2020 National Post-acute and Long-term Care Study

(Formerly called the National Study of Long-Term Care Providers)

Adult Day Services Center (ADSC) Restricted Data File

July 2022

Data Description and Usage

Table of Contents

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.

****

This document describes the data and the processes involved in creating the adult day services centers (ADSCs) provider restricted data file. NCHS recommends that data users read this document prior to working with the data.

In January 2020, the National Post-acute and Long-term Care Study (NPALS) replaced the National Study of Long-Term Care Providers (NSLTCP).

Data Files

The 2020 NPALS ADSC restricted data are at the provider-level. This document describes the ADSC provider data file. The provider data file has one record for each sampled and eligible ADSC that completed a provider questionnaire and contains characteristics about ADSCs, services they provided, types of staff employed, and aggregate participant characteristics. The ADSC provider data file contains 1,780 records and 273 variables. Each record in the ADSC data file has a primary identifier (CASEID) and records are sorted in the order of the primary identifier.

The 2020 NPALS ADSC data are provided in SAS and STATA data formats.

Documentation

There are several types of documentation available for use with the data file. These include the survey methodology documentation that provides a brief overview of the survey, the data collection procedures, and the sampling design; the survey questionnaires; this provider-specific data description and usage or readme document. A data dictionary or codebook listing the questions and response categories (without the unweighted frequencies and weighted estimates) will be available upon request.

Brief description of survey

The survey on ADSCs was conducted between November 2020 and July 2021. 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).

All 5,644 ADSCs in the final sampling frame were included in the study. Of the 5,466 ADSCs included in the study, eligibility could not be determined for 3,072 (56%). Among those cases eligibility could be determined (2,394), 1,797 (75%) were eligible and 597 (25%) were ineligible because they did not meet the survey criteria or were out of business; 17 (0.007%) eligible ADSCs did not complete the survey. However, 3,072 (56%) of ADSCs could not be contacted; therefore, the final eligibility status of these centers was unknown. Using the eligibility rate of 75%,[1] a proportion of these centers of unknown eligibility was estimated to be eligible; hence, 2,306 ADSCs of unknown eligibility were assumed to be eligible. Therefore, the total number of eligible adult day services centers was estimated as 4,103 (1,797 + 2,306). Of the 4,103 in-scope and presumed in-scope ADSCs, 1,780 ADSCs completed the provider questionnaire, for a weighted response rate (for differential probabilities of selection) of 43.0% (this is calculated by using AAPOR's Response Rate 4). 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.

---

[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.

Data dictionary

The 2020 ADSC data dictionary (codebook) for the restricted data will be available upon request. The codebook contains all five sections of information in the questionnaire: a) Background Information; b) Services Offered; c) Participant Profile; d) Staff Profile; and e) Information on COVID-19. Each variable in the 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 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=MISSING."

Questionnaire

The Questionnaire is included in the data release package and available at: https://www.cdc.gov/nchs/npals/2020-NPALS-ADSC-Questionnaire-Center.pdf. The questionnaire includes all the questions asked during the survey, along with the skip patterns for selected questions. There may be some differences in how questions were asked in the survey, and how they are coded in the restricted file. For example, the questionnaire uses "mark all that apply" questions to ask about different services that adult day services centers provide (Question 18a-i). Respondents indicated as many as five different ways that the center provided a given service. In the data file, for each service question, five binary variables were included: three separate variables corresponding to four different ways that adult day services centers provide the service (i.e., by paid center employees, by arranging for the service to be provided by outside service providers, by referring participants or family to outside service providers); one variable indicating whether the center *temporarily* does not provide, arrange, or refer for this service; and one variable indicating whether the center does not provide, arrange, or refer for this service. In addition to these five binary variables, a derived variable with three mutually exclusive response categories is included in the data file for each service. These derived variables (with collapsed categories) indicate if the center provides the service: 1) by paid center employees/ by arranging for the service to be provided by outside services providers; 2) only by referral; 3) temporarily

does not provide, arrange, or refer for this service/does not provide, arrange, or refer for the service.

