

Approximation of Variable Costs for the National Hospital Ambulatory Medical Care Survey

Iris Shimizu and Feng Lan

National Center for Health Statistics, 6525 Belcrest Rd., Rm. 915, Hyattsville, MD 20782

KEY WORDS: survey research, sample design, cost model

1. Introduction¹

The National Center for Health Statistics (NCHS) conducts the National Hospital Ambulatory Medical Care Survey (NHAMCS) to produce ambulatory medical care service information about patient visits to the emergency and outpatient departments of noninstitutional general and short-stay hospitals, exclusive of Federal hospitals, located in the 50 states and the District of Columbia. The U.S. Bureau of the Census (BOC) is the primary data collector for the NHAMCS. NHAMCS is conducted annually in a sample of hospitals that is retained across the years. The current hospital sample was first placed in service in 1991 and updated in 2000. There are plans to replace the current sample with a new and optimized sample in 2005. Among other things, research to optimize a sample requires information about the variable survey costs, which are the costs that increase with sample size at each sampling stage. This paper discusses the data and processes used to approximate annual variable costs for the 1999 NHAMCS for use in the NHAMCS sample design research.

The next section discusses the survey design and operations which affect the survey's costs. Section 3 outlines the cost model assumed for the survey while Sections 4 and 5 discuss the data used and allocation of costs to different sampling stages. Section 6 looks at cost estimates derived for NHAMCS from pilot test data. Conclusions are summarized in the last section.

2. Overview of NHAMCS Design and Operations Which Affect Costs

2.1. Sampling Design

The current NHAMCS uses a four-stage probability design with samples of geographic primary sampling units (PSUs), hospitals, clinics and emergency service areas (ESAs), and patient visits selected at the respective sampling stages. The PSUs are counties, county equivalents, townships or other minor civil divisions (for some PSUs in New England), or metropolitan statistical areas. The first-stage sample consists of 112 PSUs that comprised a probability subsample of the PSUs used in the 1985-94 National Health Interview Survey (NHIS).

Within each PSU, hospitals were selected as the secondary sampling units. The NHAMCS hospital frame was compiled from the hospitals listed in the SMG Hospital Market Database (SMG Marketing Group Inc., 1991). Hospitals with an average length of stay of less than 30 days for all patients or hospitals whose specialty was general or children's general are eligible for NHAMCS. Excluded were federal hospitals, hospital units of institutions, and hospitals with less than six beds staffed for inpatient use. Systematic random sampling with probability proportional to numbers of outpatient visits was used to select hospitals. Of the 600 hospitals in the 1999 NHAMCS sample, 550 hospitals had an emergency department (ED) and/or an outpatient department (OPD) and 50 hospitals had neither an ED nor an OPD, according to the 1991 SMG data. The sample of 600 hospitals was randomly divided into 16 subsets of approximately equal size. The subsets were assigned to four-week data collection periods on a rotating basis with one subset per period so that each hospital only participates in NHAMCS every 15 months. The hospital sample for a survey year consists of those hospitals which are assigned to the 13 four-week reporting periods for that year.

Within each hospital, either all outpatient clinics and ESAs or samples of such units were selected as the third stage sampling units. Prior to each reporting period for that hospital, a field representative visits the hospital to induct it and obtain separate lists of the ESAs in the EDs and clinics in the OPDs. For OPDs, each clinic's specialty and expected number of visits during the assigned reporting period were also collected. If there were five or fewer clinics, then all were included in the sample. If a sample hospital had more than five clinics, then a sample was randomly selected. This was done by first listing individual clinics by six categories: general medicine, surgery, pediatrics, obstetrics/gynecology, substance abuse, and other. Within each category, clinics were listed in order of clinic size, from the smallest to the largest where clinic size was defined as the expected number of patient visits during the hospital's assigned four-week reporting period. If a clinic expected fewer than 30 visits, it was grouped with one or more other clinics within its clinic group to form a sampling unit with a total of 30 or more expected visits. After grouping the clinics into sampling units, five of these sampling units were selected with probability proportional to the size of the sampling unit. In EDs, if there were five or fewer ESAs, all were selected. Otherwise, a sample of five was selected from that department. At total of 1,405

¹ The opinions expressed in this paper are those of the authors and not necessarily those of the National Center for Health Statistics.

clinics/ESAs were included in the 1999 NHAMCS.

