Sample Design Research for the 2007 National Home Health and Hospice Care Survey

Karen E. Davis (kedavis@cdc.gov) and Iris Shimizu, Ph.D. (ishimizu@cdc.gov)
National Center for Health Statistics, Centers for Disease Control and Prevention,
3311 Toledo Rd, Room 3213, Hyattsville, MD 20782

Abstract

The National Home and Hospice Care Survey (NHHCS) is a nationally representative sample survey of home health and hospice care agencies, their patients, and staff conducted by the National Center for Health Statistics (NCHS). Estimates from the NHHCS provide important information on the users and providers of this type of long term care, which is valuable for policymakers in the U.S. as the population continues to age. The NHHCS has been conducted six times since 1992 (1992, 1993, 1994, 1996, 1998 and 2000). The NHHCS was not fielded in 2002 or 2004 to allow time to conduct survey and sample design developmental work. This paper describes the research objectives, 2000 sample design, 2007 design changes and ongoing research activities. A redesigned NHHCS was fielded in summer 2007.

Keywords: Sample survey design, design effects

1. Introduction

A major goal of the 2007 NHHCS redesign is to provide policymakers, the public health community, and researchers with the data they need to address key long-term care policy issues. As the population continues to age, the use of various types of long-term care services will become an exceedingly important issue. Detailed information on utilization patterns as well as accurate assessments of the costs associated with providing individuals with home health and/or hospice care are needed given changes in the health care industry. Recent studies indicate that the elderly would prefer staying in their homes with assistance rather than submitting to residential care, increasing the importance of information on home based care (Strahan, 2006). The sample redesign addresses these research issues.

2. Survey Objectives

As in prior cycles, the purpose of the redesigned NHHCS is to collect: national baseline data on the characteristics of hospices and home health agencies, the patients they serve and the type of staff they employ; data on Medicare and Medicaid certification of agencies; data on charges to patients and the source(s) of payment for services; information about patients receiving care including functional status and diagnosis; information about the categories of people employed by home health and hospice care agencies; and information on end-of-life issues (Haupt, 1994). The survey is designed to produce statistics for current home health patients and hospice discharges with maximum precision for the available funds. In past cycles, the design was required to reduce respondent burden per sampled agency so interviewers could complete data collection in a single day. For the 2007 NHHCS design, the following research was performed:

- Investigate different scenarios for sample allocation of home health agencies, hospice care agencies, and mixed agencies (i.e., agencies that provide both home health care and hospice care);
- Investigate sample sizes necessary to produce estimates of specific subgroups of patients;
- Investigate the expansion of stratification by metropolitan statistical area (MSA) status.

With the above targeted objectives, the following priorities guided sample design development:

- Patient statistics have priority over agency statistics;
- Equal priority is to be given to statistics for home health patients and hospice patients;
- Statistics should be produced by MSA status (i.e., MSA, Micropolitan, and other non-MSA), by agency type, and by region.

Other research activities included the development of new data items to capture data of interest to researchers, creation of a supplemental survey of home health aides, use of computer assisted personal interview software, and linkage to the Outcome and Assessment Information Set used by the Centers for Medicare and Medicaid Services to expand the analytic capacity of the survey by providing descriptive and clinical information.

3. Sample Design

The 2000 sample design is a stratified two-stage probability design that is based on research undertaken for the 1992 NHHCS (Haupt, 1994). The first stage of selection is a systematic sample of the home health agencies and hospices in the sampling frame. The

second stage of sample selection is done using a sample selection table to obtain systematic probability samples of current home health patients and hospice discharges. The sampling frame for the most recent (2000) survey was derived from a frame that consisted of home health agencies listed in the 1999 SMG Marketing Group Inc. file and hospices listed in the 1999-2000 National Hospice Organization's Mailing Labels (Haupt, 2003). The universe was then stratified to draw the sample of agencies based on type of agency (home health agencies, hospices, and mixed agencies), MSA status, and region. Within these sampling strata, agencies were arrayed by ownership status, certification status, State, MSA code, county, ZIP code, and size (number of current patients).

