2022 NATIONAL POST-ACUTE AND LONG-TERM CARE STUDY

ADULT DAY SERVICES CENTER AND RESIDENTIAL CARE COMMUNITY COMPONENTS SURVEY DESIGN AND METHODOLOGY

For Public Use Data Files

Division of Health Care Statistics
National Center for Health Statistics
September 2024

Suggested Citation

National Center for Health Statistics. Division of Health Care Statistics. *2022 National Post-acute and Long-term Care Study (NPALS) Public Use Data Files Documentation*, September 2024. Hyattsville, Maryland.

Contact Information

Data users can find the latest information about NPALS on our website, at: https://www.cdc.gov/nchs/npals/index.html. Questions, suggestions, or comments concerning NPALS data may be sent to:

Mail: Division of Health Care Statistics, NCHS, 3311 Toledo Rd, Hyattsville, MD 20782

E-mail: ltcsbfeedback@cdc.gov

Phone: 301-458-4747

Contents

Please Read Carefully Before Using NCHS Public Use 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 of 2018 (CIPSEA 44 U.S.C. 3561-3583). 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. Note that provider and services user (resident or participant) files cannot be matched and merged or linked otherwise because the identifiers of facilities in the provider public use file are different from the identifiers provided in the services user (resident or participant) public use file. 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. Make no use of the identity of any person or establishment discovered inadvertently and advise the Director, NCHS, of any such discovery.
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 reidentification of individuals and establishments.

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

Report apparent errors in the data or documentation files to the Division of Health Care Statistics (DHCS).

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.

Obtaining the Data

The 2022 RCC and ADSC data files can be accessed as restricted and public use data files. The restricted data files can be accessed through the NCHS' Research Data Center (RDC): https://www.cdc.gov/rdc/index.htm. The public use data files can be accessed on the NPALS website (https://www.cdc.gov/nchs/npals/questionnaires/index.html). In addition to the RDC procedures for restricted data file access, the above terms and conditions also apply.

We appreciate users informing DHCS of any publications or presentations based on the 2022 NPALS data and cite relevant NPALS documentations/data products in their work when appropriate.

Technical Description of the National Post-acute and Long-term Care Study Survey Design and Methodology

In the following sections of the document, the introduction describes the objectives of the 2022 National Post-acute and Long-term Care Study (NPALS), the components of NPALS, and eligibility criteria for survey participants. Section two describes sampling frames, sample design, sampling procedures, data collection procedures, and scope of survey and sample outcomes. Section three describes estimation procedures including weights and other design variables, and section four briefly describes reliability of estimates.

Section 1 Introduction

The 2022 National Post-acute and Long-term Care Study (NPALS) (renamed in 2020 from the National Study of Long-Term Care Providers, or NSLTCP) is designed to provide nationally representative estimates about the supply and use of post-acute and long-term care services providers in the United States. NPALS includes seven long-term care provider sectors: residential care communities, adult day services centers, nursing homes, home health agencies, hospices, inpatient rehabilitation facilities, and long-term care hospitals. The main goals of NPALS are to: (1) Estimate the supply of paid, regulated long-term care services providers; (2) estimate key policy-relevant characteristics and practices of these providers; (3) estimate the number of long-term care services users; (4) estimate key policy-relevant characteristics of long-term care services 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. Data collection for the 2022 NPALS was conducted from September 2022 through March 2023.

Section 1.1 Components of NPALS

NPALS comprises two components: (1) primary data collected by the National Center for Health Statistics (NCHS) through surveys of residential care communities (RCCs) and adult day services centers (ADSCs), and (2) administrative data on nursing homes, home health agencies, hospices, inpatient rehabilitation facilities, and long-term care hospitals obtained from the Centers for Medicare & Medicaid Services (CMS). With the first wave of NSLTCP in 2012, NCHS conducted the study every two years. The 2022 NPALS is the sixth wave. This documentation focuses on the primary data collection component of the 2022 wave of NPALS.

Eligibility of providers was determined by a series of questions in the provider questionnaire. Providers that were screened out did not complete the rest of the questionnaire.

