DATA PAPER

The Midlife in the United States (MIDUS) Series: A National Longitudinal Study of Health and Well-being

Barry T. Radler[1]

[1] Researcher, University of Wisconsin-Madison Institute on Aging, United States

Midlife in the United States (MIDUS) is a national longitudinal study of health and well-being (http://midus.wisc.edu/). It was conceived by a multidisciplinary team of scholars interested in understanding aging as an integrated bio-psycho-social process, and as such it includes data collected in a wide array of research protocols using a variety of survey and non-survey instruments. The data captured by these different protocols (comprising around 20,000 variables) represent survey measures, cognitive assessments, daily stress diaries, clinical, biomarker and neuroscience data which are contained in separate flat or stacked data files with a common ID system that allows easy data merges among them. All MIDUS datasets and documentation are archived at the ICPSR (http://www.icpsr.umich.edu/) repository at the University of Michigan and are publicly available in a variety of formats and statistical packages. Special attention is given to providing clear user-friendly documentation; the study has embraced the Data Documentation Initiative (DDI) metadata standard and produces DDI-Lifecycle compliant codebooks. Potential for secondary use of MIDUS is high and actively encouraged. The study has become very popular with the research public as measured by data downloads and citation counts (see Reuse Potential below).

(1) Overview

Introduction

MIDUS is actually a series of related studies. MIDUS began in 1995 with a Random Digit Dial (RDD) national sample of American adults aged 24-74 [1]. The original sample included siblings of the main RDD respondents and a national sample of twin pairs. A second wave of data was collected approximately 9 years later (2004-05), and a third wave is currently being collected (2013-14). Since its inception MIDUS has continued to add new samples such that it now includes data from over 10,000 individuals. New samples include 2 sub-samples of African Americans, a longitudinal sample of Japanese adults (called MIDJA), and most recently a new national sample of Americans designed to "refresh" the core sample with an age-matched cohort (25-74). This "Refresher" sample provides opportunities to examine age, period, and cohort effects in the MIDUS data. Further, new material was added to the latest survey data collection to focus on investigating physical and mental health impacts of the Great Recession.

MIDUS is innovative on multiple levels. First, it provides ground-breaking assessment of many psychological factors in a national sample of Americans. MIDUS thus affords new directions for demography, epidemiology, and sociology via linkage of these psychological factors to key demographic variables. Second, painstaking pilot research was carried out to develop short-form assessments of many psychosocial constructs that could be used with large population samples. Third, the scientific scope of MIDUS is expansive and demonstrates that population-level inquiry can successfully encompass unprecedented scope, through a well-crafted phone interview, combined with a lengthy self-administered survey.

Finally and most importantly, MIDUS approaches the study of aging from a multi-disciplinary perspective and relies on different research protocols and disciplines to collect a broad set of data on participants. At its founding MIDUS collected data via traditional surveys, an in-person cognitive battery, and a diary study of daily stress. Since then MIDUS has added biology and neuroscience

protocols to its data collection by bringing participants into laboratory and clinical settings. These "studies within a study" provide a novel solution to disciplinary trade-offs between sampling scope and generalizability, on the one hand, and in-depth assessments of core constructs, on the other. The MIDUS design encompasses both, resulting in protocol that consists of 5 projects. The hub of the study is an initial survey (including a Phone interview and 2 Self-Administered Questionnaires) in which all respondents participate. Respondents who complete the survey become eligible for participation in four other projects: (1) a 8-day daily diary study of stressful experiences, (2) a phone-based cognitive assessment, (3) a 2-day clinic visit that collects biomarker data, and (4) affective neuroscience measures of brain images and functioning. Each of these projects is represented by its own dataset and accompanied by explanatory documentation and instruments.

Spatial coverage
Spatial coverage of MIDUS is the contiguous United States. Spatial coverage of the Japanese version of MIDUS (called MIDJA) is the City of Tokyo.