The fourth stage or ultimate sampling unit for the NHAMCS is the patient visit or encounter. Within ESAs or OPD clinics, patient visits were systematically selected over the hospital's assigned four-week reporting period. A visit is defined as a direct, personal exchange between a patient and a physician, or a staff member acting under a physician's direction, for the purpose of seeking care and rendering health services. The 1999 NHAMCS sample included a total of 50,995 visit records.

2.2. Survey Operations

About three months prior to a hospital's reporting period, NCHS sent an advance letter to the hospital administrator. About a week after the letter was mailed, the Census field representative (FR) called the hospital administrator to arrange for an appointment to further explain the study and verify hospital eligibility for the survey. During the initial visit, the FR solicited the hospital's cooperation and, where applicable, obtained information about the organization of the emergency and outpatient departments and specific information needed to sample the ESAs and clinics. After the initial visit and the development of the sampling plan, the FR contacted the hospital coordinator for the survey to arrange for induction of the sampled clinics/ESAs. During these inductions, the FR explained the data collection process, including the visit sampling procedures and instruction for completing the patient record forms (PRFs).

The actual visit sampling and data collection for the NHAMCS was primarily the responsibility of hospital staff. However, if a hospital permitted it and a sample clinic would not otherwise participate in the survey, the FRs did the visit sampling and data collection. During the data collection period, the FRs visited the sampled clinics/ESAs each week and maintained telephone contact with the hospital staff involved in the data collection effort. In these visits, the FR reviewed the log or other records used for visit sampling to determine if any cases were missing and also edited completed forms for missing data. Attempts were made to retrieve both missing cases and missing data on specific cases, either by consulting with the appropriate hospital staff or by reviewing the pertinent medical records. On the final visit the FR collected the remaining PRFs and obtained or verified the total count of visits occurring during the reporting period by reviewing the log used for sample selection or by obtaining counts directly from hospital staff. At the end of the hospital's reporting period, the FRs sent "Thank you" letters to the hospital administrator.

3. Cost Model Format for NHAMCS

A simple cost model was developed for the variable costs that depend primarily on the number of units included in the sample at each stage. An overall cost

model for the NHAMCS four-stage sample design may be expressed in the form:

$$C = C_0 + C_1 a + C_2 m + C_3 n + C_4 u$$

where

C = total survey cost.

C_0 = overhead (fixed costs).

a = number of geographic PSUs selected at the first stage.

m = number of hospitals selected at the second stage.

n = number of clinics/ESAs selected at the third stage.

u = number of visits selected at the fourth stage.

C_1 = cost per PSU for including an additional geographic PSU in the sample at the first stage (excluding costs associated with hospitals within that PSU).

C_2 = cost per hospital for including an additional hospital in the sample at the second stage within a sampled PSU (excluding costs associated with individual clinics/ESAs sampled in that hospital).

C_3 = cost per clinic/ESA for including an additional outpatient clinic or ESA in the sample at the third stage within a sampled hospital (excluding costs of PRFs for visits that might be selected in that clinic/ESA).

C_4 = cost per patient visit for including an additional visit in the sample at the fourth stage within a sampled clinic/ESA.

Fixed costs C_0 include:

- Costs of headquarters staff who oversee the data collection process, train regional office staff, and write field manuals.
- Cost of regional office staffs who train the FRs and monitor the data collection activities.

PSU level costs C_1 include:

- Costs of travel between PSUs that occur when the FR assigned to a hospital does not reside in the hospital's PSU.
- Per diem.
- Costs of training, supervising, and nurturing the FRs.

Hospital level costs C_2 include:

- Costs of travel to a hospital within PSU.
- Costs of FR time for inducting a hospital and its emergency and outpatient departments.
- Costs of FR time for miscellaneous activities (locating people, waiting before or between appointments, finding parking, etc.).

Clinic/ESA level costs C_3 include:

- Costs of FR time spent on inducting clinics/ESAs.
- Costs of travel to and between hospital clinics/ESAs when they are not on site at the hospital.

Visit level costs C_4 include:

- Costs of FR time spent on completing PRFs when clinics would not otherwise participate.
- Costs of FR time for editing and correcting PRFs completed by hospital staff.

4. Data Available for Approximating Cost Components

When the purpose for a cost model is to optimize a new sample design, the needed costs are ideally recorded by observers who shadow FRs as the FRs go about their data collection duties. In place of observers, the FRs could be asked to record the needed details of their own activities as they proceed with data collection. Both of these options are costly, however. Funds for such options are unavailable for the NHAMCS and, hence, the data used in the current model are primarily restricted to those available from administrative sources.