Section 1.2 Residential Care Community Eligibility Criteria

All RCCs that participated in the survey were 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. Additionally, RCCs had to have one or more current residents at the time of survey to participate in the 2022 survey. These communities served a predominantly adult population. RCCs licensed to exclusively serve severely mentally ill or intellectually disabled/developmentally disabled populations were excluded from NPALS.

Section 1.3 Adult Day Services Center Eligibility Criteria

To participate in the study, ADSCs had to: 1) 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); 2) have one or more average daily attendance of participants based on a typical week; and 3) have one or more participants enrolled at the center at the location at the time of the survey.

Section 2 Overview of Sample Design and Data Collection Procedures

Section 2.1 Sampling Frame

The RCC sampling frame was constructed from lists of licensed RCCs acquired from the licensing agencies in each of the 50 states and the District of Columbia. The state lists were checked for duplicate RCCs and concatenated to form a list of all RCCs, resulting in a sampling frame of 46,049 RCCs. The ADSC frame was constructed using regulatory information collected from the state regulatory agencies that license or certify ADSCs, state affiliate associations that collect administrative data about member ADSCs, and contacts in national chain provider organizations that collect administrative data about chain affiliated ADSCs. This database served as the source file for the 2022 sampling frame, further cleaned (deduplicated) and edited by NCHS to create the final frame of ADSCs consisting of 5,135 ADSCs. ADSC providers that operated multiple ADSCs at the same address were identified as separate ADSCs.

Section 2.2 Sample Design

The 2022 NPALS used a two-stage probability-based sample design. In the first stage, a stratified sample of providers were selected among RCCs and ADSCs; in the second sampling stage, current RCC services users and current ADSC services users were selected from eligible participating RCCs and ADSCs respectively. Stratified samples of 2,088 RCCs and 1,660 ADSCs were contacted to participate in the surveys. Within each eligible participating RCC and ADSC, a random sample of two services users were selected. Respondents of eligible RCCs and ADSCs were asked to prepare a list of all services users (a census) as of midnight the day before the date of the computer assisted telephone interviewing (CATI) appointment. Selection of services users was done via a telephone protocol on the day of CATI appointment. The sampling instructions for eligible and participating RCC or ADSC were described in the services user module instruments and CATI interviewers guided the respondents through the process.

Section 2.3 Data Collection Approach and Procedures

Section 2.3.1 Data Collection Approach

NPALS used a mixed-mode approach to collect provider-level data through hard-copy and Web questionnaires. A reduced CATI questionnaire was offered to nonrespondents. Person-level data were collected via CATI. The intended respondents were directors, owners, administrators, or otherwise knowledgeable staff of the RCC or ADSC provider. The RCC and ADSC survey instruments were designed to assess RCC and ADSC eligibility and to collect data on services offered, the staffing profile, and RCC resident and ADSC participant characteristics. The 2022 NPALS protocol was similar to the 2018 NSLTCP protocol except for a few changes. RCCs and ADSCs were sent hard-copy questionnaires and information on the Web questionnaires concurrently, and the reduced CATI option was offered later to those who did not submit their questionnaire via Web or mail. Unlike the 2018 wave, where RCC and ADSC eligibility was assessed via telephone through a brief screener instrument, the 2022 wave included the screener questions in the provider questionnaires. Similar to the 2018 wave, in addition to collecting provider-level data, the 2022 wave expanded the data collection to gather person-level data on residents of RCCs and participants of ADSCs. As indicated above, a mixed-mode approach was used to collect provider-level survey data through hard-copy and web questionnaires, and person-level data through CATI. Separate RCC and ADSC instruments for each mode were developed to allow for a subset of different items to be administered for each provider type.

Section 2.3.2 Data Collection Procedures

As the first step of data collection, directors or administrators of sampled RCCs and ADSCs were sent an advance notification packet that included a cover letter from the NCHS director with Frequently Asked Questions (FAQs), web survey login information, a provider-specific insert with selected results from the 2018 wave of NSLTCP, national provider association letters of support, a CDC confidentiality brochure,

and the NCHS Ethics Review Board letter of approval. In addition, various mailings and e-mails were sent, and prompting calls were made to inform directors or administrators about the study to encourage them to participate. Three questionnaire packets were sent that included the provider-specific questionnaire and a pre-addressed, postage-paid, business reply envelope.

