

Measure & Instrument Development and Support (MIDS) Contractor:

Impact Assessment of CMS Quality and Efficiency Measures

Supporting Statement B: OMB/PRA Submission Materials for the National Provider Survey of Home Health Agencies

CONTRACT NUMBER: HHSM-500-2013-13007I

TASK ORDER: HHSM-500-T0002

SUBMITTED: SEPTEMBER 18, 2018

REVISED: JUNE 26, 2019

NONI BODKIN, CONTRACTING OFFICER'S REPRESENTATIVE (COR)

HHS/CMS/OA/CCSQ/QMVIG

7500 SECURITY BOULEVARD, MAILSTOP S3-02-01

BALTIMORE, MD 21244-1850

NONI.BODKIN@CMS.HHS.GOV

TABLE OF CONTENTS

1.	Respondent Universe and Sampling Methods.....	2
2.	Procedures for Collecting Information.....	9
3.	Methods to Maximize Response Rates and Deal with Non-Response.....	10
4.	Tests of Procedures or Methods to Be Undertaken.....	11
5.	Individuals Consulted on Statistical Aspects and Individuals Collecting and/or Analyzing Data.....	12

SUPPORTING STATEMENT B – DATA COLLECTION FOR THE NATIONAL PROVIDER SURVEY OF HOME HEALTH AGENCIES

1. Respondent Universe and Sampling Methods

For both the qualitative interviews and structured survey, we will draw a sample from the universe of home health agencies (HHAs) submitting data to the Home Health Quality Reporting Program (HHQRP) in 2019. We will use a stratified random sampling approach to generate nationally representative estimates. CMS agency staff responsible for the HHQRP and quality improvement in HHAs (see Attachment I: Development of the National Provider Survey of Home Health Agencies) report that assessing differences in HHAs' responses to CMS's use of quality and efficiency measures by subgroups of HHAs is a key goal for guiding management of quality measurement programs for HHAs. We are proposing to stratify the random sample of HHAs by the following characteristics to make subgroup comparisons: (1) HHA size (as defined by the number of patient home health care episodes, using the Outcome and Assessment Information Set [OASIS]), (2) participation in the Home Health Value-Based Payment (HHVBP) model (1 = yes, 0 = no), and (3) HHA quality performance rating on the HHQRP composite quality score. A stratified random sampling approach will support the following analytic objectives:

- (1) To make national prevalence estimates of the actions that HHAs report taking in response to the CMS measures (e.g., hiring quality improvement staff or implementing clinical decision support (CDS) tools within their health information technology [health IT] systems);
- (2) To make subgroup prevalence estimates (e.g., by quality performance, HHA size, and HHVBP model participation) of the actions HHAs report taking in response to the CMS measures with adequate precision and power to detect differences between subgroups; and
- (3) To evaluate the association between the quality improvement (QI) changes HHAs report taking and performance on CMS quality measures (i.e., have QI efforts been correlated with better performance?).

Description of Sample Frame and Approach to Stratification. The sample frame (i.e., universe from which the sample will be drawn) will comprise approximately 11,000 HHAs. As discussed with CMS agency staff responsible for the HHQRP, HHAs will be grouped into three size categories (1–100 home health care episodes, 101–1,000 episodes, and 1,001 or more episodes), as derived from the count of OASIS quality measure assessments submitted by each HHA. HHAs will be defined by participation in HHVBP, based on whether their mailing address is in one of the nine states enrolled in the HHVBP model that CMS is operating.¹ HHAs will be classified into one of four quality categories using the CMS Home Health Compare Star Ratings: high (4, 4.5, or 5 stars); medium (2.5, 3, or 3.5 stars); low (1, 1.5, or 2 stars); and missing (no Star Rating available). Using these three characteristics—size, participation in HHVBP, and quality—the population will be divided into 24 sample strata. Using CMS Home Health Compare data, Tables 1a and 1b show the number of HHAs within the universe that fell into each of the 24 strata in 2015, the latest year for which the project team had complete data. Counts are for illustrative purposes; the survey, if approved, will use the latest data available for sampling, and strata with too few HHAs will be excluded.

Impact Assessment of CMS

Quality and Efficiency Measures

ⁱ As of July 2018, the HHVBP model automatically enrolled HHAs located in Arizona, Florida, Iowa, Maryland, Massachusetts, Nebraska, North Carolina, Tennessee, and Washington.

