

**2020 National Post-acute and Long-term Care
Study
Survey Methodology for the
Residential Care Community and
Adult Day Services Center Restricted Preliminary Data**

Table of Contents

Description.....	4
Sampling Design.....	6
Sampling Frame	7
Estimation Procedures	7
Reliability of Estimates.....	9
Obtaining the data.....	10
Contact Information	10

Please Read Carefully Before Using NCHS Restricted Survey Data

The National Center for Health Statistics (NCHS), Centers for Disease Control and Prevention (CDC), conducts statistical and epidemiological activities under the authority granted by the Public Health Service Act (42 U.S.C. § 242k). NCHS survey data are protected by Federal confidentiality laws including Section 308(d) Public Health Service Act [42 U.S.C. 242m(d)] and the Confidential Information Protection and Statistical Efficiency Act or CIPSEA [Pub. L. No. 115-435, 132 Stat. 5529 § 302]. These confidentiality laws state the data collected by NCHS may be used only for statistical reporting and analysis. Any effort to determine the identity of individuals and establishments violates the assurances of confidentiality provided by federal law.

Terms and Conditions

NCHS does all it can to assure that the identity of individuals and establishments cannot be disclosed. All direct identifiers, as well as any characteristics that might lead to identification, are omitted from the dataset. Any intentional identification or disclosure of an individual or establishment violates the assurances of confidentiality given to the providers of the information. Therefore, users will:

1. Use the data in this dataset for statistical reporting and analysis only.
2. Make no attempt to learn the identity of any person or establishment included in these data.
3. Not link this dataset with individually identifiable data from other NCHS or non-NCHS datasets.
4. Not engage in any efforts to assess disclosure methodologies applied to protect individuals and establishments or any research on methods of re-identification of individuals and establishments.

By using these data, you signify your agreement to comply with the above-stated statutorily based requirements.

Sanctions for Violating NCHS Data Use Agreement

Willfully disclosing any information that could identify a person or establishment in any manner to a person or agency not entitled to receive it, shall be guilty of a class E felony and imprisoned for not more than 5 years, or fined not more than \$250,000, or both.

Description

The 2020 National Post-acute and Long-term Care Study (NPALS), previously called the National Study of Long-Term Care Providers (NSLTCP) is the fifth wave in a series of biennial studies of major post-acute and long-term care providers and their services users. The main goals of the study are to: (1) estimate the supply of paid, regulated post-acute and long-term care services providers; (2) estimate key policy-relevant characteristics and practices of these providers; (3) estimate the number of post-acute and long-term care services users; (4) estimate key policy-relevant characteristics of these users; (5) produce national and state estimates where feasible within confidentiality and reliability standards; (6) compare across provider sectors; and (7) monitor trends over time.

Initiated in 2012, the first three waves of NSLTCP (2012, 2014, and 2016) included five provider sectors: residential care communities (RCCs), adult day services centers (ADSCs), nursing homes, home health agencies, and hospices, and provided national and state representative statistical information about the supply and use of long-term care services providers in the United States. In 2018, two post-acute sectors were added: long-term care hospitals and inpatient rehabilitation facilities. The 2020 wave continued to include the seven sectors included in 2018. NSLTCP collects information about providers and services users in two ways—(1) primary data collected by the National Center for Health Statistics (NCHS) through surveys of RCCs and ADSCs, and (2) by obtaining administrative data on nursing homes, home health agencies, hospices, long-term care hospitals, and inpatient rehabilitation facilities from the Centers for Medicare & Medicaid Services. This document contains the methodology and documentation for the primary data collection of the 2020 wave of NPALS. The RCC and ADSC survey data collection was conducted between November 2020 and July

2021. The data shown in these figures are based on preliminary data collected for the 2020 NPALS beginning in November 2020 through March 2021. The preliminary ADSC and RCC data files only include cases that responded by March 25, 2021. These included 3,295 eligible RCCs from a sample of 11,618, and 1,631 eligible ADSCs from a census of over 5,500 ADSCs. These data are weighted to be nationally representative. The data are considered preliminary and there may be change after the release of the final 2020 NPALS data files in 2022.

