

# 2010 National Survey of Residential Care Facilities

## Contents

Description .....	2
Sampling Design .....	2
Sampling Frame.....	3
Scope of Survey.....	3
Data Collection Procedures.....	4
Estimation Procedures .....	5
Reliability of Estimates.....	7

## Description

The 2010 National Survey of Residential Care Facilities (NSRCF) is a first-ever national probability sample survey that collects data on U.S. residential care providers, their staffs and services, and the people they serve. It is designed to provide national estimates of the number of residential care facilities operating in the United States, the number of residents receiving care, and the characteristics of both the facilities and their residents. NSRCF was conducted between March and November 2010. All residential care facilities that participated in the survey were places that were licensed, registered, listed, certified, or otherwise regulated by the state and that had 4 or more licensed, certified, or registered 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. These facilities served a predominantly adult population and had at least one current resident. Facilities licensed to serve the mentally ill or the developmentally disabled populations exclusively were excluded from the survey.

The 2010 NSRCF data were collected through in-person interviews with facility directors and their designated staffs; no interviews were conducted directly with residents. Facility data included questions on facility characteristics such as ownership, size, types of living arrangements and amenities, policies, staffing, services, and general resident characteristics. Data collected on residents included questions on the sampled residents' demographics, living arrangements, activities, health conditions, cognitive and physical functioning, and services received. The total number of facilities that participated in the 2010 NSRCF is 2,302, and data are available on 8,094 residents from these facilities. A detailed methods report on the 2010 NSRCF is available on the NSRCF website at: <http://www.cdc.gov/nchs/nsrcf.htm>.

## Sampling Design

The 2010 NSRCF used a stratified two-stage probability sampling design. The first stage, carried out by the Centers for Disease Control and Prevention's National Center for Health Statistics (NCHS), was the selection of residential care facilities from the sample frame of over 39,000 facilities, representing the universe of facilities providing residential care services in the United States. The primary sampling strata of facilities were defined by size (number of beds) and census region. Within these sampling strata, facilities were sorted by metropolitan statistical area (MSA) status and state. For the 2010 NSRCF, 3,605 facilities were systematically and randomly sampled with probability proportional to size. A detailed methods report on the 2010 NSRCF

“Design and Operation of the 2010 National Survey of Residential Care Facilities“ is available on the NSRCF web site at: <http://www.cdc.gov/nchs/nsrcf.htm>

The second stage of sample selection was completed by the interviewers during the facility in-person interviews. Current residents were randomly selected by a computer algorithm, based on a census list provided by each facility director or his/her designee. Up to 6 current residents were randomly selected per facility. Current residents were defined as residents who were on the rolls of the facility as of midnight of the day immediately before the facility interview.

## Sampling Frame

The sampling frame for the 2010 NSRCF was constructed from lists of licensed residential care facilities acquired from the licensing agencies in each of the 50 states and the District of Columbia. The state lists were checked for duplicate facilities and concatenated to form a list of all residential care facilities, resulting in a sampling frame of 39,635.

## Scope of Survey

For the 2010 NSRCF, a sample of 3,605 facilities was selected. Only facilities that were licensed, registered, certified, listed or otherwise regulated by their state; served primarily an adult population; had at least one current resident; had four or more licensed, certified, or registered residential care beds; and 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 were eligible to participate in NSRCF. Facilities that exclusively served the severely mentally ill, the mentally retarded/developmentally disabled, or both were ineligible. Of the 3,605 facilities in the sample, 2,969 (82 percent) were considered in scope. The 636 out-of-scope facilities were ineligible for one or more of the following reasons: did not meet the definition used in the survey, had gone out of business, or was combined with another sampled facility. Of the in-scope facilities, 2,302 agreed to participate, resulting in a first-stage facility unweighted response rate of 79 percent and weighted response rate of 81 percent. A total of 8,284 current residents were sampled from the responding facilities. Of these, 66 residents were considered out of scope; and resident interviews were not completed for an additional 124 sampled residents for other reasons. This resulted in 8,094 resident cases, for a second-stage unweighted response rate of 98 percent and a weighted response rate of 98.5 percent. The overall weighted survey response rate was 79.4% (i.e., facility weighted response rate \* resident weighted response rate). Weighted and unweighted response rates are reported

per Office of Management and Budget's (OMB) September 2006 Standards and Guidelines for Federal Statistics. Weighted rates measure the proportion of the total population that is represented by respondents while unweighted rates reflect only the proportion of the sample that responded. A detailed methods report on the 2010 NSRCF is available on the NSRCF website at: <http://www.cdc.gov/nchs/nsrcf.htm>

