

2022 National Post-acute and Long-term Care Study

Adult Day Services Center Services User (Participant) Public Use Data File

Data Description and Usage

September 2024

Table of Contents

Introduction	5
Data files	5
Documentation	6
Brief description of survey and outcomes	6
Data dictionary.....	7
Services User questionnaire.....	7
Data processing activities to create the public use file	8
Item nonresponse and imputed data	8
Masked variables	8
Modified variables	8
Top/bottom coded or categorized variables	9
Reliability of estimates.....	9
Analyses and Weighting of NPALS public use files	9
Preferred Reporting Items for Complex Sample Survey Analysis (PRICSSA)	12
References	14

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 or 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. 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 re-identification of individuals and establishments.

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

Data users are encouraged to report apparent errors in the ADSC services user (participant) data or documentation files to the Division of Health Care Statistics, Data Analytics and Production Branch (ltcsbfeedback@cdc.gov).

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 ADSC 2022 services user (participant) public use data file is available free of charge from the NPALS website (<https://www.cdc.gov/nchs/npals/questionnaires/index.html>).

We also appreciate that data users inform the Division of Health Care Statistics, Data Analytics and Production Branch of publications or presentations based on the 2022 NPALS data and cite relevant NPALS documentations/data products in their work when appropriate.

Introduction

This document describes the data and some of the processes involved in creating the adult day services center (ADSC) services user (or participant) public use data file. NCHS recommends that data users read this document prior to working with the data.

The National Study of Long-Term Care Providers (NSLTCP) was renamed the National Post-acute and Long-term Care Study (NPALS) in January 2020.

Data files

The 2022 NPALS ADSC public use data are released in two data files: (1) a provider-level data file and (2) services user (participant)-level data file. This document refers to the services user public use data file. The file contains one record for each sampled participant (participants were not interviewed; interviews were conducted with the ADSC directors or caregivers). This file contains participant characteristics, use of services, health status, and cognitive and physical functioning status. The provider and services user public use data file cannot be linked using the identifiers provided in these files. The provider identifier (SU_FACID) in the services user file is different from the provider identifier in the provider file and cannot be used to match and merge these files. This file contains 439 records and 98 variables. Each record has a primary identifier (PARTICIPANTID) and records in this services user file are sorted in order by the primary identifier.

The public use data are provided in ASCII format, with fixed-length records. In addition to an ASCII file, separate data files are provided in SAS, STATA, and R formats. Public use data files can be downloaded from the NPALS website as separate files. The individual files for separate download are:

Documentation files	
Survey methodology documentation	https://www.cdc.gov/nchs/data/npals/2022-Survey-Methodology-Document.pdf
Data dictionary	https://www.cdc.gov/nchs/data/npals/2022-ADSC-Services-User-PUF-codebook.pdf
Services User (Participant) Questionnaire	https://www.cdc.gov/nchs/data/npals/2022-ADSC-Services-User-Questionnaire.pdf
This document (ReadMe file)	https://www.cdc.gov/nchs/data/npals/2022-ADSC-Services-User-PUF-ReadMe.pdf

Documentation

This README file is part of the documentation package accompanying the release of the 2022 ADSC services user public use data file. The package also includes a data dictionary or codebook, the services user questionnaire, and a broader NPALS survey methodology document or technical documentation.

Brief description of survey and outcomes

The survey on ADSCs was conducted between September 2022 and March 2023. To be eligible for the study, an ADSC must (a) have been 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 Care for the Elderly (PACE); (b) have had an average daily attendance of one or more participants based on a typical week; and (c) have had one or more participants enrolled at the ADSC at the designated location at the time of the survey. Data were collected by mail, web, and computer-assisted telephone interviews (CATI).

A two-stage sampling design was used for the 2022 NPALS ADSC component. At the first stage, a stratified sample of ADSCs were selected and at the second sampling stage, samples of two participants were selected from each eligible participating provider selected at the first stage. Stratified samples of ADSCs were selected using systematic random sampling techniques within strata. From a frame of 5,135 ADSCs, 1,600 were selected for the survey. Of the 1,600 sampled ADSCs, eligibility could not be determined for 997. Among those cases where eligibility could be determined (663), 389 (59%) were eligible and 274 (17%) were ineligible because they did not meet the survey criteria or were out of business. However, 997 ADSCs (60%) could not be contacted; therefore, the final eligibility status of these centers was unknown. Using the eligibility rate of 59%,¹ a proportion of these centers of unknown eligibility was estimated to be eligible; 585 ADSCs of unknown eligibility were assumed as eligible. The total number of eligible adult day services centers was estimated as 974 (389 + 585). Of the 974 in-scope and presumed in-scope 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, 2023). To account for ADSCs of unknown eligibility, the weights of the ADSCs with

¹ The eligibility rate is calculated by the number of known eligible ADSCs divided by the total number of ADSCs with known eligibility status. ADSCs that were invalid or out of business and centers that screened out as ineligible were classified as known ineligibles.

known eligibility were adjusted upward based on the proportion of ADSCs that were actually known to be eligible. Adjustments were also made to account for non-response.

