Community Clinician Survey
Wave 1 FINAL

User Guide

May 2026

Suggested Citation:
Maust DT, Wagner J, Marcus SC, Wagner L, and Spetz J. 2025. National Dementia Workforce Study User Guide: Community Clinician Survey. Ann Arbor: University of Michigan. Available at www.ndws.org.
We thank Piotr Dworak for his assistance preparing this manual. This guide was prepared with funding from the National Institute on Aging (U54AG084520).

Table of Contents

1. Introduction and Overview

1.1. NDWS Overview

The National Dementia Workforce Study, sponsored by the National Institute on Aging (NIA) of the National Institutes of Health (NIH), is comprised of a family of surveys of the dementia care workforce in the United States. NDWS includes surveys of workers in nursing homes, assisted living communities, home care agencies, and community clinical settings.

This User Guide documents data collection for the **Community Clinician** survey. It provides an overview of the data collection protocol, survey content, information about sampling, and information necessary to analyze the data.

More information about instruments, sampling frames, other data sources available as part of the study, and instructions for accessing all NDWS data, can be found on the study website: NDWS.org.

If you have any questions, please contact info@ndws.org.

1.2. Data Access

NDWS survey data are released in two formats:

- **Restricted-use files (RUFs)** available through the NIA-funded LINKAGE platform or the restricted data enclave of the Michigan Center for the Demographic of Aging (MiCDA).
- **Public-use files (PUFs)** available through the National Archive of Computerized Data on Aging (NACDA).

Information about data access is available at: https://www.ndws.org/surveys-and-data/how-to-access-data

Questions may be directed to info@ndws.org.

1.3. Who does the NDWS Community Clinician survey represent?

The Community Clinician survey is drawn from a random sample of clinicians from all 50 states and DC. It includes primary care physicians (including geriatricians), primary care nurse practitioners, non-surgical physician assistants, psychiatrists, psychiatric-mental health nurse practitioners, and neurologists who provided outpatient care (including residential settings) or prescribed medication to Medicare beneficiaries with a dementia diagnosis in the U.S. in federal fiscal year 2023 (i.e., October 2022 through September 2023).

2. Data Collection Results and Content Documentation

2.1. Community Clinician Survey

The Community Clinician (CC) sample was derived from a comprehensive sampling frame of **492,186** eligible providers, identified via national Medicare outpatient and professional claims data. From this national cohort of clinicians providing care to Medicare beneficiaries, a representative sample of **25,000** was systematically selected for the study.

The initial phase of CC data collection occurred between August 2024 and April 2025, during which the

sample of 25,000 clinicians was invited to complete a 25-minute survey through a series of mailed and emailed invitations and reminders. All clinicians received both a web-based survey invitation and a paper-and-pencil (PAPI) version of the NDWS survey. The Wave 1 CC data release in 2025 included 4,699 respondents from this first phase of data collection.

Because of the compressed timeline during the first two NDWS project years, the second stage of data collection began in June 2025, less than 12 months after the first wave of data collection began. Rather than selecting a new sample for the survey, the incentive was increased for a non-response follow-up (NRFU) phase of data collection targeting the approximately 19,000 clinicians who had not responded to the initial phase of data collection. **This data release now provides an updated, final version of CC Wave 1 data** from that additional period of data collection with 6,554 completed surveys.

In addition, the NRFU phase included an experimental arm testing whether a short-form version of the CC questionnaire would yield a different response rate. The final CC Wave 1 data release also includes a file that includes the short-form respondents; while this includes a larger sample (n=7,378), the content is limited to the items on the short-form instrument (found here: https://www.ndws.org/surveys-and-data/surveys). Researchers should decide which version of the CC Wave 1 data are appropriate for their questions of interest.

2.2. Response Rates by Survey Wave

Data collection outcomes are summarized below.

Table 1. Survey Completion and Sample Sizes

Survey	Sample Size	Number of Completed Surveys	Survey Completion Rate
Community Clinician	25,000		
Initial Wave 1 release		4,699	18.9%
Final Wave 1 release		6,554	26.9%
Short form release		7,378	30.3%

2.3. Survey Instruments

Table 2 describes the content domains of the data collection instruments used in NDWS, with items described in the order presented to respondents. Where applicable, information on item sources and references is provided.

