The Home Health Care CAHPS Survey

Part B

Collection of Information

Employing Statistical Methods

TABLE OF CONTENTS

Section   **Page**

B. COLLECTION OF INFORMATION EMPLOYING STATISTICAL METHODS

B.1 Potential Respondent Universe and Sample Selection Method

As noted in Section A, it is necessary to consider data collection and analysis for two components of the Home Health Care CAHPS Survey. The first component involves conducting a mode experiment to develop adjusters for survey administration mode and patient mix. The second component is the national implementation, which will be voluntarily sponsored by home health agencies (HHAs); each participating HHA will contract with an independent survey vendor to conduct the survey for it. An analysis plan based on the first stage will be applied to derive comparable HHA-specific ratings for survey items and composites of items that have a common focus. The sampling plans for both the mode experiment and national implementation are described below. To maintain consistency between these two related phases of data collection, the same overall sampling methodology and data collection protocols will be used in both.

B.1.1 Sampling HHAs for the Mode Experiment

For the mode experiment, RTI will use a two-stage sampling process, starting with selection of HHAs that have voluntarily agreed to participate, followed by selection of patients from the sampled volunteer HHAs. RTI will use the most recent cost reports in the Medicare Healthcare Cost Report Information System (HCRIS) as a sample frame to select a target sample of HHAs for the mode experiment. The HCRIS contains cost reports from Medicare-certified HHAs. Because almost all HHAs are Medicare certified, the HCRIS includes the complete universe of Medicare-certified HHAs and essentially the universe of HHAs. A cost report is available for each Medicare provider number. There will be about 9,000 HHAs in the sample frame.

The data in the HCRIS include counts of patients and visits for home health care patients whose care is reimbursed by Medicare, Medicaid, and private insurance. RTI project staff has reviewed these data for their suitability as a sample frame. Generally, the data approximately match various metrics of Medicare home health care use for 2006 concerning the number of patients, visits, and episodes for Medicare patients, although the episode measure is not available for non-Medicare patients. This list of HHAs, counts of patients, and counts of visits will be used

in structuring a target sample frame of all Medicare-certified HHAs. Based on RTI's recent analysis of the cost reports in HCRIS files, approximately 30 HHAs contain potentially incorrect visit data. RTI will either delete these HHAs from the frame for the mode experiment or place them in a separate stratum for further evaluation.

There are some limitations to the use of cost report data for this purpose. Cost report data may lag current operations by a year or more. This lag may exist because some HHAs have been more recently certified, have gone out of business, or have changed ownership or characteristics. After forming the target sample frame and selecting the sample, RTI project staff will validate the information about the provider when they recruit HHAs to voluntarily participate in the mode experiment. The project staff also plans to use the home health patient assessments, the Outcome and Assessment Information Set (OASIS), to determine whether certain providers who have submitted cost reports are not currently billing. Inactive providers will be identified and removed from the frame prior to sampling for HHAs. If the project staff discovers that a provider's characteristics have changed significantly, the sample frame sorting process, described below, will allow the project statistician to choose an adjacent substitute HHA from the frame.

Given the purposes of the mode experiment (to determine potential necessary adjustments in reporting of the national comparative data for survey mode and selected patient characteristics), it is important to include a representative sample of HHAs as participants in the mode experiment. The methods to be used to select target HHAs for the mode experiment will ensure representation of home health agencies with a number of different characteristics to ensure a range of patient types. Project staff will use a combination of stratification by size (based on number of patient visits) and sorting by other HHA characteristics within strata to achieve the representative mix of HHA characteristics. The characteristics to be accounted for (in addition to size) are

- geographic location of the HHA, using a broad definition based on similarity in per capita use of home health care in contiguous states;
- whether an HHA is freestanding or hospital-based;
- whether an HHA is for-profit or not-for-profit; and
- whether the HHA is in an urban or rural area, based on county designation.

Given the likelihood that some HHAs in the initial mode experiment target sample may not agree to participate in the mode experiment, the project staff will draw a similar nearest neighbor substitute from the sorted frame.

