2007 National Home and Hospice Care Survey
and
National Home Health Aide Survey

Contents

Description

National Home and Hospice Care Survey

The 2007 National Home and Hospice Care Survey (NHHCS) is one in a continuing series of nationally representative sample surveys of U.S. home health and hospice agencies. It is designed to provide descriptive information on home health and hospice agencies, their staffs, their services, and their patients. NHHCS was first conducted in 1992 and was repeated in 1993, 1994, 1996, 1998, and 2000, and most recently in 2007.

NHHCS, conducted between August 2007 and February 2008, was reintroduced into the field in 2007 after a 7-year break. During that time, the survey was redesigned and expanded to include a computer-assisted personal interviewing (CAPI) system, many new data items, and larger sample sizes of current home health patients and hospice discharges. All agencies that participated in the survey were either certified by Medicare and/or Medicaid or were licensed by a state to provide home health and/or hospice services and currently or recently served home health and/or hospice patients. Agencies that provided only homemaker services or housekeeping services, assistance with instrumental activities of daily living (IADLs), or durable medical equipment and supplies were excluded from the survey. The 2007 NHHCS included a supplemental survey of home health aides employed by home health and/or hospice agencies, called the National Home Health Aide Survey (NHHAS).

The 2007 NHHCS data were collected through in-person interviews with agency directors and their designated staffs; no interviews were conducted directly with patients or their families and/or friends. Agency data collected, available in agency administrative records, included information on the year an agency was established, the types of services an agency provided, referral sources, specialty programs, and staffing characteristics. Data collected on home health patients and hospice discharges, available in medical records, included age, sex, race and ethnicity, services received, length of time since admission, diagnoses, medications taken, advance directives, and many other items. The total number of agencies that participated in the 2007 NHHCS is 1,036, and data are available on 9,416 current home health patients and hospice discharges from these agencies. A detailed methods report on the 2007 NHHCS will be available in the near future on the NHHCS website at http://www.cdc.gov/nchs/nhhcs.htm.

National Home Health Aide Survey

The National Home Health Aide Survey (NHHAS), the first national probability survey of home health aides, was designed to provide national estimates of home health aides employed by agencies that provide home health and/or hospice care. NHHAS was sponsored by the Office of the Assistant Secretary for Planning and Evaluation (ASPE). NHHAS, a multistage probability sample survey, was conducted as a supplement to the 2007 NHHCS. Agencies providing home health and/or hospice care were sampled into NHHCS, and then up to six home health aides were sampled from eligible participating NHHCS agencies. Home health aides were considered eligible to participate in NHHAS if they were 1) directly employed by the sampled agency; and 2) provided assistance in activities of daily living (ADLs), including bathing, dressing, transferring, eating, and toileting. NHHAS was administered to aides during their nonworking

hours by interviewers who used a computer-assisted telephone interviewing (CATI) system to collect the data. The survey instrument included sections on recruitment, training, job history, family life, management and supervision, client relations, organizational commitment and job satisfaction, workplace environment, work-related injuries, and demographics. The NHHAS questionnaire was virtually identical to the survey instrument used in the 2004 National Nursing Assistant Survey of certified nursing assistants working in nursing homes, to permit comparisons of direct care workers across long-term care workplace settings. Minor changes were made to account for differences in workplace environment and responsibilities between home health aides and certified nursing assistants. A total of 3,377 interviews of aides working in agencies providing home health and/or hospice care were completed between September 2007 and April 2008. A detailed methods report on the 2007 NHHAS will be available in the near future on the NHHCS website at http://www.cdc.gov/nchs/nhhcs.htm.