**Impact Assessment of CMS
Quality and Efficiency Measures**

Table 1a: Universe of HHAs That Participate in HHVBP

	Small (1-100 episodes)	Medium (101-1,000 episodes)	Large (1,000+ episodes)
High performance (4, 4.5, or 5 stars)	66	395	163
Medium performance (2.5, 3, or 3.5 stars)	95	482	204
Low performance (1, 1.5, or 2 stars)	78	245	42
Missing performance (no Star Rating)	275	22	1

*Categories derived from CMS Home Health Compare database and OASIS for calendar year 2015; data included location, number of OASIS assessments submitted, and quality measure Star Rating. A home health care episode is defined as the entire patient care episode on a single claim, which may encompass multiple home care visits.

Table 1b: Universe of HHAs That Do Not Participate in HHVBP

	Small (1-100 episodes)	Medium (101-1,000 episodes)	Large (1,000+ episodes)
High performance (4, 4.5, or 5 stars)	313	1,390	388
Medium performance (2.5, 3, or 3.5 stars)	466	1,947	574
Low performance (1, 1.5, or 2 stars)	655	1,610	148
Missing performance (no Star Rating)	1,757	93	0

*Categories derived from CMS Home Health Compare database and OASIS for calendar year 2015; data included location, number of OASIS assessments submitted, and quality measure Star Rating. A home health care episode is defined as the entire patient care episode on a single claim, which may encompass multiple home care visits.

Sampling Design for the Structured Survey. The project team will draw a stratified random sample of 2,272 HHAs, with the goal of achieving 1,000 responses (assuming an estimated 44% response rate). A review of prior surveys of providers indicates that an expected 44% response rate is a reasonable assumption; for example, the CMS national surveys of hospitals and nursing homes used a fielding approach similar to that of the proposed survey and achieved responses rates over 50%.ⁱⁱ The team will use multiple modes of outreach to HHAs to achieve the desired number of responses. CMS adopted a conservative estimate of response rate compared with the CMS hospital and nursing home surveys because HHAs vary substantially by size and associated QI personnel staffing, which may affect response rates.ⁱⁱⁱ

Stratifying by HHA size will help us understand differences in responses to the CMS measures between facilities with different levels of resources to invest in QI. The sample will contain 30% large HHAs and 30% small HHAs. This oversample of large and small HHAs relative to medium-sized HHAs provides greater power for evaluating differences between HHAs of different sizes. Stratifying by participation in the HHVBP model will help us determine whether

ⁱⁱ Centers for Medicare & Medicaid Services. 2018 *National Impact Assessment of the Centers for Medicare & Medicaid Services (CMS) Quality Measures Report*. Baltimore, MD: US Department of Health and Human Services; 2018. Available at: <https://www.cms.gov/Medicare/Quality-Initiatives-Patient-Assessment-Instruments/QualityMeasures/National-Impact-Assessment-of-the-Centers-for-Medicare-and-Medicaid-Services-CMS-Quality-Measures-Reports.html>.

ⁱⁱⁱ Budget constraints preclude expanding the sample beyond the original 2,272 agencies should response rates fall below 44% overall or in any stratum.

**Impact Assessment of CMS
Quality and Efficiency Measures**

the responses to the CMS measures (e.g., investments in QI actions) differ for HHAs participating in the HHVBP model and the HHQRP compared with HHAs only in the HHQRP. We propose to oversample HHVBP participants (so that 30% of the sampled HHAs are HHVBP participants) to ensure adequate precision for comparing participants with non-participants. Finally, we propose to sample proportionally with respect to HHA quality-based strata because high-, medium-, and low-quality HHAs were distributed fairly evenly across strata; as a result, oversampling was not necessary to ensure adequate power for comparisons between quality strata (see “Power Calculations for the Structured Survey” below). Tables 2a and 2b show the distribution of sampled HHAs across the 24 strata and expected number of respondents per stratum.