The project staff will first stratify by size using visit counts by HHA. Based on the visit count, they will create categories of size. After sorting on size, project staff will create strata based on HHA size. For example, if four strata were used, the largest HHAs (representing one quarter of the visits) would be the first stratum and the smallest HHAs (also with one quarter of the visits) would be the fourth. This method will produce increasing numbers of HHAs in each stratum, since fewer HHAs are large (first stratum) and more agencies are small (fourth stratum). Within strata the data will be sorted on the other HHA characteristics listed above. To avoid sampling a very large number of small HHAs that contribute few patients to a sample, the strata will be sampled disproportionately. With large numbers of agencies in each stratum, systematic sampling will select a representative sample across HHA types.

As an example of this sampling approach, using approximately 9,000 HHAs as the universe, the number of HHAs in each stratum might be 100 (the largest HHAs), 1,900, 2,500, and 4,500 (the smallest HHAs). One hundred HHAs could be selected by systematic sampling of 20, 30, 30, 20, HHAs respectively, from each of the 4 strata. In this example, 22% of the HHAs in the sample frame would be relatively large and 78% relatively small, but the larger HHAs would contribute a much larger proportion of the sample for the mode experiment—in this example, 50% of the sample would be in the two strata of larger HHAs (50 of the total 100 selected).

Within size strata, the agencies will be sorted across the other characteristics so that the systematic sampling will cover a range of the geographic and organizational characteristics mentioned above. As described below, the sampling rate can be adjusted for each stratum so that a sufficient number of patients can be drawn for the experiment.

RTI plans to recruit approximately 100 HHAs to participate in the mode experiment. The number of agencies has been chosen as a compromise between the need to cover a variety of agency types across the dimensions listed above, the intensity of effort needed to recruit HHAs to participate in the mode experiment, and recognition of the additional burden to HHAs of

participating voluntarily in this additional step toward national implementation. For the mode experiment to be valid statistically it is more important that the number of patients in the sample be adequate than the number of participating HHAs. By adjusting the sampling rate within agencies the patient sample can be achieved.

B.1.2 Sampling Patients for the Mode Experiment and the National Implementation

For both the mode experiment and the national implementation, HHAs will assemble a census of their patients (both current and discharged) for the sampling window, defined as a calendar month. Each HHA will submit a file containing patient information for all patients to whom the HHA provided home care during the sampling month to its contracted survey vendor (for the national implementation) or to RTI (for the mode experiment). The mode experiment will involve sampling and data collection for each of 3 months. The national survey will be fielded on a rolling basis, and the results for each quarter merged with data from the 3 immediately preceding quarters and analyzed. The sample frame for the national implementation will be assembled at the level of the CMS Certification Number (formerly known as the Medicare provider number), the unit of comparison for survey results reported on the Home Health Compare website.

The sample frame prepared by the HHA will contain all the patient data needed for fielding the survey and data needed for analysis. The HHA's survey vendor (for the national implementation) and RTI (for the mode experiment) will review the frame and exclude any patients who are not eligible to participate in the Home Health Care CAHPS Survey. Patients ineligible for the survey are those who

- are receiving hospice or are discharged to hospice,
- are deceased when the sample is drawn,
- are under 18 years of age at any time during their stay,
- did not have at least one skilled home health visit in the sample month and at least two home health care visits during a 60-day look-back period starting with the last day of the sampled month,
- are maternity patients,
- are "no publicity" patients,

- are receiving only nonskilled (aide) care, or
- were sampled during the last 5 months.

The requirement that a patient not be sampled more than twice a year is intended to reduce burden on individual patients and to increase the probability of response. In the mode experiment there will have been no patients sampled prior to the beginning of the experiment. Therefore, for the mode experiment, almost all the patients receiving HHA services in the sampled agencies will be eligible for the first month's patient sample frame. A small percentage of HHA patients will be lost in the mode experiment patient sampling frame due to the other exclusions (listed above). The second and third months of the experiment will have a smaller number of eligible patients because only new patients and some of those excluded from month one will become eligible.

B.1.2a   Mode Experiment Patient Sampling Specifics

For the mode experiment, the patient sample size required was computed on the basis of power to detect a difference in proportions of 0.05, from a proportion of 0.5, with 80% confidence and an alpha of 0.05. When regressions are run, predicting the proportion of patients choosing a particular answer to a survey item, the estimated coefficient of the variable indicating a particular mode is the increment in the model prediction that would arise from the reference mode in the equation. This is the difference in proportions that is targeted.

The sample size (total number of completed interviews) needed for each mode being tested in the mode experiment is about 1,570. With three data collection modes being tested, with equal sample size, the targeted number of completed surveys is 4,710. The following are estimates of response rates associated with surveys of similar patient populations and with different modes after adjustment for possible ineligibles found during the survey.