Eligible RCCs and ADSCs had the opportunity to use a scheduler to schedule the services user sampling and questionnaire CATI calls when completing the provider questionnaire by Web.

Unlike 2018 when RCCs and ADSCs could only complete the services user module if they had completed the provider module, in 2022, the services user module could be completed with respondents who had either completed the provider module or had at least answered the screening items and were deemed eligible. The approach was to conduct the services user sampling and administer the questionnaires at the time of the appointment and prompt during that call if the respondent still needed to complete their provider questionnaire. The services user module call included conducting sampling procedures to identify two RCC residents or ADSC participants and administering the services user questionnaire with the director, owner, administrator, or knowledgeable staff member for each resident or participant sampled. After data collection, data were edited to ensure that responses were accurate, consistent, logical, and complete.

The 2022 RCC and ADSC provider questionnaires and their respective services user CATI items are available at https://www.cdc.gov/nchs/npals/questionnaires/index.html.

Section 2.3.3 Scope of Survey and Sample Outcomes

For the 2022 NPALS, a sample of 2,088 RCCs was selected from the sampling frame of 46,049 RCCs. Of the 2,088 RCCs in the sample, 1,326 RCCs could not be contacted and, therefore, the eligibility status of these RCCs was unknown. Using the eligibility rate of 90% derived from RCCs that completed the

screening questions, a proportion of RCCs of unknown eligibility was estimated or "presumed" to be eligible. This estimated number along with the total number of eligible RCCs were used to estimate the total number of eligible RCCs. Of the 1,885 eligible and presumed eligible RCCs, 688 of them completed the provider questionnaire, for a weighted response rate (for differential probabilities of selection) of 34.0% calculated using AAPOR's Response Rate 4 (The American Association for Public Opinion Research [AAPOR], 2023), resulting in an estimated national total of 32,231 RCCs. The RCC services user module was completed for 549 residents for a response rate of 40%, resulting in an estimated 1,016,424 residents.

For the ADSC component, a sample of 1,660 providers were selected from a frame of 5,135 ADSCs. Of the 1,660 sampled ADSCs, 997 (60%) could not be contacted and, therefore, the eligibility status of these ADSCs was unknown. Using the eligibility rate of 59% derived from ADSCs that completed the screening questions, a proportion of ADSCs of unknown eligibility was estimated or "presumed" to be eligible. This estimated number along with the total number of eligible ADSCs were used to estimate the total number of eligible providers. Of the 974 eligible and presumed eligible ADSCs, 389 of them completed the provider questionnaire, for a weighted response rate (for differential probabilities of selection) of 40% calculated using AAPOR's Response Rate 4 (The American Association for Public Opinion Research [AAPOR], 2023), resulting in an estimated national total of 3,082 ADSCs. The ADSC services user module was completed for 439 participants for a response rate of 56%, resulting in an estimated 182,002 participants.

In total, 688 RCCs and 389 ADSCs completed the respective provider questionnaires for the 2022 NPALS survey components. The services user questionnaire was completed for 549 RCC residents and 439 ADSC participants. Restricted data from the 2022 surveys may be accessed through the NCHS Research Data Center (RDC): https://www.cdc.gov/rdc/index.htm.

Section 3 Estimation Procedures

Section 3.1 Estimation Procedures

The statistics for RCCs and ADSCs are based on samples, so they differ from data that would have been obtained if a complete census had been taken using the same definitions, instructions, and procedures. However, the probability design of the RCC and ADSC samples permit the calculation of estimates and sampling errors. The standard error of a statistic is primarily a measure of sampling variability that occurs by chance because a sample, rather than the entire population, is surveyed. The standard error also reflects part of the variation that arises in the measurement process but does not include any systematic bias that may be in the data or any other non-sampling error.