Sample Design

NHHCS
The 2007 NHHCS used a stratified two-stage probability sample 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 home health and hospice agencies from the sample frame of over 15,000 agencies, representing the universe of agencies providing home health care and hospice services in the United States. The primary sampling strata of agencies were defined by agency type and metropolitan statistical area (MSA) status. Within these sampling strata, agencies were sorted by census region, ownership, certification status, state, county, ZIP code, and size (number of employees). For the 2007 NHHCS, 1,545 agencies were systematically and randomly sampled with probability proportional to size. A detailed methods report on the 2007 NHHCS will be available in the near future on the NHHCS website at http://www.cdc.gov/nchs/nhhcs.htm.

The second stage of sample selection was completed by the interviewers during the agency interviews. The current home health patients and hospice discharges were randomly selected by a computer algorithm, based on a census list provided by each agency director or his/her designee. Up to 10 current home health patients were randomly selected per home health agency, up to 10 hospice discharges were randomly selected per hospice agency, and a combination of up to 10 current home health patients and hospice discharges were randomly selected per mixed agency. Current home health patients were defined as patients who were on the rolls of the agency as of midnight of the day immediately before the agency interview. The hospice discharges were defined as patients who were discharged from the hospice agency during the 3-month period beginning 4 months before the agency interview. Discharges that occurred because of the death of a sampled hospice patient were included.

NHHAS
NHHAS is a linked establishment and worker survey, similar to the design of the National Nursing Assistant Survey (http://www.cdc.gov/nchs/nnas2004.htm). NHHAS is based on a two-

stage probability sample design with the NHHCS agency sampled first and a random selection of aides from each of the participating sampled NHHCS agencies sampled second.

The first stage consisted of the selection of a stratified probability sample of agencies from a sample frame of over 15,000 agencies, representing the universe of agencies providing home health care and hospice services in the United States. The sample frame was stratified by type of services the agency provided and MSA status. Within these primary strata, agencies were sorted by census region, ownership, certification status, state, county, ZIP code, and size (number of employees). Then, 1,545 agencies were systematically and randomly selected with probability proportional to size. A detailed methods report on the 2007 NHHCS will be available in the near future on the NHHCS website at http://www.cdc.gov/nchs/nhhcs.htm.

In the second stage of sampling, a random sample of up to six aides was selected from each agency eligible for and participating in NHHCS. Aides were eligible for the survey if they were directly employed by the agency and provided assistance with ADLs, including eating, toileting, bathing, dressing, and transferring. The aide sampling procedure started with the NHHCS in-person agency interview. During the NHHCS in-person interview with the agency respondent, the agency provided a list of aides who met the eligibility criteria as of midnight of the day immediately before the agency interview. The interviewer numbered the list and entered the total number of aides into the CAPI system used for the NHHCS survey and sampling. The CAPI program, through systematic randomization procedures, selected up to six aides. A total sample of 4,416 aides were sampled and fielded for NHHAS; 4,279 were eligible and 3,377 aides completed NHHAS telephone interviews. A detailed methods report on NHHAS will be available in the near future on the NHHCS website at http://www.cdc.gov/nchs/nhhcs.htm.

Sampling Frame

The sampling frame for the 2007 NHHCS was constructed using three sources: (1) The Centers for Medicare & Medicaid Services Provider of Services file of home health agencies and hospices, (2) State licensing lists of home health agencies compiled by a private organization, and (3) The National Hospice and Palliative Care Organization file of hospices. The combined files were matched and identified duplicates were removed, resulting in a sampling frame of 15,488.

Scope of Survey

NHHCS

For the 2007 NHHCS, a sample of 1,545 agencies was selected. Only agencies providing home health or hospice care services to patients at the time of the survey or recently before the survey were eligible to participate in the NHHCS. Of the 1,545 agencies in the sample, 1,461 (95 percent) were considered in scope. The 84 out-of-scope agencies were ineligible for one or more of the following reasons: did not meet the definition used in the survey, had gone out of business, was a duplicate of another sampled agency, or had merged with other sampled agencies. Of the in-scope agencies, 1,036 agreed to participate, resulting in a first-stage agency unweighted