Table 2a: Sample Allocation of HHAs That Participate in VBP

	Small (1-100 episodes)	Medium (101-1,000 episodes)	Large (1,000+ episodes)
High performance (4, 4.5, or 5 stars)	<ul style="list-style-type: none"> • 19 sampled • 9 complete responses 	<ul style="list-style-type: none"> • 90 sampled • 40 complete responses 	<ul style="list-style-type: none"> • 107 sampled • 47 complete responses
Medium performance (2.5, 3, or 3.5 stars)	<ul style="list-style-type: none"> • 28 sampled • 12 complete responses 	<ul style="list-style-type: none"> • 110 sampled • 48 complete responses 	<ul style="list-style-type: none"> • 134 sampled • 59 complete responses
Low performance (1, 1.5, or 2 stars)	<ul style="list-style-type: none"> • 23 sampled • 10 complete responses 	<ul style="list-style-type: none"> • 56 sampled • 25 complete responses 	<ul style="list-style-type: none"> • 28 sampled • 12 complete responses
Missing performance (no Star Rating)	<ul style="list-style-type: none"> • 87 sampled • 38 complete responses 		

*Categories derived from CMS Home Health Compare database and OASIS for calendar year 2015; data included location, number of OASIS assessments submitted, and quality measure Star Rating. A home health care episode is defined as the entire patient care episode on a single claim, which may encompass multiple home care visits.

Table 2b: Sample Allocation of HHAs That Do Not Participate in VBP

	Small (1-100 episodes)	Medium (101-1,000 episodes)	Large (1000+ episodes)
High performance (4, 4.5, or 5 stars)	<ul style="list-style-type: none"> • 52 sampled • 23 complete responses 	<ul style="list-style-type: none"> • 179 sampled • 79 complete responses 	<ul style="list-style-type: none"> • 144 sampled • 63 complete responses
Medium performance (2.5, 3, or 3.5 stars)	<ul style="list-style-type: none"> • 77 sampled • 34 complete responses 	<ul style="list-style-type: none"> • 250 sampled • 110 complete responses 	<ul style="list-style-type: none"> • 213 sampled • 94 complete responses
Low performance (1, 1.5, or 2 stars)	<ul style="list-style-type: none"> • 109 sampled • 48 complete responses 	<ul style="list-style-type: none"> • 207 sampled • 91 complete responses 	<ul style="list-style-type: none"> • 55 sampled • 24 complete responses
Missing performance (no Star Rating)	<ul style="list-style-type: none"> • 304 sampled • 134 complete responses 		

*Categories derived from CMS Home Health Compare database and OASIS for calendar year 2015; data included location, number of OASIS assessments submitted, and quality measure Star Rating. A home health care episode is defined as the entire patient care episode on a single claim, which may encompass multiple home care visits.

Power Calculations for the Structured Survey. The team will assign weights to the survey responses to account for differential sampling probabilities. Using the design weights (and an assumed 44% response rate), we estimate that the effective sample size for national estimates will be 833 (with a design effect of 1.20). We conservatively estimate the level of precision of our national estimates and of estimates by HHA strata (based on size, HHVBP participation, and quality performance) for a survey item with a prevalence of 50%, so the standard error estimates provided in Table 3 are upper bounds. A national estimate would have a standard error of 1.7 percentage points or less.

**Impact Assessment of CMS
Quality and Efficiency Measures**

With the proposed sample size, we will have reasonable precision for estimates within one-way strata. For example, an estimate of an item that is 50% prevalent across all high-performing HHAs will have a standard error of 3.5 percentage points (with an effective sample size of 204); an analogous estimate for large HHAs will have an error of 2.98 percentage points (with an effective sample size of 281). Lastly, a corresponding estimate calculated across all HHAs that participate in HHVBP will have a standard error of 3.13 percentage points (with an effective sample of size 255). These calculations do not incorporate adjustments that will be required if response rates differ across strata.

To compare subgroups of HHAs as defined by the strata illustrated in Tables 1a and 1b (VBP participation, HHA size, and performance), we will compute the effect size as the ratio of the difference in means for the outcome variable between the two groups being compared and the standard deviation of the outcome variable (i.e., Cohen’s d). Effect sizes near 0.2 are considered a small effect; 0.5, a medium effect; and 0.8, a large effect.^{iv} Specifically, we will consider the minimum detectable effect size (MDES), which is the value of Cohen’s d that can be detected between subgroups with 80% power using an $\alpha = 0.05$ level two-sided test. The MDES for most of our analyses will represent a small to medium difference. For example, in comparisons between low- and high-performing HHAs, small and large HHAs, and HHVBP participating and non-participating HHAs, the MDES is estimated to be 0.28, 0.23, and 0.21, respectively.