Table 2. NDWS Community Clinician Survey Content-at-a-Glance[a]

Section	Key Topics Covered
1. Education, Training & Experience	Licensure, education, specialty training, board certification, preparedness for dementia care, years in practice *Includes content from:* • Survey of California Nurse Practitioners and Nurse Midwives • nccPA Health Foundation and the National Commission on Certification of Physician Assistants • Final MDS: Physicians. Federation for State Medical Boards
2. Employment Status	Number of clinical and non-clinical jobs, non-clinical roles (e.g., research, teaching)
3. Practice Settings & Characteristics	Practice setting, supervision, hours, staffing mix, team composition, EHR use, geographic location *Includes content from:*

Table 2. NDWS Community Clinician Survey Content-at-a-Glance[a]	
Section	**Key Topics Covered**
	• nccPA Health Foundation and the National Commission on Certification of Physician Assistants • 2022 National Sample Survey of Registered Nurses • Survey of California Nurse Practitioners and Nurse Midwives • nccPA Health Foundation and the National Commission on Certification of Physician Assistants • National Center for Health Workforce Analysis. 2012 National Sample Survey of Nurse Practitioners • Health Teams for Frail Older Adults. Primary Care and Geriatric Health Professionals Survey
4. Patient Panel & Scheduling	Panel size, dementia prevalence and severity, visit volume and length, caregiver involvement, interpreter services *Includes content from:* • 2021 AMA Telehealth Survey Report • Health Teams for Frail Older Adults. Primary Care and Geriatric Health Professionals Survey
5. Processes of Care: Dementia Screening, Diagnosis and Management	Screening tools, diagnostic confidence, referrals, biomarkers, medications, care priorities, community resources, barriers *Includes content from:* • Health Teams for Frail Older Adults. Primary Care and Geriatric Health Professionals Survey
6. Job Outcomes	Job satisfaction, burnout, intent to leave position *Includes content from:* • United States Department of Health and Human Services. National Center for Health Statistics. National Nursing Home Survey, 2004. National Nursing Assistant Supplement • National Center for Health Workforce Analysis. 2012 National Sample Survey of Nurse Practitioners • Maslach Burnout Inventory
7. Demographics	Age, race/ethnicity, language, household composition, caregiving responsibilities, health *Includes content from:* • California Board of Registered Nursing 2022 Survey • United States Census Bureau - Citizenship • KFF LA Times Survey of Immigrants • Centers for Disease Control and Prevention - Sexual Orientation • United States Census Bureau - Veterans • National Center for Women and Information Technology Demographics Guide • National Health Interview Survey • NPALS Direct Care Worker Pilot Study, United States Department of Health and Human Services • NIOSH Worker Well-Being Questionnaire
[a] The NRFU instrument (i.e., second phase of data collection described above in section 2.1) is identical to the original Wave 1 instrument with the exception of the addition of questions 70 (marriage status), 72 (overall health), and 73 (household income).	

Table 2. NDWS Community Clinician Survey Content-at-a-Glance[a]	
Section	**Key Topics Covered**
The CC crosswalk available here (https://www.ndws.org/surveys-and-data/surveys) identifies which items are included in the Short-Form survey.	

2.4. Codebooks

NDWS data are released in two forms: public-use files (PUFs) that have been modified for participant privacy protection (see section 2.5) and restricted-use files (RUFs). The PUFs are available through the National Archive of Computerized Data on Aging (NACDA); RUFs are available through the NIA-funded LINKAGE platform. In the PUF versions, certain data elements are suppressed to mitigate disclosure risk. For the NH and AL surveys, only the Staff surveys are available as PUFs; the Administrator surveys are only available as RUFs.

Separate data dictionary codebooks are provided for the PUF and RUF versions of each NDWS survey dataset. The PUF codebook is available on NACDA; the RUF version is available on LINKAGE for approved users. For each item, the codebooks include the variable name, label, response options, and a frequency distribution of responses, including special values used to represent missing data. The codebooks are intended to support data interpretation and analytic use of the NDWS survey files.

2.5. PUF Preparation and Disclosure Review

Protecting participant privacy is critical from both a compliance and ethics perspective. It must also be balanced with ensuring data utility for research. In accordance with best practices, we offer several layers of privacy protections for the de-identified, NDWS PUFs. First, we removed all direct identifiers under the HIPAA safe harbor method (e.g., date of birth) or variables for which we were concerned that participant identity could be considered readily ascertainable under the Human Subjects Research Regulations (e.g., geographic area). Second, to protect participant privacy while maintaining data utility, we applied an anonymity algorithm that converted continuous variables to categories, collapsing sparse categorical options, and suppressing the minimum number of identifying cells necessary to ensure that every unique combination of indirect identifiers is shared by no fewer than three respondents. Third, we removed all variables alone or in conjunction which we believed posed a potential risk to the participants' financial standing, employability, or reputation as required for exempt research under 45 CFR § § 46.104(d)(2). Additional detail is provided in the PUF documentation for each survey.

2.6. Special Values for Missing Data

NDWS survey datasets use standardized special values to represent different forms of missing or inapplicable data in the survey. These special values are documented in the data dictionary codebooks and are reflected in the frequency distributions provided for each survey item. **Table 3** below describes the special values used across NDWS surveys, and the example illustrates how these values appear in a codebook frequency table for an individual survey item.