Two participants were selected from eligible and participating ADSCs that were selected at the first stage. A questionnaire was completed (via CATI) for a total of 439 participants from 389 ADSCs. About 56% of ADSCs completed both provider and at least one participant modules and the rest of the ADSCs either completed the provider module or the participant module. Therefore, not all ADSCs in the provider public use data file are represented in the participant public use data file and vice versa. Some ADSCs only completed one participant questionnaire. During the weighting process, these participants were moved to a similar ADSC that had user records in the file.

Data dictionary

The 2022 ADSC services user data dictionary or codebook for the public use data file is provided as a single file containing all variables in the participant public use data file. Each variable in the public use data file has its own codebook entry.

If a question or a series of questions in the survey were legitimately skipped for selected respondents, then the skipped responses were coded as “-1= INAPPLICABLE” in the data dictionary. The questionnaire skip pattern is specified in the data dictionary. Data users are advised to consult the questionnaire to better understand the questionnaire skip patterns. Missing responses were coded as “-9=NOT ASCERTAINED.” The data dictionary or codebook is posted to NPALS website as part of the 2022 public use file release. Data users may also request for a copy of the data dictionary by email (ltcsbfeedback@cdc.gov).

Services User questionnaire

The services user questionnaire is included in the data release package and available at:

<https://www.cdc.gov/nchs/npals/questionnaires/index.html>

The questionnaire includes all the questions asked in the services user module. There may be some differences in how questions were asked in the questionnaire and how they were coded in the public use data file. Also, responses to some questions may not be available in the public use file. These differences are largely related to efforts to reduce disclosure risk. For instance, the public use data file may have fewer response categories (response categories collapsed) than the number of categories indicated in the questionnaire. Restricted variables and data are available to users through the NCHS Research Data Center (<http://www.cdc.gov/rdc/index.htm>).

Data processing activities to create the public use file

The raw data received from the field were reviewed and edited prior to releasing the public use data file. Data were reviewed for accuracy, logic, consistency, and completeness. Additionally, extensive disclosure risk review was conducted to prevent the identity of any facilities who participated in the survey from being made known to the public. NCHS staff used various methods to perturb the data to minimize disclosure risk, and then ensured that the perturbation did not affect the estimates. The following methods were employed on the restricted, in-house file to create the public use data file:

Item nonresponse and imputed data

Item nonresponse is a source of missing data and occurs if a respondent did not know the answer to a question or refused to answer a question, the interviewer inadvertently skipped a question due to problems relating to CATI or the interview broke off before administering the entire questionnaire. In the data file, item-nonresponse is coded as 'NOT ASCERTAINED=-9' when a respondent did not provide an answer. NCHS handled item nonresponse for age (1 case), and race-ethnicity (24 cases) by imputing.

Masked variables

To protect the confidentiality of the information respondents provided, a number of variables have been masked, or simply not included in the public use data file. In making these modifications, NCHS staff tried to maintain a balance between the need for data confidentiality and the needs of data users.

1. Direct identifiers are not included in the public use data file, such as names, addresses, and geographic information (region, state, metropolitan statistical area). There were other variables that were not included in the public use data file. Full list of restricted variables can be obtained by request (ltcsbfeedback@cdc.gov) or through the NCHS research data center (<http://www.cdc.gov/rdc/index.htm>).
2. Provider characteristics variables that are on the provider public use data file are not included in the services user public use data file (e.g. OWNERSHIP).

Modified variables

1. Payment source: The original variable had 9 response categories: Medicaid, Medicare, Older Americans Act, VA, PACE, Other government, Out of pocket, Private insurance, and Other sources. The modified variable (PAYSOURCErc) on the public use data file has 3 categories: Medicaid, Other government; and Out of pocket, Private insurance, and other sources.

2. Other categorical variables collapsed into fewer categories: Some categorical variables were collapsed into fewer response categories. These include LIVENOWrc, LIVEWITH, MEMORYrc, SIGHTrc, HEARINGrc, STAIRSrc, SELFCARErc, COMMUNICATrc, NUMDAYSrc, SYMPTOMSrc.

3. Race of resident (RACEETHrc): American Indians and Alaska Natives, Hawaiian or other Pacific Islanders, and multiple race residents were collapsed into “other race” category in the public use file.

Top/bottom coded or categorized variables

For some variables, upper and lower values were recoded for confidentiality purposes (top- or bottom-coded). Upper or lower cut-off points were used and values above the upper limit or below the lower limit included in the top and bottom categories. These included AGErc (top/bottom coded) and CHARGESrc (categorized continuous variable).

Reliability of estimates

Estimates 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 include an asterisk if the relative standard error of the estimate exceeds 30%.

Analyses and Weighting of NPALS public use files

The data collected in the 2022 NPALS were obtained through a complex, multistage sample design that involves stratification and clustering. The final weights provided for analytic purposes have been adjusted in several ways to yield valid national estimates for adult day services centers in the U.S. Users are reminded that the use of standard statistical procedures that are based on the assumption that data are generated via simple random sampling (SRS) generally will produce incorrect estimates of variances and standard errors when used to analyze data from the NPALS provider public use data file. The clustering protocols that are used in the multistage selection of the NPALS sample require other analytic procedures, as described below. Users who apply SRS techniques to the data generally will produce standard error estimates that are, on average, too small, and are likely to produce results that are subject to excessive Type I error.

