Medicare Health Outcomes Survey Field Test

Supporting Statement B

Table of Contents

List of Tables

Supporting Statement B

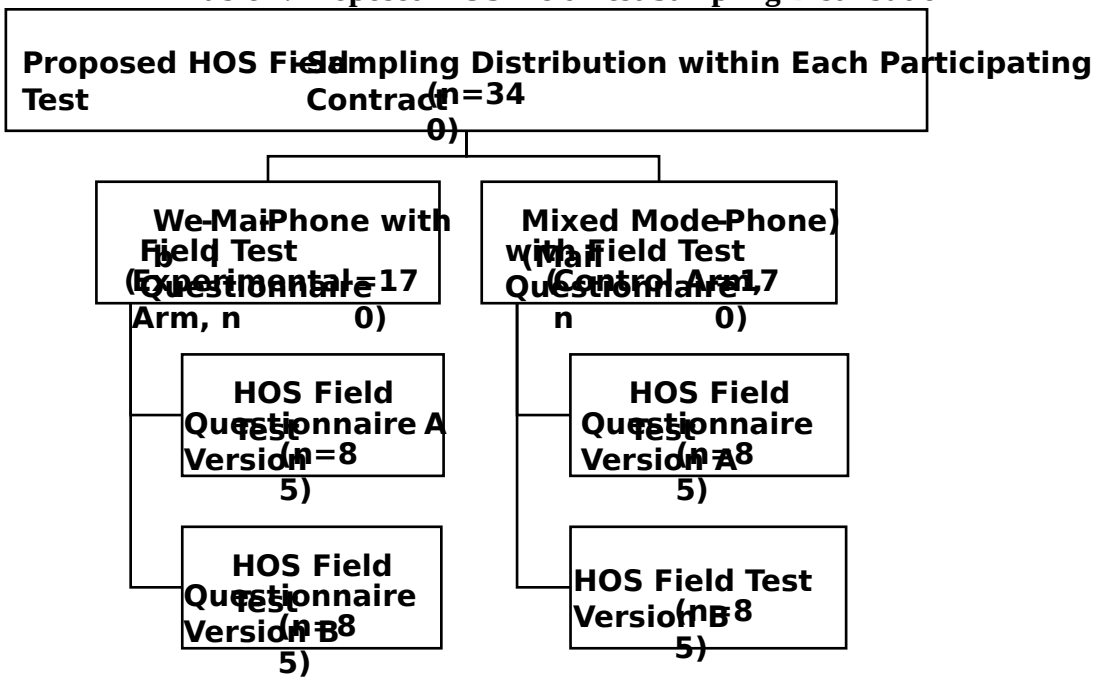
Statistical Methodology

1. Respondent Universe and Sample

The Medicare Health Outcomes Survey (HOS) is administered annually to enrollees in Medicare Advantage (MA) contracts with at least 500 members. The survey is administered to adult enrollees in the U.S. All eligible members are surveyed from MA contracts with 500-1,200 members; 1,200 members are randomly sampled from larger MA contracts. The field test sample will be drawn from the remaining eligible sample frame used to conduct the annual fielding of the HOS. This design proposes a stratified random sample of 170 enrollees for each of two study arms (experimental and control) from each of 50 MA contracts, for 8,500 enrollees total per study arm. Stratified random sampling based on contract-level characteristics (e.g., geographic region, enrollment size, level of quality measurement performance) will be used to identify 100 eligible contracts, with the goal of enrolling 50 MA contracts to participate in the field test. Contracts with < 2,000 eligible enrollees will be excluded from the study. The selection of contracts aims to create a sample that represents the diversity of the Medicare Advantage population.

Within each study arm, there will also be two versions of the field test questionnaire that will be identical except for slight differences in the wording of selected items. Three items (Question 6a-c) will have different response options between the two versions of the field test questionnaire, and three items (Question 2a, Question 8, and Question 12) will have slight differences to the item stem between the two questionnaires (see Attachment C). The goal of using two versions of the questionnaire is to estimate the differences in scores that result from the alternate item stem or response options. The two versions of the questionnaire will be administered within each contract to enrollees randomized to each questionnaire (i.e., a split-half approach where 85 enrollees will receive each questionnaire within each study arm). Refer to Table 1 for an example of the sampling distribution into each study arm and version of the questionnaire.

Table 1. Proposed HOS Field Test Sampling Distribution



2. Information Collection Procedures

Sampled enrollees will be randomly assigned into one of two study arms: an experimental arm that will receive the field test questionnaire and include the web mode and a control arm that will receive the field test questionnaire and use typical HOS survey administration procedures previously approved under OMB control number 0938-0701. Within each study arm, enrollees will be further randomized to a version of the field test questionnaire.