Data Collection Mode	Response Rate
Mail	30.0%
Phone	28.0%
Mixed	34.5%

Assuming an average response rate of about 30% based on other surveys, the fielded sample should be 15,700. However, given the nature of the HHA patient population, particularly the large proportion of aged Medicare or Medicaid patients (including many patients in a long recovery or requiring more maintenance health care), there are some concerns about achieving this response rate. We have therefore assumed in our sampling approach for a "safety margin" that allows the response to be as low as 20% while still achieving statistical power. This requires a sample of about 23,000. With a sample of this size we should be able to cover a wide variety of patients with random sampling.

This sample size can be achieved with 100 HHAs as the primary sampling units by adjusting the sampling rates on the larger and smaller HHAs recruited. The size of HHA client populations at any one time varies from fewer than 100 to many thousands. The particular mix of HHAs in different size strata may need to be adjusted depending on recruiting success but it will be easier to adjust the sampling rate. We will determine the actual stratification into size classes when the HCRIS data are analyzed, but a simple example, assuming only two size classes, shows the feasibility of this approach. Assume there are large agencies with 1,000 clients at a point in time and small agencies with 100. A target of 23,000 patients can be achieved with the following combinations of numbers and sampling rates:

- 50 large HHAs at 36% = 18,000 and 50 small at 100% = 5,000
- 25 large HHAs at 62% = 15,500 and 75 small at 100% = 7,500

RTI will aim for a distribution that is closer to the proportions of HHAs in the size categories but the more important patient target can be met with disproportionate distributions.

For the mode experiment, patients will be selected using random sampling. Within each HHA, each sampled patient will then be assigned randomly to one of three data collection modes. The project staff will not assign an entire HHA sample to one mode to avoid correlation of mode with HHA characteristics.

B.1.2b   National Implementation Sampling Specifics

For the national implementation of the Home Health Care CAHPS Survey, each participating HHA will send to its contracted survey vendor each month a patient sample frame containing information about each patient who received home health care during the sample

month, with sufficient information for the vendor to determine exclusions and with information needed for both fielding the survey and for patient-mix adjustment. The survey vendor will remove from the sample frame patients who do not meet survey eligibility requirements and then draw a random sample of the remaining patients.

Survey vendors working under contract with HHAs will be instructed to use a reliable program to generate random numbers for sampling. The Centers for Medicare & Medicaid Services (CMS) will recommend that survey vendors use the free program RATSTATS, available from the Department of Health and Human Services, Office of Inspector General website, or some other validated sample selection program such as SAS to select the sample. The sampling procedure recommended is simple random sampling, but disproportionate and proportional stratified random sampling may be allowed since some HHAs may want to analyze their own data and view survey results for individual branches. HHAs that deviate from simple random sampling (using disproportionate sampling) will be required to request an exception and obtain approval from CMS. An exception will be permitted if the minimum sample is 10 per strata and the information needed to crease weights is reported to RTI.

Although the national implementation sampling will be conducted on a monthly basis (with the survey initiated for each monthly sample within 3 weeks after the sample month ends), data from four quarters will be accrued, aggregated, analyzed, and reported on a quarterly basis, with the data from the most current quarter replacing data from the oldest of the four quarters. For 4 calendar quarters, a minimum of 300 completed surveys is the target for each participating HHA. If an HHA's patient population is too small to yield 300 completed surveys, a census will be surveyed. The 300 completed surveys needed for analysis is derived from the formula for the precision of a proportion with the estimate at .5, the confidence interval of about +/- 0.05, and a confidence level of 95%. (Many agencies, with a substantial sampling fraction, can achieve a higher precision because of the finite population correction factor.)

In the national implementation of the Home Health Care CAHPS Survey, the number of patients needed for selection each month to yield a minimum of 300 completed surveys per year (25 per month) will ultimately be determined by each HHA and its survey vendor. These will differ by HHA. The mode of administration of the survey will be an important determining factor

in response rates. Using the estimated response above, the estimated sample sizes for HHAs participating in the national implementation of the Home Health Care CAHPS Survey are the following:

Mode	Response Rate	Sample Size for 25 Responses/Month
Mail	30.0%	84
Phone	28.0%	90
Mixed	34.5%	73

Each agency survey vendor will use its experience on other surveys with home health patients and/or other similar populations, the data collection mode, and expected response rates as guides for calculating the monthly sample sizes that will be needed for the Home Health Care CAHPS Survey.