## **Data Collection Procedures**

The NSRCF data collection effort which was conducted between March and November 2010 consisted of two distinct activities—facility recruitment and facility in-person interviewing—conducted by separate staffs. Recruiters' responsibilities included: calling facilities to obtain or verify contact information for the director; sending an advance mail package; calling the director to explain the study and gain cooperation; administering a brief CAPI screening instrument to determine facility eligibility; and scheduling an in-person visit with the director for field interviewers to complete the facility interview. Interviewers' responsibilities included: making an appointment reminder call to the facility; and making an in-person visit to the facility to conduct the interview. The in-person interview consisted of administering the facility questionnaire; the resident selection questionnaire; the resident interviews; and after leaving the facility, completing the debriefing questionnaire.

During the recruitment phase, after contact information was verified by recruiters, facility directors were mailed an NSRCF folder of advance materials that contained: a personalized letter from the director of NCHS with Responses to Frequently Asked Questions (FAQs) printed on the back; an NCHS Ethics Review Board (ERB) approval letter; a brochure specifically about NSRCF; an NCHS' Confidentiality Brochure- How the National Health Care Survey Keeps Your Information Strictly Confidential; and a letter of support from the American Association of Homes and Services for the Aging (AAHSA), American Seniors Housing Association (ASHA), Assisted Living Federation of America (ALFA), National Center for Assisted Living (NCAL), and Board and Care Quality Forum. After this, recruiters called directors to explain the survey in further detail, address any questions or concerns about NSRCF, get their commitment to participate in the survey, administer the screening questionnaire to determine facility eligibility, and set a date and time for the in-person interview. Prior to the scheduled appointment, directors were mailed additional information about their interview including instructions for preparing a list of their current residents and completing the pre-interview worksheet (PIW) which contained questions asked during the in-person interview that likely would require referring to records or consulting other staff.

For the in-person interview, interviewers administered the survey using a computer-assisted personal interviewing (CAPI) instrument. CAPI consisted of three modules. The facility questionnaire, which was completed with the director or director's designee, included questions on facility characteristics such as ownership, size, types of living arrangements and amenities, policies, staffing, services, and general resident characteristics. The resident selection questionnaire was used to identify from three to six residents on the facility census lists to select for the resident questionnaires. The resident questionnaires, completed with the director or other caretaker, included questions on the sampled residents' demographics, living arrangements, activities, health conditions, cognitive and physical functioning, and services received. Respondents referred to residents' medical and other records as necessary. No residents were interviewed directly.

After the NSRCF data were collected, they were edited to ensure that all responses were accurate, consistent, logical, and complete. A detailed methods report on the 2010 NSRCF is available on the NSRCF website at: <http://www.cdc.gov/nchs/nsrcf.htm>.

## **Estimation Procedures**

Because the statistics from NSRCF are based on a sample, they differ from the data that would have been obtained if a complete census had been taken using the same definitions, instructions, and procedures. However, the probability design of the NSRCF sample permits the calculation of estimates and sampling errors. The standard error of a statistic is primarily a measure of sampling variability that occurs by chance because only a sample, rather than the entire population, is surveyed. The standard error also reflects part of the variation that arises in the measurement process but does not include any systematic bias that may be in the data, or any other nonsampling error. The chances are about 95 in 100 that an estimate from the sample differs by less than twice the standard error from the value that would be obtained from a complete census.

Standard errors can be calculated for facility and resident estimates by using any statistical software package, as long as clustering within facilities and other aspects of the complex sampling design are taken into account. Software products such as SAS, Stata, and SPSS have these capabilities. Statistics presented in NCHS publications are computed using the linearized Taylor series method of approximation as applied in SUDAAN software, which produces standard error estimates for statistics from complex sample surveys. Both of the NSRCF public-use files (facility and resident) include design variables that designate each record's stratum marker and the first-stage unit (or cluster) to which the record belongs.