response rate of 71 percent and weighted response rate of 59 percent. A total of 10,009 current home health patients and hospice discharges were sampled from the responding agencies: 5,026 current home health patients and 4,983 hospice discharges. Of these, 106 home health patients and 19 hospice discharges were considered out of scope. Furthermore, 237 current home health patients and 231 hospice discharges were excluded due to one of the following reasons: consent problems, record problems, refusals, ran out of time, and nonresponse. This resulted in 4,683 home health cases and 4,733 hospice cases, for a second-stage unweighted response rate of 95 percent (9,416/9,884) and weighted response rate of 96 percent. For the NHHCS patient health module, the overall unweighted response rate was 66 percent and the overall weighted response rate was 55 percent. 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 2007 NHHCS will be available in the near future on the NHHCS website at http://www.cdc.gov/nchs/nhhcs.htm.

NHHAS

Of the 1,036 agencies that participated in NHHCS, 52 agencies had no aides to sample, resulting in 984 agencies eligible to participate in NHHAS. Of these 984 agencies eligible for NHHAS, 22 refused to participate, and no aides were sampled at 7 additional agencies because the interviewer ran out of time or was otherwise unable to complete the aide sampling. As a result, aides were sampled from 955 eligible agencies, for an unweighted first-stage NHHAS response rate of 97 percent and a weighted response rate of 97 percent. From the 955 agencies, 4,416 home health aide cases were sampled and fielded. Of the 4,416 cases, 137 (3 percent) were ineligible for one of the following reasons: not employed on the sampling date, did not provide assistance with ADLs, were contract employees, were sampled in error, or were identified as ineligible during the aide interview because the respondent did not know whether she was an employee of the sampled agency. Thus, a total of 4,279 of the sampled cases were eligible and 3,377 aides completed the survey. At the second sampling stage, the unweighted response rate was 79 percent (3,377/4,279) and the weighted response rate was 71 percent (9,895/13,936). The overall unweighted NHHAS response rate was 54 percent (71 percent unweighted response rate for overall agency participation x 97 percent unweighted response rate for agencies participating in the NHHCS that also participated in the NHHAS, by providing a list of home health aides employed by their agency x 79 percent unweighted response rate for home health aides). The overall weighted response rate was 40 percent, using weighted response rates of the same components used to calculate the unweighted response rate (59 percent x 97 percent x 71 percent). Weighted and unweighted response rates are reported per OMB's 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 only reflect the proportion of the sample that responded. A detailed methods report on NHHAS, will be available in the near future on the NHHCS website at http://www.cdc.gov/nchs/nhhcs.htm.

Data Collection Procedures

NHHCS

The 2007 NHHCS was administered in sampled home health and hospice agencies, between August 2007 and February 2008, using a computer-assisted personal interviewing (CAPI) instrument that was loaded onto each interviewer's laptop. CAPI consisted of five modules: Agency Qualifications and Characteristics (AQ), Patient Sampling (PS), Patient Health (PH), Patient Charges and Payments (PA), and Aide Sampling (AS). A self-administered staffing questionnaire was also mailed to the agency directors who were asked to complete it before the in-person agency interview. The AQ module included agency qualifications and characteristics data items. Interviewers were instructed to complete the agency qualifications items first to ensure that the agency was eligible to participate in the survey. Interviewers were then free to administer the agency characteristics in the AQ module and the PS, PH, PA, and AS modules in any order depending on the availability of designated agency staff to answer the survey questions. The PH data items collected information about the health of current home health patients and/or hospice discharges as documented in their medical records. NHHCS also included a first-time supplemental survey of home health aides employed by home health and hospice agencies, the National Home Health Aide Survey (NHHAS).