Data processing activities to create the restricted data file

The raw data received from the field were reviewed and edited prior to releasing the restricted data file through the NCHS' Research Data Center (RDC). Data were reviewed for accuracy, logic, consistency, and completeness.

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 systems prompted respondents and interviewers, respectively, to verify if the total number of male and female participants provided by the respondent was accurate when the sum of male and female participants did not add to the total number of participants reported in an earlier question. The variables for sex and age distribution of participants were edited if the values did not add to the total number of participants (Question 5). For example, when values for the age breakdown of an ADSC (Question 20) did not equal the total number of participants currently enrolled, 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.

2. Ownership (Question 2 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 because of 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.

Edited/ Derived variables

1 . Number of full-time and part-time employees, by staff type (Question31a-e):

a . Number of full-time and part-time employees for a given staff type were edited to address the cases with missing data. Instructions were provided in the questionnaire to enter "0" if the center had no employees for a given staff type. Yet, there were cases where respondents indicated the number of staff in the response box only when specific staff categories were applicable, while leaving inapplicable response boxes blank. Thus, when editing full-time/part-time (FT/PT) variables, the missing were coded as "0" unless responses to all ten response boxes for all employee staff type were blank or missing (e.g., for employees, the number of full-time RN employees, the number of part-time RN employees, the number of full-time LPN employees, the number of part-time LPN employees, the number of full-time aide employees, the number of part-time aide employees, the number of full-time social worker employees, the number of part-time social worker employees, the number of full-time activities staff employees, and the number of part-time activities staff employees). Otherwise, the missing (-9) were kept as missing (-9). This coding scheme was similar to the scheme used in 2016 and 2018, but different from the coding scheme used in 2014. When editing the FT/PT variables in 2014, missing were coded as "0" unless responses to all four response boxes for a given staff type were blank or missing (e.g., the number of full-time RN employees, the number of part-time RN employees, the number of full-time RN contract staff, the number of part-time RN contract staff); otherwise, the missing (-9) were kept as missing (-9). In the 2014 scheme, each staff type was grouped and included both employees and contract staff.

2 . Hours per participant day, by employee staff type (i.e., RNHPPD1, LPNHPPD1,

AIDEHPPD1, SOCWHPPD1, and ACTHPPD1):

a . Hours per participant day were derived from the number of full-time equivalents for each staff type and the current number of participants (Question 5). In the 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 2018 and the previous waves of NPALS. Number of full-time and part-time employees for a given staff type were converted to number of full-time equivalents (FTEs) with an assumption that full-time is 1 FTE and part-time is 0.5 FTE.

b . Outliers for the FTE variables were defined as values that are 2 standard deviations above or below the size-specific mean for a given staff type, where size was defined as the number of participants served based on average daily attendance (1= 1-25 participants; 2=26-100 participants; 3=101 or more participants). Outliers were replaced by size-specific mean. When calculating the size-specific mean for a given staff type, cases were coded as missing if the number of FTE registered nurse employees/contract staff was greater than 999; if the number of FTE licensed practical/vocational nurse employees was greater than 999; if the number of FTE personal care aide employees was greater than 999; if the number of FTE social work employees was greater than 99; and if the number of FTE activities employees was greater than 99.

c. 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.

3. 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 to derive ANYRN_EMP) indicating whether the adult day services center had any RNs who are employees.

4. Having a computerized system that supports electronic health information exchange with physicians, pharmacies, or hospitals (ANYEX)
   a. This variable was derived from ITMDrc, ITPHARMrc, and ITHOSPrc, ITNHrc, ITLTCOTHrc (Question 15a-e).

5. Revenue source variables (Question 16) were also edited if the values reported for different revenue source categories did not add to 100%. For example, when a case had missing data for a given revenue source category, then the mean of five imputed values for that specific case was used to assess if values of the revenue source categories summed to 100%. When values did not total to 100%, values were adjusted to sum to 100% based on the proportion of values reported for different revenue source categories for the case. In addition to these edits, revenue sources were also imputed for missing cases.