The second stage of sample selection, sampling of six current patients and six discharges within each agency, was done using a sample selection table to obtain systematic probability samples of current patients and of discharges. The patients and discharges were selected from lists constructed for each agency during the interview. Current patients were defined as those patients who were on the rolls of the agency as of midnight on the day immediately before the date of the survey. Discharges referred to those patients who were discharged from care by the home health agency or hospice during a designated month between October 1999 and September 2000. Discharges due to patient death were included.

4. Redesign Goals

4.1 Agency sample allocations

An overarching goal for the redesign was to produce more precise estimates of subgroups within agency type. It was necessary to estimate the number of home health agencies, hospice agencies, and mixed agencies needed to obtain sufficient current home health patients and hospice discharges for analyzing key characteristics. From the 2000 NHHCS sampling frame, the distribution of types of agencies was as follows: 67% were home health only agencies, 11% were hospice only agencies, 16% were mixed, and 6% were unknown. The percentage of mixed agencies has changed over time and has been increasing (Remsburg, 2007). Mixed agencies tend to be more like home health only agencies, tend to be larger than hospice only agencies, and tend to have fewer hospice patients. During frame development, misclassification by agency type has been a challenge for the NHHCS. Mixed agencies are misclassified more often than the other agency types. In the 2000 NHHCS, only 65% of the agencies sampled as mixed agencies from the sampling frame were found to be mixed after the survey was completed.

A key decision was whether to stratify by 2 agency types or by 3 agency types. In the 1992-1994 designs, stratification was by 2 agency types (home health only and mixed versus hospice only and others). In the 1996-2000 designs, stratification was by 3 agency types (home health only versus hospice only versus mixed). In order to preserve the ability to track the diversification of home and hospice agencies, it was agreed to stratify by three agency types. The total number of respondent agencies required to produce the sample patient counts depends upon the average numbers of expected sample patients from each respondent agency and the number of mixed agencies in the sample. The total number of respondent sample agencies required to meet the patient level objectives can be formulated as an inverse function of the number of respondent mixed agencies in the sample (Shimizu, 2007). For NHHCS, that means the minimum sample of respondent agencies will be attained by using the maximum acceptable sample number of mixed agencies in each area cell for which patient statistics are desired. For the NHHCS sample, that means one should decide the minimum number of respondent agencies needed from home health only and/or hospice only types in order to produce planned agency level statistics.

4.2 Patient Sample Allocations

Early in the redesign process, it was decided to increase the sample size by sampling only current home health patients and sampling only hospice discharges. Note that in the 2000 NHHCS, no attempt was made to distinguish home care patients from hospice patients, as participating agencies provided a list of current patients and a list of discharges. In order to improve efforts to make subgroup analyses, up to 10 current home health patients are sampled within home health only agencies and up to 10 hospice discharges are sampled from hospice-only agencies. Hospice patients that were discharged within the three months prior to the survey are deemed eligible for the sample. If a sampled agency has fewer than 10 eligible patients, all patients are included in the sample. For mixed agencies, the patient type with the fewest eligible patients in that agency contributes minimum (all, 5) patients of that type to the sample and the remaining patient sample will be selected from the other patient type so that a total of up to 10 patients are selected from that agency.

A primary sampling goal of the NHHCS is to improve NCHS' ability to produce annual estimates with

satisfactory precision (e.g., a relative standard error (RSE) of 30% or less for prevalence estimates (p < 5%)). There is a strong interest in increasing the quality of hospice discharge estimates. In order to estimate the sample size needed for both hospice discharge estimates and home health care estimates, key subgroups were identified to facilitate the analyses of disparities.