In the facility public-use file, the variable STRATUM indicates the sampling stratum for bed size group and region, and the facility indicated by the variable FACILID is the primary sampling unit. POPFAC represents the total number of facilities for calculating the finite population correction in a stratum. The survey weight is indicated by FACFNWT. The data dictionary for the facility public-use file has a Technical Notes section that provides an example of the syntax for using these design variables to describe the sampling design in SUDAAN. The NSRCF data dictionary for the facility public-use file is available on the NSRCF website at: [ftp://ftp.cdc.gov/pub/Health\\_Statistics/NCHS/Dataset\\_Documentation/NSRCF/2010/2010NSRCF\\_FacilityPublicUseDataDictionary](ftp://ftp.cdc.gov/pub/Health_Statistics/NCHS/Dataset_Documentation/NSRCF/2010/2010NSRCF_FacilityPublicUseDataDictionary)

The resident public-use file has two stages. The stratum in the first stage is indicated by the variable RSTRATUM, in which the primary sampling unit is the facility indicated by the variable FACID. The variable for total facilities needed to calculate the finite population correction at the first stage is RPOPFAC. In the second stage, the sampling unit is the resident indicated by the variable RESNUM. In the resident public-use file, the second stage is treated as if sampling was done with replacement. In SUDAAN, to treat the second stage as if sampling was with replacement, the variable POPRES is used and has a value of -1. Many other statistical packages assume sampling with replacement if no variable for the total population at the second stage is provided. The variable for the survey weight is RESFNWT. The data dictionary for the resident public-use file has a Technical Notes section that provides an example of the syntax for using these design variables to describe the sampling design in SUDAAN. The NSRCF data dictionary for the resident public-use file is available on the NSRCF website at: [ftp://ftp.cdc.gov/pub/Health\\_Statistics/NCHS/Dataset\\_Documentation/NSRCF/2010/2010NSRCF\\_ResidentPublicUseDataDictionary](ftp://ftp.cdc.gov/pub/Health_Statistics/NCHS/Dataset_Documentation/NSRCF/2010/2010NSRCF_ResidentPublicUseDataDictionary)

The resident sample was selected from individuals residing at the sampled facility on the night before data collection began and represents residents living in residential care communities on any given day between March and November 2010.

Because NSRCF is a sample survey, data analyses must include survey weights, to inflate the sample numbers to national estimates. The weight associated with each sampled facility and each sampled resident is constructed to account for the multistage sampling design. An estimator  $\hat{X}$  for any given population total  $X$  can be expressed as a weighted sum over all sampled units, defined as

$$\hat{X} = \sum_u x(u)W(u)$$

where  $u$  represents a sampled unit,  $x(u)$  is the characteristic or response of interest for unit  $u$ , and  $W(u)$  is the final survey weight for sampled unit  $u$ . The final weight  $W(u)$  for each sampled unit is the product of two components:

1. Inverse of the probability of selection.
2. Nonresponse adjustment.

The first component of the weight for each sampled unit (facility or resident) is the inverse of the unit's selection probability. For the current resident, the selection probability is the product of two selection probabilities: the probability of selecting the facility to the NSRCF sample and the probability of selecting the current resident within the sampled NSRCF facility. The inverse of the product of these probabilities is used for weighting.

The first component was corrected to account for duplicate listings of sampled facilities in the sampling frame when duplicates were identified after the start of field work. To the extent that all duplicates of sampled facilities were identified, the corrected weights produce unbiased estimates (i.e., estimates that would be obtained if no facilities were duplicated in the sampling frame).

The second component for calculating the weight is adjustment for nonresponse. This adjustment is made for three types of nonresponse. The first two types are facility level, and the third is resident level. The first type occurs when in-scope facilities do not respond to NSRCF. In NSRCF, the second type occurs when an in-scope facility does not provide the number of current residents within the respective facility. The third type occurs when the facility does not provide information requested in the survey about the sampled resident.

Finally, the weights described above were smoothed within groups defined by census region, size and MSA status if there were outlier sampling units whose survey weights were somewhat larger than those for the remaining sample in the same group. In smoothing, total estimates for each group were preserved.

## Reliability of Estimates

NCHS bases publication of estimates for NSRCF on the relative standard error (RSE)—also known as the coefficient of variation—of the estimate and the number of sampled records on which the estimate is based. The RSE is a measure of variability and is calculated by dividing the standard error (SE) of an estimate by the estimate itself. The result is then converted into a percent by multiplying it by 100. Guidelines used by NCHS authors to determine if estimates should be presented in tables of NCHS published data reports include:

- If the estimate is based on fewer than 30 cases, then the value of the estimate is not reported. This is usually indicated with an asterisk (\*).
- If the estimate is based on a sample of 30–59 cases or on 60 or more cases and the RSE is 30 percent or more, then the estimate is reported but should not be assumed reliable. This is usually indicated with a dagger (†) preceding the figure in the tables.
- If the estimate is based on 60 or more sample cases and the RSE is less than 30 percent, then the estimate is reported and is considered reliable.