To be eligible for the study, RCCs had to be licensed, registered, listed, certified, or otherwise regulated by the state; had four or more licensed, registered, or certified beds; provided room and board with at least two meals a day, around-the-clock on-site supervision, and help with personal care, such as bathing and dressing or health related services such as medication management. RCCs had to serve a predominantly adult population. RCCs licensed to exclusively serve the mentally ill or the intellectually disabled or developmentally disabled populations were excluded from NPALS.

To be eligible for the study, ADSCs had to be licensed or certified by the state specifically to provide adult day services, or accredited by the Commission on Accreditation of Rehabilitation Facilities (CARF), or authorized or otherwise set up to participate in Medicaid (Medicaid state plan, Medicaid waiver, or Medicaid managed care), or part of a Program of All-Inclusive Center for the Elderly (PACE); have one or more participants enrolled at the ADSC at the location at the time of the survey; and in operation, serving participants onsite, at a place of residence, or both onsite or at a place of residence.

NPALS uses a multi-mode survey protocol with mail, web, and computer-assisted telephone interviews (CATI). Similar to the first 3 waves of the NPALS, the 2020 wave used a mail and web provider questionnaire with CATI follow-up for non-response. In 2018, the

NSLTCP was redesigned to include a separate screener questionnaire to determine eligibility using CATI, a provider questionnaire using mail or web, and a services user (SU) module administered through CATI to randomly select two services users (residents in RCCs and participants in ADSCs) and ask questions about each selected user. The 2020 preliminary data only includes survey submissions through mail and web who responded by March 25, 2021, excluding later mail and web submissions and all CATI survey mode respondents, which may introduce response bias and estimates may change once data are final. The 2020 questionnaires are available at <https://www.cdc.gov/nchs/npals/questionnaires.htm>.

Sampling Design

The residential care community component of the 2020 NPALS survey used a sample of residential care communities in some states and a census of residential care communities in other states. The adult day services center component of the survey used a census of adult day services centers in all states and the District of Columbia. In the residential care community component, a state was sampled if it had enough communities to enable state-level estimation, i.e., if it had a sufficient number of communities to attain at least 81 completions after inflating the sample size for the estimated ineligibility and nonresponse. In states with an insufficient number of residential care communities on the sampling frame, NCHS took a census of communities.

Among the states where a sample was selected, the primary sampling strata were defined by state and community bed size. For each primary stratum defined by state and bed size, NCHS selected residential care communities by systematic random sampling from lists of communities first sorted by metropolitan statistical area (MSA) status and then randomly ordered within each MSA status. A total of 11,618 residential care communities were sampled and all 5,482 adult day services centers in the final sampling frame were included in the study.

Sampling Frame

The residential care community sampling frame was constructed from lists of licensed residential care communities acquired from the licensing agencies in each of the 50 states and the District of Columbia. The state lists were checked for duplicate residential care communities and concatenated to form a list of all communities, resulting in a sampling frame of 44,201. For the census of adult day services centers, NCHS used a frame obtained from the National Adult Day Services Association. Adult day services providers that operated multiple centers at the same address were identified as separate centers. The master list incorporating all sources was checked for duplicate centers; these duplicates were deleted, resulting in a final sampling frame of 5,482 adult day services centers.