Standard errors can be calculated for provider and services user estimates by using any statistical software package, as long as clustering within providers and other aspects of the complex sampling design are taken into account. Software products such as SAS, Stata, R, and SPSS have these capabilities. Statistics presented in NCHS publications using 2022 NPALS data are computed using the linearized Taylor series method of approximation as applied in SAS-callable SUDAAN software or Stata, which produces standard error estimates for statistics from complex sample surveys. Both the RCC provider and services user, and ADSC provider and services user public use and restricted data files include design variables that designate each record's stratum marker and the first-stage unit (or cluster). Each of the four public use data files is accompanied by a *ReadMe* document that describes appropriate analysis procedures and approaches of implementing analysis of complex survey data in various statistical software.

Section 3.2 Weights and Other Design Variables

In the RCC and ADSC data files, statistical analysis weights were computed as the product of two components: the sampling weight and adjustment for unknown eligibility due to non-response. The sampling weights reflect the probability of selection for each selected facility (i.e., the reciprocal of its probability of selection). To account for RCCs and ADSCs 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 SAS-callable SUDAAN using a constrained logistic model to predict known eligibility and to compute the unknown eligibility adjustment factors for the weights. In both the RCC and ADSC public use data files, the variable FACWT represents the weights in the respective provider files, the variable PUFSTRATA indicates the sampling stratum, and RCCID or ADSCID, respectively, indicates the primary sampling unit in RCC and ADSC provider files. POPFAC represents the total number of RCCs and PUFPOPFAC represents the total number of ADSCs for calculating the finite population correction in a stratum.

The RCC and ADSC services user public use data files have two stages. In addition to the first stage design variables (SU_FACID, PUFSTRATA, PUFPOPFAC), in the second stage, the sampling unit is the services user indicated by the variable SUID, the finite population correction is POPSU, and the variable for weight is SUWT. The reader is encouraged to refer to the accompanying ReadMe documents included with each public use file for appropriate use of weights and other design variables in statistical analysis using provider (RCC and ADSC) and services user files.

Because the RCC and ADSC components of the 2022 NPALS are sample surveys, data analyses must include survey weights to inflate the sample numbers to national estimates. The weight associated with each sampled provider and each sampled services user is constructed to account for the multistage

sampling design. An estimator$\hat{}$ for any given population total *X* can be expressed as a weighted sum over all sampled units, defined as:

$$\hat{} X = \Sigma u x(u) W(u)$$

where *u* represents a sampled unit, *x(u)* is the characteristic or response of interest for unit *u*, and *W(u)* is the final survey weight for sampled unit *u*. The final weight *W(u)* for each sampled unit is the product of three components: (1) inverse of the probability of selection, (2) adjustment for unknown eligibility, and (3) nonresponse adjustment.

A sampled provider is deemed a respondent at the services user level if the provider completed some or all of the questions in the services user questionnaire for one or two of the services users selected. Some respondents completed the provider questionnaires but did not respond to the services user module and some providers completed the services user module but did not submit the provider questionnaire. Adjustments for non-respondents at each survey level (provider or services user) are made by shifting the sampling weights of non-respondents at that level to similar respondent providers. Adjustments were also made for sample providers whose eligibility status remained unknown at the survey end by shifting their weights to similar providers whose eligibility status was determined.

Section 4 Reliability of Estimates

Estimates published by NCHS must meet reliability criteria published in two NCHS reports: "National Center for Health Statistics Data Presentation Standards for Proportions" is available from https://ww.cdc.gov/nchs/data/series/sr_02/sr02_175.pdf and "National Center for Health Statistics Data Presentation Standards for Rates and Counts" is available from https://ww.cdc.gov/nchs/data/series/sr_02/sr02-200.pdf. Estimates not meeting NCHS standards are not presented or are flagged based on the procedure specified in these standards. Users of NPALS public use data files are encouraged to assess the reliability of estimates derived from their analyses although they are not required to use NCHS guidelines. Users are also strongly recommended to read the ReadMe documents accompanying the release of public use files and follow the instructions provided for the individual data sets.

Section 5 References

1. American Association for Public Opinion Research. 2023 Standard Definitions: Final Dispositions of Case Codes and Outcome Rates for Surveys. 10th edition. AAPOR.