For home health patients, the following characteristics were considered key to the redesign:

- Black male in non-MSA
- Black male 65 years old and over in non-MSA area
- Black with heart disease
- Black with diabetes

For hospice discharge patients, the following characteristics were considered key to the redesign:

- Black under 65 years old
- Black with cancer and without cancer under 65 years old
- Black without cancer 65 years old and over
- Black in non-MSA area

At the agency level the following characteristic was considered key to the redesign:

- Agency uses computerized medical records

In-house data from the 2000 NHHCS were used to determine the expected precision of the key home health and hospice care estimates for the redesign research (NCHS, 2002). There were 1,425 agencies that responded to the survey providing data for 4,634 home health patients, and 2,327 hospice care patients. In Tables 1 and 2, we show the agency-level estimates for computerized medical records and home health characteristics estimates along with the square root of their design effects by MSA area. The design effect is defined as the ratio of the sampling variance for a statistic computed under the sample design divided by the sampling variance that would have been obtained from a simple random sample of exactly the same size. It is widely used in the determination of sample size, and in summarizing the effect of having sampled clusters instead of elements (Groves, et al., 2004).

4.3 Mixed Agency Allocation

In each mixed agency, a sampling goal is to select 5 hospice discharges and 5 home health care current patients. We found that in mixed agencies, the distribution of hospice discharge counts (fewer than 5 and 5 or more) differed by MSA status. In order to estimate the number of responding mixed agencies required, the weighted average formula assuming a simple random sample design was used:

$$Y_{i=msa,n\_msa} = X_i p_i + 5(1 - p_i) \ .$$

Assume the weighted sample count in the 2000 NHHCS reflects the population, where

Y is the average number of sampled hospice discharges expected in all respondent sample mixed agencies;
X is the average number of hospice discharges expected in mixed agencies with fewer than 5 hospice discharges; and
p is the proportion of mixed agencies with fewer than 5 respondents.

Using 2000 NHHCS data, if the target is 10 sample patients per agency, the table below provides the average number of home health patients and hospice discharges expected in mixed agencies.

	All	**MSA**	**Non-MSA**
Mixed: Home Health	6.21	6.89	5.97
Mixed: Hospice	3.79	3.11	4.03

5. Conclusion

For the NHHCS redesign research, different scenarios for sample allocation of home health agencies, hospice care agencies, and mixed agencies were investigated. Since it was important to preserve the ability to compare agency characteristics by type of agency, the sample was stratified by 3 agency types in order to track the diversification of home health and hospice agencies. We also expanded agency stratification by MSA status in order to include micropolitan areas.

The sample sizes necessary to produce estimates of specific subgroups of patients were investigated, and a decision was made to increase the sample size from 6 patients to 10 patients in order to enhance subgroup analyses of disparities. This increase was for home health-only and hospice-only agencies.

References

Groves R, Fowler F, Couper M, Lepkowski J, Singer E, Tourangeau R. (2004), *Survey* Methodology, New York: John Wiley and Sons.

Haupt BJ. *Development of the National Home and Hospice Care Survey.* National Center for Health Statistics. Vital Health Stat 1(33). 1994.

Haupt BJ. *Characteristics of Hospice Care Discharges and Their Length of Service: United States, 2000.* National Center for Health Statistics. Vital Health Stat 13(154). 2003.

National Center for Health Statistics (2002). *Data File Documentation, National Home Health and Hospice Care Survey, 2000 (machine readable data file and documentation)*. National Center for Health Statistics, Centers for Disease Control and Prevention, Hyattsville, MD.

Remsburg R. *Summary of National Home and Hospice Care Survey (NHHCS) Sample Redesign Activities and Goals for Redesign.* Unpublished memo. National Center for Health Statistics. 2007.

Shimizu I. *The sample size/allocation game for 2007 NHHCS.* Unpublished memo. National Center for Health Statistics. 2007.

Strahan G. *National Home and Hospice Care Survey.* OMB No. 0920-0298. National Center for Health Statistics. 2006.

Table 1: Design Effects by Area and Agency Type, 2000 NHHCS[1]

Characteristic	Nation		MSA		Non-MSA	
Using a Computerized Medical Record	(%)	DEFF[2]	(%)	DEFF[2]	(%)	DEFF[2]
Home Health	32.12	2.00	35.78	1.96	25.03	2.00
Hospice	18.60	2.07	14.97	2.05	27.00	1.89
Mixed	40.32	3.55	37.45	3.57	43.30	3.25

[1]Sample of 1,425 agencies responded in 2000. Of these, 902 were in MSA and 523 were in Non-MSA.

