

2022 National Post-acute and Long-term Care Study

Adult Day Services Center Provider 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
Provider questionnaire .....	7
Data processing activities to create the public use file .....	8
Consistency checks .....	8
Changes in data due to respondent comments.....	8
Masked variables .....	9
Edited/Derived variables.....	9
Item nonresponse .....	10
List of restricted variables.....	10
Reliability of estimates.....	11
Analysis and weighting of NPALS public use data.....	11
Preferred Reporting Items for Complex Sample Survey Analysis (PRICSSA) .....	13
Reference .....	15

## **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 provider data or documentation files to the Division of Health Care Statistics, Data Analytics and Production Branch ([Itcsbfeedback@cdc.gov](mailto:Itcsbfeedback@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 provider 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 public use 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) provider 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) a services user-level (participant-level) data file. This documentation describes the ADSC provider public use file. The provider public use file contains one record for each sampled and eligible ADSC that completed a provider questionnaire. Note that the provider and services user public use data files cannot be linked using the identifiers provided in these files. The provider file covers a range of topics including characteristics about ADSCs, services they provided, types of staff employed, and aggregate participant characteristics. There are 389 records and 98 variables in the ADSC provider public use data file. Each record contains a primary identifier (ADSCID) and the records in the provider file are sorted in the order of the primary identifier.

The public use data files 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 files can be downloaded from the NPALS website as separate files. The individual files for separate download are:

Documentation files	
Survey method documentation	<a href="https://www.cdc.gov/nchs/data/npals/2022-Survey-Methodology-Document.pdf">https://www.cdc.gov/nchs/data/npals/2022-Survey-Methodology-Document.pdf</a>
Data dictionary	<a href="https://www.cdc.gov/nchs/data/npals/2022-ADSC-Provider-PUF-codebook.pdf">https://www.cdc.gov/nchs/data/npals/2022-ADSC-Provider-PUF-codebook.pdf</a>
Provider Questionnaire	<a href="https://www.cdc.gov/nchs/data/nsltcp/2022-ADSC-Provider-Questionnaire.pdf">https://www.cdc.gov/nchs/data/nsltcp/2022-ADSC-Provider-Questionnaire.pdf</a>
This document (ReadMe file)	<a href="https://www.cdc.gov/nchs/data/npals/2022-ADSC-Provider-PUF-ReadMe.pdf">https://www.cdc.gov/nchs/data/npals/2022-ADSC-Provider-PUF-ReadMe.pdf</a>

## Documentation

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

### 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 (60%). 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% derived from ADSCs that completed the screener questionnaire,<sup>1</sup> a proportion of these centers of unknown eligibility was estimated to be eligible; thus, 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,

---

<sup>1</sup>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.

2023). To account for ADSCs of unknown eligibility, the weights of the ADSCs with 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.

### **Data dictionary**

The 2022 ADSC provider data dictionary or codebook for the public use file is provided as a single file containing all four sections of information in the provider questionnaire: A) Background Information; B) Participant Profile; C) Services Offered; and D) Staff Profile. Each variable in the public use 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 beside the question text and code categories. 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 data dictionary by email ([ltcsbfeedback@cdc.gov](mailto:ltcsbfeedback@cdc.gov)).

### **Provider questionnaire**

The Provider 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 provider module. There may be some differences in how questions were asked in the questionnaire and how they were coded in the public use 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 file may provide percentages for some variables (e.g., percent of participants in various age categories, percent of participants with some or all of their long-term care services paid by Medicaid), while the questionnaire asked for specific numbers (e.g., number of participants in a given age category, number of participants with some or all of their long-term care services paid by Medicaid). The variables included in the list of restricted variables are available to users through the NCHS Research Data Center (<https://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 file. Data were reviewed for accuracy, logic, consistency and completeness. Additionally, extensive disclosure risk review was conducted to prevent the identity of any facility that 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. In brief, the following methods were employed on the restricted, in-house file to create the public use file:

