

2012 National Study of Long-Term Care Providers (NSLTCP)

Adult Day Services Center Survey

Restricted Data File

October 2013

Data Description and Usage (Readme)

Table of Contents

Data File	1
Documentation	1
Brief description of survey	1
Questionnaire	2
Data dictionary.....	2
Data Processing Activities to Create the Restricted File	3
Consistency checks.....	3
Derived variables	4
Item nonresponse	5
Imputed data.....	5
Reliability of Estimates.....	6
Accessing the Restricted Data File	7
Contact Information.....	7

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 2012 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 3,212 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) that represents one adult day services center that completed the survey. Also included in the data file are 303 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 and social work 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 September 2012 and February 2013. The ADSC 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. During data collection 42 centers were self-identified to have been in operation on or before May 31, 2012, and were added to the frame, resulting in a final sampling frame of 5,254 adult day services centers.

The survey for the ADSC component of the 2012 NSLTCP was based on a census of U.S. centers. A total of 476 (9.1%) ADSCs were identified as invalid or out of business. All remaining centers (4,778) were assumed eligible. Data were collected by three modes: self-administered hardcopy questionnaire, self-administered web questionnaires, and CATI conducted by interviewers. The questionnaire was completed for 3,212 centers, for a weighted response rate of 67.2%.

Questionnaire

The PDF for the adult day services center questionnaire of the 2012 NSLTCP is available at http://www.cdc.gov/nchs/nsltcp/nsltcp_questionnaires.htm. 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 does not have a skip instruction between Question 8 (MEDICAID) and Question 9 (MEDPAID). The intent of Question 9 is to measure the number of participants receiving long-term care services that are financed by Medicaid through direct payment to the adult day services center, as opposed to the number of participants whose care, regardless of where they receive it, is paid by Medicaid. In order to make this intent clear and be able to track changes over time, we decided to implement the skip logic in the 2012 survey data, and will add the skip logic in the future NSLTCP surveys. Thus, in the data file, if centers reported not being certified or otherwise set up to participate in Medicaid (MEDICAID=2), their responses to MEDPAID were coded as “-1=Legitimate skip.” Items following a skip pattern in the data file are identified in the data dictionary.

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 survey were identified by bold, uppercase text and were often followed by an arrow (e.g., in Question 3, the instruction reads “→ **SKIP TO QUESTION 4**”), 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 were 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. The hardcopy mail questionnaire also included skip instructions using arrows and uppercase bold text.
 - a. 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 they did not collect/track a given health information or left it unanswered (e.g., Question 40a_1; TRMDINFO=“2= No” or “-9=Not ascertained”), then their responses to whether they have the computerized capability to collect/track this health information (e.g., Question 40a_2; ITMDINFO) were coded as “-1= Legitimate skip.”
 - b. For Question 3 (OTHOWN) and Question 3a (OWNHOSP to OWNOTH), the responses to OTHOWN were edited based on responses to OWNHOSP through OWNOTH. If OTHOWN was left blank or answered as “2= No”, but the respondent answered “1=Yes” to OWNHOSP, OWNSNF, OWNHHA, OWNHOS, or OWNRCF, then OTHOWN was recoded to “1= Yes.”

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 5). 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 27).
 - 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 for that specific case or making proportional adjustments to individual categories when values did not total to TOTPART.

Derived variables

1. Number of full-time equivalents by employee staff type (i.e., Question 23: RNFTE1_RC, LPNFTE1_RC, AIDEFTE1_RC, SOCWFTE1_RC):
 - a. These variables were derived and provided in the restricted file. Number of full-time and the number of part-time employees for a given staff type 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. Outliers 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 (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 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; and if the number of FTE social work employees was greater than 99.
2. Hours per participant day, by employee staff type (i.e., RNHPPD, LPNHPPD, AIDEHPPD, SOCWHPPD):
 - 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 6). 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). For a few cases, AIDEHPPD had values greater than 24; these values were coded as 24.

3. Having the computerized capability to collect/track any health information (ANYIT):
 - a. This variable was derived from variables (e.g., ITMEDINFO, ITDEM, ITFUNC, ITORDER, ITVIEW for Question 40), indicating whether the adult day services center had the computerized capability to collect/track 19 different types of health information.
4. Having a computerized system that supports electronic health information exchange with physicians or pharmacies (ANYEX):
 - a. This variable was derived from ITMD and ITPHARM (Question 41).

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. A set of critical items were identified prior to fieldwork; these items were considered necessary for data analysis as they capture essential information about adult day services centers (e.g., ownership type, owned by another type of organization, Medicaid participation, number of enrolled participants, number of nursing employees, number of male residents, number of participants diagnosed with Alzheimer's disease and other dementias). When returned hardcopy mail questionnaires had critical items missing, interviewers conducted data retrieval via CATI to retrieve missing data for these items; 1,148 cases were sent to CATI for data retrieval. Of those cases, interviewers obtained missing data for 918 of these cases. For another 71 cases, interviewers obtained answers to some of the missing data or confirmed that the respondent wanted to leave the items as missing. Combined, this represents about 86% of the data retrieval cases.

Imputed data

In the data file, item nonresponse is coded as “-9= Not ascertained.” Missing values for EATHELP (Question 32c), BATHHELP (Question 32e), DXALZ (Question 30a), DXDEP (Question 30d), race-ethnicity (Question 26), sex (Question 27), and age (Question 28) variables, 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 3,212 respondents the percentage of imputed records ranged from 2.3% (73 missing responses) for the “Black, not Hispanic or Latino” response category for the race variable (BLACK for Question 26d) to 6.5% (209 missing responses) for the “With bathing or showering” response category for the assistance with activities of daily living variable (BATHHELP for Question 32e). 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. Class variables specified for the imputation procedure include: state, number of total participants, ownership type, chain affiliation, Medicaid participation, and metropolitan statistical area status.

Reliability of Estimates

The data collected in the 2012 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 non-response. Although the records that make up the adult day services center file were not sampled, the 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 and state estimates for adult day services centers in the United States. These weights adjust for non-response 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 2012 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 2012 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: ltcsbfeedback@cdc.gov

Phone: 301-458-4747