[2] DEFF is the design effect for the standard error of an estimate based on data from the 2000 NHHCS

Table 2: Design Effects by Area and Home Health Care Characteristics, 2000 NHHCS[1]

Characteristic	Nation		MSA		Non-MSA	
Resident	(%)	DEFF[2]	(%)	DEFF[2]	(%)	DEFF[2]
Black	12.55	2.26	12.41	2.00	13.08	3.03
Male	35.22	1.64	35.15	1.55	35.49	1.59
Female	64.78	1.64	64.85	1.55	64.51	1.59
Under age 65	29.52	2.22	29.37	2.12	30.10	2.06
65+ years	70.48	2.22	70.63	2.12	69.90	2.06
Cancer	5.14	1.88	4.81	1.80	6.37	1.84
Heart Disease	10.89	1.77	10.73	1.70	11.50	1.63
Diabetes	7.85	1.75	7.52	1.66	9.07	1.79
Black male	4.09	2.10	3.86*	1.88	4.93*	2.84
Black female	8.46	2.08	8.55	1.90	8.15	2.47
Black under age 65	4.40	1.86	4.70	1.72	3.27*	1.83
Black 65+ years	8.15	2.14	7.71	1.93	9.81	2.76
Black with Cancer	0.57*	2.62	0.65*	2.49	0.28*	1.17
Black with Heart Disease	1.39*	2.15	1.18*	2.16	2.16*	1.98
Black with Diabetes	1.39*	1.69	1.46*	1.58	1.11*	1.53
Black male under age 65	1.76	1.60	1.74	1.53	1.83*	1.75
Black male 65+ years	2.33	1.99	2.12*	1.62	3.10*	3.16
Black female under age 65	2.64	1.95	2.96	1.82	1.44*	1.56
Black female 65+ years	5.82	1.92	5.59	1.77	6.71	2.29
Black male with Cancer	0.02*	0.96	0.03*	0.94	0.00	0.00

Table 2. (cont'd)

Characteristic	Nation		MSA		Non-MSA	
Resident	(%)	DEFF[2]	(%)	DEFF[2]	(%)	DEFF[2]
Black female with Cancer	0.55*	2.67	0.62*	2.55	0.28*	1.17
Black male with Heart Disease	0.65*	1.86	0.62*	1.73	0.77*	2.22
Black female with Heart Disease	0.74*	1.67	0.56*	1.67	1.39*	1.97
Black male with Diabetes	0.43*	1.35	0.49*	1.24	0.20*	1.38
Black female with Diabetes	0.96*	1.61	0.97*	1.55	0.91*	1.56
Black male w/Cancer under 65	0.00	0.00	0.00	0.00	0.00	0.00
Black male w/Cancer 65+	0.02*	0.96	0.03*	0.94	0.00	0.00
Black female w/Cancer under 65	0.39*	2.51	0.47*	2.22	0.10*	0.91
Black female w/Cancer 65+	0.16*	1.53	0.16*	1.36	0.19*	1.23
Black male w/Heart Disease under 65	0.11*	1.64	0.12*	1.56	0.09*	1.24
Black male w/Heart Disease 65+	0.54*	1.95	0.50*	1.77	0.68*	2.26
Black female w/Heart Disease under 65	0.08*	1.20	0.04*	0.81	0.23*	1.98
Black female w/Heart Disease 65+	0.65*	1.69	0.52*	1.66	1.16*	1.58
Black male w/Diabetes under 65	0.22*	1.45	0.22*	1.39	0.20*	1.38
Black male w/Diabetes 65+	0.21*	1.34	0.26*	1.17	0.00	0.00
Black female w/Diabetes under 65	0.28*	1.80	0.33*	1.61	0.08*	1.17
Black female w/Diabetes 65+	0.68*	1.49	0.65*	1.42	0.83*	1.54

[1] Sample of 4,634 home health care current patients responded in 2000 with 2,935 in MSA and 1,699 in Non-MSA.

[2] DEFF is the design effect for the standard error of an estimate based on data from the 2000 NHHCS.

* Estimate has relative standard error greater than 30% or small sample size.