### **Consistency checks**

1. To ensure internal consistency of the data, for some questions, edit checks were programmed into the web questionnaire and CATI system and applied during data collection. These edits were programmed based on the expected range of responses for given questions and the logical consistency between questions. For instance, the web questionnaire and CATI system prompted respondents and interviewers, respectively, to verify if the total number of participants in various age groups provided by the respondent was accurate when it was not within  $\pm 10\%$  range of the total number of participants reported earlier.
2. The variables for race-ethnicity and age distribution of participants were edited if the values did not add to the total number of participants (Question 3). For example, when values for the age breakdown of an ADSC (Question 25) did not equal the total number of participants currently enrolled (Question 3), values were adjusted to sum to the total number of participants currently enrolled based on the proportion of values reported for different age categories for the case. In the public use file, a categorical variable (SIZErc) is the number of currently enrolled participants at the location. The age and race-ethnicity variables are converted into percentages using the total number of participants currently enrolled as the denominator.
3. Ownership (Question 1 OWNERSHPrc): When a case was missing a response for ownership in the survey data file, but had a value for ownership in the sampling frame, the missing value on the survey data file was recoded to the value of ownership on the sampling frame.

### **Changes in data due to respondent comments**

The NPALS Web and CATI provider questionnaires allowed respondents to enter comments by clicking an icon provided for each question on each screen. For hard-copy questionnaires, keyers entered any



notes respondents wrote in the margins or in response boxes as they keyed the data. These comments were compiled and reviewed. The original response was changed if it was determined that the comment changed the substance of the recorded answer.

### **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 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 file, such as names, addresses, and geographic information (region, state, metropolitan statistical area). Additional variables were also excluded from the public use file.
2. Modified variables:
  - a. Some variables were modified to minimize disclosure risk. For instance, (i) the maximum allowable number of participants (Question 6) and the current number of participants (Question 3) in an ADSC are not provided in the public use file but replaced by a 2-category occupancy rate variable (OCCUPANCY). (ii) Instead of providing all the different providers with which an ADSC has a computerized system that supports electronic health information exchange (Question 13a-e), a variable was derived to indicate any exchange (ANYEX) with any of the various providers.

### **Edited/Derived variables**

- 1 . Hours per participant day, by employee staff type (i.e., RNHPPD, LPNHPPD, AIDEHPPD, SOCWHPPD, and ACTHPPD):
  - a . Hours per participant day were derived from the number of full-time equivalents for each staff type (Question 27) and the current number of participants (Question 3). In previous waves of the NPALS, the average daily attendance was used instead of the current number of participants. This may lead to differences in the hours per participant day between 2020 (also 2018) and the previous waves of NPALS. The number of full-time and the number of part-time employees for a given staff type were converted into number of full-time equivalents (FTEs) with an assumption that full-time is 1 FTE and part-time is 0.5 FTE.
  - b. The number of FTEs for a given employee staff type was converted into hours by multiplying the FTEs by the average number of hours in a work week (based on a

35-hour work week), and dividing the total number of hours per staff type by the average daily attendance at the center and by the number of days in a work week (5 days). When HPPD variables had values greater than 24, these values were coded as 24.

2. Any employees (ANYRN\_EMP, ANYLPN\_EMP, ANYAIDE\_EMP, ANYSOCW\_EMP, ANYACT\_EMP), by staff type:
  - a. These variables were derived from the FTE variables for employees (e.g., RNFTE1 (which is the sum of full-time and part-time employees) to derive ANYRN\_EMP) indicating whether the adult day services center had any RNs who are employees.
3. Categorized, Top-coded or bottom-coded variables:
  - a. Required hours of training prior to providing care to participants (TRAINHOURSrc) and required continuing education hours (EDUCHOURSrc) were categorized in the public use file.
  - b. The provider public use file included several participant variables aggregated at the provider level (for example, age, race-ethnicity). Instead of providing the exact numbers, these variables were converted into percentages using the number of current participants (Question 3) as the denominator. Race-ethnicity categories were collapsed due to low numbers for some race-ethnic groups to White non-Hispanic, Black non-Hispanic, and Other race (including Hispanic, Asian, Native Hawaiian or Other Pacific Islander, American Indian or Alaska Native).