Most of the cost data for the model came from reports generated for the 1999 NHAMCS by the BOC's Cost and Response Management Network (CARMN). The FRs are requested to report to the Network every day that they work on a BOC-conducted survey and give their hours, miles, and other expenses. The CARMN reports costs separately for individual surveys conducted by BOC. For each survey, the CARMN provides cost data separately by BOC's regional offices and by survey tasks. The CARMN tasks applicable to NHAMCS are initial training, re-interview, refresher training, observation, interview, office work, and "other task codes." For each task, CARMN gives details in terms of hours and payroll charges, miles and mileage expense, per diem, other reimbursement, and phone charges. It is noted that the figures in the BOC cost record system are created for administrative management purposes and not for cost models needed in sample design research. For example, BOC does not regularly keep data on the number of sample NHAMCS hospitals handled by resident interviewers (FR lives within the hospital's PSU) as opposed to nonresident interviewers.

Postage was used for mailing induction letters from the NCHS to the sample hospitals, but CARMN does not capture this postage cost. Hence, \$0.32 (first class

postage in 1999) times the number of hospitals was added to the known variable survey costs.

The costs of recruiting FRs for the NHAMCS were ignored because the BOC includes those expenses in their overhead (fixed costs) due to the fact that BOC hires and trains FRs for potential work on all of the BOC-conducted surveys and not specifically for the NHAMCS. The costs tabulated by CARMN for the tasks of "office work" and "other tasks codes" were also assumed to be fixed costs for purposes of the NHAMCS cost model because there was insufficient information available to allocate the costs for those tasks among the different survey levels.

In addition to CARMN data, approximate times spent by hospital staff on survey inductions (and, hence, probably also spent by FRs) were provided in the supporting statement prepared by McCaig (1998) for obtaining approval from the Office of Management and Budget (OMB) for the 2000/2001 NHAMCS. Also, BOC's field operation supervisors at headquarters provided information on what was "typical" for FR times and numbers of trips in their field operations.

The files produced from "NHAMCS-101 Hospital Induction Questionnaire" (NHAMCS 101 file) and "NHAMCS-101/U Ambulatory Unit Record" (NHAMCS 101U file) were used as sources for 1999 sample unit counts (hospitals, clinics/ESAs and visits) by eligibility and response status. These questionnaire forms were completed during the inductions of hospitals, departments, and clinics/ESAs. Table 1 presents those sample counts.

5. Allocating Variable Costs

Those known costs which were associated entirely with a single survey level were allocated first. Because postage was used only for mailing advance letters to hospitals, the postage costs were allocated to the hospital level.

In the CARMN data, the tasks for initial training, re-interview, observation, and refresher training are all aimed at training, supervising, and nurturing the FRs. Hence, the expenses for those tasks were allocated entirely to the first stage sample (PSU) level.

The costs in CARMN's interview task (the only other CARMN task assumed to have variable costs) were allocated to the different survey levels because that task includes all of the data collection activities. The interview task per diem costs were allocated to the PSU level because per diem use means the FR was probably traveling away from the PSU in which he or she resided. The phone costs for the task were also allocated to the PSU level because it was believed that most, if not all, of those costs were for long distance charges incurred by FRs who called sample hospitals which were not in their residence PSUs. The task's "other reimbursement" costs were assumed to be parking fees, tolls, public transportation fares, etc., with most of the costs going for

parking at hospitals and, hence, these costs were allocated to the hospital level.

Only the costs for interview task travel and labor were split among multiple survey levels. This split was accomplished by first allocating the task miles to the PSU, hospital, and the clinic/ESA levels because travel is required to reach those units. Travel hours associated with those miles were then estimated by assuming that the FRs drove at an average speed of 50 mph when traveling between PSUs and at an average of 35 mph within PSUs (to and between hospitals and clinics). The remaining interview task hours were then allocated to the hospital, clinic, and visit levels.