Temporal coverage
Data collection for the original MIDUS sample (M1) began January 1995 and ended September 1996. The second wave of data collection (M2) began January 2004 and continued via other MIDUS projects (stress, biomarkers, and neuroscience) through May 2009. MIDJA data collection lasted from August 2008 through April 2010 [2]. All specific data collection dates (month and year) are included in each project's dataset. Because MIDUS is a longitudinal study of aging with diverse content being collected over an extended time period, data collection dates are specific to each project and somewhat open-ended. Indeed, MIDUS is currently collecting data on the third wave (M3) as well as the Refresher sample.

Species
Data collection involves human beings (Homo sapiens).

(2) Methods
Steps
The Survey data are collected using CASES software to conduct the phone interview (CATI) and perform double data entry on 2 self-administered questionnaires (SAQ). These raw datasets are processed, cleaned, and transformed according to MIDUS variable coding and naming conventions—which are documented and available with the datasets—and use SPSS and SAS. The CATI and SAQ datasets are merged into one dataset with an ID system that allows all MIDUS project datasets (cognitive, stress, biomarker, and neuroscience) to be combined at the case-level.

Sampling strategy
The MIDUS core national sample (M1) was based on a nationally representative RDD sample of non-institutionalized, English speaking adults, aged 25 to 74, selected from working telephone banks in the coterminous United

States. City-specific oversamples were also included to increase racial and geographic representativeness. Siblings of cooperating RDD respondents were also invited to participate, and a national sample of twin pairs was obtained from a national household screening project.

Specific sampling details on new subsamples (African-American, MIDJA, and the Refresher) are available with their respective publicly-available datasets.

Further, sampling and selection of participants in the non-survey projects (cognitive, daily stress, biomarker, neuroscience) is contingent upon eligibility criteria specific to each project as well as the situational exigencies of data collection protocols occurring at distributed sites across the U.S. Details of these criteria are available in each project's documentation.

Quality control
Survey instruments are pretested before fielding and extensive training is conducted by all data collection staff. Double data entry is practiced for self-administered instruments. Programmatic cleaning procedures (SAS and SPSS code) are employed to check the survey data for out-of-range values, verify skip pattern logic, apply variable and value labels, assign missing values, and create computed variables. Further data quality control was accomplished by creating DDI codebooks that verified data and metadata quality. Other projects employed quality control procedures specific to their data type. For example, the cognitive project screens its raw data for audio quality and hand codes the viability of computing processing speed and reactions times; the biomarker data collection also involves random quality checks of its tissue assays.

Privacy
MIDUS removes all geographic information and coarsens sensitive variables such as income to minimize the risk of deductive disclosure of participant identity. Further, different ID systems are used for data collection, management, and public release that maintain respondent privacy.

Ethics
MIDUS data collection is reviewed and approved by the Education and Social/Behavioral Sciences and the Health Sciences IRBs at the University of Wisconsin-Madison.

(3) Dataset description
Object name
There currently are 14 publicly-available MIDUS datasets and more are scheduled to be released. All MIDUS datasets can be found by searching the ICPSR repository using the search term "MIDUS".

Data type
All MIDUS datasets contain primary data collected directly from respondents.

Ontologies
The MIDUS biomarker data uses the Multum Lexicon to code medication data.

Format names and versions

MIDUS data are available through ICPSR in the following formats: SAS, SPSS, Stata, R, ASCII, and delimited.

Creation dates

M1 data creation began with data collection on 1/1/1995 and ended 3/1/1998 with preliminary release of dataset. M2 dataset creation (for multiple projects) began with data collection on 1/1/2004 and ended 9/1/2010 with the release of the M2 biomarker dataset.

