associated DDI codebook

Development Milestones



1. Metadata Quality Report
2. Harmonization
3. Web-based Discoverability
4. Data Extraction

Step 1. Metadata Quality Report

- Compare the harmonization spreadsheet to the Repository
- Check for:
 - ▣ Missing information
 - ▣ Inconsistent labels
 - ▣ Inconsistent data types
- Update the metadata to improve quality

Step 2. Harmonization

- Use the harmonization spreadsheet
- Create a RepresentedVariable for each row
- Store these in the repository

The screenshot shows a Google Spreadsheet titled "midus_p1_variable crosswalk_20120927". The spreadsheet is open in a browser window with multiple tabs. The main content is a table with 8 columns: M1 Variable Name, M2 Variable Name, MKE Variable Name, MR Variable Name, MKER Variable Name, M3 Variable Name, J1 Variable Name, and Comparability Notes (between MIDJA and MIDUS). The table contains numerous rows of variable mappings, such as A1PB3A, B1PB3A, BACB3A, RA1PB3A, RAACB3A, C1PB3A, and J1SE1. The "Comparability Notes" column contains text indicating whether a variable is included in the MIDJA instrument, such as "Question not included in MIDJA instrument." The spreadsheet is shared with "barryradler@gmail.com" and has "No other viewers". The bottom of the screen shows a taskbar with several open files, including "29282-0001-D....pdf", "CHANGELIST_3....xlsx", "MESAProt0002....doc", "MESAProt0002....doc", and "RIVERINVS - 1....pdf".

	A	B	C	D	E	F	G	H
1	M1 Variable Name	M2 Variable Name	MKE Variable Name	MR Variable Name	MKER Variable Name	M3 Variable Name	J1 Variable Name	Comparability Notes (between MIDJA and MIDUS)
398				RA1PB2DN	RAACB2DN	C1PB2DN		Question not included in MIDJA instrument.
399				RA1PB2DU	RAACB2DU	C1PB2DU		Question not included in MIDJA instrument.
400	A1PB3A	B1PB3A	BACB3A	RA1PB3A	RAACB3A	C1PB3A		Question not included in MIDJA instrument.
401	A1PB3B	B1PB3B	BACB3B	RA1PB3B	RAACB3B	C1PB3B		Question not included in MIDJA instrument.
402	A1PB3C	B1PB3C	BACB3C	RA1PB3C	RAACB3C	C1PB3C		Question not included in MIDJA instrument.
403	A1PB3D	B1PB3D	BACB3D	RA1PB3D	RAACB3D	C1PB3D		Question not included in MIDJA instrument.
404	A1PB3E	B1PB3E	BACB3E	RA1PB3E	RAACB3E	C1PB3E		Question not included in MIDJA instrument.
405	A1PB3F	B1PB3F	BACB3F	RA1PB3F	RAACB3F	C1PB3F		Question not included in MIDJA instrument.
406	A1PB3G	B1PB3G	BACB3G	RA1PB3G	RAACB3G	C1PB3G		Question not included in MIDJA instrument.
407	A1PB3H	B1PB3H	BACB3H	RA1PB3H	RAACB3H	C1PB3H		Question not included in MIDJA instrument.
408	A1PB3I	B1PB3I	BACB3I	RA1PB3I	RAACB3I	C1PB3I		Question not included in MIDJA instrument.
409	A1PB3J	B1PB3J	BACB3J	RA1PB3J	RAACB3J	C1PB3J		Question not included in MIDJA instrument.
410	A1PB3K	B1PB3K	BACB3K	RA1PB3K	RAACB3K	C1PB3K		Question not included in MIDJA instrument.
411		B1PBWORK		RA1PB3WK	RAACB3WK	C1PB3WK		Question not included in MIDJA instrument.
412	A1PB5	B1PB4N	BACB4N	RA1PB4N	RAACB4N	C1PB4N		Question not included in MIDJA instrument.
413	A1PB5	B1PB4M	BACB4M	RA1PB4U	RAACB4M	C1PB4U		Question not included in MIDJA instrument.
414		B1PB5	BACB5	RA1PB5	RAACB5	C1PB5		Question not included in MIDJA instrument.
415	A1PB4A	B1PB5	BACB5	RA1PB5A	RAACB5A	C1PB5A		Question not included in MIDJA instrument.
416	A1PB4B	B1PB5	BACB5	RA1PB5B	RAACB5B	C1PB5B		Question not included in MIDJA instrument.
417	A1PB4C	B1PB5	BACB5	RA1PB5C	RAACB5C	C1PB5C		Question not included in MIDJA instrument.
418	A1PB4D	B1PB5	BACB5	RA1PB5D	RAACB5D	C1PB5D		Question not included in MIDJA instrument.
419	A1PB4E	B1PB5	BACB5	RA1PB5E	RAACB5E	C1PB5E		Question not included in MIDJA instrument.
420	A1PB4F	B1PB5	BACB5	RA1PB5F	RAACB5F	C1PB5F		Question not included in MIDJA instrument.
421	A1PB4G	B1PB5	BACB5	RA1PB5G	RAACB5G	C1PB5G		Question not included in MIDJA instrument.
422				RA1PB5AB	RAACB5AB	C1PB5AB		Question not included in MIDJA instrument.
423	A1PB6	B1PB6	BACB6	RA1PB6	RAACB6	C1PB6	J1SE1	
424	A1PB8_2	B1PB7	BACB7	RA1PB7	RAACB7	C1PB7		Question not included in MIDJA instrument.
425	A1PB8_3	B1PB7A	BACB7A	RA1PB7A	RAACB7A	C1PB7A		Question not included in MIDJA instrument.

Step 3. Web-based Discoverability

- Build on top of Colectica Portal
 - ▣ Searching and information retrieval out-of-the-box
- Add cross-reference tables for easy discoverability
- Choose variables or groups of variables to include in the data extract

Step 4. Data Extraction

- Store master data in Colectica Repository
- Based on a user's selected variables, generate:
 - ▣ Datasets
 - CSV, R, SAS, SPSS, Stata
 - ▣ HTML and PDF codebooks
 - ▣ DDI XML

Progress

✓ Complete	Metadata Quality Report
In Progress	Harmonization
Upcoming	Web-based Discoverability
Upcoming	Data Extraction

Thank you

Barry Radler – UW-Madison (bradler@wisc.edu)

Jeremy Iverson – Colectica (jeremy@colectica.com)

Dan Smith-Colectica (dan@colectica.com)

midus.wisc.edu

www.colectica.com