Data were collected according to the following procedures: (1) An advance package of NHHCS information, including a letter from the NCHS director, was mailed to the director of each sampled agency, informing him/her of the purpose, content, and authorizing legislation of the survey and that he/she would be contacted by telephone to schedule an appointment. The advance package included letters of support from the National Hospice and Palliative Care Organization and from the National Association for Home Care and Hospice. Also included in the package was a copy of an NCHS report—The Use of Computerized Medical Reports in Home Health and Hospice Agencies: United States, 2000—to illustrate how the survey data can be used to present important findings in the industry. (2) After the package was mailed, the interviewer telephoned the sampled agency to speak to the director, explain the survey in further detail, address any questions or concerns about NHHCS, and schedule an in-person interview with the director. (3) After the interviewer successfully scheduled an interview, a confirmation package was mailed to the director. This package included a confirmation letter with details about agency information the interviewer would need to complete the interview, in addition to the self-administered staffing questionnaire that the director was expected to complete by the day of the agency interview. (4) At the in-person agency interview, the interviewer collected the completed staffing questionnaire and administered the AQ module of CAPI. Provided the agency was eligible to participate in the survey, the interviewer sampled up to 10 current home health patients/hospice discharges using the PS module of CAPI. In mixed agencies, a combination of up to 10 current home health patients and hospice discharges were sampled, usually 5 of each; if 5 of either group was not available, the interviewer sampled more from the group that had more than 5 on the census list. The interviewer completed the sampling exercise by cleaning (e.g., identifying and removing duplicate names on a list of current home health patients) and numbering the census lists and entering the total number of current home health patients and/or hospice discharges into CAPI. Subsequently, CAPI randomly selected 10 numbers based on the total number of current patients/hospice discharges that were entered into the computer algorithm. The sampled patients/discharges were those corresponding to the randomly generated numbers in the census list. (5) The interviewer met with designated staffs that were familiar with the sampled patients/discharges and their care and collected information

on the survey items in the PH and PA modules for each sampled patient/discharge. The respondents referred to patient medical records, administrative records, and medication administration records to answer the survey items. No patients or families/friends were interviewed directly. (6) The interviewer constructed a census list of currently employed home health aides, selecting up to six home health aides using the procedures described above for sampling patients/discharges, and requested contact information for each sampled home health aide. This information was used for NHHAS.

After the NHHCS data were collected, they were edited to ensure that all responses were accurate, consistent, logical, and complete. The medical information collected in the PH module was coded according to the *International Classification of Diseases, 9th Revision, Clinical Modification*. One primary admission diagnosis, one current primary diagnosis (or diagnosis at discharge for hospice patients), and up to 15 current secondary diagnoses (or diagnoses at discharge for hospice patients) were collected per current home health patient/hospice discharge. Up to five procedures were collected per sampled patient/discharge.

NHHAS

The 2007 NHHAS was administered by telephone using a computer-assisted telephone interviewing (CATI) system. The questionnaire included 11 modules, the first of which was a screening section to determine eligibility. In addition to the screening module, the questionnaire included modules on recruitment, education and training, job history, family life, management and supervision, client relations, job satisfaction, job rating, work-related injuries, and sociodemographics. Eligible home health aides who were no longer working at the agency when contacted for the telephone interview completed only the sections on eligibility, job history, demographics, and a section not completed by aides who were still working at the agency. This "agency leavers" section included questions on reasons for leaving the job and future plans.

Each home health aide selected for NHHAS received an advance package. The advance package included the following: a letter on NCHS letterhead that described the study, signed by the Director of NCHS with Frequently Asked Questions (FAQs) printed on the back; a $5 bill clipped to the letter signed by the NCHS Director; a welcome letter on NHHAS letterhead; a NHHAS fact sheet; a NHHAS DVD; a NHHAS gift pen; a postcard for the home health aide to indicate willingness to participate in the study and to provide name, address, telephone number and the best time and day to be reached; and a postage-paid return envelope for the postcard. These materials included a toll-free number that aides could call if they were interested in participating in NHHAS. The mode of providing the packages to the home health aides depended on whether the agency provided contact information for the sampled home health aides. If an agency provided address information for its sampled aides, the advance packages were mailed to the home health aides. If an agency did not provide address information for sampled aides, the advance packages were mailed to the agency to be distributed to the sampled aides.