5.1 Allocation of Miles

Miles for PSU (first stage) level

Travel at the PSU level consists of travel between PSUs for sample hospitals that are assigned to non-resident FRs (FR resides outside of the hospital's PSU). For estimating the miles traveled between PSUs, it was necessary to determine which hospitals were assigned to nonresident FRs. For this purpose, it was assumed that each FR resided in the PSU which contained the greatest number of hospitals assigned to that FR. Hospitals for which FR information was not available were assumed to be assigned to a resident FR. For hospitals assigned to two FRs, it was assumed that the second FR did all of the traveling to that hospital. A total of 85 hospitals were, thus, assumed to be assigned to nonresident FRs. For each of these hospitals, the distance between that hospital's PSU and the assigned FR's residence PSU was estimated by using an off-the-shelf software package to calculate the distance between the population weighted centroids of the zip code area for that hospital and the zip code area for each of the FR's hospitals which were located in the FR's residence PSU. The average of these distances was then used as a crude estimate for the distance traveled one-way between PSUs for the hospital assigned to a nonresident FR. This estimate overstates the true distance between PSUs because it also includes distances within PSUs. To compensate for the overstatement in the estimated distance, only two round trips were assumed to be made between PSUs when the visited hospital was eligible for NHAMCS instead of the five or six trips usually made to participating hospitals. For hospitals that were found to be ineligible for NHAMCS, one round trip between PSUs was assumed.

Miles for hospital and clinic/ESA (second and third stage) levels

For convenience, the residual interview task miles remaining after assignments of miles to the PSU and the clinic/ESA levels were allocated to the hospital level.

Miles were allocated to the clinic/ESA level because some (about a quarter) of the sampled clinics and ESAs were located off-site from their hospital and travel was

required to collect the data from those sites. For each trip to the sampled off-site clinics/ESAs for a hospital, it was assumed that the FR made a circuit trip that began and ended at the hospital and included a visit to every sampled off-site clinic/ESA that needed to be visited on that trip. Based on information from BOC's regional offices, an average of four miles was assumed between every pair of stops in that circuit. It was assumed that all eligible, off-site sampled clinics/ESAs were visited on the first trip and that all participating clinics/ESAs (those providing PRFs for sampled visits) were visited on each of four additional trips made for quality control and data collection.

5.2 Allocation of Non-travel Hours

Non-travel hours for the hospital level

FR time required for inducting hospitals and departments was assigned to the hospital level. From OMB clearance documentation (McCaig, 1998), hospital induction interviews take about 60 minutes for an eligible hospital and about 15 minutes for an ineligible hospital. According to the 1999 NHAMCS 101 file, there were 427 eligible hospitals and 62 ineligible hospitals. McCaig also indicated that an additional hour of induction is expected for each ED and each OPD at hospitals having those departments. There were 404 EDs and 281 OPDs in the 1999 sample.

In addition to the time for inducting hospitals and departments, whatever interview task time was not otherwise specifically allocated to a sampling stage was also assigned to the hospital level. It was assumed that this residual time was probably consumed in miscellaneous activities, such as: finding parking, locating hospital staff offices, waiting for or between appointments with hospital staff, communicating with regional offices for clinic samples (if the hospital had more than five OPD clinics) and for hospital situations not covered in field manuals or FR training, and reporting the hospital's data to regional offices or headquarters.

Non-travel hours at the clinic level

FR time to conduct inductions of eligible clinics/ESAs and to conduct clinic level "quality control" at participating clinics/ESAs was assigned to the clinic/ESA level. According to McCaig (1998), induction takes about 60 minutes of hospital staff time for each clinic/ESA. For the cost model purposes, it was assumed that only one clinic/ESA could be inducted at a time and, hence, one hour of FR time was also required for inducting each clinic/ESA. There were 1,381 eligible clinics/ESAs in the 1999 NHAMCS.

The clinic level "quality control" was assumed to require about 80 minutes (20 minutes per visit for each of four visits) of FR time at each of 1,261 participating clinics/ESAs. Quality control at the clinic level included checking the accuracy of visit sampling and taking steps to correct problems with that sampling (if any), answering

survey related questions by hospital staff, and collecting the patient record forms (PRFs) that were completed by hospital staff.

Non-travel hours at the visit level

The FR time spent on reviewing, correcting, and/or abstracting data for individual sampled visits was allocated to the visit level. For each PRF completed by the hospital staff, it was assumed that the FR spent one-half minute reviewing the PRF for completeness and that about 5 percent of those PRFs were incomplete. For those PRFs found to be incomplete, it was assumed that the FR spent an additional five minutes retrieving the missing data. For the 45 percent of PRFs abstracted by FRs, it was assumed (from regional office estimates) that the FRs required an average of eight minutes to abstract each record.

