

2014 National Study of Long-Term Care Providers (NSLTCP)

Adult Day Services Center Survey

Restricted Data File

July 2015

Data Description and Usage (Readme)

Table of Contents

| | |
|---|----|
| Data File | 4 |
| Documentation | 4 |
| Brief description of survey | 4 |
| Questionnaire | 5 |
| Data dictionary..... | 6 |
| Data Processing Activities to Create the Restricted File..... | 6 |
| Consistency checks | 6 |
| Edited/ Derived variables | 8 |
| Item nonresponse | 9 |
| Imputed data..... | 10 |
| Reliability of Estimates..... | 11 |
| Accessing the Restricted Data File | 12 |
| Contact Information | 12 |

Please Read Carefully Before Working with the Data File

The Public Health Service Act (Section 308 (d)) provides that the data collected by the National Center for Health Statistics (NCHS), Centers for Disease Control and Prevention (CDC), may be used only for the purpose of health statistical reporting and analysis. Any effort to determine the identity of any reported case is prohibited by this law. NCHS does all it can to assure that the identity of data subjects cannot be disclosed. However, the data released through the NCHS Research Data Center (RDC) include restricted variables, including geographic identifiers. Any intentional identification or disclosure of a person 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 use of the identity of any person or establishment discovered inadvertently and advise the Director, NCHS, of any such discovery.

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

This document describes the data and some of the processes involved in creating the restricted file of the 2014 National Study of Long-Term Care Providers (NSLTCP)' survey of adult day services centers. We recommend that data users read this document prior to working with the data.

Data File

The data file contains information on 2,763 adult day services centers that completed a survey in one of the three modes: a hardcopy mail questionnaire, a web questionnaire, and a computer-assisted telephone interview (CATI). Each record in the file is associated with a primary identifier (CASEID) representing one adult day services center that completed the survey. Also included in the data file are 431 variables, including CASEID and design variables, on center characteristics (e.g., ownership type, chain affiliation, number of years in operation, number of users (i.e., participants), and Medicaid participation); number of employee and contract nursing, social work and activities staff; the provision of services; demographic and health characteristics of participants in the center (e.g., race-ethnicity composition of participants, participants needing any assistance with different activities of daily living, participants' hospitalization and use of emergency department); use of electronic health records and health information exchange; derived variables; and imputed variables. The records in the file are sorted in the order of the primary identifier. The data are provided in SAS format.

Documentation

There are several types of documentation available for use with the data file. These include a data dictionary or codebook; the survey questionnaire; the survey methodology documentation that provides a brief overview of the survey, the data collection procedures, and the sampling design; and this provider-specific data description and usage or readme document. A separate readme document on data description and usage is available for the residential care community component of NSLTCP.

Brief description of survey

The survey on adult day services centers was conducted between June 2014 and January 2015. The sampling frame was 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 5,443 adult day services centers.

The survey for the adult day services center component of the 2014 NSLTCP was based on a census of U.S. centers. A total of 396 (7.3%) centers were identified as invalid or out of business. However, 2,284 centers (42.0%) could not be contacted; therefore, the final eligibility status of these communities was unknown. Using the eligibility rate,¹ a proportion of these centers of unknown eligibility was estimated to be eligible. This estimated number along with the total number of eligible communities resulting from the screening process was used to estimate the total number of eligible adult day services centers in the United States. The weights of the adult day services centers with known eligibility were adjusted upward based on the proportion of centers that were actually known to be eligible to account for the adult day services centers with unknown eligibility status. Data were collected by three modes: self-administered hardcopy questionnaire, self-administered web questionnaires, and CATI conducted by interviewers. The questionnaire was completed for 2,763 centers, for a response rate of 58.0%.

Questionnaire

The PDF for the adult day services center questionnaire of the 2014 NSLTCP is available at http://www.cdc.gov/nchs/data/nsltcp/2014_NSLTCP_Adult_Day_Services_Center_Questionnaire.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 questionnaire, 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 12a – k). Respondents indicated as many as four different ways that the center provided a given service. In the data file, for each service, five binary variables were included: four separate variables corresponding to four different ways that adult day services centers provide the service (i.e., by paid center employees, by arranging for and paying outside vendors, by arranging for outside vendors paid by others, by referral); one variable indicating whether the center provides the service in any of these ways or does not provide the 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 indicate if the center provides the service: 1) by paid center employees/ arranging for and paying outside vendors/ arranging for outside vendors paid by others; 2) only by referral; or 3) does not provide the service.

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

Data dictionary

The data dictionary or codebook is provided as a single file containing five sections in the questionnaire: Background Information; Services Offered; Staff Profile; Participant Profile; and Record Keeping. Each variable in the restricted file has its own codebook entry.