In the experimental arm (i.e., the field test questionnaires with web, mail, and telephone modes), all sampled enrollees will receive a mailed pre-notification letter in advance of survey administration. The letter will be personalized to the enrollee and will include the URL for the web version of the survey and a PIN code that is unique to the enrollee, as well as a QR code. The enrollee may enter the URL and PIN code to access the web version of the survey, which can be completed on either a computer or a mobile device such as a smartphone or tablet. Four days after the pre-notification letter is mailed, enrollees will be sent an invitation to the web survey. The invitation will be sent by email to enrollees with email addresses, and via a letter to all sampled enrollees, including those with email addresses and for whom an email address is not available.[1] A second

1 Sending an invitation by letter to those without an email address is consistent with other CMS Consumer Assessment of Health Providers and Systems (CAHPS) efforts, such as Outpatient and Ambulatory Surgery (OAS) CAHPS.

invitation email will be sent five days after the initial invitation to enrollees with email addresses. Ten days later, all enrollees who have not completed the web survey will receive the survey by mail. Enrollees who do not respond to the first mail survey will receive a second mail survey approximately three weeks later. Approximately three weeks after a second mail survey is sent, telephone administration of the survey will be attempted with all non-respondent enrollees. Mail non-respondents will receive up to five call attempts to complete the survey by telephone.

In the control arm (i.e., the HOS field test survey questionnaire with mail and telephone modes), all sampled enrollees will receive a mailed pre-notification letter, followed seven days later by the mailing of the first questionnaire. A second questionnaire will be mailed to nonrespondents approximately five weeks after the initial survey mailing. Telephone follow-up will be conducted beginning about three weeks after the mailing of the second questionnaire among non-respondents to the mail portion of the survey. Mail non-respondents will receive up to five call attempts to complete the survey by telephone. Refer to Table 2 for a comparison of the timing of fielding activities for each study arm.

Table 2. Proposed HOS Field Test Schedule of Contacts – Two-Arm Study

Survey Admin. Day	Web-Mail-Phone with Field Test Questionnaires (Experimental Arm)	Mixed Mode (Mail-Phone) with Field Test Questionnaires (Control Arm)
0	Mail pre-notification letter	Mail pre-notification letter
5	Email and mail 1st invitation	N/A
7	N/A	Mail 1st survey
10	Email 2nd invitation	N/A
15	Mail 1st survey	N/A
36	Mail 2nd survey	N/A
42	N/A	Mail 2nd survey
57	Begin phone calls	N/A
63	N/A	Begin phone calls
107	End data collection	End data collection

3. Methods to Maximize Response Rates

As described above, the data collection effort for the field test control arm will use a mixed-mode data collection protocol that includes a pre-notification letter alerting sampled members that a survey will be mailed to them shortly, a first mailing of the questionnaire,

followed by a second mailing to those who do not respond to the earlier mailing of the questionnaire. For those who also do not respond to the second mailing of the questionnaire, HOS employs a telephone follow-up which offers sampled members the opportunity to complete the survey by telephone. Mailing materials sent to all sampled members also include a toll-free telephone number that allows recipients to call in and complete the survey by telephone or ask questions about the survey. Overall, this approach has resulted in baseline response rates of between 28-47 percent on average over the last nine years of national data collection in HOS, varying somewhat by contract and region of the country (Table 3).[2]

Additionally, this field test will test the addition of a web mode to the HOS administration procedures, as described above. Recent research demonstrates the potential to improve response rates with the use of a web-based mode. In particular, in the emergency department setting, a CMS-funded test demonstrated that a higher survey response rate could be achieved by using a web survey as the initial mode of administration, followed by mail, telephone, or both.[3] Based on this research, CMS believes adding web to the existing mixed mode HOS protocol has the potential to improve response rates. In 2023, the HOS baseline response rate for MA enrollees was 32.0%.

Table 3. Historical Baseline HOS Response Rates, 2015-2023

Year	HOS Response Rate
2023	32.0%
2022	27.6%
2021	31.2%
2020	37.3%

Year	HOS Response Rate
2023	32.0%
2022	27.6%
2019	39.3%
2018	40.9%
2017	43.0%

2 Centers for Medicare & Medicaid Services, CMS-01203, Medicare Health Outcomes Survey Attachment D. Response Rates. (2025). Available at https://www.cms.gov/regulations-andguidance/legislation/paperworkreductionactof1995/pra-listing-items/cms-10203. Accessed on April 30, 2025.

3 Parast L., M. Mathews, et al. (2019). Effects of Push-to-Web Mixed Mode Approaches on Survey Response Rates: Evidence from a Randomized Experiment in Emergency Departments. *Survey Practice*, 12(2). https://doi.org/10.29115/SP-2019-0008.