Table 3. Special Values for Missing Data	
Missing value	**Represents**
"." Or " "	The item was not displayed for this type of respondent (e.g., only Nurse Practitioners saw item *Field*; other respondents would have ".")
-9	The item was displayed but they did not provide answer
-8	Respondents selected "don't know"
-7	Respondents provided an out-of-range value (the web version did not allow this; only possible where the respondent used a paper instrument)

The following example shows the distribution of the *FellowDidNotFinish* variable available in the Community Clinician Survey (Restricted Use File).

FellowDidNotFinish

Label: Fellowship training: Did not complete a fellowship

Type: numeric | **Length:** 8

Value	Count	Percent
-9=Refuse	144	7.17
0=not selected	571	28.43
1=selected	1293	64.39
Missing	2691	.

3. Sampling

3.1. Overview

This section describes sampling for the NDWS Community Clinician survey. It describes the procedures for constructing the sampling frame, implementing stratification, and selecting the sample. Procedures for applying nonresponse adjustments and poststratification factors used in developing the final weighting adjustments are presented in a Supplement. The resulting weights, along with the stratum variable, **must be included in all analyses** to correctly estimate sampling variance, and ensure valid statistical testing.

3.2. Sampling Frame

The sampling frame was constructed using national Medicare outpatient and professional claims data to identify clinicians who provided care to Medicare beneficiaries with a recorded dementia diagnosis. From these claims, National Provider Identifiers (NPIs) were extracted for eligible clinicians based on their recorded licensure and specialty. NPIs were then linked to the National Plan and Provider Enumeration System (NPPES) to obtain provider contact information and demographics. This process yielded a total of 492,186 eligible clinicians.

In addition to the clinician type (licensure and specialty), we also stored several variables determined from the CMS data, including:

- Number of patients with dementia cared for by the clinician
- Number of low-income patients (i.e., dually-eligible for Medicaid and/or the Part D low-income subsidy [LIS])
- The setting of a clinician's patient care encounters (i.e., did they practice in outpatient and/or residential settings, or were they identified only through prescription claims in Medicare Part D)
- Race/ethnicity composition of patients with dementia for each clinician's panel

Since the clinician practice addresses in NPPES may be outdated, we updated address information using a commercial vendor. We then geocoded the updated addresses and added additional variables about the characteristics of the area surrounding the practice. An example of an added variable was urbanicity, classified according to Rural–Urban Commuting Area (RUCA) codes, with RUCA codes 7–10 designated

as “rural.” More detailed information about CC sample frame construction is available in the “NDWS Wave 1 Sample Frame” documentation available on NDWS.org

We used five stratification variables. Three variables were used to create explicit strata to inform sampling:

- Clinician type (license and specialty)
- Urban/rural status of the clinician
- Whether the clinician has more/less than the overall median number of low-income patients with dementia in their panel

Certain clinician types (categories 4–6) had such a low number of rural cases that we collapsed these with the urban counterparts. **Table 4** presents the strata and the corresponding counts of clinicians from the sampling frame within each stratum.

Table 4. Community Clinician Stratification

Stratum	Clinician Type	Low-income	Rural	Frame Count	Frame Percent
1	Primary care physician	No	No	113,442	23.05
2			Yes	6,712	1.36
3		Yes	No	72,048	14.64
4			Yes	5,234	1.06
5	Primary care NP	No	No	109,516	22.25
6			Yes	9,863	2.00
7		Yes	No	41,717	8.48
8			Yes	2,651	0.54
9	PA	No	No	65,613	13.33
10			Yes	3,841	0.78
11		Yes	No	14,062	2.86
12			Yes	759	0.15
13	Psychiatrist	No	Both	14,764	3.00
14		Yes	Both	4,799	0.98
15	Psychiatric-mental health NP	No	Both	8,400	1.71
16		Yes	Both	3,423	0.70
17	Neurologist	No	Both	5,391	1.10
18		Yes	Both	9,951	2.02

NP: nurse practitioner; PA: physician assistant

Two other variables were used to create **implicit strata**:

- Setting of care (any residential, outpatient only, Part D only)
- Number of patients with dementia cared for by the clinician

“Implicit strata” were used as part of the systematic sampling procedure described in section 4.4.

3.3. Allocation

We set target numbers of respondents for each of the clinician types. Given the relatively small population sizes of psychiatric-mental health nurse practitioners (Psych NPs), psychiatrists, and neurologists, we oversampled clinicians from those categories. **Table 5** summarizes the distribution of clinicians in the sampling frame and the planned sample allocation by clinician type. Columns report the percentage of clinicians in the frame, the planned target sample size, the sample percent and the corresponding relative sampling rate (sample percent divided by frame percent). Sample weights will be constructed for use in analyses to account for differential sampling probabilities across clinician types.