Sensitivity of Results to Different Response Rate Assumptions. We performed additional power calculations to assess how a lower response rate on the standardized survey might impact our ability to examine differences between subgroups. In the computations presented in Table 3, the MDES when comparing small versus large HHAs is 0.235 with a 44% response rate and 0.312 with a 25% response rate. As the calculations below indicate, if the response rate were lower than expected, our power for detecting smaller differences between subgroups would be reduced; however, we would still have adequate power to detect medium-size differences with a 25% response rate.

We illustrate the impact of the response rate in these calculations using a hypothetical survey question: Has your home health agency implemented electronic tools to support frontline clinical staff, such as clinical decision support, condition-specific electronic alerts, or automated prompts? Based on a 44% response rate, if 90% of large HHAs have electronic tools, we would be able to detect an 8.2 percentage point difference (MDES of 0.235) between large and small HHAs (e.g., 90% for large versus 81.8% for small HHAs). We would not have sufficient power to detect smaller differences (e.g., the 5-percentage-point difference that would result if 85% of low-performing HHAs were using electronic tools). If the response rate were 25%, we would be able to detect a 11.2 percentage point difference (e.g., MDES of 0.312; 90% versus 78.8%).

Table 3. Power Calculations for Comparison of Large HHAs With Small HHAs

Response Rate	Power for Detecting Small Effect Sizes (0.235)	Power for Detecting Medium Effect Sizes (0.4)	MDES*
.25	0.558	0.948	0.312
0.35	0.704	0.989	0.264
0.44	0.800	0.997	0.235

* MDES, minimum detectable effect size, as computed using Cohen’s d. MDES calculation assumes 80% power.

^{iv} Cohen J. *Statistical Power Analysis for the Behavioral Sciences, 2nd Edition*. Hillsdale: Lawrence Erlbaum; 1988.

**Impact Assessment of CMS
Quality and Efficiency Measures**

In addition, we also performed additional power calculations to assess the ability of the standardized survey to detect differences between groups not represented in the stratified design. For example, home health agencies outside of urban areas may face greater challenges in care delivery and initiating quality improvement activities than agencies located in urban areas. Based on the current sampling design, the survey is expected to yield approximately 89 responses from home health agencies located in small towns or rural areas. We estimate that the current sampling design will have 80% power for detecting medium-sized differences between urban and rural/small-town home health agencies.

Consideration of Alternative Sampling Strategies for the Structured Survey. We considered alternative sampling strategies. For example, we considered drawing a simple random sample that would yield 1,000 respondents from the entire population (this is equivalent to sampling from each stratum at a rate that is proportional to the size of the stratum). However, such a strategy would lead to a (multiplicative) 45.3% increase in the standard error of estimates calculated across the subpopulation of large HHAs over the standard errors from our preferred strategy because the survey would include relatively few large HHAs without oversampling. Although proportional sampling will yield standard error for national estimators that are 9.1% smaller than those yielded by our preferred strategy, the standard errors for national estimators are already quite small, and any gains from reducing the standard errors further will be outweighed by the 45.3% increase in the standard errors of subgroup estimates.

Sampling Design for Qualitative Interviews. In the qualitative interview portion of the data collection, we will use purposive sampling to interview 40 HHA quality leaders across six sample strata defined by size and HHVBP enrollment as opposed to the 24 sample strata for the standardized survey. We opted not to use 24 sample strata because scheduling interviews with so many strata (and relatively few participants per stratum) would have been infeasible. (Typically, interviews are scheduled in batches to ensure that the project is conducted expeditiously; having a goal of one or two HHAs per stratum could cause an excessive number of interviews to be conducted.) The HHAs completing the qualitative interview will not be the same as those completing the standardized survey; they are distinct samples. Sampling 40 HHAs across six strata will result in as few as six and as many as seven interviews per stratum. This distribution is outlined in Table 4. The goal is to get broad representation of HHAs by size and HHVBP participation.