In this document, examples of SUDAAN computer codes are provided for illustrative purposes. Examples are provided also for the SAS, STATA, and R software packages. However, the appropriate application of these procedures is the ultimate responsibility of users. NCHS strongly recommends that NPALS data be analyzed under the direction of or in consultation with a statistician who is cognizant of sampling methodologies and techniques for the analysis of complex survey data. The ADSC provider public use data file includes design variables that designate each record’s stratum marker and the first-stage unit (or cluster) to which the record belongs. Examples follow for using these design variables with SUDAAN, STATA, SAS, and R survey procedures.

Table 1a. Computations using SUDAAN

PROC statement	NEST statement	TOTCNT statement	WEIGHT statement
PROC x FILE = y DESIGN = WOR;	NEST PUFSTRATA;	TOTCNT PUFPOPFAC; POPSU	WEIGHT SUWT;

Table 1b. Computations using STATA

Design description in STATA
svyset su_facid, strata(pufstrata) fpc(pufpopfac) vce(linearized) singleunit(missing) ParticipantID, fpc (popsu) weight (suwt)

Table 1c. Computations using SAS

PROC	STRATA	CLUSTER	WEIGHT
PROC SURVEY_ DATA = Y TOTAL = SECONDFILE;	STRATA PUFSTRATA;	CLUSTER SU_FACID;	WEIGHT SUWT;

Table 1d. Computations using R

Design description in R (with R package ‘survey’)

```
#To address lonely PSUs or see package 'survey' usage documentation for alternatives
```

```
options(survey.lonely.psu = "adjust")
```

```
#Create design object
```

```
design_object <- svydesign(id=~SU_FACID+PARTICIPANTID, strata=~PUFSTRATA, weights = ~SUWT,  
data = SUDATA, fpc=~PUFPOPFAC+~POPSU, nest = TRUE)
```

```
#Use svymean() to obtain proportions and standard errors of categorical variables as below
```

```
svymean(~VARIABLE, design_object)
```

Preferred Reporting Items for Complex Sample Survey Analysis (PRICSSA)

The table below provides a Preferred Reporting Items for Complex Survey Analysis (PRICSSA) document (Seidenberg, Moser, & West 2023) for users of the 2022 NPALS Adult day services center (ADSC) services user public use data. This information may be helpful to users when analyzing the 2022 NPALS survey data.

Table 2. Preferred Reporting Items for Complex Sample Survey Analysis

Preferred Reporting Items for Complex Sample Survey Analysis (PRICSSA)	Description
Name of survey	National Post-acute and Long-term Care Study ADSC Services User
Data collection mode	Telephone
Target population	Adult day services center (ADSC) participants (through ADSC Directors or knowledgeable staff) in the United States
Populations excluded	None
Sample design	Stratified two-stage sampling design (two ADSC participants randomly sampled from eligible/sampled ADSCs)
Variance and standard error estimation	Taylor Series Linearization
Weights	SUWT
Design variable: Stratum	PUFSTRATA
Design variable: Population correction factor	PUFPOPFAC
Design variable: Facility ID in the user file for estimation	SU_FACID
Design variable: PSU	POPSU
Presentation standards	Proportions or percentages: https://www.cdc.gov/nchs/data/series/sr_02/sr02_175.pdf Rates and counts: https://www.cdc.gov/nchs/data/series/sr_02/sr02-200.pdf
Unweighted total sample size	439 ADSC participants
Weighted total sample size	182,002 ADSC participants
Response rate (weighted)	51.0%
Location of example code	See Table 1 (a-d) above for approaches in various statistical analysis programs

Suggested citation

Adult day services center services user (participant) README file (this document):

National Center for Health Statistics. Division of Health Care Statistics. *2022 National Post-acute and Long-term Care Study (NPALS). Adult day services center services user (participant) public use data file description and usage*, September 2024. Hyattsville, Maryland.

Adult day services center services user (participant) public use data file:

National Center for Health Statistics. Division of Health Care Statistics. *2020 National Post-acute and Long-term Care Study (NPALS). Adult day services center services user (participant) public use data file*, September 2024. Hyattsville, Maryland.

Contact information

For questions, suggestions, or comments concerning NPALS data, please contact the DHCS, Data Analytics and Production Branch at:

Division of Health Care Statistics, Data Analytics and Production Branch, NCHS,
3311 Toledo Road, Hyattsville, MD 20782

E-mail: ltcsbfeedback@cdc.gov

Phone: 301-458-4747

References

- 1) The American Association for Public Opinion Research. 2023 Standard Definitions: Final Disposition of Case Codes and Outcome Rates for Surveys. 10th edition. AAPOR.
- 2) Seidenberg AB, Moser RP, West BT. Preferred Reporting Items for Complex Sample Survey Analysis (PRICSSA). *Journal of Survey Statistics and methodology* 2023; 11(4):743-757