2018 National Study of Long-Term Care Providers

Residential Care Community (RCC) Resident Restricted Data File

March 2021

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 some of the processes involved in creating the residential care community (RCC) resident restricted 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. For the remainder of this document NPALS will be referred to as NSLTCP in order to correctly match the name of the study when the 2018 surveys were fielded.

Data Files

The 2018 NSLTCP RCC restricted data are distributed in two data files: (1) a provider-level data file and (2) services user (resident)-level data file. This document refers to the resident restricted data file. The file contains one record for each sampled resident (residents were not interviewed; interviews were conducted with the directors or caregivers). This file contains resident characteristics, use of services, health status, and cognitive and physical functioning. This file contains 904 records and over 130 variables. Each record has a primary identifier (RESID). The records in the resident file are sorted in order by the primary identifier.

The restricted data are provided in SAS and STATA 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; and this provider-specific data description and usage or readme document. A separate data description and usage is available for the adult day services center component of NSLTCP. A data dictionary or codebook listing the questions and response categories (without the unweighted frequencies and weighted estimates) will be made available to data users upon request.

Brief description of survey

The survey on RCCs was conducted between July 2018 and February 2019. To be eligible for

the study an RCC had to be licensed, registered, listed, certified, or otherwise regulated by the state; had four or more licensed, registered, or certified beds; provided room and board with at least two meals a day, around-the-clock on-site supervision, and help with personal care, such as bathing and dressing or health related services such as medication management. RCCs had to serve a predominantly adult population. RCCs licensed to exclusively serve the mentally ill or the intellectually disabled/developmentally disabled populations were excluded from NSLTCP. Data were collected by mail, web, and computer-assisted telephone interviews (CATI).

From a frame of 43,770 RCCs, 2,090 were randomly selected for the survey. Of the 2,090 sampled RCCs, eligibility could not be determined for 977. Among those cases where eligibility could be determined (1,113), 857 (77%) were eligible and 354 (23%) were ineligible because they did not meet the survey criteria or were out of business. However, 977 RCCs (33%) could not be contacted; therefore, the final eligibility status of these RCCs was unknown.

Using the eligibility rate of 77%,[1] a proportion of these RCCs of unknown eligibility was estimated to be eligible; 752 RCCs of unknown eligibility were assumed as eligible. The total number of eligible RCCs was estimated as 1,609 (857 + 752). Of the 1,609 in-scope and presumed in-scope RCCs, 503 completed the provider questionnaire, for a weighted response rate (for differential probabilities of selection) of 30% (this is calculated by using AAPOR's Response Rate 4). To account for the RCCs of unknown eligibility, the weights of the RCCs with known eligibility were adjusted upward based on the proportion of communities that were actually known to be eligible. Adjustments were also made to account for non-response. Two residents were randomly selected from eligible RCCs that agreed to participate. A questionnaire was completed for a total of 904 residents from 452 RCCs. About 48.9% (419 out of 857) of RCCs completed both provider and at least one resident module and the rest of the RCCs either completed the provider module or the resident module. Therefore, not all

---

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

RCCs in the provider file are represented in the resident file and vice versa. Some RCCs only completed one resident questionnaire. During the weighting process, these residents were moved to a similar RCC that had user records in the file.

Data dictionary

The 2018 RCC resident data dictionary or codebook for the restricted data file is provided as a single file containing all variables in the resident file. Each variable in the 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."

Resident Questionnaire

The Resident Questionnaire is included in the data release package and available at: https://www.cdc.gov/nchs/data/nsltcp/2018-NSLTCP-RCC-Questionnaire-Resident.pdf

The questionnaire includes all the questions asked in the resident module. There may be some differences in how questions were asked in the questionnaire and how they were coded in the data file.

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. Data were reviewed for accuracy, logic, consistency, and completeness.

Item nonresponse and imputed data

Item nonresponse is a source of missing data and occurred 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 -9 when a respondent did not provide an answer. NCHS handled item nonresponse for age (14 cases), sex (13 cases), race and ethnicity (37 cases) by imputing using a hot-deck method. Imputed variables are flagged for the data users.

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 2018 NSLTCP were obtained through a complex, multistage sample design that involved stratification and clustering. The final weights provided for analytic purposes have been adjusted in several ways to yield valid national estimates for residential care communities 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 NSLTCP services user file. The clustering protocols that are used in the multistage selection of the NSLTCP 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 provided also for the SAS and STATA software packages. However, the appropriate application of these procedures is ultimately the responsibility of users. NCHS

strongly recommends that NSLTCP 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 RCC 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 for using these design variables with SUDAAN, STATA, and SAS survey procedures follow.

Table 1a. Computations using SUDAAN

PROC statement	NEST statement	TOTCNT statement	WEIGHT statement
PROC x FILE = y DESIGN = WOR;	NEST STRATA;	TOTCNT POPFAC; POPSU	WEIGHT SUWT;

Table 1b. Computations using STATA

Design description in STATA
svyset su_facid, strata(strata) fpc(popfac) vce(linearized) singleunit(missing)‖ResidentID, fpc (popsu) weight (suwt)

Table1c. Computations using SAS

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

Obtaining the data

The RCC 2018 resident file can be accessed through NCHS' 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 RCC provider data or documentation files to the Long-Term Care Statistics Branch (LTCSB).

NCHS also requests the user inform LTCSB of any publications or presentations produced based on the 2018 NSLTCP data and cite relevant NSLTCP documentations/data products in their work when appropriate.

Contact Information

To request a codebook or for questions, suggestions, or comments concerning NSLTCP data, please contact LTCSB at:

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

3311 Toledo Road, Hyattsville, MD 20782

E-mail: ltcsbfeedback@cdc.gov

Phone: 301-458-4747.