### **Item nonresponse**

Item nonresponse is a source of missing data that occurred when a respondent did not know the answer to a question or refused to answer a question; or if the respondent submitted the questionnaire before all the questions were answered. The variables with the highest item-nonresponse were the race-ethnicity variable (OTHERRACErc) followed by charges variable (CHARGEsrc). However, item nonresponse (weighted) was less than 10% for most other variables.

### **List of restricted variables**

Users wishing to access these restricted variables or link the provider public use file to the services public use file or non-NCHS data files (e.g., Area Resource File) need to contact the National Center for Health Statistics (NCHS) Research Data Center (<http://www.cdc.gov/rdc/index.htm>).

## 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://www.cdc.gov/nchs/data/series/sr\\_02/sr02\\_175.pdf](https://www.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://www.cdc.gov/nchs/data/series/sr\\_02/sr02-200.pdf](https://www.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 guidelines. Users of NPALS public use data files are encouraged to assess the reliability of estimates derived from their analyses, though they are not required to use NCHS presentation standards and guidelines. Users are also strongly recommended to read the README text accompanying a public use data file and follow the analysis instructions provided for the individual data sets.

## Analysis and weighting of NPALS public use data

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 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 code are provided for illustrative purposes. Examples are also provided for the SAS, STATA, and R statistical 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 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 statistical analysis software.

**Table 1a. Computations using SUDAAN**

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

**Table 1b. Computations using STATA**

Design description in STATA
svyset adscid [pweight=facwt], strata(pufstrata) fpc(pufpopfac) vce(linearized) singleunit(missing)

**Table 1c. Computations using SAS**

PROC	STRATA	CLUSTER	WEIGHT
PROC SURVEY_ DATA = Y TOTAL = SECONDFILE;	STRATA PUFSTRATA;	CLUSTER ADSCID;	WEIGHT FACWT;

**Table 1d. Computations using R**

Design description in R (with package 'survey')
# Create design object <i>design_object</i> <-svydesign(id=~ <b>ADSCID</b> , weights=~ <b>FACWT</b> , strata=~ <b>PUFSTRATA</b> , nest=TRUE, fpc=~ <b>PUFPOPFAC</b> , data = ADSCDATA)
#Use svymean() function to obtain proportions and standard errors of categorical variables: svymean(~ <b>VARIABLE</b> , <i>design_object</i> )

## 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 centers (ADSCs) provider public use data file. 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 Residential Care Community Component
Data collection mode	Mail or web with telephone follow-up
Target population	Adult day services centers (through their directors or knowledgeable staff) in the United States
Populations excluded	None
Variance and standard error estimation	Taylor Series Linearization
Sample design	Stratified random sample
Weight	FACWT
Design variable: Stratum	PUFSTRATA
Design variable: population correction factor	PUFPOPFAC
Presentation standards	Proportions or percentages: <a href="https://www.cdc.gov/nchs/data/series/sr_02/sr02_175.pdf">https://www.cdc.gov/nchs/data/series/sr_02/sr02_175.pdf</a> Rates and counts: <a href="https://www.cdc.gov/nchs/data/series/sr_02/sr02-200.pdf">https://www.cdc.gov/nchs/data/series/sr_02/sr02-200.pdf</a>
Unweighted total sample size	389 ADSCs
Weighted total sample size	3,082 ADSCs
Response rate (weighted)	40.0%
Location of example code	See Table 1a-d above for statistical analysis approaches in various analysis programs.

## Suggested citation

Adult day services center provider 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 provider public use data file description and usage*, September 2024. Hyattsville, Maryland.

Adult day services center provider 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 provider public use data file*, September 2024. Hyattsville, Maryland.

## Contact information

For questions, suggestions, or comments concerning NPALS data, please contact the Division of Health Care Statistics, 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](mailto:ltcsbfeedback@cdc.gov)

Phone: 301-458-4747.

## Reference

- 1) The American Association for Public Opinion Research. 2023 Standard Definitions: Final Disposition of Case Codes and Outcome Rates for Surveys. 10<sup>th</sup> 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