Estimation Procedures

The residential care community sample was a mix of sampled communities from states that had enough residential care communities to produce reliable state estimates and a census of residential care communities in states that did not have enough communities to produce a reliable sample. As a result, the residential care communities' estimates were subject to sampling variability and variability due to non-response. For the data on residential care communities in states where these communities were sampled, as well as for national estimates of residential care communities, the probability design of NPALS' residential care community component permits the calculation of sampling errors. The standard error of a statistic is primarily a measure of sampling variability that occurs by chance because only a sample, rather than the entire population, is surveyed. The standard error also reflects part of the variation that arises in the measurement process and non-response leading to unknown eligibility, but does not include any systematic bias that may be in the data or any other non-sampling error. The chances are about

95 in 100 that an estimate from the sample differs from the value that would be obtained from a complete census by less than twice the standard error. Point estimates and standard errors can be calculated using appropriate design and weight variables in order to account for complex sampling, when applicable. Although a census of all adult day services centers was attempted, the adult day services center estimates were subject to variability due to the amount of nonresponse, and permits the calculation of standard errors. Software products such as SAS, STATA, R, and SPSS all have these capabilities. The data files (i.e., adult day services centers and residential care communities) include design variables that can be used to calculate the standard errors. In the residential care community and adult day services center data files, statistical analysis weights were computed as the product of two components—the sampling weight (only for residential care communities in states where they were sampled) and adjustment for unknown eligibility due to non-response. For sampled states in the residential care community component, the sampling weights reflected the probability of selection for each selected facility. The sampling weight for each sample facility was the reciprocal of its probability of selection. For all the records in the adult day services center component and for all states for which we selected a census for the residential care community component, the probability of selection was equal to 1. To account for residential care communities and adult day services centers of unknown eligibility status, the weights of the facilities with known eligibility were adjusted upward based on the proportion of facilities that were actually known to be eligible. The adjustment for unknown eligibility was done in SUDAAN, using a constrained logistic model to predict known eligibility and to compute the unknown eligibility adjustment factors for the weights. In both the residential care community data file and the adult day services center data file, the variable FACSTRAT indicates the sampling stratum (bed size and

state for residential care communities and state for adult day services centers), and the facility/center indicated by the CASEID, the primary sampling unit. POPFAC represents the total number of residential care communities for calculating the finite population correction in a stratum. The survey weights FACFNWT are designed to provide national estimates. Although the adult day services centers data were not from a sample, the variability associated with the non-response was treated as if it were from a stratified (by state) sample without replacement. Thus, POPFAC represents the total number of adult day services centers for calculating the finite population correction in a stratum.

Reliability of Estimates

Estimates from sample surveys published by NCHS must meet reliability criteria based on the relative standard error (RSE or coefficient of variation) of the estimate and on the number of sampled records on which the estimate is based. Proportion estimates are not presented or are flagged based on the procedure specified in “National Center for Health Statistics Data Presentation Standards for Proportions,” available from: https://www.cdc.gov/nchs/data/series/sr_02/sr02_175.pdf. For all estimates other than estimates of proportions in the tables: estimates are not presented if they are based on fewer than 30 cases in the sample data, in which case only an asterisk (*) appears. Estimates based on 30 or more cases are replaced with an asterisk if the relative standard error of the estimate exceeds 30%.

NCHS also follows data confidentiality standards in published reports to ensure non-disclosure of respondents. Users are strongly recommended to read the readme text and follow the instructions provided for the individual data sets.

Obtaining the data

The 2020 ADSC and RCC data files can be accessed through the NCHS' Research Data Center (RDC).

In addition to following the RDC procedures for restricted data file access, there are a few conditions or restrictions for data use and they are as follows:

1. Use the data in this dataset for statistical reporting and analysis only.
2. Make no use of the identity of any person or establishment discovered inadvertently and advise the Director, NCHS, of any such discovery.
3. Report apparent errors in the data or documentation files to the Long-Term Care Statistics Branch (LTCSB).

We also request the user inform LTCSB of any publications or presentations produced based on the 2020 NPALS preliminary data cite relevant NPALS documentations/data products in their work when appropriate.

Contact Information

To request a codebook or for questions, suggestions, or comments concerning NPALS data, please contact the LTCSB at:

Long-Term Care Statistics Branch (LTCSB),
NCHS, 3311 Toledo Road, Hyattsville, MD 20782

E-mail: ltsbfeedback@cdc.gov
Phone: 301-458-4747