The sampling rate to achieve these sample sizes indicates that HHAs with monthly frame sizes of 90 or below should start with a sample equal to the sample frame. That is, all patients who meet the eligibility criteria will be included in the survey sample. For HHAs with larger sampling frames the sampling rate can be reduced, although it clearly will be higher than 50% until the frame exceeds about 180 eligible patients per month. CMS will recommend that prior to starting the national implementation, survey vendors acquire from client HHAs sample frame information for each of the 3 or 6 months prior to the beginning of the national implementation to determine an appropriate sampling rate to use during the national implementation. Monthly sample size rates should be based on the number of patients who meet survey eligibility criteria in the frames after the first test month, since that month will not have any patients who are ineligible for the survey because they would be sampled during the first month of the test file.

B.2   Information Collection Procedures

Three modes of survey administration will be allowed during the national implementation of the Home Health Care CAHPS Survey to give HHAs options in how they would like to administer the survey, based on their goals and resources. These three modes are described below:

- Mail-only mode
  - Mailing of the questionnaire and cover letter to all sampled patients.
  - Second mailing of the questionnaire with a cover letter to sample patients who do not respond to the first mailing within 3 weeks after the first questionnaire package is mailed.
- Telephone-only mode
  - A maximum of five telephone contact attempts per patient to complete the survey.
- Mixed-mode
  - Mailing of the questionnaire and cover letter to all sample patients.
  - Telephone follow-up with all sample patients who do not respond to the questionnaire mailing. A maximum of five telephone contact attempts per patient will be made to complete the survey.

Data collection for each sampled patient must be initiated no later than 3 weeks (21 days) after the close of the sample month. Once data collection begins, it must be closed out within 6 weeks. These same data collection modes will be used in the mode experiment using the same protocols.

Survey vendors who wish to become "approved" to conduct the Home Health Care CAHPS Survey on behalf of HHAs must complete the Home Health Care CAHPS survey vendor training, which will provide detailed guidance on the protocols and guidelines for all aspects of survey implementation, from sample selection to data collection and data submission. As of the date of this submission, CMS anticipates that the first training sessions for vendors will be offered in early summer 2009; however, this is contingent upon receiving OMB approval by the end of April 2009. The national implementation is expected to start in the summer of 2009.

B.3   Methods to Maximize Response Rate

Every effort will be made to maximize patient response rates, while retaining the voluntary nature of the Home Health Care CAHPS Survey. Each questionnaire mailing will include a cover letter explaining what the survey is about, who is conducting it and why, and the name and toll-free telephone number of a survey staff member that sampled patients can contact if they have questions or need additional information about the survey. For the mail-only mode of administration, both RTI (for the mode experiment) and survey vendors (for the national implementation) will use best practices in survey materials to enhance response rates. These best practices include using a simple font no smaller than 10 point size in the survey cover letters,

allowing ample white space between questions in the questionnaire, avoiding a format that displays the questions as a matrix, using a unique subject identification number on the questionnaire rather than printing the sample member's name, and displaying the OMB number and expiration date on the questionnaire. The second mailing for the mail only implementation is expected to increase the response rate, as is the telephone follow-up portion of the mixed-mode implementation.

B.4   Tests of Procedures

To achieve the purposes of the mode experiment the following analyses will be conducted:

- Analyses of individual survey items will assess missing data and item distributions.
- Hypothesis testing will detect differences in key variables between modes.
- The analysis of individual items and the hypothesis testing will form the basis for constructing an adjustor to be used for telephone and mixed-mode surveys.

Because home health care is a new area for CAHPS®, it is yet to be determined (through analyses from the mode experiment) which patient-mix adjusters may be necessary for reporting of the national survey results. Data from the field test suggest that self-reported overall health status, education and age should be included in the patient mix models. One of the purposes of the mode experiment will be to explore the usefulness of these and other variables for patient-mix adjustment. RTI will evaluate whether the ranking of home health agencies differs for adjusted and unadjusted Home Health Care CAHPS® results.

B.5   Statistical Consultation and Independent Review

This sampling and statistical plan was prepared by RTI International and reviewed by CMS. The primary statistical design was provided by Melvin Ingber of RTI International. Dr. Ingber can be reached by telephone at (410) 730-1506 or by e-mail at mingber@rti.org.