One week after the advance packages were distributed, a reminder letter was sent to the home health aides. If address information was provided, the letter was mailed directly to the home health aides. If address information was not provided, the letters were mailed to the agency to distribute to the home health aides. For home health aides for whom the agency did not provide

contact information, a second reminder letter was mailed to the agency 1 week after the first reminder letter to be distributed to the selected home health aides.

Home health aides could indicate interest in participating in NHHAS by returning the postcard, calling the toll-free number listed on the advance package information materials, or by agreeing to participate when a telephone interviewer contacted them. After the NHHAS data were collected, extensive data checking, editing, and coding were performed to ensure that the responses were accurate, consistent, logical, and complete.

Estimation Procedures

Because the statistics from NHHCS and NHHAS are based on a sample, they will differ somewhat 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 NHHCS and NHHAS permit the calculation of 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 from the value that would be obtained from a complete census by less than twice the standard error.

Standard errors can be calculated for agency, patient/discharge, and home health aide estimates using any statistical software package as long as clustering within agencies and other aspects of the complex sample design are taken into account. Software products such as SAS, STATA, and SPSS all have these capabilities. Statistics presented in NCHS publications are computed using SUDAAN software that produces standard error estimates for statistics from complex sample surveys. SUDAAN employs a first-order Taylor Series approximation of the deviation of estimates from their expected values. All three of the NHHCS public-use files (i.e., agency, patient, and medication) and the NHHAS public-use file include design variables that designate each record's stratum marker and the first-stage unit (or cluster) to which the record belongs. The design variables used to estimate characteristics in the patient/discharge file are the same design variables that should be used for the medication data, which were collected at the patient/discharge level.

In the agency public-use file, the variable indicating the stratum of the stratified sampling is STRATUM and the primary sample unit is the observation (i.e., agency) indicated by the variable AGENCYID. The variable representing the population within a stratum for the finite population correction is POPAGY. There are two sample weights: (a) SAMAGYWT for estimates not correlated with agency size, and (b) SIZAGYWT for estimates correlated with agency size (e.g., estimates of total staff across all agencies). The data dictionary for the agency public-use file has a technical section that provides an example of the syntax for using these design variables to describe the sample design in SUDAAN. The NHHCS data dictionary for the agency public-use file is available on the NHHCS website at http://www.cdc.gov/nchs/nhhcs.htm.

The patient and discharge public-use file has two stages. The stratum in the first stage is indicated by the variable PSTRATA in which the primary sample unit is the agency indicated by the variable PTAGYNUM. The variable for the finite population correction in the first stage is POPAGN. In the second stage, the stratum is the variable PHTYPE and the secondary sample unit is the observation (i.e., patient or discharge) indicated by the variable PATNUM. There is no finite population correction in the second stage with the public-use file; thereby the second stage is treated as sampling with replacement. In SUDAAN, to treat the second stage as sampling with replacement the variable POPPAT is used for which the value is -1. In many other statistical packages, not designating a variable for finite population correction at the second stage results in treatment as sampling with replacement. The sample weight is SAMWT. The data dictionary for the patient and discharge public-use file has a technical section that provides an example of the syntax for using these design variables to describe the sample design in SUDAAN. The NHHCS data dictionary for the patient and discharge public-use file is available on the NHHCS website at http://www.cdc.gov/nchs/nhhcs.htm.

The current home health patient sample describes individuals receiving home health care on the night before data collection began and represents home health care utilization on any given day between August 2007 and February 2008. The hospice discharge sample describes the annual number of discharges from hospice care. This design requires the data user to always conduct separate analyses of current home health patients and hospice discharges, using the PHTYPE variable. For current home health patients, PHTYPE=1, and for annual hospice discharges, PHTYPE=2. In order to properly account for the sample design in the calculation of standard errors, both current home health patients and hospice discharges must be used in any analysis. Any analysis should be conducted using the subpopulation command in the statistical software package.