Because these data are qualitative, the goal is not to generalize to the larger population. Rather, we aim to conduct a sufficient number of interviews per stratum to complement the quantitative data collected in the standardized survey and to provide qualitative details that can help partially explain the standardized survey results. We will release sufficient sample (50–100 HHAs per stratum) to recruit the target number of completed interviews per stratum.

Table 4: Sample allocation (n=40 total) by strata for qualitative interviews

	Small (1-100 episodes)	Medium (101-1,000 episodes)	Large (1000+ episodes)
HHVBP participant	6 completed interviews	7 completed interviews	6 completed interviews
HHVBP non-participants	7 completed interviews	7 completed interviews	7 completed interviews

*Categories derived from CMS Home Health Compare database and OASIS for calendar year 2015; data included location and number of OASIS assessments submitted. A home health care episode is defined as the entire patient care episode on single claim, which may encompass multiple home care visits.

Questionnaire Content and Design Process. The content of the survey was driven by the research question “What changes are HHAs making in response to the use of performance

**Impact Assessment of CMS
Quality and Efficiency Measures**

measures by CMS?” This overarching question was translated into five specific research questions that form the content of the surveys and interviews:

1. What types of quality improvement (QI) changes have HHAs made to improve their performance on CMS measures?
2. If a QI change was made, has it helped the HHA improve its performance on one or more CMS measures?
3. What challenges or barriers do HHAs face in reporting CMS quality measures?
4. What challenges or barriers do HHAs face in improving performance on the CMS quality measures?
5. What unintended consequences do HHAs report associated with implementation of CMS quality measures?

Attachment I details the process used to develop and test the qualitative interview guide and survey instrument. The content for the survey was informed by an environmental scan of the literature related to the five research questions, discussions with key HHA stakeholders at CMS, formative interviews with HHAs, and testing of draft survey instruments with HHAs. The formative interviews with HHAs were important for defining the structure of the survey and for identifying topics that would be more conducive to standardized questions versus questions that are open-ended in nature, as well as topics not covered in the environmental scan.

The goals of the formative interview work were to explore the following topics:

- How the CMS performance measures are changing the way in which HHAs are delivering care
- Factors that drive HHAs investments in performance improvement
- Issues HHAs face related to reporting the CMS measures
- Challenges HHAs face related to improvement on the CMS measures
- Potential unintended consequences associated with implementing the CMS quality measures.

By broadly exploring these topics, we were able to develop survey questions that addressed the research questions.

The survey development team considered including a “don’t know” option for all questions but was concerned about the effects of missing data, as respondents often default to the “don’t know” option rather than finding the answer within their organization. In cognitive tests of the instrument, respondents did not generally state that they did not know the answers to various questions.

There is also potential concern about positive response bias when fielding surveys, especially regarding reporting unintended consequences. However, in our formative and cognitive testing work (and in the CMS hospital and nursing home surveys), respondents demonstrated variation in how they responded and were willing to report negative things, such as diversion of resources to quality measurement. During interviews, they expressed frustration with the measurement programs and having to collect and report the data. They also described challenges with being able to improve their performance as well as undesired behaviors. Therefore, we do not believe the surveys as designed will lead to positive response bias among respondents.

Plan for Tabulating the Results. The analysis plan will include (1) development of survey weights, (2) response rate/nonresponse analyses, (3) psychometric evaluation of survey items,

(4) development of national and subgroup estimates (such as by level of performance and size of HHA where possible), and (5) analyses of the association between HHA performance (high/low) and HHA responses to QI changes they made in response to CMS measures, adjusting for patient and facility characteristics. All aspects of these analyses will be described in a final project report to CMS and/or in other publicly available reports, as well as in journal publications.