The web and CATI versions of the NSLTCP survey of adult day services centers used for data collection allowed respondents to answer or interviewers to only ask questions specific to the individual centers, skipping questions that did not apply. The skip instructions found in the hardcopy questionnaire were identified by bold text in a different color, which directed respondents to skip questions that did not apply. If a question or a series of questions in the survey were legitimately skipped by a particular respondent, then the response was coded as “-1= Legitimate skip” in the data dictionary. Skip patterns are specified in the data dictionary in addition to the question text and code categories. When respondents refused to answer, did not know the answer, or did not answer the question because of a breakoff (i.e., did not complete the survey), their responses to the question(s) were coded in the data dictionary as “-9= Not ascertained”. The data users are advised to consult the questionnaire before analyzing the data to better understand the question skip patterns.

Data Processing Activities to Create the Restricted File

The raw data received from the field were reviewed and edited prior to releasing the restricted data file to 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 system prompted respondents and interviewers, respectively, to verify if the total number of male and female participants provided by the respondent was accurate when it was not within $\pm 10\%$ range of the total number of participants reported earlier.
2. In most cases, the same skip logic that was applied to the web questionnaire and CATI system was used to edit the data file when the skip instruction was not followed. For instance, if respondents indicated that the adult day services center

was not authorized or set up to participate in Medicaid (Question 1b), then their responses to the number of participants in the last 30 days who used Medicaid to pay for some or all of their services received at the center (Question 7) were coded as “-1= Legitimate skip.” However, the response to Question 1b was missing and Question 7 had a response, Question 1b was recoded as ‘No’ if the response to Question 7 was ‘0’; Question 1b was recoded as ‘Yes’ if the response to Question 7 was greater than 0.

3. The variables for race-ethnicity, sex, and age distribution of participants were edited if the values did not add to the total number of participants (TOTPART for Question 3). For example, when a case had missing data for a given race-ethnicity category, then the mean of five imputed values for that specific case was used to assess if values of the race- ethnicity categories summed to TOTPART. When values did not total to TOTPART, values were adjusted to sum to TOTPART based on the proportion of values reported for different race-ethnicity categories for the case.
 - a. In addition to the original variables, edited variables for race-ethnicity, sex, and age distribution of participants are provided in the data file. The edited variables are indicated by adding “RC” as the suffix to the variable name (e.g., MALERC, FEMALERC for Question 16).
 - b. Edited values for some cases are in decimals because of the following: recoding cases with missing data to take an average of five values imputed values for that specific case or making proportional adjustments to individual categories when values did not total to TOTPART.

4. Revenue source variables (Question 9) 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.
 - a. In addition to the original variables, edited variables for revenue sources are provided in the data file. Two sets of edited variables are indicated by adding “RC” (e.g., REVMCARERC) or “RC2” (e.g., REVMCARERC2) as the suffix to the variable name. More information about imputations is provided in the later section of this document.

Edited/ Derived variables

1. Number of full-time equivalents by employee staff type (i.e., Question 14: RNFTE1, LPNFTE1, AIDEFTE1, SOCWFTE1, ACTFTE1) and contract staff type (RNFTE2, LPNFTE2, AIDEFTE2, SOCWFTE2, ACTFTE2):
 - a. These variables were derived and provided in the restricted file. Number of full-time and the number of part-time employees and contract staff for a given staff type (separately for employees and contract staff) were converted into the number of full-time equivalents (FTEs) with an assumption that full-time is 1 FTE and part-time is 0.5 FTE. Instruction was provided in the questionnaire to enter “0” if the center had no employees or contract staff 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 deriving FTE variables, we coded missing 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, we kept missing (-9) as missing (-9).
 - b. Outliers for 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 coded as the size-specific mean. When calculating the size-specific mean for a given staff type, cases were coded as missing if the number of full-time equivalent (FTE) registered nurse employees/contract staff was greater than 999; if the number of FTE licensed practical/vocational nurse employees/contract staff was greater than 999; if the number of FTE personal care aide employees/contract staff was greater than 999; if the number of FTE social work employees/contract staff was greater than 99; and if the number of FTE activities employees/contract staff was greater than 99.

2. Hours per participant day, by employee staff type (i.e., RNHPPD1, LPNHPPD1, AIDEHPPD1, SOCWHPPD1, and ACTHPPD1), and by contract staff type (i.e., RNHPPD2, LPNHPPD2, AIDEHPPD2, SOCWHPPD2, and ACTHPPD2):
 - a. Hours per participant day were derived from the number of full-time equivalents for each staff type and the average daily attendance (AVGPART for Question 2). The number of FTEs for a given employee staff type/ contract 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), any contract staff (ANYRN_CON, ANYLPN_CON, ANYAIDE_CON, ANYSOCW_CON, ANYACT_CON), and any employee or contract staff (ANYRN_EMPCON, ANYLPN_EMPCON, ANYAIDE_EMPCON, ANYSOCW_EMPCON, ANYACT_EMPCON), by staff type
 - a. These variable were derived from the FTE variables for employees and FTE variables for contract staff (e.g., RNFTE1 to derive ANYRN_EMP; RNFTE2 to derive ANYRN_CON; and both RNFTE1 and RNFTE2 to derive ANYRN_EMPCON) indicating whether the adult day services center had any RNs who are employees, any RNs who are contract staff RN, and any RNs who are employees or contract staff, respectively.