Dataset creators

The longitudinal and individual project MIDUS datasets have included a host of individuals involved in their creation over the past 20 years:

- Brim, Orville G., John D. and Catherine T. MacArthur Foundation. Research Network on Successful Midlife Development;
- Bumpass, Larry L., University of Wisconsin-Madison;
- Cleary, Paul D., Harvard Medical School;
- Kessler, Ronald C., Harvard Medical School;
- Lachman, Margie E., Brandeis University;
- Markus, Hazel Rose, Stanford University;
- Marmot, Michael G., University College London Medical School;
- Ryff, Carol D., University of Wisconsin-Madison;
- Almeida, David M., Pennsylvania State University;
- Ayanian, John S., Harvard University;
- Carr, Deborah S., University of Wisconsin-Madison;
- Coe, Christopher, University of Wisconsin-Madison;
- Davidson, Richard, University of Wisconsin-Madison;
- Krueger, Robert F., University of Minnesota;
- Love, Gayle D., University of Wisconsin-Madison;
- Mailick , Marsha, University of Wisconsin-Madison;
- Marks, Nadine F., University of Wisconsin-Madison;
- Mroczek, Daniel K., Purdue University;
- Radler, Barry, T., University of Wisconsin-Madison;
- Seeman, Teresa, University of California-Los Angeles;
- Singer, Burton H., University of Florida;
- Sloan, Richard P., Columbia University;
- Tian, Linzhu, University of Wisconsin-Madison;
- Tun, Patricia A., Brandeis University;
- Weinstein, Maxine, Georgetown University;
- Williams, David, University of Michigan

Language

MIDUS datasets use English.

Programming language

The M1 scales documentation includes some SAS code, but programming code is not generally included in the dataset or documentation.

Licence

The MIDUS data are not covered under any license and the vast majority of MIDUS data holdings are public-use files with no restrictions on their access.

Accessibility criteria

MIDUS data are freely available to the public via the Web by opening an ICPSR user account. The Milwaukee sample dataset (which contains data from a sub-sample of African Americans living in Milwaukee, Wisconsin) is the only restricted MIDUS dataset that requires special permission to access. *To obtain this file, researchers must agree to the terms and conditions of a Restricted Data Use Agreement and return it to the National Archive of Computerized Data on Aging, Inter-university Consortium for Political and Social Research, Institute for Social Research, P.O. Box 1248, University of Michigan, Ann Arbor, MI 48106-1248.* More information on this project's restrictions and a data use agreement can be found in the Access Notes section of this webpage: http://www.icpsr.umich.edu/icpsrweb/ICPSR/studies/22840

Repository location

The core M1 dataset can be found via the following DOI: http://doi.org/10.3886/ICPSR02760.v8. However, all MIDUS projects and data can be found on the ICPSR repository website (http://www.icpsr.umich.edu/) by using the search term "MIDUS".

Publication date

The core M1 dataset was originally released via ICPSR on 12/29/1999. However, other MIDUS projects have been and continue to be released since that date. All release dates for each MIDUS project dataset are available on the ICPSR website.

(4) Reuse potential

The MIDUS data have already demonstrated exceptional reuse potential as measured by ICPSR data downloads and citation counts. Since their original public release in 1999, 57,000 MIDUS datasets have been downloaded by nearly 19,000 unique users through ICPSR. MIDUS data have produced over 530 peer-reviewed articles and book chapters in over 30 different topical areas. A full bibliography of publications using MIDUS data can be accessed on the MIDUS website (http://midus.wisc.edu/) or at ICPSR (http://www.icpsr.umich.edu/).

Acknowledgements

Please add any relevant acknowledgements to anyone else who supported the project in which the data was created, but did not work directly on the data itself.

References

1. **Radler, B T** and **Ryff, C D** 2010 Who participates? Accounting for Longitudinal Retention in the MIDUS National Study of Health and Well-Being. *Journal of Aging and Health*, 22(3): 307-331. DOI: http://dx.doi.org/10.1177/0898264309358617
2. **Love, G D, Seeman, T E, Weinstein, M** and **Ryff, C D** 2010 Bioindicators in the MIDUS National Study: Protocol, Measures, Sample, and Comparative Context. *Journal of Aging and Health*, 22(8): 1059-1080. DOI: http://dx.doi.org/10.1177/0898264310374355