In addition to the above recoded/derived variables, several other variables were also recoded. Recoded variables end with 'rc' to indicate that the original variable was recoded.

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 related to the COVID-19 questions: CVDDEATHPART, CVDPPEQ1N95, CVDPPEQ4N95, CVDPPEQ2N95, CVDPPEQ3N95, CVDHOSPPARTrc ranging from 10.8%-12.9% (weighted) missing. For all other variables, item nonresponse (weighted) was less than 10%.

Imputed data

In the data file, item nonresponse is coded as "-9= Not ascertained." Missing values for revenue sources (Question 16a-h), race-ethnicity (Question 21a-i), sex (Question 19), and age (Question 20a-d) variables were imputed. After the weights were finalized, multiple imputations were created using the Cox-Iannacchione Weighted Sequential Hot Deck (WSHD)

procedure in SUDAAN. For the WSHD procedure in SUDAAN the variables used in the imputation procedure must be specified; they are referred to as the imputation class variables. Within the cross of the imputation class variables, all responding and non-responding records for a given variable were identified. The responding records were potential donors for non-responding (missing) records. In other words, respondents were selected sequentially from within the cross of the imputation class variables and became donors for missing records within that same cross of variables. For all demographic and revenue variable names ending with missing, class variables specified for the imputation procedure include: maximum number of participants allowed (MAXPART), ownership type (OWNERSHP), chain affiliation (CHAIN), Medicaid paying for long-term care services (MEDPAID), and metropolitan statistical area status (MSA). Cases with missing data were recoded as the mean of five imputed values for that specific case and cases with no missing data kept the value as respondents reported. The imputed revenue, race-ethnicity, age, and sex variables have 'rc' suffixed to the original variable name (e.g., FEMALErc, HISPANICrc, REVMCAIDrc, AG64LESSrc). Additionally, imputed variables for the revenue variables with '2rc' suffix (e.g., REVMCAID2rc, REVMCARE2rc) are provided. A flagging variable is also included to indicate cases imputed for the variable (e.g., AG64LESS_IMPFLG, FEMALE_IMPFLG, REVMCAID_IMPFLG). Among 1780 respondents, the percentage of imputed records ranged from 3.8% (69 missing responses) for the sex variables (Question 19a-b) to 9.3% (165 missing responses) for the 'Native Hawaiian and Other Pacific Islander' or NHOPI (Question 21f).

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%.

The data collected in the 2020 NPALS were obtained through a census of all adult day services centers. However, although a census was attempted, the adult day services center estimates were subject to variability due to non-response. Thus, variability associated with the non-response needs to be treated as if it were from a stratified (by state) sample without replacement. 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 United States. These weights adjust for non-response and unknown eligibility in the survey.

In this document, examples of SUDAAN, STATA and SAS analysis codes are provided for illustrative purposes. However, the appropriate application of these procedures and choice of analysis program are 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 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, and SAS survey data analysis procedures.

Table 1a. Computations using SUDAAN

PROC statement	NEST statement	TOTCNT statement	WEIGHT statement
PROC x FILE = y DESIGN = WOR;	NEST STRATAn;	TOTCNT POPFACn;	WEIGHT FACWTn;

Table 1b. Computations using STATA

Design description in STATA
svyset CASEID [pweight=FACWTn], strata(STRATAn) fpc(POPFACn) vce(linearized) singleunit(missing)

Table1c. Computations using SAS

PROC	STRATA	CLUSTER	WEIGHT
PROC SURVEY_ DATA = Y TOTAL = SECONDFILE;	STRATA STRATAn;	CLUSTER CASEID;	WEIGHT FACWTn;

Obtaining the data

The 2020 NPALS ADSC provider file can be accessed through 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 ADSC provider data or documentation files to the Long-Term Care Statistics Branch (LTCSB).

We also request that the user inform LTCSB of any publications or presentations produced based on the 2020 NPALS data and 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: ltcsbfeedback@cdc.gov

Phone: 301-458-4747.