2016	45.0%
2015	46.7%

4. Tests of Procedures or Methods

The proposed field test will test a new web-based mode of survey administration, along with proposed new HOS content recommended by a working group of researchers and experts in the field. We propose to administer this field test with an unused sample from the same MA population that is eligible to receive the annual HOS survey.[4] Sampled enrollees randomized to the experimental arm of the field test will be surveyed using a web-first protocol in which web is the initial mode of survey administration. Those who do not respond by web will receive up to two mailings of a paper survey; those who do not respond by web or mail will receive up to five call attempts to complete the survey by telephone. The control arm of the field test will use the standard HOS Survey protocol (i.e., mail, then telephone). Sampled enrollees randomized to this arm of the field test will receive up to two mailings of a paper survey; those who do not respond by mail will be contacted by telephone to attempt to complete the survey. Staff conducting telephone interviews will be debriefed to identify items that may have been frequently misunderstood by respondents (e.g., needed to be repeated, required probing, and/or resulted in requests for clarification).

The analysis of the new web-based mode of survey administration will include:

- Review of response rates overall, by mode, and within each survey arm.
- Non-response analysis adjusting for survey arm, mode, contract, and administrative characteristics of the sampled population.

- Representativeness of the population by a review of the respondent characteristics overall and by survey arm.
- Differences in patterns of response for items between survey arms. Chi-squared tests of significance for item frequencies and linear regression adjusting for case-mix, survey arm, and contracts will be examined.

The analysis of new survey content will include:

4 MAOs with at least 340 eligible members remaining after the HOS sample for national administration and any requested voluntary HOS oversample are fulfilled will be eligible for inclusion in the field test.

- Review of item frequencies, patterns of response, and item-level missingness to assess response to the new items compared to existing survey items from the national administration of the HOS.
- Representativeness of the population responding to the new items will be assessed by a review of the respondent characteristics overall and by mode of survey administration.
- Psychometric analyses, such as measurement of intraclass correlation coefficients, interunit reliability, and Cronbach's alpha to assess internal consistency reliability will be used to examine the performance of new content at the item and composite levels.

The analysis of wording updates to existing survey items will compare respondents from the control arm of the field test using standard mail and phone protocol to the national HOS implementation data for the same contracts and include:

- Review of item frequencies, the presence of ceiling or floor effects, and item-level missingness to assess response to the updated wording items compared to corresponding items in the existing HOS 3.0 questionnaire.
- Case-mix adjusted linear regression models with controls for arm and contract will assess differences in patterns of response associated with updated wording measures.
- Psychometric properties including intraclass correlation coefficients, interunit reliability, and Cronbach's alpha to assess internal consistency reliability will assess measure performance under the new wording compared to measures derived from the existing HOS 3.0 questionnaire.

To be successful, a new survey item must support reliable and valid measurement of experience with an MA contract.

Additional detail regarding statistical design modifications is presented below.

5. Statistical and Questionnaire Design Consultants

The statistical expertise for this field test design effort was coordinated by the Health Services Advisory Group (HSAG) and included statistical consultation provided by research consultants with expertise in survey methodology, case-mix adjustment, quality measurement, development of patient-reported outcome measures (PROMs), patient experience with health care, and gerontology.

The proposed design is of sufficient scale to support comparison of the effect of mode of administration on responses and response rate. The proposed design will have 80% power with a 2-sided test at alpha=0.05 to detect an effect size of 0.09 standard deviations (0.10 is "very small" by standard heuristics) between the experimental and comparison arms. The proposed design will have 80% power to detect a 2.0% difference in response rates between the two arms using a 2-sided test and alpha=0.05. Additionally, an analysis of version effects between the two questionnaires would pool data by questionnaire version across the two study arms. This approach will have the same sample sizes in each group as the comparison of the effect of mode of administration. The proposed design will have 80% power with a 2-sided test at alpha=0.05 to detect an effect size of 0.09 standard deviations between the two versions of the questionnaire.

The analysis will employ fixed effects for contracts to control for contract differences in performance, an approach used previously to assess changes in Medicare Advantage and Prescription Drug Plan (MA & PDP) CAHPS survey content.[5] Analyses will be restricted to data from the 50 contracts identified for the field test and the national administration of the HOS. If overall differences are detected using the initial model, additional models using random-effect contract intercepts and slopes will test whether changes are uniform or vary by contract.

Differences in mean responses between the experimental and the control arms will inform CMS decision-making about the inclusion of new survey items, as they will indicate survey mode and version effects on scores, respectively.

Under this design, it is proposed to strongly recommend the use of the web-first mode if the response rate for the experimental arm is significantly higher than the control arm; to recommend further study if it is not significantly different from the control arm; and to recommend against it if the experimental arm has a response rate significantly lower than the control arm.

5 Beckett, M. K., Elliott, M. N., Burkhart, Q., Cleary, P. D., Orr, N., Brown, J. A., Gaillot, S., Liu, K., & Hays, R. D. (2019). The effects of survey version on patient experience scores and plan rankings. *Health services research*, *54*(5), 1016–1022. https://doi.org/10.1111/1475-6773.13172.