Table 5. Sample and Sample Allocation by Clinician Type				
Clinician Type	**Frame Percent**	**Planned Sample Size**	**Sampled Percent**	**Relative Sampling Rate**
Primary care physician	40.11	7,925	31.70	0.79
Primary care NP	33.27	6,575	26.30	0.79
PA	17.12	3,750	15.00	0.88
Psychiatrist	3.97	2,250	9.00	2.27
Psychiatric-mental health NP	2.40	2,250	9.00	3.75
Neurologist	3.12	2,250	9.00	2.88
Total	**100%**	**25,000**	**100%**	

3.4. Sample Selection

The sample was selected using systematic selection to add implicit stratification. The list of clinicians was sorted within strata by:

- Rural status (collapsed for psychiatrists, psychiatric-mental health NPs, and neurologists)
- Setting of care
- Binary indicator for above/below median number of dementia patients
- Randomly sorted within previous sorts

Systematic selections (every k^{th} selection starting from a random start between 1 and k) were made by stratum using the targeted sample sizes for each stratum as shown above in **Table 5**. This results in the following sample selection equation:

$$\pi_h = \frac{n_h}{N_h}$$

Where:
h = stratum index
n_h = sample size allocated to each stratum
N_h = size of each stratum

All units within each stratum share the same probability of selection. The following is the formula for the corresponding sample selection weight:

$$w_h = \frac{1}{\pi_h}$$

4. Example Code for Weighted Analysis and Variance Estimation

As described above, the NDWS Community Clinician survey is based on a stratified sample of individual clinicians. Following data collection, a survey weight was developed to account for probabilities of selection, nonresponse, and poststratification. The examples below demonstrate how to incorporate the final survey weight and stratification variables in analyses of Community Clinician survey data using SAS or Stata.

To obtain correct standard errors and confidence intervals in statistical analyses, the sample design, including stratification and survey weights, must be specified in statistical analysis software. Failing to account for the design can lead to incorrect inferences.

4.1. SAS

This example demonstrates estimation of a mean and its design-adjusted standard error using PROC SURVEYMEANS, with stratification specified via the STRATA statement and survey weights specified via the WEIGHT statement. For the Community Clinician Survey, the stratification variable is `stratum` and the survey weight is `finalweight`.

As noted earlier, NDWS survey data are released in two forms: a restricted use file (RUF) available through the LINKAGE platform and a public use file (PUF) with certain elements removed to protect confidentiality. The RUF includes all design variables, while the PUF omits the stratification variable (stratum) for disclosure protection. Although the stratification variable is not available in the PUF, valid weighted point estimates can still be produced using the final survey weights alone. This will yield unbiased point estimates, however, omitting stratification information will result in slightly larger variance estimates.

- ***Producing a weighted estimate:***
  The following SAS code will generate a weighted estimate for `PracticeFT`, a continuous variable, incorporating sample design elements:

  The following code uses Community Clinician data to provide two estimates for the mean (green boxes) of `PatientPanel`: the first uses the PUF data set (i.e., without the `strata` element) and the second uses LINKAGE data and incorporates the sample design features (right table column). The only difference between the PUF and RUF means are in the respective standard errors (red boxes); those generated from the RUF data, which includes stratum, are slightly smaller (18.82 [PUF] vs. 18.25 [RUF]).

Public Use File (PUF)

```
PROC SURVEYMEANS DATA= cc_puf;
  * STRATA stratum;
  WEIGHT FinalWeight;
  VAR PatientPanel;
RUN;
```

N	Mean	Std Error of Mean	95% CL for Mean	
4617	952.532142	18.821723	915.632567	989.431716

Restricted Use File (RUF)

```
PROC SURVEYMEANS DATA= cc_ruf;
  STRATA stratum;
  WEIGHT FinalWeight;
  VAR PatientPanel;
RUN;
```

N	Mean	Std Error of Mean	95% CL for Mean	
4617	952.532142	18.252420	916.748638	988.315646

Similar survey procedures are available in SAS for other common analyses, including cross-tabulations (PROC SURVEYFREQ), linear regression (PROC SURVEYREG), and logistic regression (PROC SURVEYLOGISTIC).

4.2. Stata

In Stata, analyses must first declare the survey design. For the Community Clinician survey, this includes

the weight and stratum. In this case, weight is declared via `[pweight=FinalWeight]` and stratum as `strata(stratum)`.

```
svyset [pweight=FinalWeight], strata(stratum)
```

Then, it is necessary to reference the survey design using the `svy` prefix command. For example:

```
svy: mean PracticeFT
```