- (1) **Weighting.** We will consider three types of weights so that our analysis of survey responses will appropriately reflect the target populations of interest: design weights, nonresponse weights, and (final) survey weights. Design weights reflect the probability that each HHA is selected for the survey; nonresponse weights reflect the probability that a sampled HHA responds to the survey; the final weights make the respondent sample's characteristics similar to those of the population. Design weights are readily calculated as the ratio of eligible to sampled HHAs in particular strata (given the proposed stratified sampling design). Complex HHA-level nonresponse weights may be developed using logistic regression or more complex methods if needed. Further, final survey weights may be post-stratified using raking if deemed necessary.
- (2) **Response rate/nonresponse analyses.** We will examine response rates overall and within particular strata, including by HHVBP participation, by HHA size (number of episodes), and by quality performance. We will use logistic regression to examine the associations between known HHA characteristics and probability of nonresponse. HHA characteristics to be included in this analysis are size (e.g., number of patient episodes), HHVBP participation, quality, for-profit/nonprofit status, urban/rural, region, and clinical characteristics of the patient population.
- (3) **Psychometric evaluation of survey items.** We will evaluate missing data, item distribution (including ceiling and floor effects), internal consistency (e.g., Cronbach's alpha, which will be used to assess consistency of composites of adoption rates across several types of QI changes), and reliability. We will compute these statistics overall and by stratum.
- (4) **Subgroup estimates.** We will produce national and subgroup estimates with appropriate adjustment to account for sampling design and nonresponse. Subgroups of interest may include performance strata (low, medium, high), HHA size, patient co-morbidity, and urban/rural. The final list will be determined in consultation with CMS.
- (5) **Relationship between survey response patterns and HHA characteristics.** We will provide descriptive analyses of survey findings overall and stratified by HHA characteristics. The descriptive statistics will include the mean and median response, variation in responses, and skewness of responses by item. We will use linear and logistic regressions to examine the association between survey responses and HHA characteristics, including HHA performance, HHVBP participation, size, and region.
- (6) **Univariate and multivariate analyses.** We will conduct two main analyses. First, we will use univariate analyses within performance strata to examine associations between performance and HHA characteristics, including those obtained from the survey and from administrative data sources, such as practice size and location/region. Second, we will consider conducting multivariate regression analyses to examine associations between performance and QI changes, adjusting for potential confounding factors identified in the initial univariate analyses.

2. Procedures for Collecting Information

The first step in fielding both the qualitative interviews and the standardized survey will be to identify the most appropriate respondent. We refer to this individual as the quality leader for the HHA—that is, the individual within the organization who is most familiar with the CMS performance measures and understands the actions and quality improvement activities the organization has undertaken to improve performance on the measures. Once the project team has drawn the sample, the team will contact each HHA to identify the quality leader so that we can target the survey or interview correctly. We believe that specific identification of the appropriate respondent contributed to higher-than-expected response rates for the hospital and nursing home surveys that were conducted for the 2018 Impact Assessment Report.

The quality leader for the HHA often carries the title of administrator or chief nursing officer; however, we did not identify the HHA leader using a specific title because titles may vary between facilities. We successfully used this strategy for identifying quality leaders during formative interviewing and cognitive testing. During the interviews, these individuals demonstrated that they possessed the knowledge necessary to address the questions on the survey. The types of responses we obtained in survey development were comparable across HHAs, and the individuals did not demonstrate problems providing answers to the questions on the survey. (See Attachment I, which summarizes findings from the formative and cognitive interviews).

Qualitative Interviews. The project team will contact each HHA by telephone to confirm the mailing address and telephone number and to identify the quality leader. The respondent's name, job title, and email address will be collected at this time. Using this information, the project team will send the quality leader a letter via email that describes the study and interview and invites the quality leader or a designee to take part in the interview (Attachments IV and IX). The project team's data collection staff will follow up by telephone 3 to 5 days later to confirm interest and availability to participate in the interview. To minimize non-response bias, the project team will make up to 10 attempts, both by telephone and email, to contact the quality leaders to encourage them to participate in the interview. The project team will schedule interviews at a convenient date and time for the quality leader and will work with an administrative assistant as necessary to schedule appointments. In previous survey work, this protocol was effective at reducing non-response.^v

Standardized Survey. The project team's staff will contact each sampled HHA by telephone to confirm the mailing address and telephone number and to identify the quality leader. The respondent's name, job title, and email address will be collected to allow survey invitations to be personalized. To promote the likelihood of survey participation, the survey will use a multi-mode data collection. The team will use Web and mail as data collection or prompting modes. To allow adequate time for each mode and for U.S. Postal Service delivery of mail survey returns, the project team plans a field period of approximately 12 weeks.