The home health aide public-use file has two stages. The stratum of the first stage is indicated by the variable ASTRATA in which the primary sample unit is the agency indicated by the variable HHAAGYID. The variable for the finite population correction in the first stage is POPAGY. In the second stage, the sample unit is the observation (i.e., the home health aide) indicated by the variable HHAID; there is no stratification at this stage. There is also no finite population correction in the second stage with the public-use file; thereby the second stage is treated as sampling with replacement. In SUDAAN, to treat the second stage as sampling with replacement the variable POPHHA is used in which the value is -1. In many other statistical packages, not designating a variable for finite population correction at the second stage results in treatment as sampling with replacement. The sample weight is SAMWT. The data dictionary for the home health aide public-use file has a technical section that provides an example of the syntax for using these design variables to describe the sample design in SUDAAN.

Because NHHCS and NHHAS are sample surveys and are designed to produce national estimates for agencies, patient/discharges (NHHCS), and home health aides (NHHAS), data analyses must include sampling weights to inflate the sample numbers to national estimates. Each record in the public-use files has a weight for this purpose. By aggregating the weights, national counts can be estimated.

NHHCS estimators take into account the selection procedures of the complete survey design to develop the final sample weight for each sampled agency and each sampled patient/discharge. NHHAS estimators take into account the selection procedures of the complete survey design to develop the final sample weight for each sampled agency and each sampled home health aide. An estimator $\hat{X}$ for any given population total X can be expressed as a weighted sum over all sample units, defined as:

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

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

1. Inverse of the probability of selection (NHHCS and NHHAS)
2. Nonresponse adjustment (NHHCS and NHHAS)
3. Ratio adjustment (NHHCS)

The first component of the weight for each sampled unit (agency, home health patient, hospice discharge) is the inverse of the unit's selection probability. For the home health patient or hospice discharge, the selection probability is the product of two selection probabilities: the probability of selecting the agency to the NHHCS sample and the probability of selecting the current home health patient and/or hospice discharge within the sampled NHHCS agency. The probability of selecting a home health aide is a product of two selection probabilities: the probability of selecting an agency to the NHHCS sample and the probability the home health aide was selected within the sampled NHHCS agency. The inverse of the product of these probabilities is used for weighting.

The first component was modified for sampled agencies found to have multiple listings in the sampling frame after the agency sample was selected. For each agency found duplicated in the sampling frame, the weights of all sampled listings for the agency were summed and divided by the total number of times the agency was found in the sampling frame. To the extent that all listings of each sampled agency are identified in the sampling frame, the resulting weights produce unbiased estimates (that is, estimates that would be obtained if there were no duplicates 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 agency level and the third is person level (patient/hospice/aide). The first type occurs when in-scope agencies do not respond to NHHCS. In NHHCS, the second type occurs when an in-scope agency does not provide the number of current home health patients and/or hospice discharges within the respective agency. In NHHAS, the second type of nonresponse occurs when an in-scope agency does not permit survey of their home health aides. The third type occurs when the administrative and medical records of the sampled current home health patients and/or hospice discharges are not made available to complete the survey (NHHCS) or when the sampled home health aide fails to respond (NHHAS).

The third weight component applies only to weights used to estimate numbers of agencies. This component involves ratio adjustments that are made within groups defined by region and agency type to account for use of probability proportional to size when selecting the agency sample. The numerator of the ratio was the number of agencies in the sampling frame within each group and the denominator was the estimated number of agencies for that same group. No ratio adjustment was made to other weights (i.e., agency weights for agency-level estimates for parameters other than numbers of agencies and weights for patients, discharges, or aides). Finally, the weights described above were smoothed within groups defined by region and agency type if there were outlier sample 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 NHHCS and NHHAS 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 sample 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 an asterisk (*) 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.