5.3 Final Model for Variable Model

Applying the average cost per hour and mile to the hours and miles allocated to the different sampling stages and summing over all known costs gives a total cost for each stage in the cost model. The average cost per hour was calculated by dividing the total payroll charges recorded for the CARMN interview task by the total number of hours spent on that task. Likewise, the average cost per mile was calculated by dividing the total mileage costs recorded for that task by the task’s total miles. The total costs at each stage of the NHAMCS sample were then divided by the number of respondent units at that stage where respondents were those who actually contributed data to the 1999 NHAMCS visit estimates.

Table 2 presents the costs per respondent unit relative to that of a sample visit. For each sampling stage, the ratio numerator is the estimated cost per respondent unit at that stage and the denominator is the cost per completed PRF. It can be seen that the cost per clinic/ESA, per hospital, and per PSU are about 41, 296, and 1,142 times that for a completed PRF, respectively. If the cost per completed PRF was \$1.00, the simple model for variable costs would become:

$$\begin{aligned} \text{Variable cost} = & \$1.00 \text{ (number of PRFs)} \\ & + \$40.85 \text{ (number of clinics/ESAs)} \\ & + \$296.24 \text{ (number of hospitals)} \\ & + \$1,141.72 \text{ (number of PSUs)}. \end{aligned}$$

If one wanted to ignore OPDs and instead design a sample of hospitals with EDs, the cost of a hospital relative to that of a visit is \$318. Likewise, if one was only interested in sampling hospitals with OPDs, then the cost of a hospital relative that of a visit is \$495.

6. Model Evaluation

Because the cost estimates in the model are subject to assumptions about allocation of known costs to different

sampling levels, there was interest in comparing the model costs with those from other sources. The only other known data-based cost estimates per sample unit for NHAMCS are from the 1984 pilot study for the survey. The averages reported for that study are as follows:

Survey operation	Operation cost per unit ¹
Hospital induction	\$134 /participating hospital
Clinic/ESA induction	\$149 / participating clinic/ESA
Data collection	\$3 / completed visit record

¹ Loft, Sheatsley, and Frankel (1984).

The costs used in the 1984 averages were tallied by survey operation rather than by sampling stage as done to derive the 1999 estimates. For example, because the 1984 average cost per visit is for data collection costs, it includes costs of travel to the hospitals (a hospital level cost) and clinic-level quality control efforts (a clinic/ESA level cost). Likewise, the clinic induction costs for 1984 include costs of travel to hospitals (a hospital level cost). Hence, the figures from the 1984 pilot study are not comparable to any figures developed in this paper for the 1999 NHAMCS.

7. Conclusion

A simple model of the annual variable costs for the 1999 NHAMCS was developed primarily from administrative data. The model is subject to the accuracy of assumptions made about details of survey operations which are typically not needed in administrative records. However, for lack of resources to obtain better information, the model is assumed adequate for purposes of optimizing allocations to the different sampling stages in a four-stage NHAMCS sample design.

References:

Loft JD, Sheatsley PB, and Frankel MR (1984). Methodological Report on the Hospital Ambulatory Medical Care Evaluation Study. Contract No. 282-82-2111. Chicago, Illinois: National Opinion Research Center.

McCaig, Linda (1998). “Supporting statement” in OMB clearance package for the 2000-2001 National Hospital Ambulatory Medical Care Survey. Hyattsville, Maryland: National Center for Health Statistics.

SMG Marketing Group Inc. (1991). 1991 Hospital Market Data Base. 1342 N. LaSalle Drive, Chicago, Illinois 60610: SMG Marketing Group Inc.

Table 1: Counts of sampled units by sampling stage or hospital department and response status: 1999 NHAMCS

Units	Number sampled	Number responding
Hospitals	489	
Eligible	427	403
Department		
Emergency	404	376
Outpatient	281	241
Clinics/ESAs	1,405	
Eligible	1,381	1,261
Off-site	372	358
Visits		50,995
Hospital staff completed PRFs		28,052
FRs completed PRFs		22,943

Table 2: Estimated variable cost per unit relative to that for a sample visit by sampling stage: Derived from variable costs for the 1999 NHAMCS

Survey level	Cost ratios
	four-stage sample design
Geographic PSU	1,141.72
Hospital	296.24
With Responding Department	
Emergency	317.52
Outpatient	495.38
Clinic/ESA	40.85
Visit	1.00