^v Centers for Medicare & Medicaid Services. 2018 *National Impact Assessment of the Centers for Medicare & Medicaid Services (CMS) Quality Measures Report*. Baltimore, MD: US Department of Health and Human Services; 2018. Available at: <https://www.cms.gov/Medicare/Quality-Initiatives-Patient-Assessment-Instruments/QualityMeasures/National-Impact-Assessment-of-the-Centers-for-Medicare-and-Medicaid-Services-CMS-Quality-Measures-Reports.html>.

As recommended by Dillman et al (2009) as a best practice, we propose to contact non-responders using varying modes of data collection.^{vi}

Weeks 1–3 – Initial and follow-up email invitations to complete the survey by Web.

All home health quality leaders identified as respondents will receive a minimum of two invitations to participate in the survey via the Web. These invitations will be sent by email one week apart and will contain sufficient information for informed consent, as well as an HHA-specific personal identification number (PIN) code to access the Web survey. If no email address is available, the invitations will be sent via first class mail. Examples of emails inviting HHA quality leaders to participate in the Web are found in Attachments V: Web Survey Invitation Email and VI: Reminder Email.

Week 3–5 – Mail survey is sent to all non-responding quality leaders. To reduce non-response rates, two weeks after the initial invitation to the Web survey, non-responding HHA quality leaders will receive a paper version of the survey via first class mail (Attachments II: HHA Survey Instrument, VII: First Mail Survey Cover Letter, and X: List of Quality Measures). The project team will send a reminder letter for the paper survey (with another paper version of the survey two weeks later (Attachments II, X, and VIII: Mail Survey Reminder Letter). Both mail survey invitation letters will include instructions for completing the Web survey as well.

Weeks 7–10 – Commence telephone calls to non-responding quality leaders to prompt return of the mail survey or completion of the Web survey. Seven weeks after the initial invitation to the Web survey, non-responding HHA quality leaders will be contacted by telephone to prompt completion of the survey via the Web or by faxing the mailed survey. To minimize data collection costs related to engaging large numbers of HHA quality leaders by telephone, we will initially contact non-responders by email or by mail and reserve the more expensive telephone outreach until later in the data collection period, when the number of non-responders is smaller.

Week 11 – Non-responding HHAs will receive a final email invitation to the Web survey two weeks prior to the close of data collection.

Week 12 – Data collection closes at the end of week 12.

Throughout data collection, we will track response and cooperation within each sample stratum and use additional efforts or sample to achieve sufficient response rates in each stratum. We anticipate that the procedures outlined above and the goal of 1,000 completed surveys will result in a response rate of 40% to 50%.

3. Methods to Maximize Response Rates and Deal with Non-Response

Qualitative Interviews. For home health quality leaders who are willing to participate, we will make up to 10 attempts to schedule an appointment, both by telephone and via email. We will work with an administrative assistant to schedule the interview at a day and time most convenient for the quality leader. In previous survey work, we have found this protocol to be effective at minimizing non-response bias. In addition, 3 to 5 days after the invitation letter is

^{vi} Dillman DA, Smyth JD, Christian LM. (2009). *Internet, mail, and mixed-mode surveys: The tailored design method*. Hoboken, NJ: Wiley & Sons.

Impact Assessment of CMS Quality and Efficiency Measures

emailed, data collection staff will follow up by telephone to confirm availability and interest in participating in the interview. The quality leader may designate another individual within the organization to participate in the interview, which may further maximize participation. Quality leaders who refuse to participate in the interview or who fail to respond to the invitations will be replaced with a quality leader from another HHA with the same characteristics. During the formative development work, we generally found HHAs willing to participate in the interviews: They wanted to share their experiences with the CMS measures and explain what they are doing to improve their performance on these measures.

Structured Survey. Published surveys of nursing leaders (administrators and directors of nursing) conducted in the past 10 years report response rates of 48% to 58%.^{vii,viii,ix} In addition, surveys of organizations and/or individuals in leadership roles have experienced an overall decline in response rates similar to surveys of general populations.^{x,xi} We used these studies and the previous experience of the survey development team in conducting interviews and surveys with HHAs to arrive at our estimate of a 44% response rate.

As described in **Section 2** above, we plan to maximize response rates for the standardized survey through:

- Careful identification of the appropriate respondent.
- Use of personalization.
- Multiple attempts.
- Multiple modes of survey administration.
- Alternative modes for non-response contacts.