4. Having a computerized system that supports electronic health information exchange with physicians, pharmacies, or hospitals (ANYEX):
 - a. This variable was derived from ITMD, ITPHARM, and ITHOSP (Question 26).

5. Diagnosed conditions (DXALZ2, DXDEP2, DXSMI2, DXDD2, DXCADIO2, DXDIAB2), activities of daily living (TRANSHELP2, BATHHELP2, EATHHELP2, DRESHELP2, BEDHELP2, WALKHELP2)
 - a. Instruction was provided in the questionnaire to enter “0” if the adult day services center had no participants with diagnosed conditions listed in Question 18 or activities of daily living (ADL) limitations listed in Question 19. Yet, there were cases where respondents indicated the number of participants with a given diagnosed condition/ ADL limitation in the response box only when specific diagnosed conditions/ ADL limitation categories were applicable, while leaving inapplicable response boxes blank. We coded missing as “0” unless responses to all response boxes for Question 18 or Question 19 were blank or missing. Otherwise, we kept the missing (-9) as missing (-9).

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; the interviewer inadvertently skipped a question due to problems relating to CATI; or if the interview broke off before the entire questionnaire could be administered. The variables with the highest item nonresponse were select staffing variables: Question 14c.a AIDEFT1 (6.6%, weighted), Question 14e.b ACTPT2 (17.6%, weighted), Question 13b STAFFFT (12.2%, weighted), and Question 13a STAFFPT (33.6%, weighted).

Imputed data

In the data file, item nonresponse is coded as “-9= Not ascertained.” Missing values for revenue sources (Question 9), race-ethnicity (Question 15), sex (Question 16), and age (Question 17) variables, DXALZ (Question 18a), DXDEP (Question 18d), EATHELP (Question 19b), and BATHHELP (Question 19e) were imputed. In addition to the original variables, five sets of imputed variables are provided in the data file. Imputed variables are indicated by adding “imp” as the prefix and a numeral as the suffix to the variable name (e.g., impbathhelp_1, impbathhelp_5). A flagging variable is also included to indicate cases imputed for the variable (e.g., BATHHELP_FL). Among 2,763 respondents, the percentage of imputed records ranged from 3.7% (103 missing responses) for the “Hispanic or Latino, of any race” response category for the race variable (HISPANIC for Question 15a) to 10.1% (279 missing responses) for the “Depression” response category for the diagnosed condition variable (DXDEP for Question 18d).

After the weights were finalized, multiple imputations were created using the Cox-Iannacchione Weighted Sequential Hot Deck (WSHD) procedure in SUDAAN.

- a. 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 variable names ending with “RC” (e.g., HISPANICRC, AG85UPRC, BATTHHELPRC, DXDEPRC, REVMCARERC), class variables specified for the imputation procedure include: state, maximum number of participants allowed, ownership type, chain affiliation, Medicaid participation status, and metropolitan statistical area status. For variable names ending with “RC2” (e.g., REVMCARERC2), imputation class variables were: state, maximum number of participants allowed,

ownership type, chain affiliation, number of participants with Medicaid paying for long-term care services and metropolitan statistical area status. 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.

Reliability of Estimates

The data collected in the 2014 NSLTCP adult day services center (ADSC) survey 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 the amount of nonresponse. Although the records that make up the adult day services center file were not sampled, the variability associated with the nonresponse 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 and state estimates for adult day services centers in the United States. These weights adjust for nonresponse and unknown eligibility in the survey.

In this document, examples of SUDAAN and STATA computer code are provided for illustrative purposes. However, the appropriate application of these procedures is ultimately the responsibility of the data users. NCHS strongly recommends that users analyze the NSLTCP survey data under the direction of or in consultation with a statistician who is knowledgeable in sampling methodologies and techniques for the analysis of complex survey data.

Table 1a. Computations using SUDAAN

| PROC statement | NEST statement | TOTCNT statement | WEIGHT statement |
|----------------------------------|----------------------------------|--------------------|---------------------|
| PROC x FILE = y DESIGN = WOR; | NEST = FACSTRAT/ MISSUNIT; | TOTCNT= POPFAC; | WEIGHT= FACFNWT; |

Table 1b. Computations using STATA

| |
|---|
| Design description in STATA |
| svyset facid [pweight=facfnwt], strata(facstrat) fpc(popfac) vce(linearized) singleunit(missing) |

Accessing the Restricted Data File

The 2014 NSLTCP ADSC survey restricted data file 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 only for statistical reporting and analysis.
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 file or documentation to the Long-Term Care Statistics Branch (LTCSB).

We also request the user to inform LTCSB of any publications or presentations produced based on the 2014 NSLTCP ADSC survey data, and cite relevant NSLTCP documentations/ data products in their work when appropriate.

Contact Information

For questions, suggestions, or comments concerning the NSLTCP data, please contact the LTCSB at:

Long-term Care Statistics Branch (LTCSB), NCHS,

3311 Toledo Road, Hyattsville, MD 20782

E-mail: lcsbfeedback@cdc.gov

Phone: 301-458-4747