We anticipate that the data collection procedures for the structured survey will result in a response rate of 44% and achieve 1,000 completed surveys. We will track both facility characteristics and titles of HHA quality leaders among non-responding HHAs to better adjust for non-response in analyses of results, to examine possible response bias, and to describe the characteristics of non-responders.

4. Tests of Procedures or Methods to Be Undertaken

We developed and tested the data collection protocol and draft qualitative interview guide and draft standardized survey with a small number of HHA quality leaders (please refer to Attachment I of the clearance package). Findings from the formative interviews and cognitive testing helped to determine the structure of the qualitative interview protocol and the standardized survey and the approach needed to identify the appropriate respondent(s) to the survey in the HHA.

^{vii} Castle N. Measuring staff turnover in home health agencies. *The Gerontologist*. 2006;46(2):210–219.

^{viii} Mukamel DB, Spector WD, Zinn JS, Huang L, Weimer DL, Dozier A. Nursing homes' response to the Home Health Compare report card. *Journals of Gerontology Series B: Psychological Sciences and Social Sciences*. 2007;62(4):S218–S225.

^{ix} Centers for Medicare & Medicaid Services. 2018 National Impact Assessment of the Centers for Medicare & Medicaid Services (CMS) Quality Measures Report. Baltimore, MD: US Department of Health and Human Services; 2018. Available at: <https://www.cms.gov/Medicare/Quality-Initiatives-Patient-Assessment-Instruments/QualityMeasures/National-Impact-Assessment-of-the-Centers-for-Medicare-and-Medicaid-Services-CMS-Quality-Measures-Reports.html>.

^x Cocyota CS, Harrison DA. What (not) to expect when surveying executives: a meta-analysis of top manager response rates and techniques over time. *Organizational Research Methods*. 2006;9:133–160.

^{xi} Baruch Y, Holton BC. Survey response rate levels and trends in organizational research. *Human Relations*. 2008;61:1139–1160.

**Impact Assessment of CMS
Quality and Efficiency Measures**

Formative interviews were used to guide the development of the structured survey and qualitative interview protocol. Nine HHAs participated in the formative interviews, which were conducted by telephone. The formative interviews with HHAs were designed to:

- Assess whether respondents could understand the information we sought to collect to address the five research questions.
- Assess whether respondents would provide biased (i.e., only favorable) responses with regard to CMS programs or their actions taken in response to performance measurement.
- Explore the language potential respondents might use to describe the topics.
- Identify potential response options or areas to probe.

HHAs included in the formative interviews were purposively sampled to represent variation in the size of the provider entity, the region of the country and location (urban versus rural) of the provider, and performance on CMS measures. Additionally, HHAs were sampled based on whether they had a relationship with a hospital (e.g., hospital-based HHAs). The individuals interviewed were senior leaders who were responsible for the overall quality and safety of clinical care within the HHAs. Interviewees were asked to provide feedback on lessons learned related to the use of the performance measures and on any other concerns not covered in the qualitative interview guide.

The draft standardized survey was tested with nine HHAs via cognitive interviews conducted by telephone. A range of HHAs (size, quality performance, and region) were selected for the cognitive interviews to capture variation in the expected range of responses. The cognitive interviews were designed to assess respondents' understanding of the draft survey items and key concepts and to identify problematic terms, items, or response options. During this time, the draft instruments were also reviewed by the RAND Corporation and Health Services Advisory Group, Inc. (HSAG) project teams and CMS staff. The draft survey was revised based on findings from the cognitive interviews and feedback received from the various reviewers to produce the final version of the home health survey to be used in 2019–2020.

5. Individuals Consulted on Statistical Aspects and Individuals Collecting and/or Analyzing Data

The survey, sampling approach, and data collection procedures were designed by the RAND Corporation under contract to HSAG under the leadership of:

Cheryl Damberg, PhD
RAND Corporation
1776 Main Street
Santa Monica, CA 90407

Kanaka Shetty, MD

Key input to the design of the data collection was received from the following individuals:

- Cheryl Damberg, RAND Project Director
- Kanaka Shetty, RAND Co-Project Director
- Michael Robbins, Lead Statistician
- Matthew Cefalu, Statistician
- Deborah Kim, RAND Survey Director